Osband's principle for identification functions

SHORT COMMUNICATION

Osband's principle for identification functions

Timo Dimitriadis[1,2] · Tobias Fissler[3,4] · Johanna Ziegel[5]

Abstract

Given a statistical functional of interest such as the mean or median, a (strict) identification function is zero in expectation at (and only at) the true functional value. Identification functions are key objects in forecast validation, statistical estimation and dynamic modelling. For a possibly vector-valued functional of interest, we fully characterise the class of (strict) identification functions subject to mild regularity conditions.

Keywords Calibration · Characterisation · Identification function · Point forecasts · Z-estimation

Mathematics Subject Classification 62C07 · 62F10 · 62J20

✉ Timo Dimitriadis
timo.dimitriadis@awi.uni-heidelberg.de

Tobias Fissler
tobias.fissler@math.ethz.ch

Johanna Ziegel
johanna.ziegel@stat.unibe.ch

[1]   Alfred Weber Institute of Economics, Heidelberg University, Bergheimer Str. 58, 69115 Heidelberg, Germany

[2]   Heidelberg Institute for Theoretical Studies, 69118 Heidelberg, Germany

[3]   Department of Finance, Accounting and Statistics, Vienna University of Economics and Business (WU), Welthandelsplatz 1, 1020 Vienna, Austria

[4]   Present Address: RiskLab, Department of Mathematics, ETH Zurich, Rämistrasse 101, 8092 Zurich, Switzerland

[5]   Department of Mathematics and Statistics, Institute of Mathematical Statistics and Actuarial Science, University of Bern, Alpeneggstrasse 22, 3012 Bern, Switzerland

1 Introduction and informal statement of main result

Consider a statistical functional T of the random variable $Y \sim F$, that is, a mapping $F \mapsto T(F)$, such as the mean or the median. In the theory of forecast validation, a corresponding strict identification function $V(x, y)$ takes the forecast x and the realisation y of Y as arguments and its expectation with respect to $Y \sim F$ is zero if and only if x equals the true functional value $T(F)$. This defining property makes identification functions a central tool in forecast validation through calibration tests (Nolde and Ziegel 2017), often referred to as backtests in finance, and to forecast rationality (or optimality) tests in economics (Elliott et al. 2005; Dimitriadis et al. 2021b). Furthermore, these functions are fundamental to zero (Z) or generalised method of moments (GMM) estimation (Huber 1967; Hansen 1982; Newey and McFadden 1994), where they are often called moment functions or moment conditions. However, their statistical applications go much beyond these two fields and among others, they influence dynamic modelling through generalised autoregressive score (GAS) models (Creal et al. 2013), isotonic regression estimates (Jordan et al. 2022), or the derivation of anytime valid sequential tests (Casgrain et al. 2022). A complete understanding of the full class of (strict) identification functions for a given functional is crucial in these applications. Our main contribution, Theorem 4, provides such a full characterisation result.

