ECONSTOR
Make Your Publications Visible.

A Service of

ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Dimitriadis, Timo; Patton, Andrew J.; Schmidt, Patrick W.

Working Paper

Testing forecast rationality for measures of central tendency

Hohenheim Discussion Papers in Business, Economics and Social Sciences, No. 12-2020

Provided in Cooperation with:
Faculty of Business, Economics and Social Sciences, University of Hohenheim

Suggested Citation: Dimitriadis, Timo; Patton, Andrew J.; Schmidt, Patrick W. (2020) : Testing forecast rationality for measures of central tendency, Hohenheim Discussion Papers in Business, Economics and Social Sciences, No. 12-2020, Universität Hohenheim, Fakultät Wirtschafts- und Sozialwissenschaften, Stuttgart, https://nbn-resolving.de/urn:nbn:de:bsz:100-opus-18143

This Version is available at:
https://hdl.handle.net/10419/225630

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

UNIVERSITY OF HOHENHEIM

Hohenheim Discussion Papers in Business, Economics and Social Sciences

TESTING FORECAST RATIONALITY FOR MEASURES OF CENTRAL TENDENCY

Timo Dimitriadis
University of Hohenheim/Heidelberg Institute for Theoretical Studies

Andrew J. Patton
Duke University

Patrick W. Schmidt
University of Zurich

Institute of Economics

12-2020

wiso.uni-hohenheim.de

Discussion Paper 12-2020

Testing Forecast Rationality for Measures of Central Tendency

Timo Dimitriadis, Andrew J. Patton, Patrick W. Schmidt

Download this Discussion Paper from our homepage:

https://wiso.uni-hohenheim.de/papers

ISSN 2364-2084

Die Hohenheim Discussion Papers in Business, Economics and Social Sciences dienen der schnellen Verbreitung von Forschungsarbeiten der Fakultät Wirtschafts- und Sozialwissenschaften. Die Beiträge liegen in alleiniger Verantwortung der Autoren und stellen nicht notwendigerweise die Meinung der Fakultät Wirtschafts- und Sozialwissenschaften dar.

Hohenheim Discussion Papers in Business, Economics and Social Sciences are intended to make results of the Faculty of Business, Economics and Social Sciences research available to the public in order to encourage scientific discussion and suggestions for revisions. The authors are solely responsible for the contents which do not necessarily represent the opinion of the Faculty of Business, Economics and Social Sciences.

Testing Forecast Rationality for Measures of Central Tendency*

Timo Dimitriadis† Andrew J. Patton‡ Patrick W. Schmidt§

This Version: September 28, 2020
First Version: February 15, 2019

Abstract

Rational respondents to economic surveys may report as a point forecast any measure of the central tendency of their (possibly latent) predictive distribution, for example the mean, median, mode, or any convex combination thereof. We propose tests of forecast rationality when the measure of central tendency used by the respondent is unknown. We overcome an identification problem that arises when the measures of central tendency are equal or in a local neighborhood of each other, as is the case for (exactly or nearly) symmetric distributions. As a building block, we also present novel tests for the rationality of mode forecasts. We apply our tests to survey forecasts of individual income, Greenbook forecasts of U.S. GDP, and random walk forecasts for exchange rates. We find that the Greenbook and random walk forecasts are best rationalized as mean, or near-mean forecasts, while the income survey forecasts are best rationalized as mode forecasts.

Keywords: forecast evaluation, weak identification, survey forecasts, mode forecasts
J.E.L. Codes: C53, D84, E27

*We thank Jonas Brehmer, Tobias Fissler, Rafael Frongillo, Tilmann Gneiting, Arnaud Maurel, Adam Rosen, Melanie Schienle, Jörg Stoye, as well as seminar participants at Duke University, KIT Karlsruhe, HITS Heidelberg, Universität Hohenheim, Universität Heidelberg, CSS Workshop in Zürich, HKMEtrics Workshop in Mannheim, ISI World Statistics Congress in Kuala Lumpur, the 2019 Statistische Woche in Trier, the EC^2 2019 Conference in Oxford, and the 2020 Econometric Society World Congress. Code to replicate the empirical work in this paper in the form of the `R` package `fcrat` is available at https://github.com/Schmidtpk/fcrat.

†Heidelberg Institute for Theoretical Studies, Heidelberg, Germany and University of Hohenheim, Stuttgart, Germany, e-mail: timo.dimitriadis@h-its.org

‡Department of Economics, Duke University, Durham, USA, e-mail: andrew.patton@duke.edu

§University of Zurich, Zurich, Switzerland, e-mail: patrickwolfgang.schmidt@uzh.ch

1 Introduction

Economic surveys are a rich source of information about future economic conditions, yet most economic surveys are vague about the specific statistical quantity the respondent should report. For example, the New York Federal Reserve's labor market survey asks respondents "What do you believe your annual earnings will be in four months?" A reasonable response to this question is the respondent reporting her mathematical expectation of future earnings, or her median, or her mode; all common measures of the central tendency of a distribution. When these measures coincide, as they do for symmetric unimodal distributions, this ambiguity does not affect the information content of the forecast. When these measures differ, the specific measure adopted by the respondent can influence its use in other applications, and testing rationality of forecasts becomes difficult.

We consider a class of central tendency measures defined by the set of convex combinations of the mean, median, and mode,[1] and we propose tests of forecast rationality that can be employed when the specific measure of central tendency used by the respondent is unknown. Similar to Elliott et al. (2005), we propose a testing framework that nests the mean as a special case, but unlike that approach we allow for alternative forecasts within the class of measures of central tendency, rather than measures that represent other aspects of the predictive distribution (such as non-central quantiles or expectiles). Our approach faces an identification problem: for symmetric distributions, the combination weight vector is unidentified, for "mildly" asymmetric distributions, the weight vector is only weakly identified, and even for strongly asymmetric distributions the weight vector may only be partially identified. Economic variables may fall into any of these cases, and a valid testing approach must accommodate these measures of central tendency being equal, unequal, or in a local neighborhood of each other. We use the work of Stock and Wright (2000) to obtain asymptotically valid confidence sets for the combination weights and to test forecast rationality.

Before implementing the above test for rationality for a general forecast of central tendency, we must first overcome a lack of rationality tests for mode forecasts.[2] Rationality tests for mean

[1] These are the three measures of central tendency described in standard introductory statistics textbooks, e.g. McClave et al. (2017). The Bank of England's quarterly Inflation Report, for example, reports these three measures as distinct point forecasts for future GDP growth and inflation. Our approach can be extended to consider a broader set of measures of centrality; we discuss this in Section 3.

[2] We use the phrase "mode forecasts," or similar, as shorthand for the forecaster reporting the mode of her predictive

forecasts go back to at least Mincer and Zarnowitz (1969), see Elliott and Timmermann (2016) for a recent survey, while rationality tests for quantile forecasts (nesting median forecasts as a special case) are considered in Christoffersen (1998) and Gaglianone et al. (2011). A critical impediment to similar tests for mode forecasts is that the mode is not an "elicitable functional" (Heinrich, 2014), meaning that it cannot be obtained as the solution to an expected loss minimization problem.[3] We obtain a test for mode forecast rationality by first proposing novel results on the *asymptotic elicitability* of the mode. We define a functional to be asymptotically elicitable if there exists a sequence of elicitable functionals that converges to the target functional. We consider the (elicitable) "generalized modal interval," defined in detail in Section 2.2, and show that it converges to the mode for a general class of probability distributions. We combine these results with recent work on mode regression (Kemp and Silva, 2012; Kemp et al., 2020) and nonparametric kernel methods to obtain mode forecast rationality tests analogous to well-known tests for mean and median forecasts.

We apply our proposed new tests in three important economic applications. We firstly analyze nearly 4,000 individual income survey responses from the Survey of Consumer Expectations conducted by the Federal Reserve Bank of New York. We find the intriguing result that we can reject rationality with respect to the mean or median, however we cannot reject rationality when interpreting these as mode forecasts, suggesting that survey participants report the anticipated *most likely* outcome rather than the average or median. Interestingly, only convex combinations very close to the mode are rationalizable; these survey respondents appear to report the mode or a functional very close to it. When allowing for cross-respondent heterogeneity, we find that forecasts from younger survey respondents with income below the median cannot be rationalized using *any* measure of central tendency, while forecasts from high-income respondents, regardless of their age, are rationalizable for many different measures of centrality.

We next analyze the Federal Reserve staff's "Greenbook" forecasts of quarterly U.S. GDP. Consistent with these forecasts being constructed using econometric models, which generally focus

distribution as her point forecast.

[3] Gneiting (2011) provides an overview of elicitability and identifiability of statistical functionals and shows that several important functionals such as variance, Expected Shortfall, and mode are not elicitable. Fissler and Ziegel (2016) introduce the concept of higher-order elicitability, which facilitates the elicitation of vector-valued (stacked) functionals such as the variance and Expected Shortfall, though not the mode (Dearborn and Frongillo, 2020).

on the mean, we find that we cannot reject rationality with respect to the mean, however we can reject with respect to the median and mode. Finally, we revisit the famous result of Meese and Rogoff (1983) that exchange rates are approximately unpredictable when evaluated by the squared-error loss function, i.e. the lagged exchange rate is an optimal mean forecast. For the JPY/EUR and AUD/EUR exchange rates we find evidence consistent with this: the lagged exchange rate is not rejected as a mean forecast, while it is rejected when taken as a mode or median forecast. For the USD/EUR exchange rate, on the other hand, we cannot reject rationality with respect to any convex combination of these measures of central tendency; the random walk forecast is consistent with rationality under any of these measures.

We evaluate the finite sample performance of the new mode rationality test and of the proposed method for obtaining confidence sets for measures of centrality through an extensive simulation study. We use cross-sectional and time-series data generating processes with a range of levels of asymmetry. We find that our proposed mode forecast rationality test has satisfactory size properties, even in small samples, and exhibits strong power across different misspecification designs. Our simulation design allows us to consider the four identification cases that can arise in practice: strongly identified (skewed data), where the mean, median and mode differ; unidentified (symmetric unimodal data), where all centrality measures coincide; weakly identified (mildly skewed data), where the centrality measures differ but are close to each other; and partially identified (skewed location-scale data), where one centrality measure is a convex combination of the other two. We find that in the symmetric case, the resulting confidence sets contain, correctly, the entire set of convex combinations of mean, median and mode. In the asymmetric cases, our rationality test is able to identify the combination weights corresponding to the issued centrality forecast.

Our paper is related to the large literature on forecasting under asymmetric loss, see Granger (1969), Christoffersen and Diebold (1997), Patton and Timmermann (2007) and Elliott et al. (2008) amongst others. The work in these papers is motivated by the fact that forecasters may wish to use a loss function other than the omnipresent squared-error loss function. The use of asymmetric loss functions generally leads to point forecasts that differ from the mean (though this is not always true, see Gneiting, 2011 and Patton, 2020), and generally these point forecasts are not interpretable as

measures of central tendency. For example, Christoffersen and Diebold (1997) show that the linex loss function implies an optimal point forecast that is a weighted sum of the mean and variance, while Elliott et al. (2008) find that their sample of macroeconomic forecasters report an expectile with asymmetry parameter around 0.4. Instead of moving from the mean to a point forecast that is not a measure of location, our novel approach considers moving only *within* the general set of location measures defined by convex combinations of the mean, median, and mode.[4]

Our paper is also related to experimental work on eliciting centrality measures from survey respondents. In an early study, Peterson and Miller (1964) found that respondents could accurately predict the mode and median, if incentivized correctly, but had difficulty reporting accurate estimates of the mean. Other work on eliciting centrality measures includes Hossain and Okui (2013) for the mean, Dufwenberg and Gneezy (2000) and Kirchkamp and Reiß (2011) for the median, and Charness and Dufwenberg (2006) and Sapienza et al. (2013) for the modal interval. Kröger and Pierrot (2019) find the mode is the measure most frequently adopted by experiment participants asked to summarize their predictive distribution as a point forecast.

The remainder of the paper is structured as follows. In Section 2 we propose new forecast rationality tests for the mode based on the concepts of asymptotic elicitability and identifiability. Section 3 presents forecast rationality tests for general measures of central tendency, allowing for weak and partial identification. Section 4 presents simulation results on the finite-sample properties of the proposed tests, and Section 5 presents empirical applications using survey forecasts of earnings, Greenbook forecasts of U.S. GDP growth, and the random-walk forecast for exchange rates. Proofs are presented in Appendix A and additional results are presented in the Supplemental Appendix.

2 Eliciting and Evaluating Mode Forecasts

2.1 General Forecast Rationality Tests

Let $Z_t = \left(Y_t, X_t, \tilde{\boldsymbol{h}}_t\right)$ be a stochastic process defined on a common probability space $\left(\Omega, \mathcal{F}, \mathbb{P}\right)$. Y_{t+1} denotes the (scalar) variable of interest, $\tilde{\boldsymbol{h}}_t$ denotes a vector of variables known to the forecaster at

[4] Bauer and Rudebusch (2016) suggest using the *difference* between the mean and the mode of the predictive distribution for future interest rates as a measure of the likelihood that interest rates will hit the zero lower bound.

the time she issues her point forecast for Y_{t+1}, which is denoted X_t. We define the information set $\mathcal{F}_t = \sigma\{Y_s, X_s, \tilde{\boldsymbol{h}}_s; s \leq t\}$ as the σ-field containing all information known to the forecaster at time t. We assume the econometrician observes an $\mathcal{F}_t$-measurable $(k \times 1)$ vector $\boldsymbol{h}_t$. We denote the distribution of Y_{t+1} given $\mathcal{F}_t$ by F_t, and with corresponding density f_t. Expectations are denoted $E_t[\cdot] = E[\cdot|\mathcal{F}_t]$. We use $\mathcal{P}$ to denote a class of distributions. All limits are taken as $T \to \infty$.

We start by considering rationality tests (also known as calibration tests; Nolde and Ziegel, 2017) for the mean, i.e. we assume that the forecasts X_t are one-step ahead *mean* forecasts for Y_{t+1}. We are interested in testing if these forecasts are rational, which would imply the null hypothesis:

$$\mathbb{H}_0 : X_t = \mathbb{E}[Y_{t+1}|\mathcal{F}_t] \ \forall \ t \text{ a.s.} \tag{2.1}$$

This is carried out by considering the "identification function," which for the mean is simply the difference between the forecast and the realized value, i.e., the forecast error:[5]

$$V_{\text{Mean}}\big(X_t, Y_{t+1}\big) = X_t - Y_{t+1} \equiv \varepsilon_t. \tag{2.2}$$

Under the above null hypothesis and subject to standard regularity conditions, it is straight forward to show that $T^{-1/2}\sum_{t=1}^{T} V_{\text{Mean}}\big(X_t, Y_{t+1}\big)\boldsymbol{h}_t \xrightarrow{d} \mathcal{N}\big(0, \Omega_{\text{Mean}}\big)$, and that

$$J_T = \frac{1}{T}\left(\sum_{t=1}^{T} V_{\text{Mean}}\big(X_t, Y_{t+1}\big)\boldsymbol{h}_t^\top\right)\widehat{\Omega}_{T,\text{Mean}}^{-1}\left(\sum_{t=1}^{T} V_{\text{Mean}}\big(X_t, Y_{t+1}\big)\boldsymbol{h}_t\right) \xrightarrow{d} \chi_k^2 \tag{2.3}$$

as $T \to \infty$, where $\widehat{\Omega}_{T,\text{Mean}}$ is a consistent estimator of Ω_{Mean}. This result facilitates testing whether given forecasts X_t are rational mean forecasts for the realizations Y_{t+1}. Specifically, the statistic in equation (2.3) can be used to test whether the identification function $V_{\text{Mean}}\big(X_t, Y_{t+1}\big)$ is correlated with the instrument vector $\boldsymbol{h}_t$. Under the null of forecast rationality, this correlation should be zero. As in most other tests in the literature, this is of course only a test of a necessary condition for forecast rationality, and the conclusion may be sensitive to the choice of instruments, $\boldsymbol{h}_t$.

[5] In econometrics the forecast error is usually defined as $Y_{t+1} - X_t$, that is, as the negative of the identification function in equation (2.2). Given the important role that forecast identification functions play in this paper, we adopt the definition for ε_t given in equation (2.2), and we refer to ε_t as the forecast error.

The test statistic in equation (2.3) is only informative about rationality if the forecasts are interpreted as being for the *mean* of Y_{t+1}. The decision-theoretic framework of identification functions and consistent loss functions is fundamental for generalizations to other measures of central tendency, such as the median and the mode. For a general real-valued functional $\Gamma : \mathcal{P} \to \mathbb{R}$, a strict identification function $V_\Gamma(Y, x)$ is defined by being zero in expectation if and only if x equals the functional $\Gamma(F)$. Strict identification functions are generally obtained as the derivatives of strictly consistent loss functions, which are defined as having the functional $\Gamma(F)$ as their unique minimizer (in expectation). A functional is called *identifiable* if a strict identification function exists, and is called *elicitable* if a strictly consistent loss function exists. See Gneiting (2011) for a general introduction to elicitability and identifiability.

The forecast error $X_t - Y_{t+1}$ is a strict identification function for the mean, and a strict identification function for the median is given by the step function

$$V_{\text{Med}}\big(X_t, Y_{t+1}\big) = \mathbb{1}_{\{Y_{t+1} > X_t\}} - \mathbb{1}_{\{Y_{t+1} \leq X_t\}}. \tag{2.4}$$

We obtain a test of median forecast rationality by replacing V_{Mean} and $\widehat{\Omega}_{T,\text{Mean}}$ by V_{Med} and $\widehat{\Omega}_{T,\text{Med}}$ in equation (2.3).

2.2 The Mode Functional

In contrast to the mean and the median, rationality tests for mode forecasts are more challenging to consider. The underlying reason is that there do not exist identification functions for the mode for random variables with continuous Lebesgue densities (Heinrich, 2014; Dearborn and Frongillo, 2020). In this section we simplify notation and refer to the target variable and forecast as Y and x. We define the mode for random variables with continuous Lebesgue densities as the global maxima of the density function.[6] We make the following distinction in the notion of *unimodality*.

[6] More generally, the mode is often defined as the limit, as $\delta \to 0$, of the modal midpoint functional MMP_δ, given in equation (2.5) below (Gneiting, 2011; Dearborn and Frongillo, 2020). This definition coincides with the global maxima of the density function for distributions with continuous Lebesgue density; and it coincides with the points of maximal probability for discrete distributions (Heinrich, 2014). Note that our definition of the mode as the global maxima of a density function is only valid for distributions with *continuous* Lebesgue densities as otherwise the density function can be modified on singletons (null sets in the distribution) without altering the underlying probability measure.

Definition 2.1. *An absolutely continuous distribution is defined as (a) weakly unimodal if it has a unique and well-defined mode, and (b) strongly unimodal if it is weakly unimodal and does not have further local modes.*

Heinrich (2014) and Dearborn and Frongillo (2020) show that strictly consistent loss functions or strict identification functions do not exist for the mode for the class of weakly unimodal distributions. While this does not imply that there do not exist any such functions for the class of strongly unimodal distributions, none have yet been found.

Gneiting (2011) notes that it is sometimes stated informally that the mode is an optimal point forecast under the loss function $L_\delta(x, Y) = \mathbb{1}_{\{|x-Y|\leq\delta\}}$ for some fixed $\delta > 0$. In fact, this loss function elicits the midpoint of the modal interval (also known as the modal midpoint, or MMP) of length 2δ. The MMP is defined as the midpoint of the interval of length 2δ that contains the highest probability,

$$\mathrm{MMP}_\delta(P) = \sup_{x\in\mathbb{R}} P\big(Y \in [x-\delta, x+\delta]\big). \tag{2.5}$$

More formally, it holds that for any $\delta > 0$ small enough, the modal midpoint is well-defined for all distributions with unique and well-defined mode and it holds that $\lim_{\delta\downarrow 0} \mathrm{MMP}_\delta(P) = \mathrm{Mode}(P)$ (Gneiting, 2011). In a similar manner, Eddy (1980), Kemp and Silva (2012) and Kemp et al. (2020) propose estimation of the mode by estimating the modal interval with an asymptotically shrinking length. We formalize these ideas in the decision-theoretical framework in the following definition.

Definition 2.2. *The functional $\Gamma : \mathcal{P} \to \mathbb{R}$ is asymptotically elicitable (identifiable) relative to $\mathcal{P}$ if there exists a sequence of elicitable (identifiable) functionals $\Gamma_k : \mathcal{P} \to \mathbb{R}$, $k \in \mathbb{N}$, such that $\Gamma_k(P) \to \Gamma(P)$ for all $P \in \mathcal{P}$.*

As the modal midpoint converges to the mode, this establishes *asymptotic elicitability* for the mode functional for the class of weakly unimodal probability distributions with continuous Lebesgue densities. Unfortunately, this does not directly allow for asymptotic *identifiability* of the mode as any pseudo-derivative of the loss function L_δ equals zero. We establish asymptotic identifiability of the mode through the *generalized modal midpoint*.

Definition 2.3. *Given a kernel function $K(\cdot)$ and bandwidth parameter δ, the generalized modal midpoint, $\Gamma_\delta^K(P)$, of $Y \sim P$ is defined as*

$$\Gamma_\delta^K(P) = \arg\min_{x \in \mathbb{R}} \mathbb{E}\left[L_\delta^K(x, Y)\right], \qquad \text{where} \qquad L_\delta^K(x, Y) = -\frac{1}{\delta} K\left(\frac{x - Y}{\delta}\right). \tag{2.6}$$

The familiar modal midpoint is nested in this definition by using a rectangular kernel for $K(u) = \mathbb{1}_{\{|u| \leq 1\}}$, and this definition allows for smooth generalizations. As this definition involves the argmin of a function, we first establish that this is well-defined and that it converges to the mode functional. The following proposition also considers identifiability of the generalized modal midpoint.

Theorem 2.4. *Let K be a strictly positive kernel function on the real line that is log-concave, i.e. $\log(K(u))$ is a concave function, and additionally let $\int K(u)\mathrm{d}u = 1$ and $\int |u| K(u)\mathrm{d}u < \infty$. Let $\mathcal{P}$ be the class of absolutely continuous and weakly unimodal distributions with bounded and Lipschitz-continuous density and let $\tilde{\mathcal{P}} \subset \mathcal{P}$ be the subclass of strongly unimodal distributions.*

(a) The functional Γ_δ^K induced by the loss function (2.6) is well-defined for all $\delta > 0$ and $P \in \mathcal{P}$.

(b) For $\delta \to 0$, it holds that $\Gamma_\delta(P) \to \mathrm{Mode}(P)$ for all $P \in \mathcal{P}$.

(c) If K is differentiable, for all (fixed) $\delta > 0$ and $P \in \tilde{\mathcal{P}}$, it holds that the function

$$V_\delta^K(x, Y) = \frac{\partial}{\partial x} L_\delta^K(x, Y) = -\frac{1}{\delta^2} K'\left(\frac{x - Y}{\delta}\right) \tag{2.7}$$

is a strict identification function for Γ_δ^K. In particular, the generalized modal midpoint is identifiable and the mode is asymptotically identifiable with respect to $\tilde{\mathcal{P}}$.

This theorem shows that for the classes of weakly (strongly) unimodal distributions, the generalized modal midpoint is elicitable (and identifiable), and consequently, the mode is asymptotically elicitable (and identifiable). While $V_\delta^K(x, Y)$ being an identification function for the generalized modal midpoint is obvious from Definition 2.3, Theorem 2.4(c) establishes its strictness.

For a fixed $\delta > 0$, strict identifiability is doomed to fail when both the underlying distribution and the kernel function have bounded support as the expected identification function equals zero

for values far outside both supports. Theorem 2.4(c) shows that employing log-concave kernels with infinite support circumvents this problem. While kernels with bounded support generally exhibit a superior performance in nonparametric statistics, this proposition motivates the use of kernels with infinite support like the Gaussian density function. Furthermore, the Gaussian density function, among many others, satisfies the required log-concavity of the kernel function.[7]

2.3 Forecast Rationality Tests for the Mode

Having established the asymptotic identifiability of the mode in the previous section, we now consider rationality testing of mode forecasts, i.e. testing the following null hypothesis,

$$\mathbb{H}_0 : X_t = \text{Mode}(Y_{t+1}|\mathcal{F}_t) \ \forall \ t \text{ a.s.} \tag{2.8}$$

While classical, $\sqrt{T}$-consistent rationality tests based on strict identification functions are unavailable for the mode due to its non-identifiability, we next propose a rationality test for mode forecasts based on an asymptotically shrinking bandwidth parameter δ_T. Consider the (asymptotically valid) identification function V_δ with bandwidth δ_T, and multiplied by the instruments $\boldsymbol{h}_t$,

$$\psi(Y_{t+1}, X_t, \boldsymbol{h}_t, \delta_T) = V^K_{\delta_T}(X_t, Y_{t+1})\boldsymbol{h}_t = -\frac{1}{\delta_T^2} K'\left(\frac{X_t - Y_{t+1}}{\delta_T}\right)\boldsymbol{h}_t. \tag{2.9}$$

We make the following assumptions.

Assumption 2.5.

(A1) The sequence $\left(Y_t, X_t, \boldsymbol{h}_t\right)$ *is stationary and ergodic.*

(A2) It holds that $\mathbb{E}\left[||\boldsymbol{h}_t||^{2+\delta}\right] < \infty$ *for some* $\delta > 0$.

(A3) The matrix $\mathbb{E}\left[\boldsymbol{h}_t\boldsymbol{h}_t^\top\right]$ *has full rank for all* $t = 1, \ldots, T$.

(A4) The conditional distribution of $\varepsilon_t = X_t - Y_{t+1}$ *given* $\mathcal{F}_t$ *is absolutely continuous with density* $f_t(\cdot)$ *which is three times continuously differentiable with bounded derivatives.*

[7] Theorem 2.4(c) also holds if log-concavity of the underlying density, instead of the kernel function, holds. This illustrates that kernels with bounded support can be employed at the cost of restricting the class of distributions.

(A5) $K : \mathbb{R} \to \mathbb{R}$, $u \mapsto K(u)$ *is a non-negative and continuously differentiable kernel function such that: (i)* $\int K(u)\mathrm{d}u = 1$, *(ii)* $\int uK(u)\mathrm{d}u = 0$, *(iii)* $\sup K(u) \leq c < \infty$, *(iv)* $\sup K'(u) \leq c < \infty$, *(v)* $\int u^2 K(u)\mathrm{d}u < \infty$, *(vi)* $\int \left|K'(u)\right|\mathrm{d}u < \infty$, *(vii)* $\int uK'(u)\mathrm{d}u < \infty$.

(A6) δ_T *is a strictly positive and deterministic sequence such that (i)* $T\delta_T \to \infty$, *and (ii)* $T\delta_T^7 \to 0$.

The above assumptions are a combination of standard assumptions from rationality testing and nonparametric statistics. Condition (A1) facilitates the use of a standard central limit theorem and allows for both time-series and cross-sectional applications. For non-stationary data, this can be replaced by the assumption of an α-mixing process (see e.g. White (2001), Section 5.4) by slightly strengthening the moment condition (A2). In cross-sectional applications with independent observations, this assumption can be replaced (and weakened) by the classical Lindeberg condition (see e.g. White (2001), Section 5.2). Condition (A2) is a standard moment assumption in time-series applications. Notice that as the kernel function K' is bounded, we do not require existence of any moments of Y_t or X_t, which makes this more flexible than rationality testing for mean forecasts. The full rank condition (A3) prevents the instruments from being perfectly colinear which in turn prevents the asymptotic covariance matrix from being singular. Assumption (A4) assumes a relatively smooth behavior of the conditional density function which is required to apply a Taylor expansion common to the nonparametric literature. Conditions (A5) and (A6) are standard kernel and bandwidth conditions from the nonparametric literature. We discuss specific kernel and bandwidth choices in Supplemental Appendices S.2 and S.3 respectively.