In the jargon of decision theory (Gneiting 2011), the quantity of interest Y attains values in an *observation domain* $\mathsf{O} \subseteq \mathbb{R}^d$, which is equipped with the Borel-σ-algebra. The class of potential probability distributions F of Y is denoted by $\mathcal{F}$. Forecasts are elements of an *action domain* $\mathsf{A} \subseteq \mathbb{R}^k$. Formally, the functional of interest T is a potentially set-valued mapping from $\mathcal{F}$ to A, denoted by $T : \mathcal{F} \twoheadrightarrow \mathsf{A}$, where the notation $\twoheadrightarrow$ indicates that the values of T are subsets of A, with the convention that we identify point-valued functionals such as the mean with the singleton containing this value. For $\mathsf{O} = \mathsf{A} = \mathbb{R}$, prime examples for T are the mean or the α-quantile $q_\alpha(F) = \{x \in \mathbb{R} \mid \lim_{t \uparrow x} F(t) \leq \alpha \leq F(x)\}$, $\alpha \in (0, 1)$, where the latter is interval-valued. Prime examples for multivariate functionals are the mean-functional in case of multivariate observations ($\mathsf{O} = \mathsf{A} = \mathbb{R}^k$). For univariate observations, examples are multiple quantiles at different levels, the pair (mean, variance) with the natural action domain $\mathsf{A} = \mathbb{R} \times [0, \infty)$ or the pair consisting of the quantile and the Expected Shortfall (ES) at the same level with natural action domain $\mathsf{A} = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1 \geq x_2\}$, see Examples 2 and 3 for details. To present the formal definition of an identification function $V : \mathsf{A} \times \mathsf{O} \to \mathbb{R}^k$, let us introduce the convention that V is called $\mathcal{F}$-*integrable* if for each of its components V_i the integral $\int_\mathsf{O} V_i(x, y) \, dF(y)$ exists and is finite for all $x \in \mathsf{A}$ and $F \in \mathcal{F}$. Moreover, we shall use the shorthand $\bar{V}(x, F) = \int_\mathsf{O} V(x, y) \, dF(y)$ for any $x \in \mathsf{A}$, $F \in \mathcal{F}$, where the integral is understood componentwise.

Definition 1 (*Identification function and identifiability*)

(i) An $\mathcal{F}$-integrable map $V : \mathsf{A} \times \mathsf{O} \to \mathbb{R}^k$ is an $\mathcal{F}$-*identification function* for a functional $T : \mathcal{F} \twoheadrightarrow \mathsf{A} \subseteq \mathbb{R}^k$ if for all $x \in \mathsf{A}$ and for all $F \in \mathcal{F}$

$$x \in T(F) \implies \bar{V}(x, F) = 0.$$

(ii) An $\mathcal{F}$-integrable map $V : \mathsf{A} \times \mathsf{O} \to \mathbb{R}^k$ is a *strict $\mathcal{F}$-identification function* for a functional $T : \mathcal{F} \twoheadrightarrow \mathsf{A} \subseteq \mathbb{R}^k$ if for all $x \in \mathsf{A}$ and for all $F \in \mathcal{F}$

$$x \in T(F) \iff \bar{V}(x, F) = 0.$$

(iii) A functional $T : \mathcal{F} \twoheadrightarrow \mathsf{A} \subseteq \mathbb{R}^k$ is called *$\mathcal{F}$-identifiable* if there exists a strict $\mathcal{F}$-identification function for it.

On the class of distributions on $\mathbb{R}$ with a finite mean, $\mathcal{F}^1(\mathbb{R})$, the mean is identifiable with strict $\mathcal{F}^1(\mathbb{R})$-identification function $V(x, y) = x - y$. Likewise, the τ-expectile, $\tau \in (0, 1)$, possesses a strict $\mathcal{F}^1(\mathbb{R})$-identification function $V(x, y) = 2|\mathbb{1}\{y \leq x\} - \tau|(x - y)$. On the class $\mathcal{F}_\alpha(\mathbb{R})$ of distributions on $\mathbb{R}$ such that there exists an x with $F(x) = \alpha$, the α-quantile admits the strict $\mathcal{F}_\alpha(\mathbb{R})$-identification function $V(x, y) = \mathbb{1}\{y \leq x\} - \alpha$. Functionals failing to be identifiable on practically relevant classes of distributions are the variance and Expected Shortfall. On such classes $\mathcal{F}$, both of them violate the selective convex level sets property, which is necessary for identifiability (Osband 1985; Fissler et al. 2021).[1] However, the pairs (mean, variance) and (quantile, ES) turn out to be identifiable with corresponding two-dimensional strict identification functions, see Examples 2 and 3.