Theorem 2.6. *Under Assumption 2.5 and* $\mathbb{H}_0 : X_t = \mathrm{Mode}(Y_{t+1}|\mathcal{F}_t)\ \forall\ t$ *a.s., it holds that*

$$\delta_T^{3/2} T^{-1/2} \sum_{t=1}^{T} \psi(Y_{t+1}, X_t, \boldsymbol{h}_t, \delta_T) \xrightarrow{d} \mathcal{N}\big(0, \Omega_{\mathrm{Mode}}\big), \tag{2.10}$$

where $\Omega_{\mathrm{Mode}} = \mathbb{E}\left[\boldsymbol{h}_t \boldsymbol{h}_t^\top f_t(0)\right] \int K'(u)^2 \mathrm{d}u$.

We obtain a test for the rationality of mode forecasts by drawing on the literature on nonparametric estimation. Unsurprisingly, therefore, the rate of convergence is slower than $\sqrt{T}$; Assumption (A6) requires $\delta_T \propto T^{-\kappa}$ with $\kappa \in (1/7, 1)$, which implies that the fastest convergence rate approaches

$T^{2/7}$ and is obtained by setting $\delta_T \approx T^{-1/7}$.[8]

We utilize a standard Gaussian kernel in our analysis. While kernels with infinite support yield slightly less efficient estimates (Li and Racine, 2006), $V_\delta(x, Y)$ is a *strict* identification function if and only if K has infinite support and is strictly increasing (decreasing) left (right) of its mode, conditions that are satisfied by the Gaussian kernel.[9]

Following Kemp and Silva (2012) and Kemp et al. (2020), we estimate the covariance matrix by its sample counterpart,

$$\widehat{\Omega}_{T,\text{Mode}} = \frac{1}{T}\sum_{t=1}^{T} \delta_T^{-1} K'\left(\frac{X_t - Y_{t+1}}{\delta_T}\right)^2 \boldsymbol{h}_t \boldsymbol{h}_t^\top. \tag{2.11}$$

The following theorem shows consistency of the asymptotic covariance estimator, its proof is presented in Supplemental Appendix S.1.

Theorem 2.7. *Given Assumption 2.5, it holds that* $\widehat{\Omega}_{T,\text{Mode}} \xrightarrow{P} \Omega_{Mode}$.

We can now define the Wald test statistic:

$$J_T = \left(\delta_T^{3/2} T^{-1/2} \sum_{t=1}^{T} \psi\big(Y_{t+1}, X_t, \boldsymbol{h}_t, \delta_T\big)\right)^\top \widehat{\Omega}_{T,\text{Mode}}^{-1} \left(\delta_T^{3/2} T^{-1/2} \sum_{t=1}^{T} \psi\big(Y_{t+1}, X_t, \boldsymbol{h}_t, \delta_T\big)\right). \tag{2.12}$$

The following statement follows directly from Theorem 2.6 and Theorem 2.7.

Corollary 2.8. *Under Assumption 2.5 and the null hypothesis* $\mathbb{H}_0 : X_t = \text{Mode}(Y_{t+1}|\mathcal{F}_t)\ \forall\ t$ *a.s., it holds that* $J_T \xrightarrow{d} \chi^2_k$.

This corollary justifies an asymptotic test at level $\alpha \in (0,1)$ which rejects $\mathbb{H}_0$ when $J_T > Q_k(1-\alpha)$, where $Q_k(1-\alpha)$ denotes the $(1-\alpha)$ quantile of the χ^2_k distribution.

We now turn to the behavior of our test statistic J_T under the alternative hypothesis,

$$\mathbb{H}_A : \mathbb{E}\big[f_t'(0)\boldsymbol{h}_t\big] \neq 0 \qquad \text{for all } t = 1, \ldots, T. \tag{2.13}$$

[8] Under additional assumptions, the speed of convergence of a nonparametric estimator may be increased via the use of higher-order kernel functions, see e.g. Li and Racine (2006), however the generalized modal midpoint introduced in Definition 2.3 requires the log-concave kernel to be well-defined and unique, and as this assumption is automatically violated for higher-order kernels we do not consider them here.

[9] We also considered the relatively efficient quartic (or biweight) kernel but did not observe a change in power relative to the Gaussian kernel. See Supplemental Appendix S.2 for a further discussion of the kernel choice.

The following theorem characterizes the behavior of our mode rationality test under the alternative.

Theorem 2.9. *If $T\delta_T^3 \to \infty$, then under Assumption 2.5 and the alternative hypothesis, $\mathbb{H}_A$, it holds that $\mathbb{P}(J_T \geq c) \to 1$ for any constant $c \in \mathbb{R}$.*

The assumption $T\delta_T^3 \to \infty$ is fulfilled using our bandwidth choice, described in Supplemental Appendix S.3. While our mode rationality test has asymptotically correct size for the class of weakly unimodal distributions, it only has power when the instruments are correlated with the conditional density slope at zero, $f_t'(0)$. This is analogous to the case for standard mean and median rationality tests which have power against the alternative that $\mathbb{E}\big[V(X_t, Y_{t+1})\boldsymbol{h}_t\big] \neq 0$, where $V(X_t, Y_{t+1})$ is the identification function for the mean or median.[10] When the instruments include a constant, a sufficient condition for our alternative hypothesis is $f_t'(0) > 0$ (or $f_t'(0) < 0$) for all $t = 1, \ldots, T$, i.e. when all forecasts are issued to the left (to the right) of the mode. This can be interpreted as uniform power against the class of strongly unimodal distributions. Our test will have low power to reject forecasts that are far in the tail, where the density is (almost) flat, however this test is intended for forecasts of some measure of central tendency, and such forecasts will lie broadly in the central region of the distribution where this concern does not arise.

Beyond the rationality tests proposed here, the concept of asymptotic elicitability is of interest in its own right. Asymptotic elicitability may facilitate forecast comparison and elicitation of novel measures of uncertainty. See for example Eyting and Schmidt (2018) for an elicitation procedure for the maximum, a functional that generally is not elicitable (Bellini and Bignozzi, 2015).

3 Evaluating Forecasts of Measures of Central Tendency

We define the class of measures of central tendency as the set of convex combinations of the mean, median and mode, and we next propose identification-robust tests of whether *any* element of the class is consistent with forecast rationality.[11]

[10] See e.g. Theorem 2 of Giacomini and White (2006), where their Comment 6 and Nolde and Ziegel (2017) point out that the theory can be directly adapted to rationality testing.

[11] We focus on these three measures as foundational centrality measures, but our testing approach can easily be extended to include other measures, e.g. the trimmed mean, Huber's (1964) robust centrality measure, Barendse's (2020) "interquantile expectation," and others.

Our approach is related to, but distinct from, Elliott et al. (2005). These authors consider the case that a respondent's point forecast corresponds to some quantile (or expectile) of her predictive distribution. They employ a parametric loss function ("lin-lin" for quantiles, "quad-quad" for expectiles), $L(Y, X; \tau)$ with a scalar unknown parameter (τ) characterizing the asymmetry of the loss. Elliott et al. (2005) use GMM to estimate the τ that best describes the sequence of forecasts and realizations, and test whether forecast rationality holds at the estimated value for τ. Our approach differs economically in that we only consider forecasts within a general class of measures of centrality, while the classes of forecasts considered by Elliott et al. (2005) contains only a single centrality measure. Statistically, our approach differs as we work with weighted averages of identification functions, not parametric loss functions, and we are forced to address the feature that our parameter may be partially, weakly, or un-identified, which precludes point estimation.

Our analysis begins by considering the $(3 \times k)$ matrix:

$$\boldsymbol{\psi}_{t,T} = T^{-1/2} \begin{pmatrix} \boldsymbol{h}_t^\top \widehat{\boldsymbol{W}}_{T,\text{Mean}} & \times & \varepsilon_t \\ \boldsymbol{h}_t^\top \widehat{\boldsymbol{W}}_{T,\text{Med}} & \times & \left(\mathbb{1}_{\{\varepsilon_t > 0\}} - \mathbb{1}_{\{\varepsilon_t < 0\}}\right) \\ \boldsymbol{h}_t^\top \widehat{\boldsymbol{W}}_{T,\text{Mode}} & \times & \delta_T^{-1/2} K'\left(\frac{-\varepsilon_t}{\delta_T}\right) \end{pmatrix}, \tag{3.1}$$

where $\varepsilon_t = X_t - Y_{t+1}$ denotes the forecast error. The rows of (3.1) are comprised of the identification functions for the mean and median, and the asymptotically valid choice for the mode (as discussed in the previous section), multiplied by the $\mathcal{F}_t$-measurable $(k \times 1)$ vector of instruments $\boldsymbol{h}_t$. We normalize each of these using symmetric and positive definite matrices, $\widehat{\boldsymbol{W}}_{T,\text{Mean}}$, $\widehat{\boldsymbol{W}}_{T,\text{Med}}$ and $\widehat{\boldsymbol{W}}_{T,\text{Mode}}$, which may depend on the sample. For example, one could use the identity matrix, the square-root of the inverse sample covariance matrix, or some other matrix.

For any weight vector $\theta \in \Theta = \{\theta \in \mathbb{R}^3 : ||\theta||_1 = 1, \theta \geq 0\}$, we define the k-dimensional random variable $\phi_{t,T}(\theta)$ as a combination of the three individual (normalized and interacted with the instrument vector) identification functions:

$$\phi_{t,T}(\theta) = \theta^\top \psi_{t,T}. \tag{3.2}$$

It is important to note that the convex combination of identification functions defined by the vector θ does not necessarily lead to a forecast that is the *same* convex combination of the underlying measures of central tendency. (For example, an equal-weighted combination of the identification functions will *not* generally lead to a forecast that is an equal-weighted combination of the mean, median and mode; it will generally be some other convex combination.) At the vertices the weights are clearly identical, but for non-degenerate combinations the weights will generally differ, though they will also lie in the three-dimensional unit simplex, Θ.

Consider the GMM objective function based on the above convex combination:

$$S_T(\theta) = \left[\sum_{t=1}^{T} \phi_{t,T}(\theta)\right]^{\top} \widehat{\Sigma}_T^{-1}(\theta) \left[\sum_{t=1}^{T} \phi_{t,T}(\theta)\right], \tag{3.3}$$

where $\widehat{\Sigma}_T^{-1}(\theta)$ denotes an $O_p(1)$ positive definite weighting matrix, which may depend on the parameter θ. Unlike the problem in Elliott et al. (2005), the unknown parameter in our framework (the weight vector θ) cannot be assumed to be well identified. For symmetric distributions, the combination weights are unidentified. For distributions that exhibit only mild asymmetry a weak identification problem arises. For asymmetric distributions where one measure is a convex combination of the other two we have partial identification of the weight vector. The distribution of economic variables may or may not exhibit asymmetry, and so addressing this identification problem is a first-order concern.

The possibility that the true parameter θ_0 is unidentified, partially identified, or weakly identified implies that the objective function $S_T(\theta)$ may be flat or almost flat in a neighborhood of θ_0, ruling out consistent estimation of θ_0. Stock and Wright (2000) show that, under regularity conditions, we can nevertheless construct asymptotically valid confidence bounds for θ_0, by showing that the objective function S_T evaluated at θ_0 continues to exhibit an asymptotic χ^2 distribution. This facilitates the construction of asymptotically valid confidence bounds even in a setting where the parameter vector may be strongly identified, weakly identified, or unidentified.

For the technical treatment of this approach, we define the counterpart of $\phi_{t,T}(\theta)$ that depends on the probability limits of the weighting matrices: $\tilde{\phi}_{t,T}(\theta) = \theta^{\top}\tilde{\boldsymbol{\psi}}_{t,T}$, where

$$\tilde{\boldsymbol{\psi}}_{t,T} = T^{-1/2} \begin{pmatrix} \boldsymbol{h}_t^\top \boldsymbol{W}_{\text{Mean}} & \times & \varepsilon_t \\ \boldsymbol{h}_t^\top \boldsymbol{W}_{\text{Med}} & \times & \left(\mathbb{1}_{\{\varepsilon_t > 0\}} - \mathbb{1}_{\{\varepsilon_t < 0\}}\right) \\ \boldsymbol{h}_t^\top \boldsymbol{W}_{\text{Mode}} & \times & \delta_T^{-1/2} K'\left(\frac{-\varepsilon_t}{\delta_T}\right) \end{pmatrix}, \tag{3.4}$$

and where $\boldsymbol{W}_{\text{Mean}}$, $\boldsymbol{W}_{\text{Med}}$, and $\boldsymbol{W}_{\text{Mode}}$ denote the (symmetric and positive definite) probability limits of $\widehat{\boldsymbol{W}}_{T,\text{Mean}}$, $\widehat{\boldsymbol{W}}_{T,\text{Med}}$ and $\widehat{\boldsymbol{W}}_{T,\text{Mode}}$. We make the following assumption.

Assumption 3.1. *There exists a $\theta_0 \in \Theta$ and sequences $\phi^*_{t,T}(\theta_0)$ and $u_{t,T}(\theta_0)$, such that*

$$\tilde{\phi}_{t,T}(\theta_0) := \phi^*_{t,T}(\theta_0) + u_{t,T}(\theta_0), \tag{3.5}$$

and

(a) $\left\{\phi^*_{t,T}(\theta_0), \mathcal{F}_{t+1}\right\}$ *is a martingale difference sequence,*

(b) $\sum_{t=1}^T ||u_{t,T}(\theta_0)||^2 \xrightarrow{P} 0$ *and* $\sum_{t=1}^T \mathbb{E}\left[||u_{t,T}(\theta_0)||^{2+\delta}\right] \to 0$,

(c) $\sum_{t=1}^T u_{t,T}(\theta_0)\tilde{\phi}_{t,T}(\theta_0) \xrightarrow{P} 0$ *and* $\sum_{t=1}^T \mathbb{E}\left[u_{t,T}(\theta_0)\tilde{\phi}_{t,T}(\theta_0)\right] \to 0$.

The decomposition in equation (3.5) implies that the sequence $\tilde{\phi}_{t,T}(\theta_0)$ is an *approximate* martingale difference sequence (MDS) in the sense that $\tilde{\phi}_{t,T}(\theta_0)$ can be decomposed into a MDS $\phi^*_{t,T}(\theta_0)$ and some asymptotically vanishing sequence $u_{t,T}(\theta_0)$. The assumption of an approximate MDS is weaker than the standard MDS assumption,

$$\exists\ \theta_0 \in \Theta \text{ s.t. } \left\{\tilde{\phi}_t(\theta_0), \mathcal{F}_{t+1}\right\} \text{ is a MDS}, \tag{3.6}$$

but stricter than the classically imposed assumption

$$\exists\ \theta_0 \in \Theta \text{ s.t. } \mathbb{E}\left[\tilde{\phi}_t(\theta_0)\right] = 0. \tag{3.7}$$

An exact MDS assumption as in (3.6) does not hold for the baseline case that X_t is an optimal mode forecast, as in that case the MDS assumption only holds asymptotically (see the proof of Theorem

2.6 for details). The unconditional moment condition in (3.7) is too weak for our purpose as in that case a CLT for stationary and ergodic (or strong mixing) processes requires that moments of order $r > 2$ are finite, which is not fulfilled for the mode case as these moments diverge (arbitrarily slowly) through the bandwidth parameter δ_T.

Assumption 3.1 is an intermediate case of classically-imposed conditions (3.6) and (3.7) and can easily be shown to hold at the three vertices, where the forecast is the mean, median or mode. Specifically, when X_t is a mean or median forecast (i.e. $\theta_0 = (1,0,0)$ or $\theta_0 = (0,1,0)$), set $u_{t,T}(\theta_0) = 0$ and $\{\tilde{\phi}_{t,T}(\theta_0), \mathcal{F}_{t+1}\}$ is obviously a MDS. When X_t is the true conditional mode of Y_{t+1}, (i.e. $\theta_0 = (0,0,1)$), set

$$u_{t,T}(\theta_0) = \mathbb{E}_t\left[\tilde{\phi}_{t,T}(\theta_0)\right] = (T\delta_T)^{-1/2}\mathbb{E}_t\left[K'\left(\frac{-\varepsilon_t}{\delta_T}\right)\right]\boldsymbol{h}_t^\top, \tag{3.8}$$

as in the proof of Theorem 2.6. The conditions on $u_{t,T}(\theta_0)$ in Assumption 3.1 are fulfilled by the arguments given in the proof of Theorem 2.6 and Lemma S.1.1 - Lemma S.1.4 in the Supplemental Appendix. When X_t is a convex combination of a mean and median forecast, i.e., $\theta_0 = (\xi, 1-\xi, 0)$ for some $\xi \in [0,1]$, we set $u_{t,T}(\theta_0) = 0$ and $\{\tilde{\phi}_{t,T}(\theta_0), \mathcal{F}_{t+1}\}$ is again a MDS. When X_t is a convex combination with non-zero weight on the mode, Assumption 3.1 is difficult to verify.

We further impose the following regularity conditions on our process.

Assumption 3.2. *(a)* $\mathbb{E}\left[|\varepsilon_t|^{2+\delta}\right] < \infty$ *and* $\mathbb{E}\left[||\boldsymbol{h}_t||^{2+\delta}|\varepsilon_t|^{2+\delta}\right] < \infty$, *(b)* $\widehat{\boldsymbol{W}}_{T,\text{Mean}} \xrightarrow{P} \boldsymbol{W}_{\text{Mean}}$, $\widehat{\boldsymbol{W}}_{T,\text{Med}} \xrightarrow{P} \boldsymbol{W}_{\text{Med}}$, *and* $\widehat{\boldsymbol{W}}_{T,\text{Mode}} \xrightarrow{P} \boldsymbol{W}_{\text{Mode}}$ *for some symmetric and positive definite matrices* $\boldsymbol{W}_{\text{Mean}}$, $\boldsymbol{W}_{\text{Med}}$ *and* $\boldsymbol{W}_{\text{Mode}}$.

Theorem 3.3 below presents the asymptotic distribution of the process $\sum_{t=1}^{T}\phi_{t,T}(\theta_0)$ at the true parameter θ_0.

Theorem 3.3. *Given Assumption 2.5, Assumption 3.1 and Assumption 3.2, it holds that*

$$\sum_{t=1}^{T}\phi_{t,T}(\theta_0) \xrightarrow{d} \mathcal{N}\big(0, \Sigma(\theta_0)\big), \tag{3.9}$$

where

$$\begin{aligned}\Sigma(\theta_0) = \mathbb{E}\Big[&\theta_{10}^2 \boldsymbol{W}_{\text{Mean}} \boldsymbol{h}_t \boldsymbol{h}_t^\top \boldsymbol{W}_{\text{Mean}} \varepsilon_t^2 + \theta_{20}^2 \boldsymbol{W}_{\text{Med}} \boldsymbol{h}_t \boldsymbol{h}_t^\top \boldsymbol{W}_{\text{Med}} \big(\mathbb{1}_{\{\varepsilon_t > 0\}} - \mathbb{1}_{\{\varepsilon_t < 0\}}\big)^2 \\ &+ \theta_{30}^2 \boldsymbol{W}_{\text{Mode}} \boldsymbol{h}_t \boldsymbol{h}_t^\top \boldsymbol{W}_{\text{Mode}} f_t(0) \int K'(u)^2 \mathrm{d}u \\ &+ 2\theta_{10}\theta_{20} \boldsymbol{W}_{\text{Mean}} \boldsymbol{h}_t \boldsymbol{h}_t^\top \boldsymbol{W}_{\text{Med}} \varepsilon_t \big(\mathbb{1}_{\{\varepsilon_t > 0\}} - \mathbb{1}_{\{\varepsilon_t < 0\}}\big)\Big].\end{aligned} \tag{3.10}$$

Under the null hypothesis, Assumption 3.1 implies that $\phi_{t,T}(\theta_0)$ is an approximate MDS, i.e. this sequence is approximately (as $T \to \infty$) uncorrelated. Consequently, we do not need to rely on HAC covariance estimation, and can instead estimate the asymptotic covariance matrix using the simple sample covariance matrix:

$$\widehat{\Sigma}_T(\theta) = \frac{1}{T} \sum_{t=1}^{T} \phi_{t,T}(\theta) \phi_{t,T}(\theta)^\top. \tag{3.11}$$

The next theorem shows consistency of the outer product covariance estimator, its proof is presented in Supplemental Appendix S.1.

Theorem 3.4. *Given Assumptions 2.5, 3.1, and 3.2, it holds that* $\widehat{\Sigma}_T(\theta) \xrightarrow{P} \Sigma(\theta_0)$.

Corollary 3.5. *Given Assumptions 2.5, 3.1, and 3.2, it holds that* $S_T(\theta_0) \xrightarrow{d} \chi^2_k$.

Following Stock and Wright (2000), this corollary allows to construct asymptotically valid confidence regions for θ_0 with coverage probability $(1-\alpha)\%$ by considering the set

$$\big\{\theta_0 \in \Theta : S_T(\theta_0) \leq Q_k(1-\alpha)\big\}, \tag{3.12}$$

where $Q_k(1-\alpha)$ denotes the $(1-\alpha)$ quantile of the χ^2_k distribution.

Given the above results, we obtain a test for forecast rationality for a general measure of central tendency by evaluating the GMM objective function using a dense grid of convex combination parameters $\theta_j \in \Theta$ for $j = 1, \ldots, J$. An asymptotically valid confidence set is given by the values of θ_j for which $S_T(\theta_j) \leq Q_k(1-\alpha)$. These values represent the centrality measures that "rationalize" the observed sequence of forecasts and realizations, in that rationality cannot be rejected for these

measures of centrality. It is possible that the confidence set is empty, in which case we reject rationality at the α significance level for the *entire class* of general centrality measures.

4 Simulation Study

This section studies the finite-sample performance of the methods proposed above. Section 4.1 presents simulations for the mode rationality test and Section 4.2 analyzes the test for rationality for general measures of central tendency.

4.1 Rationality tests for mode forecasts

To evaluate the finite-sample properties of our mode rationality test, we simulate data from the following data generating processes (DGPs). We consider two cross-sectional DGPs, one homoskedastic and the other heteroskedastic, and two time-series DGPS, a simple AR(1) and an AR(1)-GARCH(1,1). We simulate data using the following unified framework:

$$Y_{t+1} = Z_t^\top \zeta + \sigma_{t+1}\xi_{t+1}, \quad \text{where} \quad \xi_{t+1} \overset{iid}{\sim} \mathcal{SN}(0,1,\gamma), \tag{4.1}$$

where $\mathcal{SN}(0,1,\gamma)$ is a skewed standard Normal distribution, and Z_t denotes a vector of covariates (possibly including lagged values of Y_{t+1}), ζ denotes a parameter vector and σ_{t+1} denotes a conditional variance process. Using this general formulation, we consider four cases:

(1) Homoskedastic cross-sectional data: $Z_t \overset{iid}{\sim} \mathcal{N}\big((1,1,-1,2) , \text{diag}(0,1,1,0.1) \big)$, where $\zeta = (1,1,1,1)$ and $\sigma_{t+1} = 1$.

(2) Heteroskedastic cross-sectional data: As in case (1), but $\sigma_{t+1} = 0.5 + 1.5(t+1)/T$.

(3) AR(1) data: $Z_t = Y_t$ with $\zeta = 0.5$ and $\sigma_{t+1} = 1$.

(4) AR(1)-GARCH(1,1) data: As in case (3), but $\sigma^2_{t+1} = 0.1 + 0.8\sigma^2_t + 0.1\sigma^2_t\xi^2_t$.

All the above DGPs are based on a skewed Gaussian residual distribution with skewness parameter γ. This choice nests the case of a standard Gaussian distribution at $\gamma = 0$, and in this case all

measures of centrality coincide. As the skewness parameter increases in magnitude, the measures of centrality increasingly diverge.

For the general DGP in (4.1), optimal one-step ahead mode forecasts are given by

$$X_t = \text{Mode}(Y_{t+1}|\mathcal{F}_t) = Z_t^\top \zeta + \sigma_{t+1} \,\text{Mode}(\xi_t), \tag{4.2}$$

where $\text{Mode}(\xi_t)$ depends on the skewness parameter γ.

We consider a range of skewness parameters, $\gamma \in \{0, 0.1, 0.25, 0.5\}$, and sample sizes $T \in \{100, 500, 2000, 5000\}$. In all cases we use $10,000$ replications. To evaluate the size of our test in finite samples, we generate optimal mode forecasts through equation (4.2) and apply the mode forecast rationality test based on three choices of instruments: we use the instrument choices $\boldsymbol{h}_{t,1} = 1$ and $\boldsymbol{h}_{t,2} = (1, X_t)$ for all DGPs; for the two cross-sectional DGPs, our third choice of instruments is $\boldsymbol{h}_{t,3} = (1, X_t, Z_{t,1})$, while for the two time-series DGPs we use $\boldsymbol{h}_{t,3} = (1, X_t, Y_{t-1})$.[12]

Table 1 presents the finite-sample sizes of the test under the different DGPs, sample sizes, instrument choices, and skewness parameters. In all cases we use a Gaussian kernel and set the nominal size to 5%. Results for nominal test sizes of 1% and 10% are given in Table S.1 and Table S.2 in the Supplemental Appendix.

We find that our mode rationality test leads to finite-sample rejection rates close to the nominal test size, across all of the different choices of DGPs, instruments, and skewness parameters. (In the Supplemental Appendix we present similar results for different significance levels and kernel choices.) Table 1 reveals that an increasing degree of skewness in the underlying conditional distribution negatively influences the tests' performance. This is explained by the fact that for more skewed data we choose a smaller bandwidth parameter (following the rule of thumb described in Supplemental Appendix S.3), resulting in less efficient estimates. Consequently, for highly skewed distributions, the mode rationality test requires larger sample sizes in order to converge to the nominal test size. Supplemental Appendix S.4 presents simulation results for two studies of the finite-sample power of this test, confirming that the test has non-trivial power in reasonable scenarios.

[12] Notice that the choice of instruments $(1, X_t, Y_t)$ is invalid for the AR-GARCH DGP for $\gamma = 0$ as in that case X_t and Y_t are perfectly colinear. Consequently, we use the lag Y_{t-1} as instrument here.

Table 1: Size of the mode rationality test in finite samples

	Instrument Set 1				Instrument Set 2				Instrument Set 3			
Skewness	0	0.1	0.25	0.5	0	0.1	0.25	0.5	0	0.1	0.25	0.5
Sample size												
	Panel A: Homoskedastic iid data											
100	4.7	4.7	5.9	9.1	5.1	5.1	6.1	8.2	5.5	5.3	6.1	7.7
500	5.2	5.9	8.0	8.9	5.3	5.7	7.2	7.8	5.2	5.7	7.3	7.2
2000	5.3	6.9	8.7	7.2	5.3	6.2	7.4	6.8	5.4	6.1	6.8	6.3
5000	5.3	6.1	7.2	6.4	5.1	5.8	6.4	5.7	5.1	5.5	6.4	5.6
	Panel B: Heteroskedastic data											
100	5.1	5.1	6.5	9.3	5.3	5.2	6.2	8.4	5.6	5.7	6.1	7.6
500	6.0	5.9	8.4	8.2	5.7	5.5	7.6	7.4	5.6	5.9	7.2	6.8
2000	5.1	6.8	8.4	6.9	5.2	6.2	7.4	6.3	5.3	6.0	7.1	6.2
5000	5.3	7.2	9.2	6.5	5.1	6.1	7.7	6.0	4.8	5.9	7.1	5.9
	Panel C: Autoregressive data											
100	4.8	4.8	6.0	8.3	5.5	5.1	5.9	7.5	5.8	5.2	5.8	7.1
500	5.6	5.9	7.8	9.3	5.4	5.7	7.1	8.0	5.3	5.2	6.4	7.6
2000	5.5	6.3	8.0	6.8	5.3	5.8	7.1	6.2	5.2	5.7	6.5	5.7
5000	5.1	6.4	7.2	6.5	5.2	6.5	6.5	5.9	5.2	5.9	6.3	5.9
	Panel D: AR-GARCH data											
100	4.7	4.8	5.7	9.7	5.0	4.9	5.8	8.6	5.4	5.4	5.8	8.2
500	5.5	6.4	8.6	9.3	5.7	6.1	7.2	8.4	5.6	5.4	6.8	7.8
2000	5.3	6.5	8.3	6.7	4.8	5.5	7.2	6.2	5.0	5.6	6.6	6.2
5000	5.3	6.6	7.7	6.3	5.2	5.9	6.7	6.1	5.3	6.0	6.6	5.6

Notes: This table presents the empirical rejection rates (in percent) of the mode rationality test using various sample sizes, various levels of skewness in the residual distribution, different choices of instruments, and the four DGPs described in equation (4.1). The nominal significance level is 5%.