Regarding the flexibility of the class of identification functions, the following observation is immediate: If $V(x, y)$ is a strict $\mathcal{F}$-identification function for $T : \mathcal{F} \twoheadrightarrow \mathsf{A} \subseteq \mathbb{R}^k$, it can be multiplied with any $\mathbb{R}^{k \times k}$-valued function $h(x)$ of full rank and remains a strict identification function for T. Intriguingly, Theorem 4 formally states that, subject to mild regularity conditions, the reverse is also true, and the entire class of strict identification functions is given by

$$\{h(x)V(x, y) \, | \, h : \mathsf{A} \to \mathbb{R}^{k \times k}, \ \det(h(x)) \neq 0 \text{ for all } x \in \mathsf{A}\}. \tag{1}$$

Besides its theoretical appeal, this characterisation result opens the way for diverse applications. First, it can be used to optimise power of (conditional) calibration (forecast rationality or optimality) tests studied in Nolde and Ziegel (2017). It is further related to efficient Z- or GMM-estimation based on conditional moment conditions in the sense of Chamberlain (1987) and Newey (1993), where the matrix h is submerged in the choice of an optimal instrument matrix; see Theorem 3.1 and especially Remark 3.2 in Dimitriadis et al. (2021a) for details. Based on the choice of an identification function (called score by these authors) as their forcing variable, dynamic GAS models of Creal et al. (2013) determine an autoregressive model structure for a corresponding functional of interest that nests classical ARMA and GARCH models for the mean and variance. In these models, the so-called scaling matrix takes the place of the matrix h and, as already called for by Creal et al. (2013, p. 779), this choice "warrants separate inspection".

The following examples discuss interesting applications of our characterisation result in (1) to vector-valued functionals.

---

[1] T satisfies the selective convex level sets property of $\mathcal{F}$ if for any $F, G \in \mathcal{F}$ and for any $\lambda \in (0, 1)$ such that $(1 - \lambda)F + \lambda G \in \mathcal{F}$ it holds that $T(F) \cap T(G) \subseteq T((1 - \lambda)F + \lambda G)$.

Example 2 (Mean and variance) The pair (mean, variance) is identifiable on the class $\mathcal{F}^2(\mathbb{R})$ of distributions with finite variance with the two-dimensional strict $\mathcal{F}^2(\mathbb{R})$-identification function

$$V(x_1, x_2, y) = \begin{pmatrix} x_1 - y \\ x_2 - (y - x_1)^2 \end{pmatrix}.$$

One can use the characterisation result (1) to produce a multitude of other strict $\mathcal{F}^2(\mathbb{R})$-identification functions. Motivated by the decomposition of the variance into the difference of the second moment the squared expectation, a comparably intuitive one is

$$V'(x_1, x_2, y) = \begin{pmatrix} x_1 - y \\ x_2 + x_1^2 - y^2 \end{pmatrix}, \tag{2}$$

which arises by choosing the full rank matrix $h(x_1, x_2) = \begin{pmatrix} 1 & 0 \\ 2x_1 & 1 \end{pmatrix}$.

Example 3 (Quantile and ES) In financial mathematics, Value-at-Risk at level $\alpha \in (0, 1)$ (VaR_α) denotes the lower α-quantile, $\mathrm{VaR}_\alpha(F) = \inf q_\alpha(F) = \inf\{x \in \mathbb{R} \mid \alpha \leq F(x)\}$. Then, the ES at level $\alpha \in (0, 1)$ of a distribution F is formally defined as

$$\mathrm{ES}_\alpha(F) = \frac{1}{\alpha} \int_0^\alpha \mathrm{VaR}_\beta(F) \, d\beta = \frac{1}{\alpha} \int y \mathbb{1}\{y \leq \mathrm{VaR}_\alpha(F)\} \, dF(y)$$
$$- \frac{\mathrm{VaR}_\alpha(F)}{\alpha} \big(F(\mathrm{VaR}_\alpha(F)) - \alpha\big). \tag{3}$$