4.2 Rationality tests for an unknown measure of central tendency

In this section we examine the small sample behavior of the asymptotic confidence sets for the measures of central tendency, described in Section 3. As in the previous section, we consider the four DGPs described in and after equation (4.1) and the same varying sample sizes T, skewness parameters γ, and instruments $\boldsymbol{h}_t$. We generate optimal one-step ahead forecasts for the mean, median and mode as the true conditional mean, median and mode of Y_{t+1} given $\mathcal{F}_t$:

$$X_t^{\text{Mean}} = \zeta^\top Z_t + \sigma_{t+1}\,\text{Mean}(\xi_{t+1}), \tag{4.3}$$

$$X_t^{\text{Med}} = \zeta^\top Z_t + \sigma_{t+1}\,\text{Median}(\xi_{t+1}), \tag{4.4}$$

$$X_t^{\text{Mode}} = \zeta^\top Z_t + \sigma_{t+1}\,\text{Mode}(\xi_{t+1}). \tag{4.5}$$

We use the notation $\boldsymbol{X}_t = \left(X_t^{\text{Mean}}, X_t^{\text{Med}}, X_t^{\text{Mode}}\right)^\top$ and consider convex combinations of these functionals $X_t = \boldsymbol{X}_t^\top \beta$, using the following specifications: (a) Mean: $\beta = (1,0,0)^\top$, (b) Mode: $\beta = (0,0,1)^\top$, (c) Median: $\beta = (0,1,0)^\top$, (d) Mean-Mode: $\beta = (1/2,0,1/2)^\top$, (e) Mean-Median: $\beta = (1/2,1/2,0)^\top$, (f) Median-Mode: $\beta = (0,1/2,1/2)^\top$, (g) Mean-Median-Mode: $\beta = (1/3,1/3,1/3)^\top$.

For the interpretation of simulation results below, notice that the functional identified by a convex combination of identification functions for the mean, median and the mode (with weights θ) is some convex combination of the mean, median and mode forecasts, but with possibly different combination weights, β. An equal-weighted combination of the mean and the mode forecasts, for example, is not necessarily identified by an equal-weighted combination of the mean and mode identification functions, rather it will generally be some other weighted combination of these functions.

Note also that for any fixed forecast combination weight vector β, there are possibly infinitely many corresponding identification function weights θ. For example, in the right-skewed DGP used here, the three centrality measures are ordered Mode $<$ Median $<$ Mean, and given the functional forms of the optimal forecasts for this DGP presented in equations (4.3)-(4.5), this implies that when the true forecast is the median (and so the forecast combination weight vector is $\beta = (0,1,0)^\top$), any identification function combination weight vector $\theta = [\theta_{\text{Mean}}, \theta_{\text{Med}}, \theta_{\text{Mode}}] \in \Theta$ that satisfies

$$\theta_{\text{Mean}} = \frac{\text{Median}(\xi_{t+1}) - \text{Mode}(\xi_{t+1})}{\text{Mean}(\xi_{t+1}) - \text{Median}(\xi_{t+1})} \cdot \theta_{\text{Mode}} \tag{4.6}$$

will lead to a combination forecast that coincides with the median forecast.[13] This again highlights the identification problems that can arise in our analysis of forecasts of measures of central tendency. In contrast, the mean and mode centrality measures in this DGP each have a unique identification function combination weight vector, equal to the associated forecast combination weight vector.

[13] When the DGP is conditionally location-scale, one of the three centrality measures can always be expressed as a (constant) convex combination of the other two. In this DGP, this is the median, but in other applications it may be any of the three measures. When the conditional distribution exhibits variation in higher-order moments or other "shape" parameters, this restriction will generally not hold, and the variation may or may not be sufficient to separately identify the three centrality measures.

Analyzing the coverage properties of this method requires knowledge of the set of identification function weights θ corresponding to the forecast weights β used to construct the forecast. In general, θ is not known in closed-form; we use one thousand draws from each DGP, sample size, and skewness level to numerically obtain the identification function weights corresponding to each set of forecast combination weights.

Table 2 shows the empirical coverage rates for the confidence sets of centrality measures for the seven simulated convex combinations of functionals, different sample sizes, skewness parameters, and DGPs. When the set of identification function weights corresponding to a particular forecast combination weight vector is not a singleton, we choose the mid-point of the line that defines this set.[14] In the left panel of Table 2 the data is unimodal and symmetric, and the measures of central tendency coincide, making all seven forecasts identical. The test outcomes, however, can differ as each row uses a different set of moment conditions to evaluate forecast rationality. We see that in all cases the coverage rates are very close to the nominal 90% level. In the right panel of Table 2 the data is asymmetric and the measures of central tendency differ. The coverage rates remain close to the nominal 90% level, especially for larger sample sizes.

Figure 1 illustrates the average rejection (coverage) rates based on 90% confidence sets for the central tendency measures across a richer set of combination weights. (We omit the mean-median combination forecast from this figure in the interest of space.) This figure uses the AR-GARCH DGP; equivalent results for the homoskedastic cross-sectional DGP are shown in Figure S.4 in the Supplemental Appendix. Each point in the triangles corresponds to a tested centrality measure, i.e. to one value of θ, and for each point we compute how often it is contained in the 90% confidence set. We depict a coverage rate between 85% and 100% by a black point, a coverage rate between 50% and 85% by a dark grey point and anything below 50% by a light grey point. We use a cut-off of 85% to include points with coverage rates "close" to the nominal rate of 90%. We use a sample size $T = 2000$, instruments $\boldsymbol{h}_t = (1, X_t)$ and a Gaussian kernel. The upper panel presents results for the DGPs with zero skewness and the lower panel considers skewness of $\gamma = 0.5$.

[14]For this DGP, θ is a singleton only when the forecast is the mean or the mode. Table S.3 and Table S.4 in the Supplemental Appendix show finite-sample rejection rates for three values of θ: the end-points of the line defining the set, and the mid-point reported in Table 2. In all cases we find very similar results for all three points.

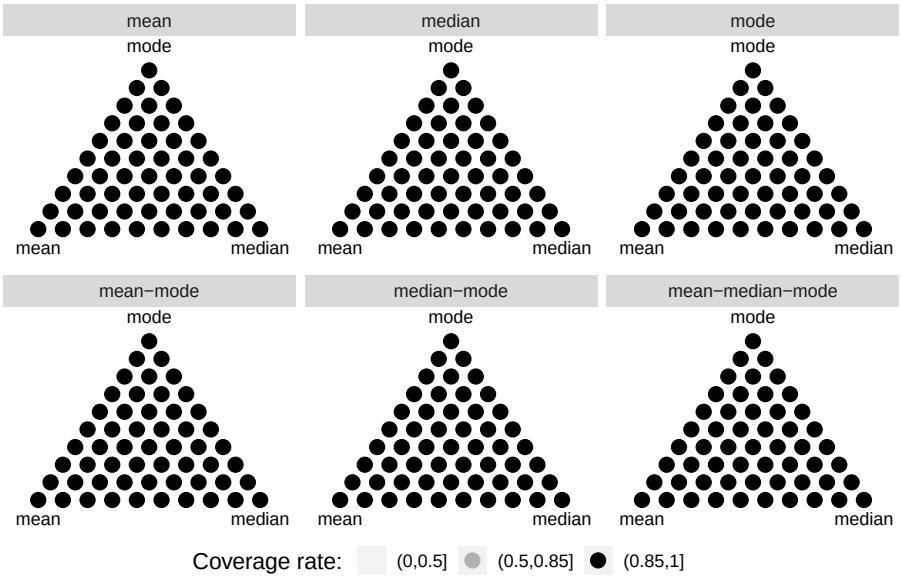
Table 2: Coverage of the central tendency confidence sets in finite samples

	Symmetric data				Skewed data			
Sample size	100	500	2000	5000	100	500	2000	5000
Centrality measure								
	Panel A: Homoskedastic iid data							
Mean	89.3	90.0	89.8	89.2	89.4	90.2	89.6	90.2
Mode	90.1	89.1	89.8	90.1	85.7	86.0	87.6	88.8
Median	89.5	89.0	89.6	89.6	90.3	91.9	91.4	91.2
Mean-Mode	90.1	89.0	89.4	90.0	90.7	92.2	92.0	91.9
Mean-Median	89.8	89.6	89.6	89.5	90.4	91.1	90.3	90.4
Median-Mode	89.9	89.0	89.6	89.8	90.1	90.9	90.7	91.4
Mean-Median-Mode	90.1	88.9	89.5	89.8	90.9	92.3	92.3	91.8
	Panel B: Heteroskedastic data							
Mean	89.2	89.9	89.6	90.3	88.7	89.6	89.4	89.5
Mode	89.5	89.5	89.9	89.8	85.8	86.4	87.9	88.7
Median	89.0	89.8	90.1	89.9	90.0	90.9	90.8	90.7
Mean-Mode	89.2	89.5	89.8	90.0	90.0	91.1	91.0	90.9
Mean-Median	89.2	89.6	90.0	90.5	89.3	90.2	89.7	90.0
Median-Mode	89.3	89.4	89.8	90.0	89.0	90.0	90.3	90.8
Mean-Median-Mode	89.2	89.7	89.9	90.1	90.3	91.2	91.1	91.3
	Panel C: Autoregressive data							
Mean	89.2	90.1	89.9	90.3	89.4	89.4	89.8	90.3
Mode	89.2	89.8	89.3	89.9	85.4	86.2	88.0	88.7
Median	89.0	89.8	89.1	89.7	90.4	90.7	92.3	90.7
Mean-Mode	88.9	89.9	89.2	89.7	90.5	91.6	92.5	91.3
Mean-Median	88.8	90.2	89.3	89.9	90.0	89.8	91.2	89.7
Median-Mode	89.1	89.9	89.1	89.9	89.4	90.5	91.5	91.1
Mean-Median-Mode	89.0	89.7	89.2	89.6	90.7	91.7	92.7	91.7
	Panel D: AR-GARCH data							
Mean	89.2	90.4	89.9	90.0	89.5	89.9	89.9	89.8
Mode	90.0	89.0	89.6	89.9	85.9	86.1	88.4	88.9
Median	88.8	89.4	89.3	89.5	90.1	91.2	91.3	90.8
Mean-Mode	89.6	89.1	89.6	89.6	90.3	91.3	92.3	91.3
Mean-Median	88.7	89.7	89.3	89.4	90.1	90.7	90.8	89.9
Median-Mode	90.1	89.0	89.5	89.7	89.4	90.6	92.0	91.1
Mean-Median-Mode	89.6	89.2	89.3	89.7	90.4	91.7	92.3	91.2

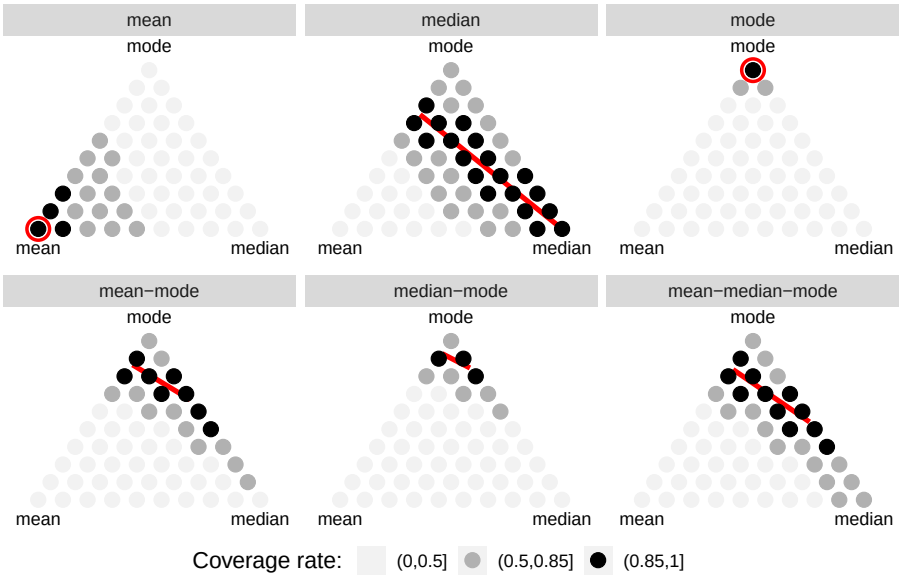
Notes: This table presents empirical coverage rates of nominal 90% confidence sets for the forecasts of central tendency. We report the results for symmetric and skewed data ($\gamma = 0$ or 0.5), for various sample sizes and the four DGPs described in equation (4.1). As instruments we use $\boldsymbol{h}_t = (1, X_t)$.

Figure 1: Coverage rates of the confidence regions for central tendency measures

(a) AR-GARCH DGP with skewness $\gamma = 0$



(b) AR-GARCH DGP with skewness $\gamma = 0.5$



This figure shows coverage rates of 90% confidence regions for the measures of central tendency for the AR-GARCH DGP. The true forecasted functional is given in the text above each triangle. The points inside the triangle correspond to convex combinations of the vertices, which are the mean, median and mode functionals. The color of the points indicates how often a specific point is contained in the 90% confidence regions. The upper panel shows results for the symmetric DGP, where all central tendency measures are equal. The lower panel uses a skewed DGP, with $\gamma = 0.5$. We use a red circle or a red line to indicate the (set of) central tendency measure(s) that correspond(s) to the forecast. We consider the sample size $T = 2000$, the instruments $\boldsymbol{h}_t = (1, X_t)$ and use a Gaussian kernel.

Panel A of Figure 1 uses a unimodal DGP with zero skewness, and so all measures of central tendency and all convex combinations thereof coincide. This implies that all six of these triangles are identical; we include them here for ease of comparison with the lower panel, where the optimal forecasts differ. (The points inside each triangle need not be identical, as they represent tests for optimality of different centrality measures.) Under symmetry and unimodality, every point in the triangle should be contained in the confidence set with probability 90%, and this figure is consistent with this, thus confirming the procedure's coverage level in this simulation design.

Panel B of Figure 1 considers nonzero skewness, and the measures of central tendency differ. In the two point-identified cases (mean and mode), we use a red circle to highlight the single centrality measure that corresponds to the forecast. The remaining cases are only partially identified and we use a red line to indicate the set of centrality measures that correspond to the forecast.

The upper left plot of Panel B considers optimal mean forecasts and exhibits the expected behavior: the mean, and convex combinations close to the mean, are usually contained in the confidence set, whereas points far away are usually excluded. A similar, but more pronounced, picture can be observed for optimal mode forecasts in the upper right plot.

The median plot in Panel B of Figure 1 reveals, as expected, that the optimal median forecast is generally not rejected when testing using the identification function for the median (revealed by the dot at the median vertex being black). This plot further shows that the convex combinations of mean, median and mode that coincide with the median (see equation 4.6) are also generally included in the confidence set.

For the interpretation of the three convex combination plots in the bottom row of Panel B, recall that the functional identified by a convex combination of identification functions for the mean, median and the mode is some convex combination of these three measures, but with possibly different combination weights. Further, as the median is itself a convex combination of the mean and mode in this simulation design, the set of combination weights consistent with rationality is a line from one edge of the triangle to another. In this design the line always connects the mean-mode edge to the mode-median edge. In all three cases we note that this line is nicely contained inside the region of dark dots, indicating correct coverage in these partially identified cases.

5 Evaluating Economic Forecasts

We now apply the tests developed above in three economic forecasting applications. Open source code to implement the tests proposed in this paper is available in the `R` package *fcrat* posted at https://github.com/Schmidtpk/fcrat.

5.1 Survey forecasts of individual income

Firstly, we consider data from the Labor Market Survey component of the Survey of Consumer Expectations,[15] conducted by the Federal Reserve Bank of New York in March, July, and November of each year. Our sample runs from March 2015 to March 2018.[16] In this survey participants are asked a variety of questions, including about their current earnings and their beliefs about their earnings in four months' time (i.e., the date of the next survey). Using adjacent surveys we obtain a sample 3,916 pairs of forecasts (X_t) and realizations (Y_{t+1}).[17] In testing the rationality of these forecasts we assume that all participants report the same, unknown, measure of centrality as their forecast. In the next subsection we explore potential heterogeneity in the measure of centrality used by different respondents.

Table 3 presents the results of rationality tests for three measures of central tendency, and for a variety of instrument sets. The first instrument set includes just a constant, and the rationality test simply tests whether the forecast errors have unconditional mean, median or mode, respectively, of zero. The other instrument sets additionally include the forecast (X_t) itself, and other information about the respondent collected in the survey. We consider the respondent's income at the time of making the forecast, indicators for the respondent's type of employer,[18] and whether the respondent received any job offers in the past four months.

[15] Source: Survey of Consumer Expectations, © 2013-2019 Federal Reserve Bank of New York (FRBNY). The SCE data are available without charge at http://www.newyorkfed.org/microeconomics/sce and may be used subject to license terms posted there. FRBNY disclaims any responsibility or legal liability for this analysis and interpretation of Survey of Consumer Expectations data.

[16] This data is made available with an 18-month lag.

[17] We drop observations that include forecasts or realizations of annualized income below \$1,000 or above \$1 million, which represent less than 1% of the initial sample. We also drop observations where the ratio of the realization to the forecast, or its inverse, is between 9 and 13, to avoid our results being affected by misplaced decimal points or by the failure to report annualized income (leading to proportional errors of around 10 to 12 respectively).

[18] The survey includes the categories government, private (for-profit), non-profit, family business, and "other." The first two categories dominate the responses and so we only consider indicators for those.

Table 3: Evaluating income survey forecasts

	Centrality measure		
Instruments	Mean	Median	Mode
1	0.008	0.401	0.737
$1, X$	0.014	0.000	0.768
$1, X$, lag income	0.031	0.000	0.970
$1, X$, government	0.064	0.000	0.949
$1, X$, private	0.047	0.000	0.898
$1, X$, job offers	0.006	0.000	0.913

Notes: This table presents p-values from tests of rationality of individual income forecasts from the New York Federal Reserve's Survey of Consumer Expectations. The columns present test results when interpreting the point forecasts as forecasts of the mean, median or mode. The rows present results for different choices of instruments used in the test: 1 is the constant, X is the forecast, "lag income" is the respondent's income at the time of the forecast, "government" and "private" are indicators for the self-reported industry in which the respondent works, "job offers" is an indicator for whether the respondent received any job offers in the previous four months.

The first row of Table 3 shows that when only a constant is used, rationality of the survey forecasts can be rejected for the mean, but cannot be rejected for the other two measures of central tendency. When we additionally include the forecast as an instrument, we can reject rationality as mean or median forecasts, but we cannot reject them as rational mode forecasts. We are similarly able to reject rationality as mean and median forecasts when we include additional covariates, but find no evidence against rationality when these forecasts are interpreted as mode forecasts.[19]

Figure 2 shows the convex combinations of mean, median and mode forecasts that lie in the confidence set constructed using the methods for weakly-identified GMM estimation in Stock and Wright (2000).[20] In the left panel we see that when using only a constant as the instrument, we are able to reject rationality for the mean, and for measures of centrality "close" to the mean, but we are unable to reject rationality for the median or mode and measures in a neighborhood of these. This is consistent with the entries in the first row of Table 3, which correspond directly to the three vertices in Figure 2. In the right panel of Figure 2, when the instrument set includes a constant

[19] We find that we are able to reject rationality of the mode in some other applications, discussed below, and so it is not the case that forecasts are always rationalizable as mode forecasts; i.e., the mode test power is not zero.

[20] The interpretation of this figure is slightly different to that of Figure 1: in that figure the shade of each dot indicated the proportion of times, across simulations, that point was included in the confidence set, allowing us to study the finite-sample coverage rates of our procedure. In Figure 2 the shade of each dot indicates whether, for this sample, that point is included in the 90% confidence set, the 95% confidence set, or is outside the 95% confidence set, the latter indicating a rejection of rationality at the 5% significance level.

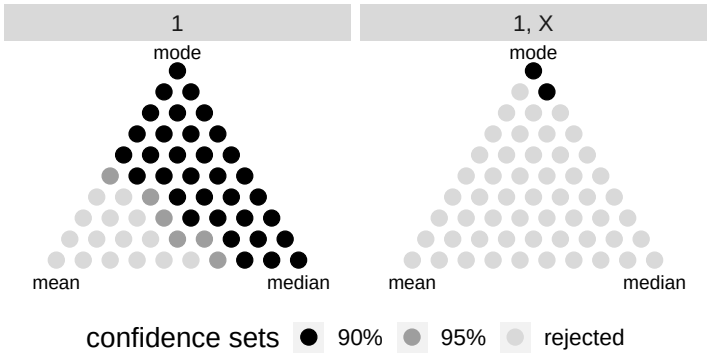


Figure 2: Confidence sets for income survey forecasts. This figure shows the measures of centrality that rationalize the New York Federal Reserve income survey forecasts. Black dots indicate that the measure is inside the Stock-Wright 90% confidence set, dark grey dots indicate that the measure is inside the 95% confidence set, light grey dots indicate that rationality for that measure can be rejected at the 5% level. The left panel uses just a constant as the instrument; the right panel uses a constant and the forecast.

and the forecast, we see that only the mode and centrality measures very close to the mode are included in the confidence set; all other forecasts can be rejected at the 5% level.

When interpreting the results in Figure 2, and similar figures below, it is worth keeping in mind that the power to detect sub-optimal forecasts is not uniform across values of θ: the mean and median are estimable at rate $T^{1/2}$, while the mode is only estimable at rate approximately $T^{2/7}$. This implies that for comparably sub-optimal forecasts, power will be lower at the mode vertex than at the mean or median vertices. This unavoidable variation in power means that the information conveyed by inclusion in the confidence set differs across values of θ.[21]

Overall, we conclude that the responses to the New York Fed's income survey, taken on aggregate, are consistent with rationality when interpreted as mode forecasts, but not when interpreted as forecasts of the mean, median, or convex combinations of these measures.

5.2 Heterogeneity in individual income forecasts

The analysis of individual income survey forecasts above used $3,916$ pairs of forecasts and realizations from a total of $2,628$ unique survey respondents. This naturally raises the question of whether

[21] Stock and Wright (2000) suggest caution when interpreting a small but nonempty confidence set, as such an outcome is consistent with both a correctly-specified model (a rational forecast, in our case) estimated precisely and also with a misspecified model (irrational forecast) facing low power. These two interpretations clearly have very different economic implications, but cannot be disentangled empirically. Given the lower power at the mode vertex, the latter explanation may be relevant here.

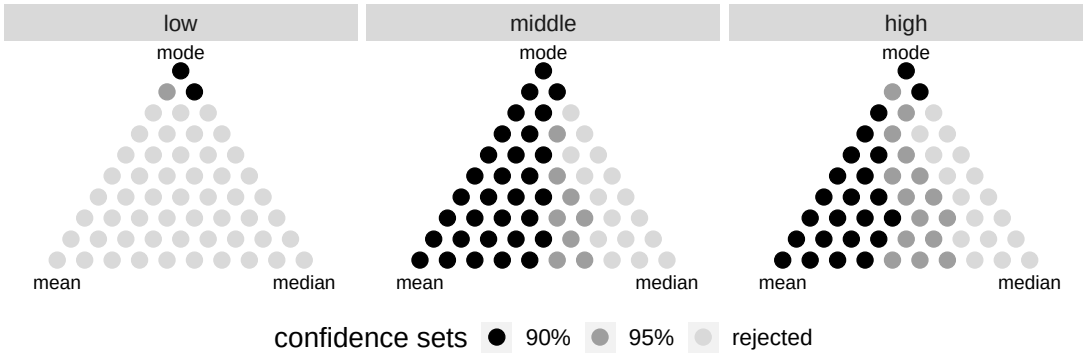


Figure 3: Confidence sets for income survey forecasts, stratified by income. This figure shows the measures of centrality that rationalize the New York Federal Reserve income survey forecasts, for low-, middle- and high-income respondents. Groups are formed using terciles of lagged reported income. Black dots indicate that the measure is inside the Stock-Wright 90% confidence set, dark grey dots indicate that the measure is inside the 95% confidence set, light grey dots indicate that rationality for that measure can be rejected at the 5% level. All panels use a constant and the forecast as test instruments.

there is heterogeneity in the measure of centrality used by different respondents. Given that our survey respondents generally only appear once or twice in our sample, allowing for arbitrary heterogeneity is not empirically feasible, however by exploiting other information on covariates contained in the survey we may shed some light on this question.

Firstly, we consider stratifying our sample by income. This is motivated by the possibility that, in addition to a different *level* of future income, low-income respondents face a different *shape* of future income, compared with high-income respondents. This analysis may also reveal that respondents at different income levels use different centrality measures to summarize their predictive distribution. Figure 3 presents confidence sets for forecast rationality of measures of centrality for low-, middle-, and high-income respondents based on terciles of the distribution of reported income.[22] We see that for low-income respondents, only the mode and measures very close to the mode are contained in the confidence set. For middle- and high-income respondents, the mode, mean, and centrality measures "close" to the mean and mode are included in the confidence set. This finding is consistent with all respondents using the mode, and only the distribution for low-income respondents allowing for separate identification of the mode and the mean. It is also consistent with low-income respondents reporting the mode as their forecast, while other respondents report the mean.

[22] Qualitatively similar results are found if we stratify the sample into just two groups based on median income.

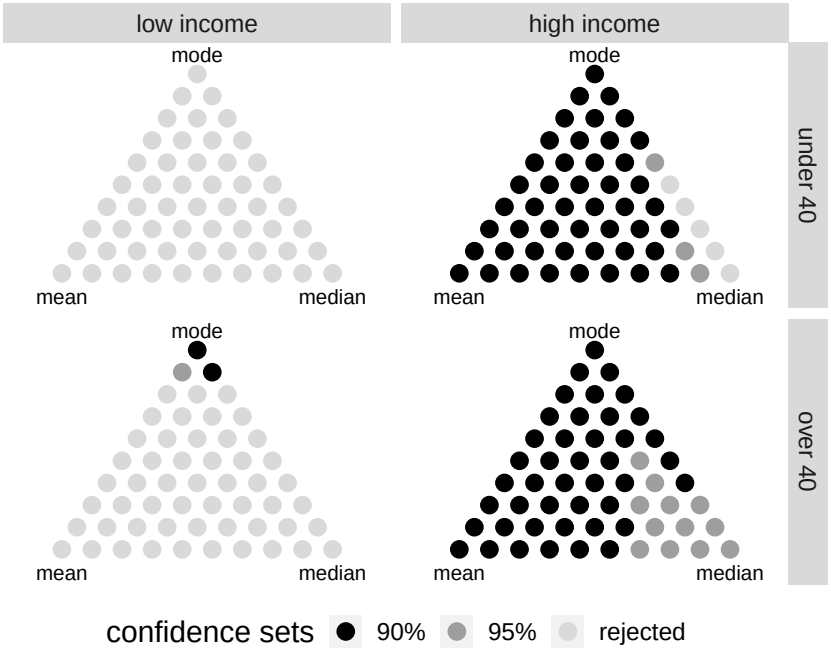


Figure 4: Confidence sets for income survey forecasts, stratified by income and age. This figure shows the measures of centrality that rationalize the New York Federal Reserve income survey forecasts, for low- and high-income respondents who are below or above the age of 40. Income groups are formed using the median lagged reported income. Black dots indicate that the measure is inside the Stock-Wright 90% confidence set, dark grey dots indicate that the measure is inside the 95% confidence set, light grey dots indicate that rationality for that measure can be rejected at the 5% level. All panels use a constant and the forecast as test instruments.

Next we consider a two-way sort, where we stratify the sample both by income and by age.[23] Stratifying by age is motivated by the possibility that younger respondents have less experience in the workforce, and may be less able to predict future earnings. Figure 4 presents the striking result that forecasts from younger low-income workers cannot be rationalized using *any* measure of centrality; all points lie outside the 95% confidence set. Forecasts from low-income older respondents can be rationalized as only mode forecasts. In contrast, we find that high-income workers, regardless of age, can be rationalized using almost all measures of centrality; we reject rationality only for measures close to the median. This figure suggests that younger low-income workers have difficulty predicting their income over the coming four months, and make systematic errors when doing so.

[23]The latest versions of survey contain a field for whether the respondent is above or below 40 years of age, while earlier versions asked for the respondent's specific age. To maximize the coverage, we adopt the age 40 split contained in the later versions of the survey. We have 2,457 (1,332) respondents over (under) age 40. We use the median income within each age category to define "low" and "high" income groups.

Finally we consider a sort based on whether the respondent reported having received a job offer in the previous four months. Such respondents are more likely to change jobs in the coming period, and therefore face a more uncertain distribution of future income.[24] Figure 5 reveals that forecasts from low-income respondents who reported receiving a job offer in the past four months can only be rationalized as mode forecasts; rationality for all other centrality measures is rejected. Forecasts from low-income respondents who did not receive a job offer in the previous four months, thereby presumably having more predictable future earnings, are rationalizable as mean, mode and centrality measures between these vertices. Similar results are found for high-income respondents who received a job offer, while for high-income respondents with no recent job offer, nearly all measures of central tendency can be rationalized.

Overall, the results in this section indicate some important heterogeneity in both the rationality of point forecasts and the measure of central tendency employed. We find that forecasts from younger, low-income survey respondents cannot be rationalized using any measure of central tendency, and forecasts from low-income survey respondents who are likely to change jobs in the coming period can only be rationalized as mode forecasts. In contrast, forecasts from respondents with income above the median, regardless of their age or likelihood of changing jobs, can be rationalized using many different centrality measures.

5.3 Greenbook forecasts of U.S. GDP

We now consider one-quarter-ahead forecasts of U.S. GDP growth produced by the staff of the Board of Governors of the Federal Reserve (the so-called "Greenbook" forecasts), from 1967Q2 until 2015Q2, a total of 192 observations.[25] These forecasts are prepared in preparation for each meeting of the Federal Open Market Committee, and substantial resources are devoted to them, see e.g. Romer and Romer (2000). Greenbook forecasts are available several times each quarter; for this analysis we take the single forecast closest to the middle date in each quarter. Broadly similar results are found when using the first, or last, forecast within each quarter.

[24] We find very similar results when we segment respondents by their estimated probability of receiving a job offer in the next four months, or by their estimated probability of staying in the same job in the next four months.

[25] Greenbook forecasts are only available to the public after a five-year lag.

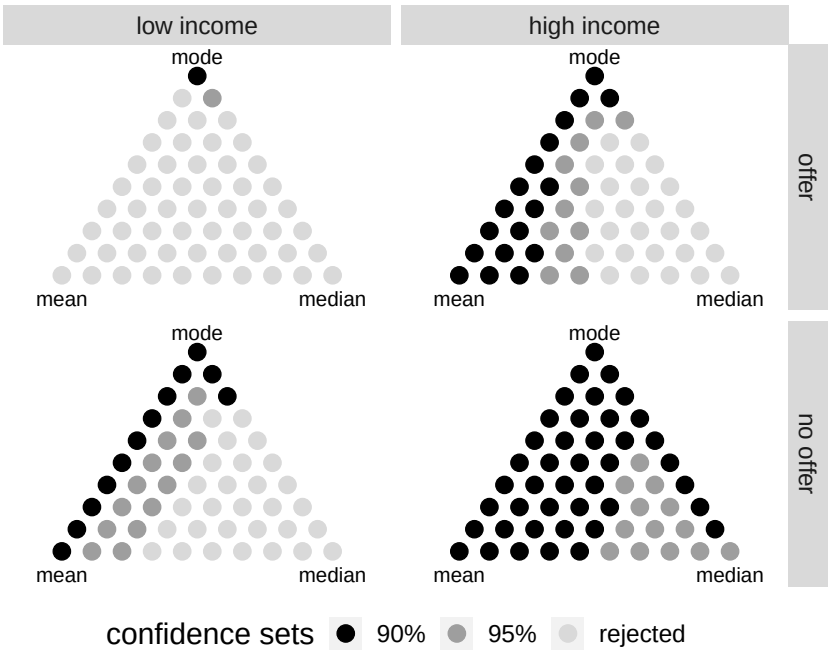


Figure 5: Confidence sets for income survey forecasts, stratified by income and job offer. This figure shows the measures of centrality that rationalize the New York Federal Reserve income survey forecasts, for low- and high-income respondents in the private sector or not. Groups are formed using the median lagged reported income, and whether or not the respondent reported receiving at least one job offer in the past four months. Black dots indicate that the measure is inside the Stock-Wright 90% confidence set, dark grey dots indicate that the measure is inside the 95% confidence set, light grey dots indicate that rationality for that measure can be rejected at the 5% level. All panels use a constant and the forecast as test instruments.

Figure 6 presents the confidence set for the measures of centrality that can be rationalized for these forecasts. As GDP growth is measured with error and official values are often revised, we present results for three different "vintages" of the realized value: the first, second and most recent release. For the first and second vintages, we see that only measures of centrality "close to" the mean can be rationalized as optimal, while the mode, median and similar measures can all be rejected. This is particularly noteworthy given the known lower power at the mode vertex. Using the most recent vintage for GDP growth, both the mean and median, and centrality measures between and near those, are included in the confidence set. That the Greenbook GDP forecasts are rational when interpreted as mean forecasts, but not when taken as mode or median forecasts, is consistent with the Fed staff using econometric models for these forecasts, as such models almost invariably focus on the mean.[26]

[26] Reifschneider and Tulip (2019) discuss the ambiguity in the specific centrality measure reported in the Greenbook

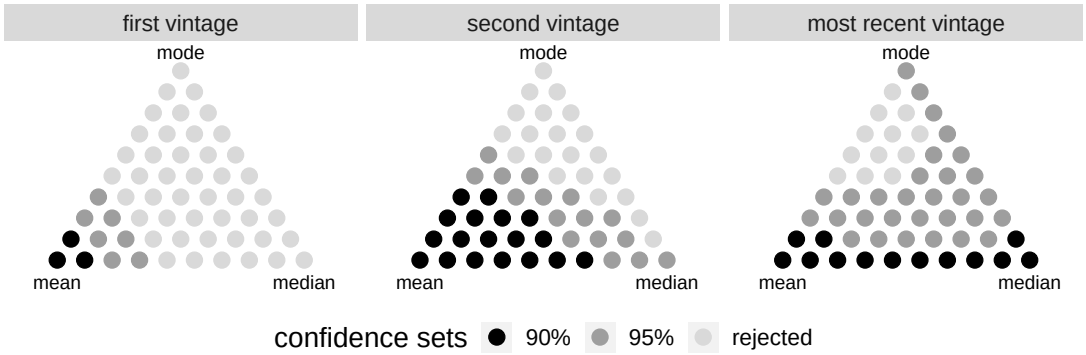


Figure 6: Confidence sets for Greenbook GDP forecasts. This figure shows the measures of centrality that rationalize the Federal Reserve Board's "Greenbook" forecasts of U.S. GDP growth. The three panels use three different measures of GDP growth in a given quarter. Black dots indicate that the measure is inside the Stock-Wright 90% confidence set, dark grey dots indicate that the measure is inside the 95% confidence set, light grey dots indicate that rationality for that measure can be rejected at the 5% level. All panels use a constant and the forecast as test instruments.

5.4 Random walk forecasts of exchange rates

For our final empirical application we revisit the famous result of Meese and Rogoff (1983), that exchange rate movements are approximately unpredictable when evaluated by the squared-error loss function, implying that the lagged exchange rate is an optimal mean forecast. See Rossi (2013) for a more recent survey of the literature on forecasting exchange rates. We use daily data from the European Central Bank's "Statistical Data Warehouse" on the USD/EUR, JPY/EUR and AUD/EUR exchange rates, over the period January 2000 to July 2020, a total of $5,265$ trading days. Note that our sample period has no overlap with that of Meese and Rogoff (1983), and so their conclusions about the mean-optimality of the random walk forecast need not hold in our data.

Figure 7 presents the results of our tests for rationality, all of which use a constant and the forecast as the instrument set. The middle and right panels reveal that for the JPY/EUR and AUD/EUR exchange rates the lagged exchange rate is not rejected as a mean forecast, while it is rejected when taken as a mode or median forecast. Thus the rationality of the random walk forecast critically depends, for these exchange rates, on whether it is interpreted as a mean, median or mode forecast. For the USD/EUR exchange rate we cannot reject rationality with respect to *any* of convex combination of these measures of central tendency, implying that the random walk forecast

forecasts, but write that they are "typically viewed as modal forecasts" by the Federal Reserve staff. Our results suggest that they are better interpreted as mean forecasts.

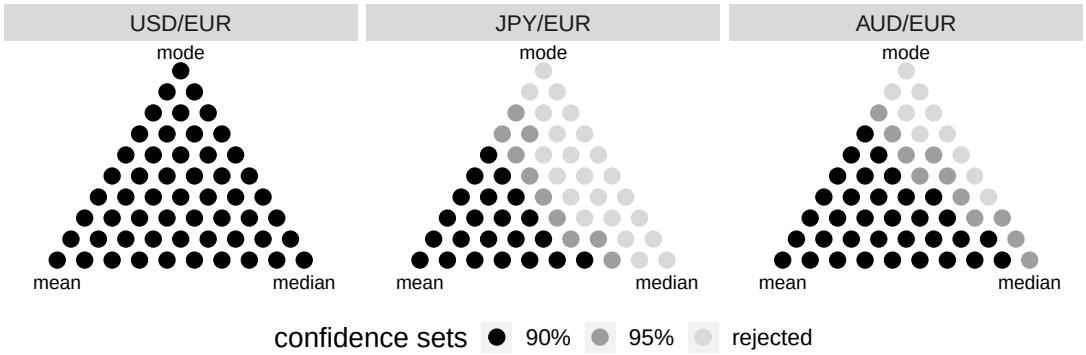


Figure 7: Confidence sets for random walk forecasts of exchange rates. This figure shows the measures of centrality that rationalize the random walk forecast of daily exchange rates movements. Black dots indicate that the measure is inside the Stock-Wright 90% confidence set, dark grey dots indicate that the measure is inside the 95% confidence set, light grey dots indicate that rationality for that measure can be rejected at the 5% level. All panels use a constant and the forecast as test instruments.

is consistent with rationality under any of these measures.[27] The mean vertex being included in the confidence set for all three exchange rates, indicating no evidence against rationality of the random walk model when interpreted as a mean forecast, is consistent with the conclusion of Meese and Rogoff (1983).

6 Conclusion

Reasonable people can interpret a request for their prediction of a random variable in a variety of ways. Some, including perhaps most economists, will report their expectation of the value of the variable (i.e., the mean of their predictive distribution), others might report the value such that the observed outcome is equally likely to be above or below it (i.e., the median), and others may report the value most likely to be observed (i.e, the mode). Still others might solve a loss minimization problem and report a forecast that is not a measure of central tendency. Economic surveys generally request a *point* forecast, despite calls for surveys to solicit distributional forecasts, see Manski (2004) for example, and the specific type of point forecast (mean, median, etc.) to be reported is generally not made explicit in the survey.

This paper proposes new methods to test the rationality of forecasts of some unknown measure

[27] Results for the GBP/EUR and CAD/EUR exchange rates are identical to those for the USD/EUR.

of central tendency. Similar to Elliott et al. (2005), we propose a testing framework that nests the mean forecast as a special case, but unlike that paper we allow for alternative forecasts within a general class of measures of central tendency, rather than measures that represent other aspects of the predictive distribution (such as non-central quantiles or expectiles). We consider the class of central tendency measures generated by convex combinations of the mean, median and mode.

We face an identification problem in economic applications, as the weights in the convex combination of centrality measures may be strongly, partially, weakly, or un-identified, depending on the presence and strength of asymmetry in the predictive distribution of the target variable. Even strongly asymmetric variables admit only partial identification of the weight vector if they belong to a location-scale family, which is a common assumption for economic time series. We overcome this using the work of Stock and Wright (2000), and obtain confidence sets that contain the measures of centrality that "rationalize" a given sequence of forecasts and realizations.

As a building block for the above tests, we also present new tests for the rationality of mode forecasts. Mode regression has received some attention in the recent literature (see, e.g., Kemp and Silva, 2012 and Kemp et al., 2020), however tests for mode forecast rationality similar to those available for the mean and median (e.g., Mincer and Zarnowitz, 1969 and Gaglianone et al., 2011) are lacking. Direct analogs of existing tests are infeasible because the mode is not elicitable (Heinrich, 2014). We introduce the concept of *asymptotic elicitability* and show it applies to the mode by considering a generalized modal midpoint with asymptotically vanishing length. We then present results that allow for tests similar to the famous Mincer-Zarnowitz regression for mean forecasts.

We apply our tests in three economic forecasting applications. Using individual income expectations survey data collected by the New York Federal Reserve, we reject forecast rationality with respect to the mean or median, however we cannot reject rationality when interpreting these as mode forecasts. We also find evidence of heterogeneity in this sample: for example, forecasts from younger, low-income respondents are not rationalizable using *any* measure of centrality, while forecasts from high-income respondents, regardless of their age, are rationalizable for many, though not all, measures of centrality. Next we study the Federal Reserve's "Greenbook" forecasts of U.S. GDP, and we find that we cannot reject rationality with respect to the mean, however we can reject

with respect to the median and mode. This is consistent with the Greenbook forecasts being made using econometric models, which almost invariably focus on the mean as the measure of central tendency. Finally, we revisit the famous result of Meese and Rogoff (1983) that random walk forecasts for exchange rates are rational mean forecasts. For the USD/EUR exchange rate, we find we cannot reject rationality with respect to *any* of convex combination of the mean, median and mode, indicating the random walk forecast is rational under any of these measures. For the JPY/EUR and AUD/EUR exchange rates, however, we find the random walk forecast is only rational for centrality measures "close" to the mean; rationality with respect to the median and mode is rejected.

A Proofs

Proof of Theorem 2.4. For the proof of statement (a), let $Y \sim P \in \mathcal{P}$ with density f, define $\tilde{K}_\delta(e) = \frac{1}{\delta}K\left(\frac{e}{\delta}\right)$ and introduce the notation $\bar{L}_\delta^K(x,P) = \mathbb{E}_{Y\sim P}\left[L_\delta^K(x,Y)\right]$. Then, it holds that

$$\bar{L}_\delta^K(x,P) = -\int \frac{1}{\delta}K\left(\frac{x-y}{\delta}\right) f(y)\,\mathrm{d}y = -\int \tilde{K}_\delta\left(x-y\right) f(y)\,\mathrm{d}y = -(f * \tilde{K}_\delta)(x),$$

where $f * \tilde{K}_\delta$ denotes the convolution of the functions f and $\tilde{K}_\delta$. Ibragimov (1956) shows that for any log-concave density, its convolution with any other unimodal distribution function is again unimodal.[28] Thus, $\bar{L}_\delta^K(x,P)$ exhibits a unique minima which shows that Γ_δ^K is well-defined by (2.6).

We continue with statement (b) and show that $\Gamma_\delta(P) \to \mathrm{Mode}(P)$ for all $P \in \mathcal{P}$. Notice that

$$\bar{L}_\delta^K(x,P) = -\int \frac{1}{\delta}K\left(\frac{x-y}{\delta}\right) f(y)\,\mathrm{d}y = -\int f(x+u\delta)K\left(u\right)\,\mathrm{d}u = -\int f(x+u\delta)\,\mathrm{d}\mathcal{K}(u), \quad \text{(A.1)}$$

by applying integration by substitution and by interpreting the kernel $K(\cdot)$ as the density of the probability measure $\mathcal{K}$. It holds that

$$\sup_{x\in\mathbb{R}}\left|\bar{L}_\delta^K(x,P) - \big(-f(x)\big)\right| = \sup_{x\in\mathbb{R}}\left|f(x) - \int f(x+u\delta)\mathrm{d}\mathcal{K}(u)\right| \quad \text{(A.2)}$$

$$\leq \sup_{x\in\mathbb{R}}\int \left|f(x) - f(x+u\delta)\right|\mathrm{d}\mathcal{K}(u) \leq \sup_{x\in\mathbb{R}}\int \left|c\delta u\right|\mathrm{d}\mathcal{K}(u) = c\delta\int |u|K(u)\mathrm{d}u \to 0 \quad \text{(A.3)}$$

[28] Ibragimov (1956) calls densities satisfying this property *strongly unimodal*. It is important to note that his notion of strong unimodality is different from ours introduced in Definition 2.1.

as $\delta \to 0$ as f is Lipschitz continuous with constant $c \geq 0$ and $\int |u|K(u)\mathrm{d}u < \infty$. Hence, $\bar{L}_\delta^K(x, P)$ converges uniformly (for all $x \in \mathbb{R}$) to $-f(x)$ as $\delta \to 0$ and it holds that $\arg\min_x \bar{L}_\delta^K(x, P) \to \arg\min_x(-f(x))$ as $\delta \to 0$. Consequently,

$$\lim_{\delta \to 0} \Gamma_\delta(P) = \lim_{\delta \to 0} \left(\arg\min_{x \in \mathbb{R}} \bar{L}_\delta^K(x, P) \right) = \arg\min_{x \in \mathbb{R}} \left(\lim_{\delta \to 0} \bar{L}_\delta^K(x, P) \right) = \arg\min_{x \in \mathbb{R}} \left(-f(x) \right),$$

which equals the mode for distributions with continuous Lebesgue density by definition.

For the proof of statement (c), we consider a fixed $\delta > 0$ and define $\bar{V}_\delta^K(x, P) := -\mathbb{E}_{Y \sim P}[V_\delta(x, Y)] = \frac{1}{\delta^2} \int K'\left(\frac{x-y}{\delta}\right) f(y)\mathrm{d}y$ and $\tilde{K}'_\delta(e) = \frac{1}{\delta^2} K'\left(\frac{e}{\delta}\right)$. Then,

$$\begin{aligned}
\bar{V}_\delta^K(x, P) &= \int \tilde{K}'_\delta(x - y) f(y) dy = \left(\tilde{K}'_\delta * f\right)(x) = \left(\tilde{K}_\delta * f'\right)(x) = \int \tilde{K}_\delta(x - y) f'(y) dy && \text{(A.4)} \\
&= \int \tilde{K}_\delta(x + \mathrm{Mode}(P) - y) f'(y - \mathrm{Mode}(P)) dy && \text{(A.5)} \\
&= \int \tilde{K}_\delta(x + \mathrm{Mode}(P) - y) g'(y) dy && \text{(A.6)}
\end{aligned}$$

for some shifted density g with mode at zero. As the kernel $\tilde{K}_\delta$ is log-concave it has a monotone likelihood ratio (Proposition 2.3 (b) in Saumard and Wellner, 2014), i.e. for any $a \leq b$ and $y \geq 0$, it holds that $\frac{\tilde{K}_\delta(a-y)}{\tilde{K}_\delta(a)} \leq \frac{\tilde{K}_\delta(b-y)}{\tilde{K}_\delta(b)}$ and as $g'(y) \leq 0$ for $y \geq 0$, this implies that

$$\frac{\tilde{K}_\delta(a - y)}{\tilde{K}_\delta(b - y)} g'(y) \geq \frac{\tilde{K}_\delta(a)}{\tilde{K}_\delta(b)} g'(y). \tag{A.7}$$

Analogously, the same inequality follows for any $a \leq b$ and $y \leq 0$, where the monotone likelihood ratio above holds with the reverse inequality but at the same time $g'(y) \geq 0$. Thus, for any $x > z$,

$$\begin{aligned}
\bar{V}_\delta^K(x, P) &= \int \tilde{K}_\delta(x + \mathrm{Mode}(P) - y) g'(y) dy && \text{(A.8)} \\
&= \int \frac{\tilde{K}_\delta(x + \mathrm{Mode}(P) - y)}{\tilde{K}_\delta(z + \mathrm{Mode}(P) - y)} \tilde{K}_\delta(z + \mathrm{Mode}(P) - y) g'(y) dy && \text{(A.9)} \\
&\geq \frac{\tilde{K}_\delta(x + \mathrm{Mode}(P))}{\tilde{K}_\delta(z + \mathrm{Mode}(P))} \int \tilde{K}_\delta(z + \mathrm{Mode}(P) - y) g'(y) dy && \text{(A.10)} \\
&= \frac{\tilde{K}_\delta(x + \mathrm{Mode}(P))}{\tilde{K}_\delta(z + \mathrm{Mode}(P))} \bar{V}_\delta^K(z, P). && \text{(A.11)}
\end{aligned}$$

For any root x^* such that $\bar{V}_\delta^K(x^*, P) = 0$, it follows that $\bar{V}_\delta^K(x_u, P) \geq 0$ for all $x_u > x^*$ and $\bar{V}_\delta^K(x_l, P) \leq 0$ for all $x_l < x^*$ as the kernel $\tilde{K}_\delta$ has infinite support. If there exist two distinct roots $x_l^* < x_u^*$ of $\bar{V}_\delta^K(\cdot, P)$, the above implies that $\bar{V}_\delta^K(x, P) = 0$ for all $x \in [x_l^*, x_u^*]$, which implies that the roots of $\bar{V}_\delta^K(\cdot, P)$ constitute a closed interval. As $\bar{L}_\delta^K(x, P)$ has a unique maximum for all $P \in \tilde{\mathcal{P}} \subset \mathcal{P}$ by part (a) and $\bar{V}_\delta^K(x, P) = -\frac{\partial}{\partial x}\bar{L}_\delta^K(x, P)$, it follows that $\bar{V}_\delta^K(x, P)$ must have a unique root, which implies that $V_\delta^K(x, Y)$ is a strict identification function for the generalized modal midpoint. □

Proof of Theorem 2.6. We define

$$g_{t,T} := \delta_T^{3/2} T^{-1/2} \psi(Y_{t+1}, X_t, \boldsymbol{h}_t, \delta_T) = -(T\delta_T)^{-1/2} K'\left(\frac{X_t - Y_{t+1}}{\delta_T}\right)\boldsymbol{h}_t, \tag{A.12}$$

$$g_{t,T}^e := \mathbb{E}_t[g_{t,T}], \qquad \text{and} \qquad g_{t,T}^* := g_{t,T} - g_{t,T}^e, \tag{A.13}$$

such that

$$\delta_T^{3/2} T^{-1/2} \sum_{t=1}^T \psi(Y_{t+1}, X_t, \boldsymbol{h}_t, \delta_T) = \sum_{t=1}^T g_{t,T}^e + \sum_{t=1}^T g_{t,T}^*. \tag{A.14}$$

Lemma S.1.1 shows that $\sum_{t=1}^T g_{t,T}^e \xrightarrow{P} 0$. Thus, it remains to shows that $\sum_{t=1}^T g_{t,T}^* \xrightarrow{d} \mathcal{N}\big(0, \Omega_{\text{Mode}}\big)$. For some arbitrary, but fixed $\lambda \in \mathbb{R}^k, ||\lambda||_2 = 1$, we define

$$z_{t,T} := \lambda^\top g_{t,T}^*, \qquad \bar{\omega}_T^2 := \sum_{t=1}^T \text{Var}(z_{t,T}), \qquad \omega^2 := \lambda^\top \Omega_{\text{Mode}} \lambda, \qquad \text{and} \qquad h_{t,T} := \frac{z_{t,T}}{\bar{\omega}_T} \tag{A.15}$$

and show that a univariate CLT for martingale difference sequences (MDS) holds for $\sum_{t=1}^T h_{t,T}$. It obviously holds that $\big(g_{t,T}^*, \mathcal{F}_{t+1}\big)$ is a MDS as $g_{s,T}^* \in \mathcal{F}_{t+1}$ for all $s \leq t$ and $\mathbb{E}_t\big[g_{t,T}^*\big] = 0$ a.s. by definition. Thus, $\big(z_{t,T}, \mathcal{F}_{t+1}\big)$ and $\big(h_{t,T}, \mathcal{F}_{t+1}\big)$ are also MDS.