On any subclass of $\mathcal{F}_\alpha(\mathbb{R})$ where ES_α is finite, e.g. on $\mathcal{F}_\alpha(\mathbb{R}) \cap \mathcal{F}^1(\mathbb{R})$, there is the following strict identification function for $(q_\alpha, \mathrm{ES}_\alpha)$

$$V(x_1, x_2, y) = \begin{pmatrix} \mathbb{1}\{y \leq x_1\} - \alpha \\ x_2 - \frac{y}{\alpha}\mathbb{1}\{y \leq x_1\} \end{pmatrix},$$

where the second component naturally corresponds to a truncated expectation. Applying (1) with the full rank matrix $h(x_1, x_2) = \begin{pmatrix} 1 & 0 \\ x_1/\alpha & 1 \end{pmatrix}$, one obtains the alternative strict identification function

$$V'(x_1, x_2, y) = \begin{pmatrix} \mathbb{1}\{y \leq x_1\} - \alpha \\ x_2 - \frac{y}{\alpha}\mathbb{1}\{y \leq x_1\} + \frac{x_1}{\alpha}(\mathbb{1}\{y \leq x_1\} - \alpha) \end{pmatrix}. \tag{4}$$

The advantage of V' over V is that when evaluating V' on a discontinuous distribution with $F(\mathrm{VaR}_\alpha(F)) > \alpha$, even though the first components of V and V' fail to be an identification function for q_α,[2] the second component of V' still vanishes in expectation

---

[2] To obtain a better understanding of identifiability for the possibly set-valued α-quantile and its lower endpoint VaR_α, one can distinguish three cases. First, if F is strictly increasing and continuous at its α-quantile, the latter is singleton-valued and $V(x, y) = \mathbb{1}\{y \leq x\} - \alpha$ is a strict identification function both for q_α and for VaR_α. Second, if F is flat at its set-valued α-quantile, V is still a strict identification function

when plugging in the correct values for $q_\alpha(F)$ and $\mathrm{ES}_\alpha(F)$ for x_1 and x_2. Intuitively, the second component of V' adds a correction term corresponding to the one in the lower line of (3). The choice (4) is already utilised by Dimitriadis and Bayer (2019, Eq. (4)) for Z-estimation of a joint quantile and ES regression model and naturally shows up in consistent scoring functions for $(q_\alpha, \mathrm{ES}_\alpha)$, see Fissler and Ziegel (2016, Corollary 5.5). Finally notice that the $\mathrm{ES}_\alpha(F)$ is sometimes also defined as the upper average quantile over VaR_β with $\beta \in (\alpha, 1)$. Then, our results apply *mutatis mutandis*.

2 Formal statement of main result

The assertion of Theorem 4, and in particular its proof, parallels Osband's principle for consistent scoring functions Fissler and Ziegel (2016, Theorem 3.2), see also Osband (1985), Gneiting (2011). Up to our knowledge, the assertion has first been stated in the PhD thesis Fissler (2017, Proposition 3.2.1). We need the following assumptions.

Assumption 1 Let $\mathcal{F}$ be a convex class of distributions on O such that for every $x \in \mathrm{int}(A) \subseteq \mathbb{R}^k$ there are $F_1, \ldots, F_{k+1} \in \mathcal{F}$ satisfying $0 \in \mathrm{int}\big(\mathrm{conv}(\{\bar{V}(x, F_1), \ldots, \bar{V}(x, F_{k+1})\})\big)$, where for any set $B \subseteq \mathbb{R}^k$, $\mathrm{int}(B)$ denotes the interior of B and $\mathrm{conv}(B)$ denotes the convex hull of B.

Assumption 2 For every $y \in \mathbb{R}^d$ there exists a sequence $(F_n)_{n \in \mathbb{N}}$ of distributions $F_n \in \mathcal{F}$ that converges weakly to the Dirac-measure δ_y and a compact set $K \subset \mathbb{R}^d$ such that the support of F_n is contained in K for all n.

Assumption 3 Suppose that for Lebesgue almost all $x \in \mathrm{int}(A)$ the maps $V(x, \cdot)$ and $V'(x, \cdot)$ are locally bounded. Moreover, suppose that the complement of the set

$$C := \{(x, y) \in \mathrm{int}(A) \times O \mid V(x, \cdot) \text{ and } V'(x, \cdot) \text{ are continuous at the point } y\}$$

has $(k + d)$-dimensional Lebesgue measure zero.