In the following, we verify the following three conditions of Theorem 24.3 of Davidson (1994): (a) $\sum_{t=1}^T \text{Var}(h_{t,T}) = 1$, (b) $\sum_{t=1}^T h_{t,T}^2 \xrightarrow{P} 1$, and (c) $\max_{1 \leq t \leq T} |h_{t,T}| \xrightarrow{P} 0$. Lemma S.1.2 shows that $\bar{\omega}_T^2 = \sum_{t=1}^T \text{Var}[z_{t,T}] \to \bar{\lambda}^\top \Omega_{\text{Mode}} \lambda = \omega^2$. Thus, for all T sufficiently large, $\bar{\omega}_T^2$ is strictly positive and hence, $h_{t,T}$ is well-defined and $\sum_{t=1}^T \text{Var}(h_{t,T}) = 1$, which shows condition (a). Lemma S.1.3 shows

that $\sum_{t=1}^{T} z_{t,T}^2 \xrightarrow{P} \bar{\omega}^2$ as $T \to \infty$ which implies condition (b), i.e. $\sum_{t=1}^{T} h_{t,T}^2 = \sum_{t=1}^{T} \frac{z_{t,T}^2}{\bar{\omega}_T^2} \xrightarrow{P} 1$. Eventually, Lemma S.1.4 shows condition (c) and we can apply Theorem 24.3 of Davidson (1994) in order to conclude that for all $\lambda \in \mathbb{R}^k, ||\lambda||_2 = 1$, it holds that $\sum_{t=1}^{T} h_{t,T} \xrightarrow{d} \mathcal{N}(0,1)$. As $\bar{\omega}_T^2 \to \bar{\omega}^2$, Slutsky's theorem implies that $\sum_{t=1}^{T} z_{t,T} = \sum_{t=1}^{T} \lambda^\top g_{t,T}^* \xrightarrow{d} \mathcal{N}\big(0, \bar{\omega}^2\big)$, and as this holds for all $\lambda \in \mathbb{R}^k, ||\lambda||_2 = 1$, we apply the Cramér-Wold theorem and get that $\sum_{t=1}^{T} g_{t,T}^* \xrightarrow{d} \mathcal{N}\big(0, \Omega_{\text{Mode}}\big)$, which concludes the proof of this theorem. □

Proof of Theorem 2.9. As in the proof of Theorem 2.6, we split $\sum_{t=1}^{T} g_{t,T} = \sum_{t=1}^{T} g_{t,T}^e + \sum_{t=1}^{T} g_{t,T}^*$. It holds again that $\sum_{t=1}^{T} g_{t,T}^* \xrightarrow{d} \mathcal{N}(0, \Omega_{\text{Mode}})$, as this part of the proof does not depend on assumptions on $f_t'(0)$ imposed in the respective null and alternative hypotheses. As in the proof of Lemma S.1.1, we obtain that

$$\sum_{t=1}^{T} g_{t,T}^e = T^{-1/2} \delta_T^{3/2} \sum_{t=1}^{T} \boldsymbol{h}_t \int K(u)\, f_t'(\delta_T u)\, \mathrm{d}u. \tag{A.16}$$

From a Taylor expansion of $f_t'(\delta_T u)$ around zero, we obtain that for some $\zeta \in [0,1]$, it holds that $f_t'(\delta_T u) = f_t'(0) + (\delta_T u) f_t''(0) + \frac{(\delta_T u)^2}{2} f_t'''(\zeta \delta_T u)$ and hence,

$$\begin{aligned}
\sum_{t=1}^{T} g_{t,T}^e \;=\; & \frac{1}{T} \sum_{t=1}^{T} (T\delta_T^3)^{1/2} \boldsymbol{h}_t f_t'(0) \int K(u)\, \mathrm{d}u && \text{(A.17)} \\
& + \frac{1}{T} \sum_{t=1}^{T} (T\delta_T^5)^{1/2} \boldsymbol{h}_t f_t''(0) \int u K(u)\, \mathrm{d}u && \text{(A.18)} \\
& + \frac{1}{T} \sum_{t=1}^{T} (T\delta_T^7)^{1/2} \boldsymbol{h}_t \int u^2 K(u)\, f_t'''(\zeta \delta_T u)\, \mathrm{d}u. && \text{(A.19)}
\end{aligned}$$

The second term equals zero as $\int u K(u)\, \mathrm{d}u = 0$ by assumption and the third term converges to zero as $\boldsymbol{h}_t$ is stationary and ergodic and thus a weak law of large numbers applies and furthermore, $T\delta_T^7 \to 0$ by assumption. In contrast, for the first term we get that $\int K(u)\, \mathrm{d}u = 1$ and as $|f_t'(0)|$ is bounded from above, $\frac{1}{T} \sum_{t=1}^{T} f_t'(0) \boldsymbol{h}_t \xrightarrow{P} \mathbb{E}[f_t'(0) \boldsymbol{h}_t]$. As $(T\delta_T^3)^{1/2} \to \infty$ by assumption and $\mathbb{E}[f_t'(0) \boldsymbol{h}_t] \neq 0$ under the alternative, we obtain that the first term diverges in probability. This implies that for any $c \in \mathbb{R}$, $\mathbb{P}\left(\left|\sum_{t=1}^{T} g_{t,T}^e\right| \geq c\right) \to 1$, and consequently also $\mathbb{P}\left(\left|\sum_{t=1}^{T} g_{t,T}\right| \geq c\right) \to 1$. As furthermore $\widehat{\Omega}_{T,\text{Mode}} \xrightarrow{P} \Omega_{\text{Mode}}$ by Theorem 2.7, which is uniformly positive definite and $J_T =$

$\left(\sum_{t=1}^{T} g_{t,T}\right)^{\top} \widehat{\Omega}_{T,\text{Mode}}^{-1} \left(\sum_{t=1}^{T} g_{t,T}\right)$, the conditions of Theorem 8.13 of White (1994) are satisfied and we can conclude that for any $c \in \mathbb{R}$, $\mathbb{P}\left(|J_T| \geq c\right) \to 1$, which concludes the proof of this theorem. □

Proof of Theorem 3.3. For all fixed $\lambda \in \mathbb{R}^k$ such that $||\lambda||_2 = 1$, we define

$$\begin{aligned}\sigma^2 = \lambda^{\top}\Sigma(\theta_0)\lambda = \mathbb{E}\Big[&\theta_{10}^2 \left(\boldsymbol{h}_t^{\top}\boldsymbol{W}_{\text{Mean}}\lambda\right)^2 \varepsilon_t^2 + \theta_{20}^2 \left(\boldsymbol{h}_t^{\top}\boldsymbol{W}_{\text{Med}}\lambda\right)^2 \left(\mathbb{1}_{\{\varepsilon_t>0\}} - \mathbb{1}_{\{\varepsilon_t<0\}}\right)^2 \\ &+ \theta_{30}^2 \left(\boldsymbol{h}_t^{\top}\boldsymbol{W}_{\text{Mode}}\lambda\right)^2 f_t(0) \int K'(u)^2 \mathrm{d}u \\ &+ 2\theta_{10}\theta_{20} \left(\boldsymbol{h}_t^{\top}\boldsymbol{W}_{\text{Mean}}\lambda\right)\left(\boldsymbol{h}_t^{\top}\boldsymbol{W}_{\text{Med}}\lambda\right) \varepsilon_t \left(\mathbb{1}_{\{\varepsilon_t>0\}} - \mathbb{1}_{\{\varepsilon_t<0\}}\right)\Big],\end{aligned} \tag{A.20}$$

and $\sigma_T^2 = \sum_{t=1}^{T} \text{Var}\left(\phi_{t,T}^*(\theta_0)\lambda\right)$ for the MDS $\phi_{t,T}^*(\theta_0)$. Lemma S.1.5 shows that $\sigma_T^2 \to \sigma^2$ and thus, σ_T^2 is strictly positive for T large enough and consequently, $\sigma_T^{-1}\sum_{t=1}^{T}\phi_{t,T}^*(\theta_0)\lambda$ is well-defined. In the following, we first show that $\sigma_T^{-1}\sum_{t=1}^{T}\phi_{t,T}^*(\theta_0)\lambda \xrightarrow{d} \mathcal{N}\left(0,1\right)$ by applying Theorem 24.3 in Davidson (1994) and by verifying that the respective conditions hold (with $X_{t,T} = \sigma_T^{-1}\phi_{t,T}^*(\theta_0)\lambda$). Lemma S.1.6 shows that $\sum_{t=1}^{T}\left(\phi_{t,T}^*(\theta_0)\lambda\right)^2 \xrightarrow{P} \sigma^2$, which implies condition (a) of Theorem 24.3 of Davidson (1994), i.e. $\sigma_T^{-2}\sum_{t=1}^{T}\left(\phi_{t,T}^*(\theta_0)\lambda\right)^2 \xrightarrow{P} 1$. Lemma S.1.7 shows condition (b), i.e. $\max_{t=1,\ldots,T}\left|\sigma_T^{-1}\phi_{t,T}^*(\theta_0)\lambda\right| \xrightarrow{P} 0$. Thus, we can apply Theorem 24.3 of Davidson (1994) and conclude that $\sigma_T^{-1}\sum_{t=1}^{T}\phi_{t,T}^*(\theta_0)\lambda \xrightarrow{d} \mathcal{N}\left(0,1\right)$. As $\sigma_T^2 \to \sigma^2$, Slutsky's theorem implies that $\sum_{t=1}^{T}\phi_{t,T}^*(\theta_0)\lambda \xrightarrow{d} \mathcal{N}(0,\sigma^2)$ and as this holds for all $\lambda \in \mathbb{R}^k$ such that $||\lambda||_2 = 1$, we can conclude that $\sum_{t=1}^{T}\phi_{t,T}^*(\theta_0) \xrightarrow{d} \mathcal{N}\left(0,\Sigma(\theta_0)\right)$. Furthermore, $||\sum_{t=1}^{T}\phi_{t,T}^*(\theta_0) - \tilde{\phi}_{t,T}(\theta_0)|| = ||\sum_{t=1}^{T} u_{t,T}(\theta_0)|| \xrightarrow{P} 0$ by Assumption 3.1 and

$$\sum_{t=1}^{T}||\tilde{\phi}_{t,T}(\theta) - \phi_{t,T}(\theta)|| = \sum_{t=1}^{T}\left|\left|\theta \cdot \begin{pmatrix} T^{-1/2}\boldsymbol{h}_t^{\top}\left(\widehat{\boldsymbol{W}}_{T,\text{Mean}} - \boldsymbol{W}_{\text{Mean}}\right)\varepsilon_t \\ T^{-1/2}\boldsymbol{h}_t^{\top}\left(\widehat{\boldsymbol{W}}_{T,\text{Med}} - \boldsymbol{W}_{\text{Med}}\right)\left(\mathbb{1}_{\{\varepsilon_t>0\}} - \mathbb{1}_{\{\varepsilon_t<0\}}\right) \\ T^{-1/2}\boldsymbol{h}_t^{\top}\left(\widehat{\boldsymbol{W}}_{T,\text{Mode}} - \boldsymbol{W}_{\text{Mode}}\right)\delta_T^{-1/2}K'\left(\frac{-\varepsilon_t}{\delta_T}\right) \end{pmatrix}\right|\right| \xrightarrow{P} 0$$

as it holds that $\widehat{\boldsymbol{W}}_{T,\text{Mean}} \xrightarrow{P} \boldsymbol{W}_{\text{Mean}}$, $\widehat{\boldsymbol{W}}_{T,\text{Med}} \xrightarrow{P} \boldsymbol{W}_{\text{Med}}$ and $\widehat{\boldsymbol{W}}_{T,\text{Mode}} \xrightarrow{P} \boldsymbol{W}_{\text{Mode}}$ by assumption. Hence we can conclude that $\sum_{t=1}^{T}\phi_{t,T}(\theta_0) \xrightarrow{d} \mathcal{N}\left(0,\Sigma(\theta_0)\right)$, which concludes the proof of this theorem. □

References

Bank of England (2019). *Inflation Report, August 2019*. Available at https://www.bankofengland.co.uk/inflation-report/2019/august-2019.

Barendse, S. (2020). Efficiently weighted estimation of tail and interquartile expectations. Working Paper, Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2937665.

Bauer, M. D. and Rudebusch, G. D. (2016). Monetary policy expectations at the zero lower bound. *Journal of Money, Credit and Banking*, 48(7):1439–1465.

Bellini, F. and Bignozzi, V. (2015). On elicitable risk measures. *Quantitative Finance*, 15(5):725–733.

Charness, G. and Dufwenberg, M. (2006). Promises and partnership. *Econometrica*, 74(6):1579–1601.

Christoffersen, P. F. (1998). Evaluating interval forecasts. *International Economic Review*, 39(4):841–62.

Christoffersen, P. F. and Diebold, F. X. (1997). Optimal prediction under asymmetric loss. *Econometric Theory*, 13(06):808–817.

Davidson, J. (1994). *Stochastic Limit Theory: An Introduction for Econometricians*. Advanced Texts in Econometrics. OUP Oxford.

Dearborn, K. and Frongillo, R. (2020). On the indirect elicitability of the mode and modal interval. *Annals of the Institute of Statistical Mathematics*, 72:1095–1108.

Dufwenberg, M. and Gneezy, U. (2000). Measuring beliefs in an experimental lost wallet game. *Games and Economic Behavior*, 30(2):163–182.

Eddy, W. F. (1980). Optimum kernel estimators of the mode. *Annals of Statistics*, 8(4):870–882.

Elliott, G., Komunjer, I., and Timmermann, A. (2005). Estimation and testing of forecast rationality under flexible loss. *Review of Economic Studies*, 72(4):1107–1125.

Elliott, G., Komunjer, I., and Timmermann, A. (2008). Biases in macroeconomic forecasts: Irrationality or asymmetric loss? *Journal of the European Economic Association*, 6(1):122–157.

Elliott, G. and Timmermann, A. (2016). *Economic Forecasting*. Princeton University Press.

Eyting, M. and Schmidt, P. (2018). Belief elicitation with multiple point predictions. Working paper, Available at https://download.uni-mainz.de/RePEc/pdf/Discussion_Paper_1818.pdf.

Fissler, T. and Ziegel, J. F. (2016). Higher order elicitability and Osband's principle. *Annals of Statistics*, 44(4):1680–1707.

Gaglianone, W. P., Lima, L. R., Linton, O., and Smith, D. R. (2011). Evaluating value-at-risk models via quantile regression. *Journal of Business & Economic Statistics*, 29(1):150–160.

Giacomini, R. and White, H. (2006). Tests of conditional predictive ability. *Econometrica*, 74(6):1545–1578.

Gneiting, T. (2011). Making and evaluating point forecasts. *Journal of the American Statistical Association*, 106(494):746–762.

Granger, C. W. J. (1969). Prediction with a generalized cost of error function. *Journal of the Operational Research Society*, 20:199–207.

Heinrich, C. (2014). The mode functional is not elicitable. *Biometrika*, 101(1):245–251.

Hossain, T. and Okui, R. (2013). The binarized scoring rule. *Review of Economic Studies*, 80(3):984–1001.

Huber, P. J. (1964). Robust estimation of a location parameter. *Annals of Statistics*, 53(1):73–101.

Ibragimov, I. A. (1956). On the composition of unimodal distributions. *Theory of Probability and Its Applications*, 1:255–260.

Kemp, G. C., Parente, P. M., and Silva, J. S. (2020). Dynamic vector mode regression. *Journal of Business & Economic Statistics*, 38(3):647–661.

Kemp, G. C. and Silva, J. S. (2012). Regression towards the mode. *Journal of Econometrics*, 170(1):92 – 101.

Kirchkamp, O. and Reiß, J. P. (2011). Out-of-equilibrium bids in first-price auctions: Wrong expectations or wrong bids. *Economic Journal*, 121(557):1361–1397.

Kröger, S. and Pierrot, T. (2019). What point of a distribution summarises point predictions? Working Paper, Available at http://hdl.handle.net/10419/206533.

Li, Q. and Racine, J. S. (2006). *Nonparametric Econometrics: Theory and Practice*. Princeton University Press.

Manski, C. F. (2004). Measuring expectations. *Econometrica*, 72(5):1329–1376.

McClave, J. T., Benson, P. G., and Sincich, T. (2017). *Statistics for Business and Economics, 13th Edition*. Pearson, New York.

Meese, R. and Rogoff, K. (1983). Empirical exchange rate models of the seventies: Do they fit out of sample? *Journal of International Economics*, 14:3–24.

Mincer, J. and Zarnowitz, V. (1969). The Evaluation of Economic Forecasts. In *Economic Forecasts and Expectations: Analysis of Forecasting Behavior and Performance*, pages 3–46. National Bureau of Economic Research, Inc.

Nolde, N. and Ziegel, J. F. (2017). Elicitability and backtesting: Perspectives for banking regulation. *Annals of Applied Statistics*, 11(4):1833–1874.

Patton, A. J. (2020). Comparing possibly misspecified forecasts. *Journal of Business & Economic Statistics*, 38(4):796–809.

Patton, A. J. and Timmermann, A. (2007). Testing forecast optimality under unknown loss. *Journal of the American Statistical Association*, 102(480):1172–1184.

Peterson, C. and Miller, A. (1964). Mode, median, and mean as optimal strategies. *Journal of Experimental Psychology*, 68(4):363–367.

Reifschneider, D. and Tulip, P. (2019). Gauging the uncertainty of the economic outlook using historical forecasting errors: The Federal Reserve's approach. *International Journal of Forecasting*, 35:1564–1582.

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

Rossi, B. (2013). Exchange rate predictability. *Journal of Economic Literature*, 51(4):1063–1119.

Sapienza, P., Toldra-Simats, A., and Zingales, L. (2013). Understanding trust. *Economic Journal*, 123(573):1313–1332.

Saumard, A. and Wellner, J. A. (2014). Log-concavity and strong log-concavity: A review. *Statistics Surveys*, 8:45–114.

Stock, J. H. and Wright, J. H. (2000). GMM with weak identification. *Econometrica*, 68(5):1055–1096.

von Bahr, B. and Esseen, C.-G. (1965). Inequalities for the rth absolute moment of a sum of random variables, $1 \leq r \leq 2$. *Annals of Mathematical Statistics*, 36(1):299–303.

White, H. (1994). *Estimation, Inference and Specification Analysis*. Econometric Society Monographs. Cambridge University Press.

White, H. (2001). *Asymptotic Theory for Econometricians*. Academic Press, San Diego.

SUPPLEMENTARY MATERIAL FOR

Testing Forecast Rationality for Measures of Central Tendency

Timo Dimitriadis and Andrew J. Patton and Patrick W. Schmidt

This Version: September 28, 2020

S.1 Technical Proofs

Lemma S.1.1. *Given Assumption 2.5 and the null hypothesis in (2.8), it holds that* $\sum_{t=1}^{T} g_{t,T}^{e} \xrightarrow{P} 0$.

Proof. Applying integration by parts yields that

$$g_{t,T}^{e} = -\mathbb{E}_t\left[(T\delta_T)^{-1/2}K'\left(\frac{\varepsilon_t}{\delta_T}\right)\boldsymbol{h}_t\right] = -(T\delta_T)^{-1/2}\boldsymbol{h}_t\int K'\left(\frac{e}{\delta_T}\right)f_t(e)\,\mathrm{d}e \tag{S.1.1}$$

$$= T^{-1/2}\delta_T^{1/2}\boldsymbol{h}_t\int K\left(\frac{e}{\delta_T}\right)f_t'(e)\,\mathrm{d}e - T^{-1/2}\delta_T^{1/2}\boldsymbol{h}_t\left[K\left(\frac{e}{\delta_T}\right)f_t(e)\right]_{e=-\infty}^{e=\infty}. \tag{S.1.2}$$

As $\lim_{e\to\pm\infty}K(e)=0$ and f_t is bounded from above, the latter term is zero a.s. for all $T\in\mathbb{N}$. By transformation of variables, it further holds that

$$g_{t,T}^{e} = T^{-1/2}\delta_T^{1/2}\boldsymbol{h}_t\int K\left(\frac{e}{\delta_T}\right)f_t'(e)\,\mathrm{d}e = T^{-1/2}\delta_T^{3/2}\boldsymbol{h}_t\int K(u)\,f_t'(\delta_T u)\,\mathrm{d}u. \tag{S.1.3}$$

A Taylor expansion of $f_t'(\delta_T u)$ around zero is given by

$$f_t'(\delta_T u) = f_t'(0) + (\delta_T u)f_t''(0) + \frac{(\delta_T u)^2}{2}f_t'''(\zeta\delta_T u), \tag{S.1.4}$$

for some $\zeta\in[0,1]$ and $f_t'(0)=0$ holds under the null hypothesis specified in (2.8). Consequently,

$$\sum_{t=1}^{T} g_{t,T}^{e} = T^{-1/2}\delta_T^{5/2}\sum_{t=1}^{T} f_t''(0)\boldsymbol{h}_t\int uK(u)\,\mathrm{d}u \tag{S.1.5}$$

$$+\frac{1}{2}T^{-1/2}\delta_T^{7/2}\sum_{t=1}^{T}\boldsymbol{h}_t\int u^2K(u)\,f_t'''(\zeta\delta_T u)\,\mathrm{d}u. \tag{S.1.6}$$

As $\int uK(u)\,\mathrm{d}u=0$ by assumption (A5), the first term is zero for all $T\in\mathbb{N}$. As $\sup_x f_t'''(x)\le c$ by

Assumption (A4) and $\int u^2 K(u)\,\mathrm{d}u \leq c < \infty$ by Assumption (A5), we obtain

$$\frac{1}{2}T^{-1/2}\delta_T^{7/2}\sum_{t=1}^{T}\boldsymbol{h}_t\int u^2 K(u)\, f_t'''(\zeta\delta_T u)\,\mathrm{d}u \leq \frac{1}{2}c^2\left(T\delta_T^7\right)^{1/2}\frac{1}{T}\sum_{t=1}^{T}\boldsymbol{h}_t \xrightarrow{P} 0, \tag{S.1.7}$$

as $T\delta_T^7 \to 0$ for $T \to \infty$ by Assumption (A6) and $\frac{1}{T}\sum_{t=1}^{T}\boldsymbol{h}_t \xrightarrow{P} \mathbb{E}[\boldsymbol{h}_t]$ by a law of large numbers for stationary and ergodic sequences. The result of the lemma follows. □

Lemma S.1.2. *Given Assumption 2.5 and under the null hypothesis in (2.8), it holds that* $\sum_{t=1}^{T}\mathrm{Var}(z_{t,T}) \to \bar{\omega}^2 = \lambda^\top \Omega_{\mathrm{Mode}}\lambda$.

Proof. We first observe that $\mathrm{Var}(z_{t,T}) = \mathbb{E}\left[\left(\lambda^\top\left(g_{t,T} - g_{t,T}^e\right)\right)^2\right]$ as $\mathbb{E}\left[\lambda^\top\left(g_{t,T} - g_{t,T}^e\right)\right] = 0$. Hence,

$$\mathrm{Var}[z_{t,T}] = \mathbb{E}\left[\left(\lambda^\top g_{t,T}\right)^2\right] - \mathbb{E}\left[\left(\lambda^\top g_{t,T}^e\right)^2\right], \tag{S.1.8}$$

as $\mathbb{E}\left[\left(\lambda^\top g_{t,T}^e\right)\cdot\left(\lambda^\top g_{t,T}\right)\right] = \mathbb{E}\left[\left(\lambda^\top g_{t,T}^e\right)\cdot\mathbb{E}_t\left[\lambda^\top g_{t,T}\right]\right] = \mathbb{E}\left[\left(\lambda^\top g_{t,T}^e\right)^2\right]$. For the first term in (S.1.8), we obtain

$$\mathbb{E}\left[\left(\lambda^\top g_{t,T}\right)^2\right] = \mathbb{E}\left[(T\delta_T)^{-1}(\lambda^\top\boldsymbol{h}_t)^2\mathbb{E}_t\left[K'\left(\frac{X_t - Y_{t+1}}{\delta_T}\right)^2\right]\right] \tag{S.1.9}$$

$$= \mathbb{E}\left[(T\delta_T)^{-1}(\lambda^\top\boldsymbol{h}_t)^2\int K'\left(\frac{e}{\delta_T}\right)^2 f_t(e)\,\mathrm{d}e\right] \tag{S.1.10}$$

$$= \frac{1}{T}\,\mathbb{E}\left[(\lambda^\top\boldsymbol{h}_t)^2\int K'(u)^2\, f_t(\delta_T u)\,\mathrm{d}u\right]. \tag{S.1.11}$$

As $\delta_T \to 0$ when $T \to \infty$, we find

$$\sum_{t=1}^{T}\mathbb{E}\left[\left(\lambda^\top g_{t,T}\right)^2\right] \to \mathbb{E}\left[(\lambda^\top\boldsymbol{h}_t)^2 f_t(0)\right]\int K'(u)^2\,\mathrm{d}u = \lambda^\top\mathbb{E}\left[f_t(0)\boldsymbol{h}_t\boldsymbol{h}_t^\top\right]\lambda\int K'(u)^2\,\mathrm{d}u. \tag{S.1.12}$$

For the second term in (S.1.8), inserting the equality in (S.1.3) yields

$$\left(\lambda^\top g_{t,T}^e\right)^2 = \left(\delta_T^{3/2}T^{-1/2}(\lambda^\top\boldsymbol{h}_t)\int K'(u)\, f_t'(\delta_T u)\mathrm{d}u\right)^2 \leq \delta_T^3 T^{-1}||\lambda||^2||\boldsymbol{h}_t||^2\left|\int K'(u)\, f_t'(u\delta_T)\mathrm{d}u\right|^2.$$

As $\sup_x |f_t'(x)| \leq c$ and $\left|\int K'(u)\,\mathrm{d}u\right| \leq c$ by assumption, it holds that

$$\sum_{t=1}^{T} \mathbb{E}\left[\left(\lambda^\top g_{t,T}^e\right)^2\right] \leq \delta_T^3 c^2 ||\lambda||^2 \left(\frac{1}{T}\sum_{t=1}^{T} \mathbb{E}\left[||\boldsymbol{h}_t||^2\right]\right) \to 0, \tag{S.1.13}$$

as $\delta_T^3 \to 0$ as $T \to \infty$. The result of the lemma follows by combining (S.1.12) and (S.1.13). □

Lemma S.1.3. *Given Assumption 2.5 and under the null hypothesis in (2.8), it holds that* $\sum_{t=1}^{T} z_{t,T}^2 \xrightarrow{P} \bar{\omega}^2 = \lambda^\top \Omega_{\text{Mode}} \lambda$.

Proof. We define

$$h_{1,T} := \sum_{t=1}^{T} \left(z_{t,T}^2 - \mathbb{E}_t\left[z_{t,T}^2\right]\right) \qquad \text{and} \qquad h_{2,T} := \sum_{t=1}^{T} \mathbb{E}_t\left[z_{t,T}^2\right] - \bar{\omega}^2, \tag{S.1.14}$$

such that $\sum_{t=1}^{T} z_{t,T}^2 - \bar{\omega}^2 = h_{1,T} + h_{2,T}$. We first show that $h_{1,T} \xrightarrow{L_p} 0$ for some $1 < p < 2$ sufficiently small enough and thus $h_{1,T} \xrightarrow{P} 0$. For this, first notice that $z_{t,T}^2 - \mathbb{E}_t\left[z_{t,T}^2\right]$ is a $\mathcal{F}_{t+1}$-MDS by definition. Thus, we can apply the von Bahr and Esseen (1965)-inequality for some $p \in (1,2)$ for MDS (in the first line) in order to conclude that

$$\mathbb{E}\left[|h_{1,T}|^p\right] = \mathbb{E}\left[\left|\sum_{t=1}^{T} z_{t,T}^2 - \mathbb{E}_t\left[z_{t,T}^2\right]\right|^p\right] \leq 2\sum_{t=1}^{T} \mathbb{E}\left[\left|z_{t,T}^2 - \mathbb{E}_t\left[z_{t,T}^2\right]\right|^p\right] \tag{S.1.15}$$

$$\leq 2\sum_{t=1}^{T} 2^{p-1}\left(\mathbb{E}\left[\left|z_{t,T}^2\right|^p\right] + \mathbb{E}\left[\left|\mathbb{E}_t\left[z_{t,T}^2\right]\right|^p\right]\right) \leq 2^{p+1}\sum_{t=1}^{T} \mathbb{E}\left[\left|z_{t,T}\right|^{2p}\right], \tag{S.1.16}$$

where we use in the second inequality that $|a+b|^p \leq 2^{p-1}\left(|a|^p + |b|^p\right)$ for any $a, b \in \mathbb{R}$. Using the same inequality again yields

$$\mathbb{E}\left[\left|z_{t,T}\right|^{2p}\right] = \mathbb{E}\left[\left|\lambda^\top g_{t,T} - \lambda^\top g_{t,T}^e\right|^{2p}\right] \leq 2^{2p-1}\left(\mathbb{E}\left[\left|\lambda^\top g_{t,T}\right|^{2p}\right] + \mathbb{E}\left[\left|\lambda^\top g_{t,T}^e\right|^{2p}\right]\right). \tag{S.1.17}$$

Thus,

$$\sum_{t=1}^{T} \mathbb{E}\left[\left|\lambda^\top g_{t,T}\right|^{2p}\right] = (T\delta_T)^{1-p}\frac{1}{T}\sum_{t=1}^{T} \mathbb{E}\left[\left|\lambda^\top \boldsymbol{h}_t\right|^{2p} \int \left|K'(u)\right|^{2p} f_t(\delta_T u)\mathrm{d}u\right] \to 0, \tag{S.1.18}$$

as $(T\delta_T)^{1-p} \to 0$ for all $p \in (1,2)$, $\mathbb{E}\left[\left|\lambda^\top \boldsymbol{h}_t\right|^{2p}\right] < \infty$ for $p > 1$ sufficiently small such that $2p \leq 2+\delta$ (for the $\delta > 0$ from Assumption (A2)), and $\int |K'(u)|^{2p} f_t(\delta_T u)\mathrm{d}u \leq cc^{2p-1} \int |K'(u)|\,\mathrm{d}u < \infty$ as $\int |K'(u)|\,\mathrm{d}u < \infty$, $\sup_u |K'(u)| \leq c$ and $\sup_x f_t(x) \leq c$ a.s. by assumption. Similarly, we obtain

$$\sum_{t=1}^{T} \mathbb{E}\left[\left|\lambda^\top g_{t,T}^e\right|^{2p}\right] = \delta_T^{2p-1}(T\delta_T)^{1-p}\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}\left[\left|\lambda^\top \boldsymbol{h}_t\right|^{2p}\left|\int K'(u)\, f_t(\delta_T u)\mathrm{d}u\right|^{2p}\right] \to 0, \tag{S.1.19}$$

which shows that $h_{1,T} \xrightarrow{L_p} 0$ for some $p > 1$ sufficiently small which implies that $h_{1,T} \xrightarrow{P} 0$.