Assumptions 1, 2, and 3 basically correspond to Assumptions (V1), (F1), and (VS1) in Fissler and Ziegel (2016), respectively. Assumption 1 ensures that the class $\mathcal{F}$ is sufficiently rich, implying in particular the surjectivity of T onto $\mathrm{int}(A)$ and the fact that there are no redundancies in V in the sense that all its components are needed; see Remark 5 for some further comments. Assumptions 2 and 3 ensure that $V(x, y)$ can be approximated by a sequence of integrals $\bar{V}(x, F_n)$.

Theorem 4 *Let* $T : \mathcal{F} \twoheadrightarrow A \subseteq \mathbb{R}^k$ *be a functional with a strict $\mathcal{F}$-identification function* $V : A \times O \to \mathbb{R}^k$. *Then the following two assertions hold:*

(i) *If* $h \colon A \to \mathbb{R}^{k \times k}$ *is a matrix-valued function with* $\det(h(x)) \neq 0$ *for all* $x \in A$, *then* $V'(x, y) = h(x)V(x, y)$ *is also a strict $\mathcal{F}$-identification function for T.*

---

for the set-valued q_α, but it is only a (non-strict) identification function for the singleton-valued VaR_α. Third, if F is discontinuous at $\mathrm{VaR}_\alpha(F)$ such that $F(\mathrm{VaR}_\alpha(F)) > \alpha$ (that is, if $F \notin \mathcal{F}_\alpha(\mathbb{R})$), neither q_α nor VaR_α are identified by V.

(ii) *Let V satisfy Assumption 1 and let $V' : \mathsf{A} \times \mathsf{O} \to \mathbb{R}^k$ be an $\mathcal{F}$-identification function for T. Then there is a matrix-valued function $h : \text{int}(\mathsf{A}) \to \mathbb{R}^{k \times k}$ such that*

$$\bar{V}'(x, F) = h(x)\bar{V}(x, F)$$

for all $x \in \text{int}(\mathsf{A})$ and for all $F \in \mathcal{F}$.

If V' is a strict $\mathcal{F}$-identification function for T and it also satisfies Assumption 1, then additionally $\det(h(x)) \neq 0$ for all $x \in \text{int}(\mathsf{A})$. If the integrated identification functions $\bar{V}(\cdot, F)$ and $\bar{V}'(\cdot, F)$ are continuous, then also h is continuous, which implies that either $\det(h(x)) > 0$ for all $x \in \text{int}(\mathsf{A})$ or $\det(h(x)) < 0$ for all $x \in \text{int}(\mathsf{A})$.

Moreover, if $\mathcal{F}$ satisfies Assumption 2 and V, V' satisfy Assumption 3 it even holds that

$$V'(x, y) = h(x)V(x, y) \qquad (5)$$

for Lebesgue almost all $(x, y) \in \text{int}(\mathsf{A}) \times \mathsf{O}$.

Proof of Theorem 4 Part (i) is a direct consequence of the linearity of the expectation. For (ii), the proof of the existence of h follows along the lines of Theorem 3.2 in Fissler and Ziegel (2016). One just needs to replace $\nabla \bar{S}(x, F)$ with $\bar{V}'(x, F)$. If V' satisfies Assumption 1 as well, one directly obtains that h must have full rank on $\text{int}(\mathsf{A})$ by exchanging the roles of V and V'. If the expected identification functions are both continuous, the continuity of h follows again exactly like in the proof of Theorem 3.2 in Fissler and Ziegel (2016).