We continue by showing that $h_{2,T} \xrightarrow{P} 0$. Using the same argument as in (S.1.8), we split

$$h_{2,T} = \sum_{t=1}^{T}\mathbb{E}_t\left[z_{t,T}^2\right] - \bar{\omega}^2 = \sum_{t=1}^{T}\mathbb{E}_t\left[(\lambda^\top g_{t,T})^2\right] - \sum_{t=1}^{T}(\lambda^\top g_{t,T}^e)^2 - \bar{\omega}^2. \tag{S.1.20}$$

Applying a transformation of variables yields

$$\sum_{t=1}^{T}(\lambda^\top g_{t,T}^e)^2 = \delta_T \frac{1}{T}\sum_{t=1}^{T}(\lambda^\top \boldsymbol{h}_t)^2\left(\int K'(u) f_t(\delta_T u)\mathrm{d}u\right)^2 \tag{S.1.21}$$

$$\leq \delta_T\left(\frac{1}{T}\sum_{t=1}^{T}(\lambda^\top \boldsymbol{h}_t)^2\right)\left(c\int |K'(u)|\mathrm{d}u\right)^2 \xrightarrow{P} 0, \tag{S.1.22}$$

as $\delta_T \to 0$, $\frac{1}{T}\sum_{t=1}^{T}(\lambda^\top \boldsymbol{h}_t)^2 = \mathbb{E}\left[(\lambda^\top \boldsymbol{h}_t)^2\right] + o_P(1)$ and $\left(\int |K'(u)|\mathrm{d}u\right)^2 \leq \int |K'(u)|^2\mathrm{d}u < \infty$ by assumption. Furthermore,

$$\sum_{t=1}^{T}\mathbb{E}_t\left[\left(\lambda^\top g_{t,T}\right)^2\right] = (T\delta_T)^{-1}\sum_{t=1}^{T}(\lambda^\top \boldsymbol{h}_t)^2\mathbb{E}_t\left[K'\left(\frac{\varepsilon_t}{\delta_T}\right)^2\right] \tag{S.1.23}$$

$$= \left(\frac{1}{T}\sum_{t=1}^{T}(\lambda^\top \boldsymbol{h}_t)^2\right)\int K'(u)^2 f_t(\delta_T u)\mathrm{d}u \xrightarrow{P} \mathbb{E}\left[f_t(0)(\lambda^\top \boldsymbol{h}_t)^2\right]\int K'(u)^2\mathrm{d}u = \bar{\omega}^2, \tag{S.1.24}$$

as for all $u \in \mathbb{R}$, $\frac{1}{T}\sum_{t=1}^{T}(\lambda^\top \boldsymbol{h}_t)^2 f_t(\delta_T u) \xrightarrow{P} \mathbb{E}\left[f_t(0)(\lambda^\top \boldsymbol{h}_t)^2\right]$. Thus, we find $h_{2,T} \xrightarrow{P} 0$ and consequently $\sum_{t=1}^{T} z_{t,T}^2 - \bar{\omega}^2 \xrightarrow{P} 0$, which concludes this proof. □

Lemma S.1.4. *Given Assumption 2.5 and the null hypothesis in (2.8), it holds that* $\max_{1\leq t\leq T}|h_{t,T}| \xrightarrow{P} 0$.

Proof. Let $\zeta > 0$ and $\delta > 0$ (sufficiently small such that $\mathbb{E}\left[||\boldsymbol{h}_t||^{2+\delta}\right] < \infty$). Then,

$$\begin{aligned}\mathbb{P}\left(\max_{1\leq t\leq T}|h_{t,T}|>\zeta\right) &= \mathbb{P}\left(\max_{1\leq t\leq T}|h_{t,T}|^{2+\delta}>\zeta^{2+\delta}\right) \leq \sum_{t=1}^{T}\mathbb{P}\left(|h_{t,T}|^{2+\delta}>\zeta^{2+\delta}\right) \\ &\leq \zeta^{-2-\delta}\sum_{t=1}^{T}\mathbb{E}\left[|h_{t,T}|^{2+\delta}\right] = \zeta^{-2-\delta}\bar{\omega}_T^{-2-\delta}\sum_{t=1}^{T}\mathbb{E}\left[|z_{t,T}|^{2+\delta}\right],\end{aligned} \tag{S.1.25}$$

by Markov's inequality. Employing the same steps as in the proof of Lemma S.1.3 following Equation (S.1.17) and replacing the exponent "$2p$" by "$2+\delta$" yields that $\sum_{t=1}^{T}\mathbb{E}\left[|z_{t,T}|^{2+\delta}\right] \to 0$. As $\bar{\omega}_T \to \bar{\omega}^2 > 0$, this directly implies that $\mathbb{P}\left(\max_{1\leq t\leq T}|h_{t,T}| > \zeta\right) \to 0$.

□

Lemma S.1.5. *Given Assumption 2.5, Assumption 3.1 and Assumption 3.2, for all $\lambda \in \mathbb{R}^k$ such that $||\lambda||_2 = 1$, it holds that $\sum_{t=1}^{T}\text{Var}\left(\phi_{t,T}^*(\theta_0)\lambda\right) \to \sigma^2$.*

Proof. As $\phi_{t,T}^*$ is a $\mathcal{F}_{t+1}$-MDS, it holds that $\mathbb{E}\left[\phi_{t,T}^*(\theta_0)\lambda\right] = 0$ and thus, $\text{Var}\left(\phi_{t,T}^*(\theta_0)\lambda\right) = \mathbb{E}\left[\left(\phi_{t,T}^*(\theta_0)\lambda\right)^2\right]$. We further find

$$\begin{aligned}\sum_{t=1}^{T}\mathbb{E}\left[\left(\phi_{t,T}^*(\theta_0)\lambda\right)^2\right] &= \frac{1}{T}\sum_{t=1}^{T}\mathbb{E}\left[\theta_{10}^2\left(\lambda^\top \boldsymbol{W}_{\text{Mean}}\boldsymbol{h}_t\right)^2\varepsilon_t^2\right] \\ &+ \frac{1}{T}\sum_{t=1}^{T}\mathbb{E}\left[\theta_{20}^2\left(\lambda^\top \boldsymbol{W}_{\text{Med}}\boldsymbol{h}_t\right)^2\left(\mathbb{1}_{\{\varepsilon_t>0\}}-\mathbb{1}_{\{\varepsilon_t<0\}}\right)^2\right] \\ &+ \frac{1}{T}\sum_{t=1}^{T}\mathbb{E}\left[\theta_{30}^2\left(\lambda^\top \boldsymbol{W}_{\text{Mode}}\boldsymbol{h}_t\right)^2\delta_T^{-1}K'\left(\frac{-\varepsilon_t}{\delta_T}\right)^2\right] \\ &+ \frac{2}{T}\sum_{t=1}^{T}\mathbb{E}\left[\theta_{10}\theta_{20}\left(\lambda^\top \boldsymbol{W}_{\text{Mean}}\boldsymbol{h}_t\right)\left(\lambda^\top \boldsymbol{W}_{\text{Med}}\boldsymbol{h}_t\right)\varepsilon_t\left(\mathbb{1}_{\{\varepsilon_t>0\}}-\mathbb{1}_{\{\varepsilon_t<0\}}\right)\right] \\ &+ \frac{2}{T}\sum_{t=1}^{T}\mathbb{E}\left[\theta_{10}\theta_{30}\left(\lambda^\top \boldsymbol{W}_{\text{Mean}}\boldsymbol{h}_t\right)\left(\lambda^\top \boldsymbol{W}_{\text{Mode}}\boldsymbol{h}_t\right)\varepsilon_t\delta_T^{-1/2}K'\left(\frac{-\varepsilon_t}{\delta_T}\right)\right] \\ &+ \frac{2}{T}\sum_{t=1}^{T}\mathbb{E}\left[\theta_{20}\theta_{30}\left(\lambda^\top \boldsymbol{W}_{\text{Med}}\boldsymbol{h}_t\right)\left(\lambda^\top \boldsymbol{W}_{\text{Mode}}\boldsymbol{h}_t\right)\left(\mathbb{1}_{\{\varepsilon_t>0\}}-\mathbb{1}_{\{\varepsilon_t<0\}}\right)\delta_T^{-1/2}K'\left(\frac{-\varepsilon_t}{\delta_T}\right)\right] \\ &+ \sum_{t=1}^{T}\mathbb{E}\left[(u_{t,T}(\theta_0)\lambda)^2\right] - 2\sum_{t=1}^{T}\mathbb{E}\left[\left(u_{t,T}(\theta_0)\lambda\right)\left(\tilde{\phi}_{t,T}(\theta_0)\lambda\right)\right].\end{aligned} \tag{S.1.26}$$

For the last two terms, we have $\sum_{t=1}^{T}\mathbb{E}\left[\left(u_{t,T}(\theta_0)\lambda\right)^2\right]\to 0$ and $\sum_{t=1}^{T}\mathbb{E}\left[\left(u_{t,T}(\theta_0)\lambda\right)\left(\tilde{\phi}_{t,T}(\theta_0)\lambda\right)\right]\to 0$ by assumption. For the fifth term,

$$\frac{2}{T}\sum_{t=1}^{T}\mathbb{E}\left[\theta_{10}\theta_{30}\left(\lambda^\top \boldsymbol{W}_{\text{Mean}}\boldsymbol{h}_t\right)\left(\lambda^\top \boldsymbol{W}_{\text{Mode}}\boldsymbol{h}_t\right)\delta_T^{-1/2}\mathbb{E}_t\left[\varepsilon_t K'\left(\frac{-\varepsilon_t}{\delta_T}\right)\right]\right] \tag{S.1.27}$$

$$= -\frac{2}{T}\sum_{t=1}^{T}\mathbb{E}\left[\theta_{10}\theta_{30}\left(\lambda^\top \boldsymbol{W}_{\text{Mean}}\boldsymbol{h}_t\right)\left(\lambda^\top \boldsymbol{W}_{\text{Mode}}\boldsymbol{h}_t\right)\delta_T^{3/2}\int uK'(u)\,f_t(\delta_T u)\mathrm{d}u\right]\to 0, \tag{S.1.28}$$

as $\delta_T^{3/2}\to 0$, $\int uK'(u)\,\mathrm{d}u<\infty$ and the respective moments are finite. The sixth term converges to zero by a similar argument by bounding $\left|\mathbb{1}_{\{\varepsilon_t>0\}}-\mathbb{1}_{\{\varepsilon_t<0\}}\right|\leq 1$. For the third term, it holds that

$$\begin{aligned}\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}\left[\theta_{30}^2\left(\lambda^\top \boldsymbol{W}_{\text{Mode}}\boldsymbol{h}_t\right)^2\delta_T^{-1}K'\left(\frac{-\varepsilon_t}{\delta_T}\right)^2\right] &= \frac{1}{T}\sum_{t=1}^{T}\mathbb{E}\left[\theta_{30}^2\left(\lambda^\top \boldsymbol{W}_{\text{Mode}}\boldsymbol{h}_t\right)^2\int K'(u)^2 f_t(\delta_T u)\mathrm{d}u\right]\\ &\to \mathbb{E}\left[\theta_{30}^2\left(\lambda^\top \boldsymbol{W}_{\text{Mode}}\boldsymbol{h}_t\right)^2 f_t(0)\int K'(u)^2\,\mathrm{d}u\right].\end{aligned}$$

The remaining first, second and fourth terms obviously converge to the equivalent quantities of σ^2 by employing a standard CLT, which concludes the proof of this lemma. □

Lemma S.1.6. *Given Assumption 2.5, Assumption 3.1 and Assumption 3.2, for all $\lambda\in\mathbb{R}^k$ such that $||\lambda||_2=1$, it holds that $\sum_{t=1}^{T}\left(\phi_{t,T}^*(\theta_0)\lambda\right)^2\xrightarrow{P}\sigma^2$.*

Proof. We apply the same factorization as in (S.1.26) (however without the expectation operator). By applying a law of large numbers for stationary and ergodic sequences (Theorem 3.34 in White (2001)), we obtain that

$$\begin{aligned}&\frac{1}{T}\sum_{t=1}^{T}\left(\theta_{10}\left(\boldsymbol{h}_t^\top \boldsymbol{W}_{\text{Mean}}\lambda\right)\varepsilon_t\right)^2\xrightarrow{P}\mathbb{E}\left[\left(\theta_{10}\left(\boldsymbol{h}_t^\top \boldsymbol{W}_{\text{Mean}}\lambda\right)\varepsilon_t\right)^2\right],\\ &\frac{1}{T}\sum_{t=1}^{T}\left(\theta_{20}\left(\boldsymbol{h}_t^\top \boldsymbol{W}_{\text{Med}}\lambda\right)\left(\mathbb{1}_{\{\varepsilon_t>0\}}-\mathbb{1}_{\{\varepsilon_t<0\}}\right)\right)^2\xrightarrow{P}\mathbb{E}\left[\left(\theta_{20}\left(\boldsymbol{h}_t^\top \boldsymbol{W}_{\text{Med}}\lambda\right)\left(\mathbb{1}_{\{\varepsilon_t>0\}}-\mathbb{1}_{\{\varepsilon_t<0\}}\right)\right)^2\right],\\ &\frac{2}{T}\sum_{t=1}^{T}\left(\theta_{10}\theta_{20}\left(\boldsymbol{h}_t^\top \boldsymbol{W}_{\text{Mean}}\lambda\right)\left(\boldsymbol{h}_t^\top \boldsymbol{W}_{\text{Med}}\lambda\right)\varepsilon_t\left(\mathbb{1}_{\{\varepsilon_t>0\}}-\mathbb{1}_{\{\varepsilon_t<0\}}\right)\right)^2\\ &\quad\xrightarrow{P}2\mathbb{E}\left[\left(\theta_{10}\theta_{20}\left(\boldsymbol{h}_t^\top \boldsymbol{W}_{\text{Mean}}\lambda\right)\left(\boldsymbol{h}_t^\top \boldsymbol{W}_{\text{Med}}\lambda\right)\varepsilon_t\left(\mathbb{1}_{\{\varepsilon_t>0\}}-\mathbb{1}_{\{\varepsilon_t<0\}}\right)\right)^2\right].\end{aligned}$$

Furthermore, a slight modification of Lemma S.1.3 (multiplying with $\theta_{30}^2 \left(\boldsymbol{h}_t^\top \boldsymbol{W}_{\text{Mode}} \lambda\right)^2$ instead of $\left(\boldsymbol{h}_t^\top \lambda\right)^2$) yields that

$$\frac{1}{T}\sum_{t=1}^{T} \theta_{30}^2 \left(\boldsymbol{h}_t^\top \boldsymbol{W}_{\text{Mode}} \lambda\right)^2 \delta_T^{-1} K'\left(\frac{-\varepsilon_t}{\delta_T}\right)^2 \xrightarrow{P} \mathbb{E}\left[\theta_{30}^2 \left(\boldsymbol{h}_t^\top \boldsymbol{W}_{\text{Mode}} \lambda\right)^2 f_t(0) \int K'(u)^2 \,\mathrm{d}u\right]. \quad \text{(S.1.29)}$$

We now show that the remaining four terms vanish asymptotically (in probability). For the mixed mean/mode term, we apply a similar addition of a zero (adding and subtracting $\mathbb{E}_t[\dots]$) as in the proof of Lemma S.1.3. For this, we first note that

$$\frac{2}{T}\sum_{t=1}^{T} \theta_{10}\theta_{30} \left(\boldsymbol{h}_t^\top \boldsymbol{W}_{\text{Mean}} \lambda\right)\left(\boldsymbol{h}_t^\top \boldsymbol{W}_{\text{Mode}} \lambda\right) \delta_T^{-1/2} \mathbb{E}_t\left[\varepsilon_t K'\left(\frac{-\varepsilon_t}{\delta_T}\right)\right] \quad \text{(S.1.30)}$$

$$= -\frac{2}{T}\sum_{t=1}^{T} \theta_{10}\theta_{30} \left(\boldsymbol{h}_t^\top \boldsymbol{W}_{\text{Mean}} \lambda\right)\left(\boldsymbol{h}_t^\top \boldsymbol{W}_{\text{Mode}} \lambda\right) \delta_T^{3/2} \int u K'(u)\, f_t(\delta_T u) \mathrm{d}u \xrightarrow{P} 0, \quad \text{(S.1.31)}$$

as $\delta_T^{3/2} \to 0$. In the following, we further show that

$$\frac{2}{T}\sum_{t=1}^{T} \theta_{10}\theta_{30} \left(\boldsymbol{h}_t^\top \boldsymbol{W}_{\text{Mean}} \lambda\right)\left(\boldsymbol{h}_t^\top \boldsymbol{W}_{\text{Mode}} \lambda\right) \delta_T^{-1/2} \left\{\varepsilon_t K'\left(\frac{-\varepsilon_t}{\delta_T}\right) - \mathbb{E}_t\left[\varepsilon_t K'\left(\frac{-\varepsilon_t}{\delta_T}\right)\right]\right\} \xrightarrow{L_p} 0,$$

for any $p \in (1,2)$ small enough. As in the proof of Lemma S.1.3, we apply the von Bahr and Esseen (1965) inequality and Minkowski's inequality in order to conclude that

$$\begin{aligned}
&\mathbb{E}\left[\left|\frac{2}{T}\sum_{t=1}^{T} \theta_{10}\theta_{30} \left(\boldsymbol{h}_t^\top \boldsymbol{W}_{\text{Mean}} \lambda\right)\left(\boldsymbol{h}_t^\top \boldsymbol{W}_{\text{Mode}} \lambda\right) \delta_T^{-1/2} \left\{\varepsilon_t K'\left(\frac{\varepsilon_t}{\delta_T}\right) - \mathbb{E}_t\left[\varepsilon_t K'\left(\frac{\varepsilon_t}{\delta_T}\right)\right]\right\}\right|^p\right] \\
&\leq \frac{2^{p+2}}{T^p}\sum_{t=1}^{T}\left\{\mathbb{E}\left[\left|\theta_{10}\theta_{30} \left(\boldsymbol{h}_t^\top \boldsymbol{W}_{\text{Mean}} \lambda\right)\left(\boldsymbol{h}_t^\top \boldsymbol{W}_{\text{Mode}} \lambda\right) \delta_T^{-1/2} \varepsilon_t K'\left(\frac{\varepsilon_t}{\delta_T}\right)\right|^p\right]\right. \\
&\qquad \left. + \mathbb{E}\left[\left|\theta_{10}\theta_{30} \left(\boldsymbol{h}_t^\top \boldsymbol{W}_{\text{Mean}} \lambda\right)\left(\boldsymbol{h}_t^\top \boldsymbol{W}_{\text{Mode}} \lambda\right) \delta_T^{-1/2} \mathbb{E}_t\left[\varepsilon_t K'\left(\frac{\varepsilon_t}{\delta_T}\right)\right]\right|^p\right]\right\},
\end{aligned}$$

where the first term is bounded from above by

$$\leq \frac{2^{p+2}}{T^p}\sum_{t=1}^{T} \mathbb{E}\left[\left|\theta_{10}\theta_{30} \left(\boldsymbol{h}_t^\top \boldsymbol{W}_{\text{Mean}} \lambda\right)\left(\boldsymbol{h}_t^\top \boldsymbol{W}_{\text{Mode}} \lambda\right)\right|^p \delta_T^{-p/2} \int \left|e K'\left(\frac{e}{\delta_T}\right)\right|^p f_t(e) \mathrm{d}e\right]$$

$$= 2^{p+2}\delta_T^{1+p/2}T^{1-p}\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}\left[\left|\theta_{10}\theta_{30}\left(\boldsymbol{h}_t^\top \boldsymbol{W}_{\text{Mean}}\lambda\right)\left(\boldsymbol{h}_t^\top \boldsymbol{W}_{\text{Mode}}\lambda\right)\right|^p \int \left|uK'(u)\right|^p f_t(\delta_T u)\mathrm{d}u\right] \to 0,$$

as $\delta_T^{1+p/2} \to 0$, $T^{1-p} \to 0$ for any $p > 1$ and as the respective moments are bounded by assumption. Similar arguments also yield that the second term converges to zero (compare to (S.1.19)). Applying the same line of reasoning for the mixed median/mode terms shows that

$$\frac{2}{T}\sum_{t=1}^{T}\theta_{20}\theta_{30}\left(\boldsymbol{h}_t^\top \boldsymbol{W}_{\text{Med}}\lambda\right)\left(\boldsymbol{h}_t^\top \boldsymbol{W}_{\text{Mode}}\lambda\right)\delta_T^{-1/2}\mathbb{E}_t\left[\left(\mathbb{1}_{\{\varepsilon_t>0\}} - \mathbb{1}_{\{\varepsilon_t<0\}}\right)K'\left(\frac{-\varepsilon_t}{\delta_T}\right)\right] \xrightarrow{P} 0. \quad \text{(S.1.32)}$$

For the fourth and last term, $\sum_{t=1}^{T}\left(u_{t,T}(\theta_0)\lambda\right)^2 \xrightarrow{P} 0$ and $\sum_{t=1}^{T}\left(u_{t,T}(\theta_0)\lambda\right)\left(\tilde{\phi}_{t,T}(\theta_0)\lambda\right) \xrightarrow{P} 0$ by assumption, which concludes this proof. □

Lemma S.1.7. *Given Assumption 2.5, Assumption 3.1 and Assumption 3.2, for all $\lambda \in \mathbb{R}^k$ such that $||\lambda||_2 = 1$, it holds that $\max_{1\le t\le T}\left|\sigma^{-1}\phi^*_{t,T}(\theta_0)\lambda\right| \xrightarrow{P} 0$.*

Proof. Let $\zeta > 0$ and $\delta > 0$ (sufficiently small such that $\mathbb{E}\left[||\boldsymbol{h}_t||^{2+\delta}\right] < \infty$ holds). Then, as in (S.1.25) in the proof of Lemma S.1.4, we get that

$$\mathbb{P}\left(\max_{1\le t\le T}\left|\sigma_T^{-1}\phi^*_{t,T}(\theta_0)\lambda\right| > \zeta\right) \le \zeta^{-2-\delta}\sigma_T^{-2-\delta}\sum_{t=1}^{T}\mathbb{E}\left[\left|\phi^*_{t,T}(\theta_0)\lambda\right|^{2+\delta}\right], \quad \text{(S.1.33)}$$

by Markov's inequality. Furthermore, we get that

$$4^{-2-\delta}\sum_{t=1}^{T}\mathbb{E}\left[\left|\phi^*_{t,T}(\theta_0)\lambda\right|^{2+\delta}\right] \le \sum_{t=1}^{T}\mathbb{E}\left[\left|u_{t,T}(\theta_0)\right|^{2+\delta}\right] \quad \text{(S.1.34)}$$

$$+\theta_{10}^{2+\delta}T^{-\frac{\delta}{2}}\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}\left[\left|\boldsymbol{h}_t^\top \boldsymbol{W}_{\text{Mean}}\lambda\right|^{2+\delta}\left|\varepsilon_t\right|^{2+\delta}\right] \quad \text{(S.1.35)}$$

$$+\theta_{20}^{2+\delta}T^{-\frac{\delta}{2}}\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}\left[\left|\boldsymbol{h}_t^\top \boldsymbol{W}_{\text{Med}}\lambda\right|^{2+\delta}\left|\mathbb{1}_{\{\varepsilon_t>0\}} - \mathbb{1}_{\{\varepsilon_t<0\}}\right|^{2+\delta}\right] \quad \text{(S.1.36)}$$

$$+\theta_{30}^{2+\delta}T^{-\frac{\delta}{2}}\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}\left[\left|\boldsymbol{h}_t^\top \boldsymbol{W}_{\text{Mode}}\lambda\right|^{2+\delta}\delta_T^{-\frac{2+\delta}{2}}\left|K'\left(\frac{-\varepsilon_t}{\delta_T}\right)\right|^{2+\delta}\right]. \quad \text{(S.1.37)}$$

The first term converges to zero by Assumption 3.1. The second and third term converge to zero as $T^{-\frac{\delta}{2}} \to 0$ and the respective moments are bounded by assumption. For the last term, we obtain

convergence equivalently to the proof of Lemma S.1.4,

$$\theta_{30}^{2+\delta} T^{-\frac{\delta}{2}} \frac{1}{T} \sum_{t=1}^{T} \mathbb{E}\left[\left| \boldsymbol{h}_t^\top \boldsymbol{W}_{\text{Mode}} \lambda \right|^{2+\delta} \delta_T^{-\frac{2+\delta}{2}} \left| K'\left(\frac{-\varepsilon_t}{\delta_T} \right) \right|^{2+\delta} \right] \tag{S.1.38}$$

$$\leq \theta_{30}^{2+\delta} (T\delta_T)^{-\frac{\delta}{2}} \frac{1}{T} \sum_{t=1}^{T} \mathbb{E}\left[\left| \boldsymbol{h}_t^\top \boldsymbol{W}_{\text{Mode}} \lambda \right|^{2+\delta} \int \left| K'(u) \right|^{2+\delta} f_t(\delta_T u) \mathrm{d}u \right], \tag{S.1.39}$$

which converges to zero as $(T\delta_T)^{-\frac{\delta}{2}} \to 0$ and the respective moments are bounded by assumption.

□

Proof of Theorem 2.7. Let $\lambda \in \mathbb{R}^k$, $||\lambda||_2 = 1$ be a fixed and deterministic vector. Then,

$$\begin{aligned}
&\lambda^\top \widehat{\Omega}_{T,\text{Mode}} \lambda - \lambda^\top \Omega_{\text{Mode}} \lambda \\
&= \frac{1}{T} \sum_{t=1}^{T} \delta_T^{-1} K'\left(\frac{X_t - Y_{t+1}}{\delta_T} \right)^2 (\lambda^\top \boldsymbol{h}_t)^2 - \frac{1}{T} \sum_{t=1}^{T} \mathbb{E}_t \left[\delta_T^{-1} K'\left(\frac{X_t - Y_{t+1}}{\delta_T} \right)^2 (\lambda^\top \boldsymbol{h}_t)^2 \right] \\
&+ \frac{1}{T} \sum_{t=1}^{T} \mathbb{E}_t \left[\delta_T^{-1} K'\left(\frac{X_t - Y_{t+1}}{\delta_T} \right)^2 (\lambda^\top \boldsymbol{h}_t)^2 \right] - \mathbb{E}\left[(\lambda^\top \boldsymbol{h}_t)^2 f_t(0) \int K'(u)^2 \mathrm{d}u \right].
\end{aligned} \tag{S.1.40}$$

We start by showing that the last line in (S.1.40) is $o_P(1)$,

$$\frac{1}{T} \sum_{t=1}^{T} \mathbb{E}_t \left[\delta_T^{-1} K'\left(\frac{X_t - Y_{t+1}}{\delta_T} \right)^2 (\lambda^\top \boldsymbol{h}_t)^2 \right] = \frac{1}{T} \sum_{t=1}^{T} (\lambda^\top \boldsymbol{h}_t)^2 \delta_T^{-1} \int K'\left(\frac{e}{\delta_T} \right)^2 f_t(e) \mathrm{d}e \tag{S.1.41}$$

$$= \frac{1}{T} \sum_{t=1}^{T} (\lambda^\top \boldsymbol{h}_t)^2 \int K'(u)^2 f_t(\delta_T u) \mathrm{d}u \xrightarrow{P} \mathbb{E}\left[(\lambda^\top \boldsymbol{h}_t)^2 f_t(0) \int K'(u)^2 \mathrm{d}u \right], \tag{S.1.42}$$

as $f_t(\delta_T u) \to f_t(0) \leq c$, and by further applying a weak law of large numbers for stationary and ergodic data as $\mathbb{E}\left[||\boldsymbol{h}_t||^{2+\delta} \right] < \infty$.

We further show that the penultimate line in (S.1.40) converges to zero in L_p (p-th mean) for some $p > 1$ small enough. By applying the von Bahr and Esseen (1965) inequality for MDS, we get

$$\begin{aligned}
&\mathbb{E}\left[\left| \frac{1}{T} \sum_{t=1}^{T} \delta_T^{-1} K'\left(\frac{\varepsilon_t}{\delta_T} \right)^2 (\lambda^\top \boldsymbol{h}_t)^2 - \frac{1}{T} \sum_{t=1}^{T} \mathbb{E}_t \left[\delta_T^{-1} K'\left(\frac{\varepsilon_t}{\delta_T} \right)^2 (\lambda^\top \boldsymbol{h}_t)^2 \right] \right|^p \right] \\
&\leq 2T^{-p} \sum_{t=1}^{T} \mathbb{E}\left[\left| \delta_T^{-1} K'\left(\frac{\varepsilon_t}{\delta_T} \right)^2 (\lambda^\top \boldsymbol{h}_t)^2 \right|^p \right] + 2T^{-p} \sum_{t=1}^{T} \mathbb{E}\left[\left| \mathbb{E}_t \left[\delta_T^{-1} K'\left(\frac{\varepsilon_t}{\delta_T} \right)^2 (\lambda^\top \boldsymbol{h}_t)^2 \right] \right|^p \right].
\end{aligned}$$

For the first term, we get that

$$T^{-p}\sum_{t=1}^{T}\mathbb{E}\left[\left|\delta_T^{-1}K'\left(\frac{\varepsilon_t}{\delta_T}\right)^2(\lambda^\top \boldsymbol{h}_t)^2\right|^p\right] = (T\delta_T)^{-p}\sum_{t=1}^{T}\mathbb{E}\left[\left|\lambda^\top \boldsymbol{h}_t\right|^{2p}\int\left|K'\left(\frac{e}{\delta_T}\right)\right|^{2p} f_t(e)\mathrm{d}e\right]$$
$$= (T\delta_T)^{1-p}\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}\left[\left|\lambda^\top \boldsymbol{h}_t\right|^{2p}\int\left|K'(u)\right|^2 f_t(\delta_T u)\mathrm{d}u\right] \to 0,$$

as $(T\delta_T)^{1-p}\to 0$ for any $p>1$, $\mathbb{E}\left[||\boldsymbol{h}_t||^{2p}\right]<\infty$ for $p>1$ small enough, the density f_t is bounded from above, and $\int|K'(u)|^2\mathrm{d}u<\infty$ by assumption. The second term converges by a similar argument as further detailed in (S.1.19) in the proof of Lemma S.1.3. As L_p convergence for any $p>1$ implies convergence in probability, the result of the theorem follows. □

Proof of Theorem 3.4. For notational simplicity, we show consistency of the covariance estimator by considering the bilinear forms $\lambda^\top\left(\frac{1}{T}\sum_{t=1}^{T}\phi_{t,T}(\theta_0)\phi_{t,T}(\theta_0)^\top\right)\lambda$ and $\sigma^2 := \lambda^\top\Sigma(\theta_0)\lambda$, given in (A.20), for some arbitrary but fixed $\lambda\in\mathbb{R}^k$ such that $||\lambda||_2=1$. Then, we get that

$$\lambda^\top\left(\frac{1}{T}\sum_{t=1}^{T}\phi_{t,T}(\theta_0)\phi_{t,T}(\theta_0)^\top\right)\lambda = \frac{1}{T}\sum_{t=1}^{T}\theta_{10}^2\left(\boldsymbol{h}_t^\top\widehat{\boldsymbol{W}}_{T,\mathrm{Mean}}\lambda\right)^2\varepsilon_t^2 \tag{S.1.43}$$
$$+\theta_{20}^2\left(\boldsymbol{h}_t^\top\widehat{\boldsymbol{W}}_{T,\mathrm{Med}}\lambda\right)^2\left(\mathbb{1}_{\{\varepsilon_t>0\}}-\mathbb{1}_{\{\varepsilon_t<0\}}\right)^2+\theta_{30}^2\left(\boldsymbol{h}_t^\top\widehat{\boldsymbol{W}}_{T,\mathrm{Mode}}\lambda\right)^2\delta_T^{-1}K'\left(\frac{-\varepsilon_t}{\delta_T}\right)^2 \tag{S.1.44}$$
$$+2\theta_{10}\theta_{20}\left(\boldsymbol{h}_t^\top\widehat{\boldsymbol{W}}_{T,\mathrm{Mean}}\lambda\right)\left(\boldsymbol{h}_t^\top\widehat{\boldsymbol{W}}_{T,\mathrm{Med}}\lambda\right)\varepsilon_t\left(\mathbb{1}_{\{\varepsilon_t>0\}}-\mathbb{1}_{\{\varepsilon_t<0\}}\right) \tag{S.1.45}$$
$$+2\theta_{10}\theta_{30}\left(\boldsymbol{h}_t^\top\widehat{\boldsymbol{W}}_{T,\mathrm{Mean}}\lambda\right)\left(\boldsymbol{h}_t^\top\widehat{\boldsymbol{W}}_{T,\mathrm{Mode}}\lambda\right)\varepsilon_t\ \delta_T^{-1/2}K'\left(\frac{-\varepsilon_t}{\delta_T}\right) \tag{S.1.46}$$
$$+2\theta_{20}\theta_{30}\left(\boldsymbol{h}_t^\top\widehat{\boldsymbol{W}}_{T,\mathrm{Med}}\lambda\right)\left(\boldsymbol{h}_t^\top\widehat{\boldsymbol{W}}_{T,\mathrm{Mode}}\lambda\right)\left(\mathbb{1}_{\{\varepsilon_t>0\}}-\mathbb{1}_{\{\varepsilon_t<0\}}\right)\delta_T^{-1/2}K'\left(\frac{-\varepsilon_t}{\delta_T}\right). \tag{S.1.47}$$

We show convergence in probability for the individual matrix components for the first term,

$$\frac{1}{T}\sum_{t=1}^{T}\theta_{10}^2\left(\boldsymbol{h}_t^\top\widehat{\boldsymbol{W}}_{T,\mathrm{Mean}}\lambda\right)^2\varepsilon_t^2 = \sum_{i,j,\iota,l}\widehat{\boldsymbol{W}}_{T,\mathrm{Mean},ij}\widehat{\boldsymbol{W}}_{T,\mathrm{Mean},\iota l}\frac{1}{T}\sum_{t=1}^{T}\theta_{10}^2\boldsymbol{h}_{t,i}\lambda_j\boldsymbol{h}_{t,\iota}\lambda_l\varepsilon_t^2 \tag{S.1.48}$$
$$\xrightarrow{P}\sum_{i,j,\iota,l}\boldsymbol{W}_{\mathrm{Mean},ij}\boldsymbol{W}_{\mathrm{Mean},\iota l}\,\mathbb{E}\left[\theta_{10}^2\boldsymbol{h}_{t,i}\lambda_j\boldsymbol{h}_{t,\iota}\lambda_l\varepsilon_t^2\right] = \mathbb{E}\left[\theta_{10}^2\left(\boldsymbol{h}_t^\top\boldsymbol{W}_{\mathrm{Mean}}\lambda\right)^2\varepsilon_t^2\right]. \tag{S.1.49}$$

Convergence of the remaining terms follows analogously by considering the terms component-wisely and by applying similar arguments as in Lemma S.1.6. □

S.2 Kernel Choice

The asymptotic results presented in Section 2.3 rely on the chosen kernel K satisfying Assumption (A5). Besides the normalization $\int K(u)\mathrm{d}u = 1$ and boundedness assumptions, we impose the *first-order kernel* condition $\int uK(u)\mathrm{d}u = 0$ (and $\int u^2K(u)\mathrm{d}u > 0$ follows from the non-negativity of K). As discussed in Li and Racine (2006), higher-order kernels allow one to apply a Taylor expansion of higher order and can thereby obtain a faster rate of convergence, which could in theory be made arbitrarily close to $\sqrt{T}$, at the cost of stronger smoothness assumptions on the underlying density function. However, in our application of kernel functions to the generalized modal midpoint in Definition 2.3, we need to ensure that the limit of this quantity is well-defined and unique, and that the identification is *strict*. For this, we assume in Theorem 2.4 that the kernel function is log-concave which is automatically violated for higher-order kernels. Consequently, we do not consider higher-order kernels in this work.

It is well-known in the literature on nonparametric statistics that kernels with bounded support can be more efficient. However, *strict* identifiability of the generalized modal midpoint only holds for kernel functions with unbounded support, which motivates our usage of unbounded kernel functions such as the Gaussian kernel. Figure S.3 further illustrates that the test power does not increase by employing a biweight kernel, which has bounded support.

S.3 Bandwidth Choice

We follow the rule-of-thumb proposed by Kemp and Silva (2012) and Kemp et al. (2020) in setting the bandwidth parameter, with one modification to deal with skewness. Specifically, as discussed in Section 2.3, in order to obtain an optimal convergence for our nonparametric test (for first-order kernels), we choose $\delta_T \approx T^{-1/7}$. Following Kemp and Silva (2012), we choose δ_T proportional to $T^{-0.143}$, which is almost $T^{-1/7}$:

$$\delta_T = k_1 \cdot k_2 \cdot T^{-0.143}. \tag{S.3.1}$$

As in Kemp and Silva (2012) and Kemp et al. (2020), we choose k_1 proportional to the median

absolute deviation of the forecast error, a robust measure for the variation in the data,

$$k_1 = 2.4 \times \widehat{\text{Med}}_t\big(\big|(X_t - Y_{t+1}) - \widehat{\text{Med}}_s(X_s - Y_{s+1})\big|\big). \tag{S.3.2}$$

The choice of the bandwidth parameter should be proportional to the scale of the underlying data such that test results are robust to linear re-scaling. Using preliminary simulations, we found better finite-sample results when this measure is scaled by 2.4.

Following early simulation analyses, we introduce a second constant, k_2, to adjust the bandwidth for the skewness of the forecast error, measured by the absolute value of Pearson's second skewness coefficient, $\hat{\gamma}$.

$$k_2 = \exp(-3\,|\hat{\gamma}|), \qquad \text{where } \hat{\gamma} = \frac{3\left(\frac{1}{T}\sum_t (X_t - Y_{t+1}) - \widehat{\text{Med}}_t[X_t - Y_{t+1}]\right)}{\hat{\sigma}(X_t - Y_{t+1})}. \tag{S.3.3}$$

For symmetric distributions, $k_2 = 1$ and this term vanishes from the bandwidth formula. For such distributions, and assuming a symmetric kernel as in our empirical work, the generalized modal midpoint equals the mode and employing a larger bandwidth increases efficiency. As skewness increases in magnitude the distance between the mode and the generalized modal midpoint increases for a fixed bandwidth, and to ensure satisfactory finite-sample properties a smaller bandwidth is needed. Our simple expression for k_2 achieves this.

S.4 Power study for the mode forecast rationality test

To analyze the power of the mode forecast rationality test introduced in Section 2.3 of the main paper we use the same DGPs as in Section 4.1 and consider two forms of sub-optimal forecasts:

(a) **Bias**: $\tilde{X}_t = X_t + \kappa\sigma_X$, where $\sigma_X = \sqrt{\text{Var}(X_t)}$ for $\kappa \in (-1, 1)$, and

(b) **Noise**: $\tilde{X}_t = X_t + \mathcal{N}(0, \kappa\sigma_X^2)$, where $\sigma_X = \sqrt{\text{Var}(X_t)}$, for $\kappa \in (0, 1)$.

The first type of misspecification introduces deterministic bias, where the degree of misspecification depends on the misspecification parameter κ. We standardize the bias using the unconditional standard deviation of the optimal forecasts, $\sqrt{\text{Var}(X_t)}$. The second type of misspecification introduces

independent noise, and the magnitude of the noise is regulated through the parameter κ: for $\kappa = 1$, the signal-to-noise ratio is one, as the standard deviation of the signal equals the standard deviation of the independent noise, and as κ shrinks to zero the noise vanishes.

Figure S.1 presents power plots for the "biased" forecasts and Figure S.2 presents power plots for the "noisy" forecasts. In each of the plots, we plot the rejection rate against the degree of misspecification κ. For all plots, we use the instrument choice $(1, X_t)$, a Gaussian kernel, and a nominal level of 5%. Notice that for $\kappa = 0$ the figures reveal the empirical test size.

Figure S.1 and Figure S.2 reveal that the proposed mode rationality test exhibits, as expected, increasing power for an increasing degree of misspecification. Also as expected, larger sample sizes lead to tests with greater power, although even the two smaller sample sizes exhibit reasonable power, particular in the case of biased forecasts. The figures also reveal that increasing degree of skewness yields to a slight loss of power. This is driven through the bandwidth choice, where larger values of the (empirical) skewness result in a smaller bandwidth, and consequently a lower test power (analogous to the bias-variance trade-off in the nonparametric estimation literature).

Results corresponding to the "biased" forecast case when using a biweight kernel are presented in Figure S.3. That figure reveals that the finite-sample size and power are very similar to those for the Gaussian kernel.

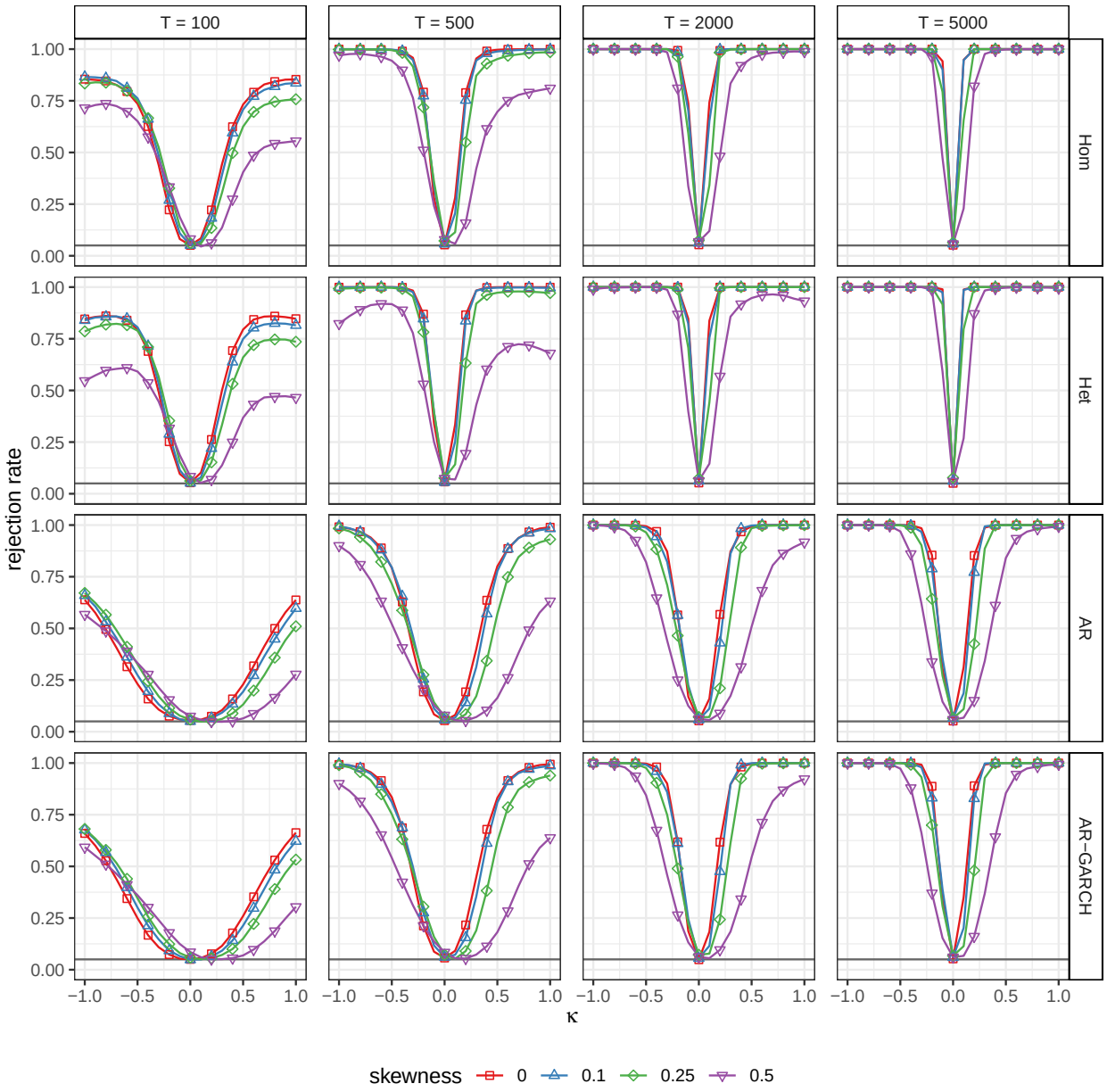


Figure S.1: Test power for the bias simulation setup. This figure plots the empirical rejection frequencies against the degrees of misspecification κ for different sample sizes in the vertical panels and for the four DGPs in the horizontal panels. The misspecification follows the "bias" setup and we utilize the instrument vector $(1, X_t)$ and a nominal significance level of 5%.

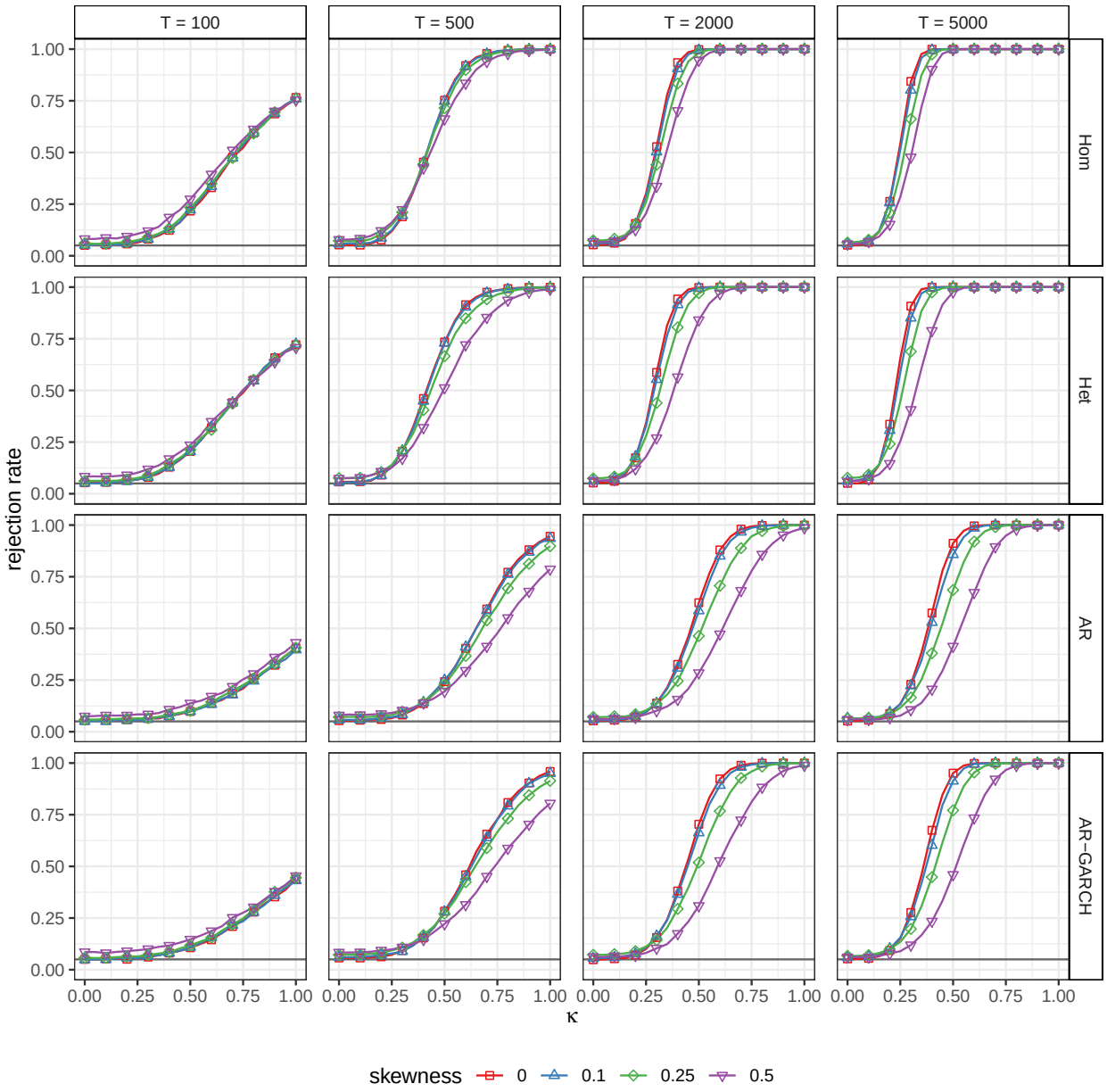


Figure S.2: Test power for the noise simulation setup. This figure plots the empirical rejection frequencies against the degrees of misspecification κ for different sample sizes in the vertical panels and for the four DGPs in the horizontal panels. The misspecification follows the "noise" setup and we utilize the instrument vector $(1, X_t)$ and a nominal significance level of 5%.

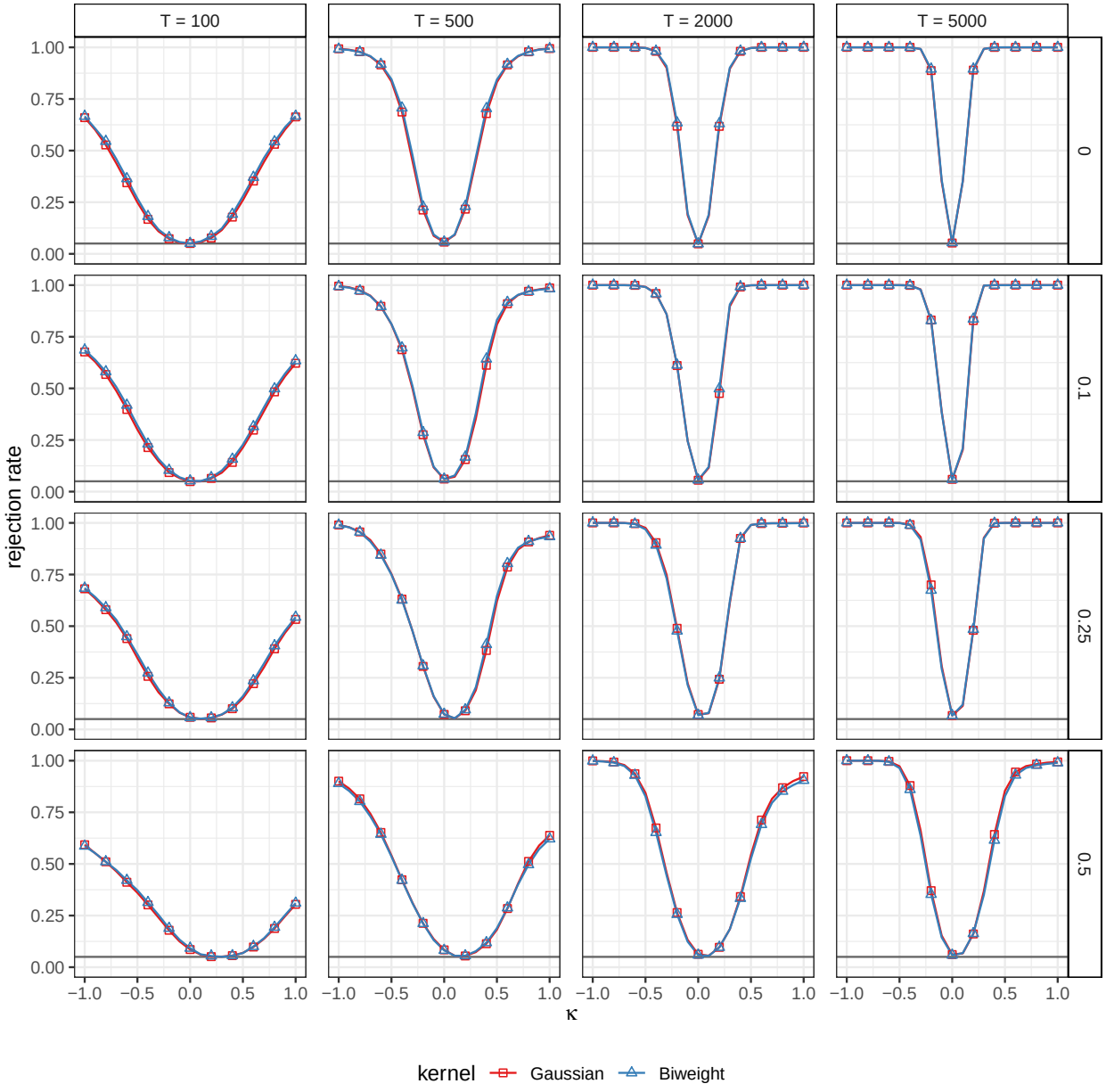


Figure S.3: Test power for different kernel functions. This figure plots the empirical rejection frequencies for the Gaussian and the biweight kernels against the degrees of misspecification κ for different sample sizes in the vertical panels and for four skewness levels in the horizontal panels. We simulate data from the AR-GARCH process, the misspecification follows the *bias* setup and we utilize the instrument vector $(1, X_t)$ and a nominal significance level of 5%.

S.5 Additional Plots and Tables

Table S.1: Empirical size of the mode rationality test: 1% significance level

	Instrument set 1				Instrument set 2				Instrument set 3			
Skewness	0	0.1	0.25	0.5	0	0.1	0.25	0.5	0	0.1	0.25	0.5
Sample size												
	Panel A: Homoskedastic iid data											
100	0.9	1.1	1.2	2.4	1.0	1.2	1.3	1.9	1.2	1.2	1.2	1.8
500	1.0	1.2	2.0	2.8	1.0	1.2	1.6	2.2	1.0	1.2	1.7	1.8
2000	1.2	1.5	2.2	1.9	1.0	1.3	1.6	1.5	0.9	1.2	1.6	1.3
5000	0.9	1.6	1.8	1.4	1.0	1.2	1.7	1.3	1.1	1.0	1.5	1.2
	Panel B: Heteroskedastic data											
100	1.0	1.1	1.5	2.6	1.2	1.0	1.4	2.3	1.3	1.1	1.2	2.0
500	1.4	1.2	2.3	2.2	1.1	1.1	1.9	1.8	1.0	1.2	1.6	1.7
2000	1.0	1.6	2.5	1.8	1.1	1.4	1.9	1.6	1.0	1.3	1.8	1.4
5000	1.0	1.6	2.7	1.4	0.9	1.3	2.0	1.2	1.0	1.1	1.7	1.2
	Panel C: Autoregressive data											
100	0.9	0.8	1.3	2.6	1.2	0.8	1.1	1.7	1.4	1.1	1.1	1.6
500	1.1	1.2	2.0	3.1	1.1	1.2	1.5	2.4	1.1	1.1	1.4	2.1
2000	1.1	1.2	2.2	1.8	1.2	1.1	1.8	1.5	1.0	1.1	1.6	1.4
5000	1.0	1.5	1.7	1.6	1.0	1.5	1.6	1.2	1.1	1.4	1.6	1.4
	Panel D: AR-GARCH data											
100	0.8	0.8	1.1	2.6	0.9	1.1	1.1	2.0	1.0	1.1	1.2	1.9
500	1.0	1.4	2.1	2.8	1.2	1.3	1.6	2.4	1.1	1.2	1.5	2.2
2000	1.0	1.4	2.3	1.8	1.0	1.1	1.8	1.4	0.9	1.2	1.7	1.3
5000	1.1	1.6	2.0	1.5	1.0	1.4	1.7	1.3	0.9	1.2	1.5	1.1

Notes: This table presents the empirical size of the mode rationality test for a Gaussian kernel, varying sample sizes, varying levels of skewness in the residual distribution and different instrument choices for a nominal significance level of 1%.

Table S.2: Empirical size of the mode rationality test: 10% significance level

	Instrument set 1				Instrument set 2				Instrument set 3			
Skewness	0	0.1	0.25	0.5	0	0.1	0.25	0.5	0	0.1	0.25	0.5
Sample size												
	Panel A: Homoskedastic iid data											
100	9.5	9.8	11.3	15.3	10.4	11.0	11.4	14.7	10.5	10.9	12.2	14.1
500	10.9	12.0	14.3	14.7	10.4	10.9	13.3	13.8	10.8	10.9	13.2	13.2
2000	11.1	12.6	14.7	12.7	10.6	11.8	13.2	12.5	10.3	11.7	12.8	12.0
5000	10.2	11.4	12.8	11.6	10.3	11.0	11.9	11.2	10.0	10.8	11.9	10.6
	Panel B: Heteroskedastic data											
100	10.0	10.4	12.1	15.7	10.6	10.6	11.9	14.8	11.1	10.9	11.9	14.2
500	11.4	11.6	14.4	14.8	11.5	11.1	13.6	13.1	10.9	11.1	13.4	12.5
2000	10.1	12.4	15.1	13.0	10.3	12.1	13.7	11.6	10.2	11.7	12.9	11.5
5000	10.3	13.0	15.4	12.0	10.1	11.9	13.7	11.8	10.1	11.4	12.8	11.1
	Panel C: Autoregressive data											
100	9.6	10.2	11.4	14.7	10.8	10.6	11.5	13.6	11.3	10.9	11.6	13.2
500	11.2	12.0	14.2	15.1	10.7	11.2	13.1	14.1	10.8	10.9	12.5	13.7
2000	10.5	12.3	13.9	12.2	10.4	11.1	12.4	11.7	10.3	11.3	12.2	11.6
5000	10.2	12.1	13.3	11.9	10.6	11.8	12.0	11.6	10.4	11.2	11.9	11.1
	Panel D: AR-GARCH data											
100	9.6	9.7	11.1	16.6	10.2	10.5	11.4	15.2	10.7	10.6	11.8	14.8
500	11.2	12.2	14.6	15.3	11.1	11.6	13.7	14.5	11.2	10.9	13.0	13.9
2000	10.5	12.2	14.0	12.3	10.3	11.1	13.0	11.8	10.0	10.7	12.4	11.4
5000	10.0	12.2	13.9	12.0	10.5	11.4	12.5	11.5	10.3	11.3	12.0	11.2

Notes: This table presents the empirical size of the mode rationality test for a Gaussian kernel, varying sample sizes, varying levels of skewness in the residual distribution and different instrument choices for a nominal significance level of 10%.