For the pointwise assertion (5), consider $(x, y) \in \text{int}(\mathsf{A}) \times \mathsf{O}$ such that both $V(x, \cdot)$ and $V'(x, \cdot)$ are continuous at y. (Due to Assumption 3, this holds for Lebesgue almost all (x, y).) Let $(F_n)_{n \in \mathbb{N}} \subseteq \mathcal{F}$ be a sequence as specified in Assumption 2. That is, $(F_n)_{n \in \mathbb{N}}$ converges weakly to δ_y and the supports of all F_n are contained in some compact set $K \subset \mathbb{R}^d$. We claim that $\bar{V}(x, F_n)$ and $\bar{V}'(x, F_n)$ converge to $V(x, y)$ and $V'(x, y)$, respectively, providing the arguments for the former convergence only. By Skorohod's theorem, there is a sequence of random variables $(\xi_n)_{n \in \mathbb{N}}$ on some probability space with distributions F_n, such that ξ_n converges to y almost surely. By the continuous mapping theorem, $V(x, \xi_n)$ converges to $V(x, y)$ almost surely. Since $V(x, \cdot)$ is assumed to be locally bounded and since $\xi_n \in K$ almost surely, $V(x, \xi_n)$ is bounded almost surely. Hence, we can apply the dominated convergence theorem to conclude that $\bar{V}(x, F_n) = \mathbb{E}V(x, \xi_n) \to V(x, y)$. □

Remark 5 For part (i) of Theorem 4, no surjectivity assumption is necessary. In fact, the identification functions at (2) and (4) are also strict identification functions for (mean, variance) and $(q_\alpha, \text{ES}_\alpha)$, respectively, when considering the action domain $\mathsf{A} = \mathbb{R}^2$. However, it is obvious that part (ii) of Theorem 4 cannot hold without a surjectivity assumption. In fact, $V''(x_1, x_2, y) = V'(x_1, x_2, y)\mathbb{1}\{x_2 \geq 0\} + \mathbb{1}\{x_2 < 0\}$ would also be a strict identification function for (mean, variance) on the action domain $\mathbb{R}^2$.

On the other hand, also the richness, in particular, the convexity of $\mathcal{F}$ are needed. Just recall that on the class of symmetric distributions with strictly increasing distribution function, the mean and the median coincide. Hence, both $V(x, y) = x - y$ and

$V'(x, y) = \mathbb{1}\{y \leq x\} - 1/2$ are strict identification functions, but do not fulfil (5). The reason is that the class of symmetric distributions fails to be convex, unless all distributions have the same mean, in which case the interior of the action domain would be empty under surjectivity.

Remark 6 One may wonder about the flexibility concerning the dimension of an identification function. Suppose that $V(x, y)$ is a strict $\mathcal{F}$-identification function for some functional T, which takes values in $\mathbb{R}^k$. Clearly, for any matrix-valued function $h(x) \in \mathbb{R}^{\ell \times k}$ where possibly $\ell \neq k$, the product $V'(x, y) = h(x)V(x, y)$ is an $\mathcal{F}$-identification function for T. If $\ell > k$ and the rank of $h(x)$ is k for all x, V' is still a strict $\mathcal{F}$-identification function. However, V' will not satisfy Assumption 1, thus containing redundancies (in fact, the easiest way to construct such a V' is by simply copying some components of V). On the other hand, if $\ell < k$, the proof of Theorem 4 (ii) implies that V' cannot be a strict $\mathcal{F}$-identification function.

The latter statement can be exemplified by considering the systemic risk measure $\text{CoVaR}_{\alpha|\beta}$, which, given a two-dimensional observation (Y_1, Y_2), it is defined as the VaR_α of the conditional distribution of Y_2, given that Y_1 exceeds its VaR_β. Then, the pair $(\text{VaR}_\beta, \text{CoVaR}_{\alpha|\beta})$ is identifiable on the class of absolutely continuous distributions with positive density on $\mathbb{R}^2$ with a corresponding strict identification function

$$V(x_1, x_2, y_1, y_2) = \begin{pmatrix} \mathbb{1}\{x_1 \leq y_1\} - \beta \\ \mathbb{1}\{x_1 > y_1\}\big(\mathbb{1}\{x_2 \leq y_2\} - \alpha\big) \end{pmatrix},$$

see Fissler and Hoga (2022, Theorem 4.2). Due to the argument above, the one-dimensional identification function

$$V'(x_1, x_2, y_1, y_2) = \mathbb{1}\{x_1 > y_1\}\mathbb{1}\{x_2 > y_2\} - (1 - \alpha)(1 - \beta)$$

suggested in Banulescu-Radu et al. (2021) cannot be a strict identification function for $(\text{VaR}_\beta, \text{CoVaR}_{\alpha|\beta})$ on the class of absolutely continuous distributions with positive density, see Fissler and Hoga (2022, Remark 4.3).

References

Banulescu-Radu D, Hurlin C, Leymarie J, Scaillet O (2021) Backtesting marginal expected shortfall and related systemic risk measures. Manag Sci 67:5730–5754

Casgrain P, Larsson M, Ziegel J (2022) Anytime-valid sequential testing for elicitable functionals via supermartingales. Preprint. arXiv:2204.05680

Chamberlain G (1987) Asymptotic efficiency in estimation with conditional moment restrictions. J Economet 34(3):305–334

Creal D, Koopman SJ, Lucas A (2013) Generalized autoregressive score models with applications. J Appl Economet 28(5):777–795

Dimitriadis T, Bayer S (2019) A joint quantile and expected shortfall regression framework. Electron J Stat 13(1):1823–1871

Dimitriadis T, Fissler T, Ziegel JF (2021a) The efficiency gap. Preprint, (version v2). arXiv:2010.14146v2

Dimitriadis T, Patton AJ, Schmidt PW (2021b) Testing forecast rationality for measures of central tendency. Preprint. arXiv:1910.12545

Elliott G, Komunjer I, Timmermann A (2005) Estimation and testing of forecast rationality under flexible loss. Rev Econ Stud 72(4):1107–1125

Fissler T (2017) On higher order elicitability and some limit theorems on the Poisson and Wiener space. PhD thesis, University of Bern. http://biblio.unibe.ch/download/eldiss/17fissler_t.pdf

Fissler T, Hoga Y (2022) Backtesting systemic risk forecasts using multi-objective elicitability. Preprint, (version v4). arXiv:2104.10673v4

Fissler T, Ziegel JF (2016) Higher order elicitability and Osband's principle. Ann Stat 44(4):1680–1707

Fissler T, Frongillo R, Hlavinová J, Rudloff B (2021) Forecast evaluation of quantiles, prediction intervals, and other set-valued functionals. Electron J Stat 15(1):1034–1084

Gneiting T (2011) Making and evaluating point forecasts. J Am Stat Assoc 106:746–762

Hansen LP (1982) Large sample properties of generalized method of moments estimators. Econometrica 50(4):1029–54

Huber PJ (1967) The behavior of maximum likelihood estimates under nonstandard conditions. In: Proceedings of the Fifth Berkeley symposium on mathematical statistics and probability. University of California Press, Berkeley, pp 221–233

Jordan AI, Mühlemann A, Ziegel JF (2022) Characterizing the optimal solutions to the isotonic regression problem for identifiable functionals. Ann Inst Stat Math 74(3):489–514

Newey WK (1993) Efficient estimation of models with conditional moment restrictions. In: Maddala G, Rao C, Vinod H (eds) Handbook of statistics, vol 11: econometrics

Newey WK, McFadden D (1994) Large sample estimation and hypothesis testing. In: Engle RF, McFadden D (eds) Handbook of econometrics, chapter 36, vol 4. Elsevier, Amsterdam, pp 2111–2245

Nolde N, Ziegel JF (2017) Elicitability and backtesting: perspectives for banking regulation. Ann Appl Stat 11(4):1833–1874

Osband KH (1985) Providing incentives for better cost forecasting. PhD thesis, University of California, Berkeley. https://doi.org/10.5281/zenodo.4355667

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.