Table S.3: Empirical coverage of the confidence sets for central tendency: Cross-sectional data

Centrality measure	θ_{Mean}	θ_{Med}	θ_{Mode}	Symmetric data 100	500	2000	5000	Skewed data 100	500	2000	5000
				Panel A: Homoskedastic iid data							
Mean	1.00	0.00	0.00	89.3	90.0	89.8	89.2	89.4	90.2	89.6	90.2
Mode	0.00	0.00	1.00	90.1	89.1	89.8	90.1	85.7	86.0	87.6	88.8
Median	0.28	0.00	0.72	90.0	88.9	89.4	89.7	91.3	93.0	93.3	92.2
Median	0.15	0.50	0.35	89.5	89.0	89.6	89.6	90.3	91.9	91.4	91.2
Median	0.00	1.00	0.00	89.5	89.5	89.7	89.6	89.5	90.2	89.8	90.4
Mean-Mode	0.15	0.00	0.85	90.1	88.9	89.3	89.9	91.1	92.2	92.2	92.1
Mean-Mode	0.08	0.18	0.74	90.1	89.0	89.4	90.0	90.7	92.2	92.0	91.9
Mean-Mode	0.00	0.37	0.63	89.9	88.7	89.6	89.8	90.6	91.9	91.8	91.9
Mean-Median	0.50	0.00	0.50	89.8	89.2	89.2	89.6	91.0	92.0	92.3	90.4
Mean-Median	0.49	0.29	0.22	89.8	89.6	89.6	89.5	90.4	91.1	90.3	90.4
Mean-Median	0.41	0.59	0.00	89.5	89.8	89.8	89.2	89.7	90.3	89.4	89.7
Median-Mode	0.08	0.00	0.92	90.1	88.9	89.6	89.9	90.2	91.0	90.8	91.3
Median-Mode	0.04	0.08	0.88	89.9	89.0	89.6	89.8	90.1	90.9	90.7	91.4
Median-Mode	0.00	0.17	0.83	89.9	88.9	89.6	89.9	90.0	90.9	90.7	91.4
Mean-Median-Mode	0.18	0.00	0.82	90.2	88.9	89.2	89.9	91.3	92.5	92.7	92.6
Mean-Median-Mode	0.10	0.24	0.66	90.1	88.9	89.5	89.8	90.9	92.3	92.3	91.8
Mean-Median-Mode	0.00	0.51	0.49	89.7	88.8	89.6	89.6	90.5	91.8	91.8	91.4
				Panel B: Heteroskedastic data							
Mean	1.00	0.00	0.00	89.2	89.9	89.6	90.3	88.7	89.6	89.4	89.5
Mode	0.00	0.00	1.00	89.5	89.5	89.9	89.8	85.8	86.4	87.9	88.7
Median	0.28	0.00	0.72	89.3	89.2	89.9	90.1	90.4	91.1	91.1	91.8
Median	0.15	0.50	0.35	89.0	89.8	90.1	89.9	90.0	90.9	90.8	90.7
Median	0.00	1.00	0.00	89.3	89.9	90.1	90.2	89.3	89.9	89.7	89.5
Mean-Mode	0.15	0.00	0.85	89.3	89.4	89.8	89.9	90.2	91.3	91.1	91.4
Mean-Mode	0.08	0.18	0.74	89.2	89.5	89.8	90.0	90.0	91.1	91.0	90.9
Mean-Mode	0.00	0.37	0.63	89.3	89.7	89.9	90.0	89.9	90.8	90.7	91.0
Mean-Median	0.50	0.00	0.50	89.1	89.4	90.1	90.3	89.5	90.0	89.6	90.5
Mean-Median	0.49	0.29	0.22	89.2	89.6	90.0	90.5	89.3	90.2	89.7	90.0
Mean-Median	0.41	0.59	0.00	89.2	89.8	89.9	90.6	89.0	89.8	89.2	88.1
Median-Mode	0.08	0.00	0.92	89.3	89.3	89.8	89.9	89.2	90.0	90.3	90.8
Median-Mode	0.04	0.08	0.88	89.3	89.4	89.8	90.0	89.0	90.0	90.3	90.8
Median-Mode	0.00	0.17	0.83	89.4	89.5	89.9	90.0	89.1	90.1	90.1	90.4
Mean-Median-Mode	0.18	0.00	0.82	89.2	89.4	89.7	90.0	90.6	91.3	91.1	91.5
Mean-Median-Mode	0.10	0.24	0.66	89.2	89.7	89.9	90.1	90.3	91.2	91.1	91.3
Mean-Median-Mode	0.00	0.51	0.49	89.2	89.6	89.8	90.0	90.1	90.9	90.6	90.9

Notes: This tables presents the empirical coverage rates of the confidence sets for the forecasts of central tendency with a nominal coverage rate of 90%. We report the results for symmetric ($\gamma = 0$) and skewed data ($\gamma = 0.5$), for four sample sizes ($T = 100, 500, 2000, 5000$) and the two cross-sectional DGPs. We fix the instruments $\boldsymbol{h}_t = (1, X_t)$ and use a Gaussian kernel. In this application the set of identification function weights (θ) corresponding to a particular forecast combination weight vector is either a singleton (for the mean and mode) or a line. For the cases where the set is a line we present results for the end-points and the mid-point of this line.

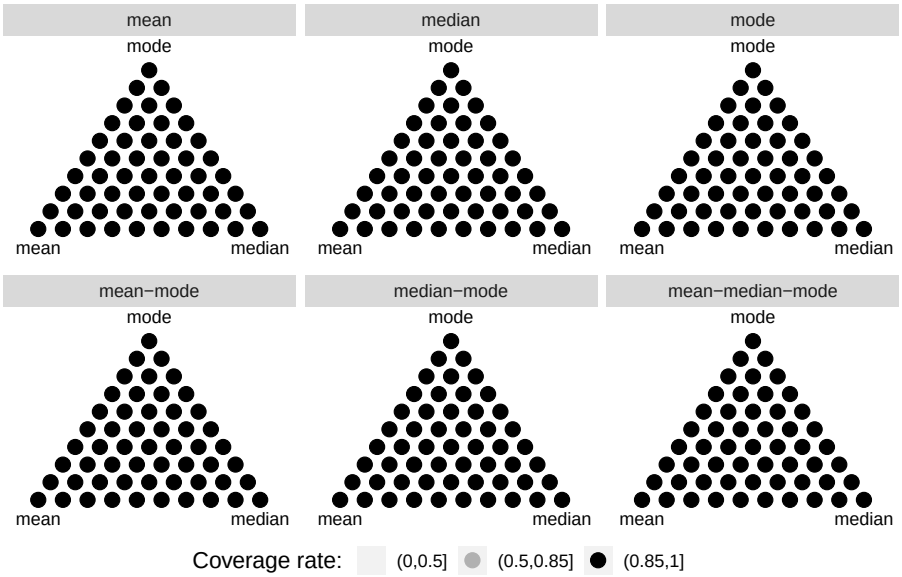
Table S.4: Empirical coverage of the confidence sets for central tendency: Time series data

Centrality measure	θ_{Mean}	θ_{Med}	θ_{Mode}	Symmetric data				Skewed data			
				100	500	2000	5000	100	500	2000	5000
				Panel A: Autoregressive data							
Mean	1.00	0.00	0.00	89.2	90.1	89.9	90.3	89.4	89.4	89.8	90.3
Mode	0.00	0.00	1.00	89.2	89.8	89.3	89.9	85.4	86.2	88.0	88.7
Median	0.28	0.00	0.72	89.1	89.7	89.0	89.6	91.5	92.3	93.5	91.7
Median	0.15	0.50	0.35	89.0	89.8	89.1	89.7	90.4	90.7	92.3	90.7
Median	0.00	1.00	0.00	88.9	90.0	89.2	89.9	89.0	89.2	90.7	89.9
Mean-Mode	0.15	0.00	0.85	89.1	89.6	89.1	89.8	90.5	92.0	92.7	91.7
Mean-Mode	0.08	0.18	0.74	88.9	89.9	89.2	89.7	90.5	91.6	92.5	91.3
Mean-Mode	0.00	0.37	0.63	89.0	89.9	89.4	89.6	90.2	91.4	92.3	91.5
Mean-Median	0.50	0.00	0.50	88.9	89.8	89.3	89.9	90.8	91.0	92.8	90.1
Mean-Median	0.49	0.29	0.22	88.8	90.2	89.3	89.9	90.0	89.8	91.2	89.7
Mean-Median	0.41	0.59	0.00	88.8	90.5	89.6	90.0	89.2	88.8	90.0	89.4
Median-Mode	0.08	0.00	0.92	89.2	89.7	89.1	89.8	89.4	90.8	91.4	91.1
Median-Mode	0.04	0.08	0.88	89.1	89.9	89.1	89.9	89.4	90.5	91.5	91.1
Median-Mode	0.00	0.17	0.83	89.1	90.0	89.1	89.8	89.3	90.6	91.5	91.0
Mean-Median-Mode	0.18	0.00	0.82	89.0	89.7	89.1	89.7	91.2	92.2	93.2	92.2
Mean-Median-Mode	0.10	0.24	0.66	89.0	89.7	89.2	89.6	90.7	91.7	92.7	91.7
Mean-Median-Mode	0.00	0.51	0.49	89.1	89.7	89.2	89.6	90.2	91.1	92.0	91.2
				Panel B: AR-GARCH data							
Mean	1.00	0.00	0.00	89.2	90.4	89.9	90.0	89.5	89.9	89.9	89.8
Mode	0.00	0.00	1.00	90.0	89.0	89.6	89.9	85.9	86.1	88.4	88.9
Median	0.28	0.00	0.72	89.5	89.1	89.1	89.4	91.1	92.5	93.1	92.0
Median	0.15	0.50	0.35	88.8	89.4	89.3	89.5	90.1	91.2	91.3	90.8
Median	0.00	1.00	0.00	88.6	89.5	89.8	89.6	89.1	89.7	89.8	90.1
Mean-Mode	0.15	0.00	0.85	89.9	88.9	89.4	89.7	90.4	91.5	92.7	91.5
Mean-Mode	0.08	0.18	0.74	89.6	89.1	89.6	89.6	90.3	91.3	92.3	91.3
Mean-Mode	0.00	0.37	0.63	89.5	89.1	89.6	89.6	90.2	91.2	92.1	91.1
Mean-Median	0.50	0.00	0.50	89.0	89.3	89.1	89.3	90.9	91.6	92.1	90.5
Mean-Median	0.49	0.29	0.22	88.7	89.7	89.3	89.4	90.1	90.7	90.8	89.9
Mean-Median	0.41	0.59	0.00	88.8	89.8	89.6	89.6	89.2	89.7	89.9	89.3
Median-Mode	0.08	0.00	0.92	90.2	89.0	89.5	89.7	89.4	90.5	92.1	91.0
Median-Mode	0.04	0.08	0.88	90.1	89.0	89.5	89.7	89.4	90.6	92.0	91.1
Median-Mode	0.00	0.17	0.83	90.0	89.2	89.6	89.7	89.3	90.5	91.8	90.8
Mean-Median-Mode	0.18	0.00	0.82	89.9	88.9	89.3	89.7	91.0	92.0	92.8	92.4
Mean-Median-Mode	0.10	0.24	0.66	89.6	89.2	89.3	89.7	90.4	91.7	92.3	91.2
Mean-Median-Mode	0.00	0.51	0.49	89.4	89.2	89.3	89.6	89.9	91.4	91.9	91.1

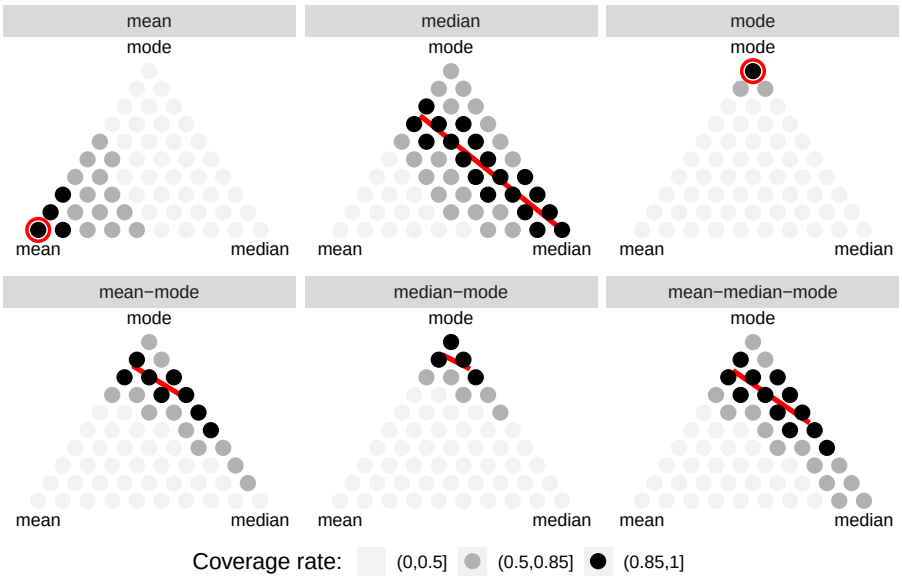
Notes: This tables presents the empirical coverage rates of the confidence sets for the forecasts of central tendency with a nominal coverage rate of 90%. We report the results for symmetric ($\gamma = 0$) and skewed data ($\gamma = 0.5$), for four sample sizes ($T = 100, 500, 2000, 5000$) and the two time-series DGPs. We fix the instruments $\boldsymbol{h}_t = (1, X_t)$ and use a Gaussian kernel. In this application the set of identification function weights (θ) corresponding to a particular forecast combination weight vector is either a singleton (for the mean and mode) or a line. For the cases where the set is a line we present results for the end-points and the mid-point of this line.

Figure S.4: Coverage rates of the confidence regions for central tendency measures for the homoskedastic DGP.

(a) Cross Sectional Homoskedastic DGP with skewness $\gamma = 0$



(b) Cross Sectional Homoskedastic DGP with skewness $\gamma = 0.5$



This figure shows coverage rates of 90% confidence regions for the measures of central tendency for the homoskedastic DGP. The true forecasted functional is given in the text above the triangle. The points inside the triangle correspond to convex combinations of the vertices, which are the mean, median and mode functionals. The color of the points indicates how often a specific point is contained in the 90% confidence regions. The upper panel shows results for the symmetric DGP, where all central tendency measures are equal. The lower panel uses a skewed DGP, with $\gamma = 0.5$. We use a red circle or a red line to indicate the (set of) central tendency measure(s) that correspond(s) to the forecast. We consider sample size $T = 2000$, instruments $\boldsymbol{h}_t = (1, X_t)$ and use a Gaussian kernel.

Hohenheim Discussion Papers in Business, Economics and Social Sciences

This paper series aims to present working results of researchers of all disciplines from the Faculty of Business, Economics and Social Sciences and their cooperation partners since 2015.

Institutes

510 Institute of Financial Management
520 Institute of Economics
530 Institute of Health Care & Public Management
540 Institute of Communication Science
550 Institute of Law and Legal Sciences
560 Institute of Education, Labour and Society
570 Institute of Marketing & Management
580 Institute of Interorganizational Management & Performance

Research Areas (since 2017)

INEPA "Inequality and Economic Policy Analysis"
TKID "Transformation der Kommunikation – Integration und Desintegration"
NegoTrans "Negotiation Research – Transformation, Technology, Media and Costs"
INEF "Innovation, Entrepreneurship and Finance"

The following table shows recent issues of the series. A complete list of all issues and full texts are available on our homepage: https://wiso.uni-hohenheim.de/papers

No.	Author	Title	Inst
01-2018	Michael D. Howard Johannes Kolb	FOUNDER CEOS AND NEW VENTURE MEDIA COVERAGE	INEF
02-2018	Peter Spahn	UNCONVENTIONAL VIEWS ON INFLATION CONTRAOL: FORWARD GUIDANCE, THE NEO-FISHERIAN APPROACH, AND THE FISCAL THEORY OF THE PRICE LEVEL	520
03-2018	Aderonke Osikominu Gregor Pfeifer	PERCEIVED WAGES AND THE GENDER GAP IN STEM FIELDS	INEPA
04-2018	Theresa Grafeneder-Weissteiner Klaus Prettner Jens Südekum	THREE PILLARS OF URBANIZATION: MIGRATION, AGING, AND GROWTH	INEPA
05-2018	Vadim Kufenko Vincent Geloso Klaus Prettner	DOES SIZE MATTER? IMPLICATIONS OF HOUSEHOLD SIZE FOR ECONOMIC GROWTH AND CONVERGENCE	INEPA
06-2018	Michael Trost	THE WHOLE IS GREATER THAN THE SUM OF ITS PARTS – PRICING PRESSURE INDICES FOR MERGERS OF VERTICALLY INTEGRATED FIRMS	520
07-2018	Karsten Schweikert	TESTING FOR COINTEGRATION WITH TRESHOLD ADJUSTMENT IN THE PRESENCE OF STRUCTURAL BREAKS	520
08-2018	Evanthia Fasoula Karsten Schweikert	PRICE REGULATIONS AND PRICE ADJUSTMENT DYNAMICS: EVIDENCE FROM THE AUSTRIAN RETAIL FUEL MARKET	520

No.	Author	Title	Inst
09-2018	Michael Ahlheim Jan Neidhardt Ute Siepmann Xiaomin Yu	WECHAT – USING SOCIAL MEDIA FOR THE ASSESSMENT OF TOURIST PREFERENCES FOR ENVIRONMENTAL IMPROVEMENTS IN CHINA	520
10-2018	Alexander Gerybadze Simone Wiesenauer	THE INTERNATIONAL SALES ACCELERATOR: A PROJECT MANAGEMENT TOOL FOR IMPROVING SALES PERFORMANCE IN FOREIGN TARGET MARKETS	570
11-2018	Klaus Prettner Niels Geiger Johannes Schwarzer	DIE WIRTSCHAFTLICHEN FOLGEN DER AUTOMATISIERUNG	INEPA
12-2018	Martyna Marczak Thomas Beissinger	COMPETITIVENESS AT THE COUNTRY-SECTOR LEVEL: NEW MEASURES BASED ON GLOBAL VALUE CHAINS	520
13-2018	Niels Geiger Klaus Prettner Johannes Schwarzer	AUTOMATISIERUNG, WACHSTUM UND UNGLEICHHEIT	INEPA
14-2018	Klaus Prettner Sebastian Seiffert	THE SIZE OF THE MIDDLE CLASS AND EDUCATIONAL OUTCOMES: THEORY AND EVIDENCE FROM THE INDIAN SUBCONTINENT	INEPA
15-2018	Marina Töpfer	THE EFFECT OF WOMEN DIRECTORS ON INNOVATION ACTIVITY AND PERFORMANCE OF CORPORATE FIRMS - EVIDENCE FROM CHINA –	INEF
16-2018	Timo Walter	TRADE AND WELFARE EFFECTS OF A POTENTIAL FREE TRADE AGREEMENT BETWEEN JAPAN AND THE UNITED STATES	INEPA
17-2018	Jonas Frank	THE EFFECTS OF ECONOMIC SANCTIONS ON TRADE: NEW EVIDENCE FROM A PANEL PPML GRAVITY APPROACH	INEPA
18-2018	Jonas Frank	THE EFFECT OF CULTURE ON TRADE OVER TIME – NEW EVIDENCE FROM THE GLOBE DATA SET	520
19-2018	Dario Cords Klaus Prettner	TECHNOLOGICAL UNEMPLOYMENT REVISITED: AUTOMATION IN A SEARCH AND MATCHING FRAMEWORK	INEPA
20-2018	Sibylle Lehmann-Hasemeyer Andreas Neumayer	THE PERSISTENCE OF OWNERSHIP INEQUALITY – INVESTORS ON THE GERMAN STOCK EXCHANGES, 1869-1945	INEPA
21-2018	Nadja Dwenger Lukas Treber	SHAMING FOR TAX ENFORCEMENT: EVIDENCE FROM A NEW POLICY	520
22-2018	Octavio Escobar Henning Mühlen	THE ROLE OF FDI IN STRUCTURAL CHANGE: EVIDENCE FROM MEXICO	520

No.	Author	Title	Inst
24-2018	Peng Nie Lanlin Ding Alfonso Sousa-Poza	OBESITY INEQUALITY AND THE CHANGING SHAPE OF THE BODYWEIGHT DISTRIBUTION IN CHINA	INEPA
25-2018	Michael Ahlheim Maike Becker Yeniley Allegue Losada Heike Trastl	WASTED! RESOURCE RECOVERY AND WASTE MANAGEMENT IN CUBA	520
26-2018	Peter Spahn	WAS WAR FALSCH AM MERKANTILISMUS?	520
27-2018	Sophie Therese Schneider	NORTH_SOUTH TRADE AGREEMENTS AND THE QUALITY OF INSTITUTIONS: PANEL DATA EVIDENCE	INEPA
01-2019	Dominik Hartmann Mayra Bezerra Beatrice Lodolo Flávio L. Pinheiro	INTERNATIONAL TRADE, DEVELOPMENT TRAPS, AND THE CORE-PERIPHERY STRUCTURE OF INCOME INEQUALITY	INEPA
02-2019	Sebastian Seiffert	GO EAST: ON THE IMPACT OF THE TRANSIBERIAN RAILWAY ON ECONOMIC DEVELOPMENT IN EASTERN RUSSIA	INEPA
03-2019	Kristina Bogner	KNOWLEDGE NETWORKS IN THE GERMAN BIOECONOMY: NETWORK STRUCTURE OF PUBLICLY FUNDED R&D NETWORKS	520
04-2019	Dominik Hartmann Mayra Bezerra Flávio L. Pinheiro	IDENTIFYING SMART STRATEGIES FOR ECONOMIC DIVERSIFICATION AND INCLUSIVE GROWTH IN DEVELOPING ECONOMIES. THE CASE OF PARAGUAY	INEPA
05-2019	Octavio Escobar Henning Mühlen	DECOMPOSING A DECOMPOSITION: WITHIN-COUNTRY DIFFERENCES AND THE ROLE OF STRUCTURAL CHANGE IN PRODUCTIVITY GROWTH	INEPA
06-2019	Dominik Hartmann Cristian Figueroa Mary Kaltenberg Paolo Gala	MAPPING STRATIFICATION: THE INDUSTRY-OCCUPATION SPACE REVEALS THE NETWORK STRUCTURE OF INEQUALITY	INEPA
07-2019	Stephan Fichtner Herbert Meyr	BIOGAS PLANT OPTIMIZATION BY INCREASING ITS FLEXIBILITY CONSIDERING UNCERTAIN REVENUES	580
08-2019	Annika Lenz Muhammed Kaya Philipp Melzer Andreas Schmid Josepha Witt Mareike Schoop	DATA QUALITY AND INFORMATION LOSS IN STANDARDISED INTERPOLATED PATH ANALYIS – QUALITY MEASURES AND GUIDELINES	NegoTrans
09-2019	Thilo R. Huning Fabian Wahl	THE FETTERS OF INHERITANCE? EQUAL PARTITION AND REGIONAL ECONOMIC DEVELOPMENT	520

10-2019	Peter Spahn	KEYNESIAN CAPITAL THEORY, DECLINING INTEREST RATES AND PERSISTING PROFITS	520
11-2019	Thorsten Proettel	INTERNATIONAL DIGITAL CURRENCIES AND THEIR IMPACT ON MONETARY POLICY – AN EXPLORATION OF IMPLICATIONS AND VULNERABILITY	520
12-2019	Franz X. Hof Klaus Prettner	RELATIVE CONSUMPTION, RELATIVE WEALTH, AND LONG-RUN GROWTH: WHEN AND WHY IS THE STANDARD ANALYSIS PRONE TO ERRONEOUS CONCLUSIONS?	INEPA
13-2019	Vadim Kufenko Vincent Geloso	WHO ARE THE CHAMPIONS? INEQUALITY, ECONOMIC FREEDOM AND THE OLYMPICS	INEPA
14-2019	Laura-Kristin Baric Niels Geiger	POLITICAL IMPLICATIONS OF ECONOMIC INEQUALITY: A LITERATURE SURVEY	INEPA
15-2019	Jens Grüb	MERGERS AND PARTIAL TACIT COLLUSION	520
01-2020	David E. Bloom Victoria Y. Fan Vadim Kufenko Osondu Ogbuoji Klaus Prettner Gavin Yamey	GOING BEYOND GDP WITH A PARSIMONIOUS INDICATOR: INEQUALITY-ADJUSTED HEALTHY LIFETIME INCOME	INEPA
02-2020	Michael Kuhn Klaus Prettner	RISING LONGEVITY, INCREASING THE RETIREMENT AGE, AND THE CONSEQUENCES FOR KNOWLEDGE-BASED LONG-RUN GROWTH	520
03-2020	Simiao Chen Zhangfeng Jin Klaus Prettner	THE RETIREMENT MIGRATION PUZZLE IN CHINA	520
04-2020	Andreas Orland Davud Rostam-Afschar	FLEXIBLE WORK ARRANGEMENTS AND PRECAUTIONARY BEHAVIOUR: THEORY AND EXPERIMENTAL EVIDENCE	520
05-2020	Michael Fritsch Martin Obschonka Fabian Wahl Michael Wyrwich	THE DEEP IMPRINT OF ROMAN SANDALS: EVIDENCE OF LONG-LASTING EFFECTS OF ROMAN RULE ON PERSONALITY, ECONOMIC PERFORMANCE, AND WELL-BEING IN GERMANY	520
06-2020	Sibylle Lehmann-Hasemeyer Klaus Prettner Paul Tscheuschner	THE SCIENTIFIC EVOLUTION AND ITS ROLE IN THE TRANSITION TO SUSTAINED ECONOMIC GROWTH	520
07-2020	Vladan Ivanovic Vadim Kufenko	IT'S A MAN'S WORLD? THE RISE OF FEMALE ENTREPRENEURSHIP DURING PRIVATIZATION IN SERBIA	INEPA

08-2020	Stephanie Briel Aderonke Osikominu Gregor Pfeifer Mirjam Reuter Sascha Satlukal	OVERCONFIDENCE AND GENDER DIFFERENCES IN WAGE EXPECTATIONS	520
09-2020	Thomas Beissinger Joël Hellier Martyna Marczak	DIVERGENCE IN LABOUR FORCE GROWTH: SHOULD WAGES AND PRICES GROW FASTER IN GERMANY?	520
10-2020	Benjamin Jung	THE HETEROGENEOUS EFFECTS OF CHINA‘S ACCESSION TO THE WORLD TRADE ORGANIZATION	INEPA
11-2020	Timo Dimitriadis Xiaochun Liu Julie Schnaitmann	ENCOMPASSING TESTS FOR VALUE AT RISK AND EXPECTED SHORTFALL MULTI-STEP FORECASTS BASED ON INFERENCE ON THE BOUNDARY	520
12-2020	Timo Dimitriadis Andrew J. Patton Patrick W. Schmidt	TESTING FORECAST RATIONALITY FOR MEASURES OF CENTRAL TENDENCY	520

IMPRINT

University of Hohenheim
Dean’s Office of the Faculty of Business, Economics and Social Sciences
Palace Hohenheim 1 B
70593 Stuttgart | Germany
Fon +49 (0)711 459 22488
Fax +49 (0)711 459 22785
wiso@uni-hohenheim.de
wiso.uni-hohenheim.de