Econometric Theory, **39**, 2023, 27–69.
doi:10.1017/S0266466621000360

UNIFORM INFERENCE IN A GENERALIZED INTERVAL ARITHMETIC CENTER AND RANGE LINEAR MODEL

YANQIN FAN
University of Washington

XUETAO SHI
University of Sydney

Via generalized interval arithmetic, we propose a Generalized Interval Arithmetic Center and Range (GIA-CR) model for random intervals, where parameters in the model satisfy linear inequality constraints. We construct a constrained estimator of the parameter vector and develop asymptotically uniformly valid tests for linear equality constraints on the parameters in the model. We conduct a simulation study to examine the finite sample performance of our estimator and tests. Furthermore, we propose a coefficient of determination for the GIA-CR model. As a separate contribution, we establish the asymptotic distribution of the constrained estimator in Blanco-Fernández (2015, Multiple Set Arithmetic-Based Linear Regression Models for Interval-Valued Variables) in which the parameters satisfy an increasing number of random inequality constraints.

1. INTRODUCTION

1.1. Motivation and Main Contributions

Interval data have become prevalent in empirical research in diverse disciplines. Examples include: (i) The U.S. Energy Information Administration provides the state level minimum and maximum retail prices of electricity (García-Ascanio and Maté (2010)), and the U.S. Department of Agriculture provides daily low and high prices on agricultural commodities and livestock (Lin and González-Rivera (2016)); (ii) under the Health and Retirement Study (HRS) questionnaire protocol, a respondent is asked to report her wealth. If she does not comply, then the

We are grateful to Peter C. B. Phillips, Patrik Guggenberger, and two anonymous referees for their insightful comments on the previous version of this paper. We also thank Aman Ullah, participants of the 8th International Symposium on Econometric Analysis and Forecasting at Dongbei University of Finance and Economics, statistics seminar at Northeast Normal University, the 2018 International Symposium of Quantitative Economics at Jilin University, the 2018 Seattle–Vancouver Econometrics Conference at Simon Fraser University, and seminars at University of Alberta, Korea University, Seoul National University, UC Irvine, and UC Riverside for helpful discussions. This work was facilitated by the Hyak supercomputer system at the University of Washington. Address correspondence to Yanqin Fan, Department of Economics, University of Washington, Seattle, Washington 98195, USA; e-mail: fany88@uw.edu

respondent is asked to report if her wealth falls within a sequence of brackets. The HRS thus yields a wealth interval for each respondent; see Manski and Tamer (2002) and references therein; (iii) in markets with microstructural frictions, the researcher observes two prices for an asset at the same time: the bid price at which an investor could sell and the ask price at which an investor could buy. The price pair constitutes the bid-ask price interval, which measures the liquidity of the market and the size of the transaction cost; see Demsetz (1968) and Lee (1993); and (iv) in epidemiological studies, researchers focus on interval data frequently (Cowling et al. (2009); Kenah, Lipsitch, and Robins (2008); Vynnycky and Fine (2000)). During the recent coronavirus disease (COVID-19) outbreak, the study of key disease transmission parameters has drawn lots of attention by researchers; see Nishiura, Linton, and Akhmetzhanov (2020), Du et al. (2020), and Ganyani et al. (2020). Examples of the key parameters include the generation interval defined as time between infection events in an infector–infectee pair, and the (clinical-onset) serial interval defined as the time between symptom onsets in an infector–infectee pair.

Depending on the specific application, interval data may represent precise observations on random intervals of interest such as the bid-ask price interval in Example (iii) and the (clinical-onset) serial interval in Example (iv); or they may represent incomplete observations on random variables which are not always observed such as in Example (ii). In the former case, interval data are used to estimate and conduct inference on parameters in models for random intervals or simply interval models, while in the latter case, models of interest are for random variables and parameters in such models are typically only partially identified with interval data. See Aumann (1965), Debreu (1967), Hukuhara (1967), McShane (1969), and Artstein and Vitale (1975) for some early studies on set valued stochastic processes and sequences of random sets, and Manski and Tamer (2002), Beresteanu and Molinari (2008), Beresteanu, Molchanov, and Molinari (2011), and Bontemps, Magnac, and Maurin (2012) for the research on partially identified models.

This paper studies construction, estimation, and inference in interval models, where interval data represent precise observations on random intervals. Existing interval models for random intervals include model M_G proposed in Blanco-Fernández et al. (2015) and the Center and Range (CR) model in Neto and de Carvalho (2010); see the review of related works below. The CR model and its constrained estimation have proven to be useful for forecasting random intervals. However, formal statistical inference procedures and goodness-of-fit measures for the CR model are lacking in the current literature. All the existing measures of goodness-of-fit for the CR model are based on ad hoc combinations of goodness-of-fit measures for the center and range regressions, respectively; see Neto and de Carvalho (2010) and references therein. In addition, the CR model only allows the center (range) of the dependent interval to depend on the centers (ranges) of the covariate intervals. This could be restrictive in some applications. In contrast to the CR model, model M_G allows the center/range of the

https://doi.org/10.1017/S0266466621000360 Published online by Cambridge University Press

dependent interval to depend on both centers and ranges of the covariate intervals. However, the additional random constraints on the parameters in model M_G substantially complicate the asymptotic distribution of the estimators rendering inference extremely difficult if not impossible.

The first main contribution of this paper is to construct a new class of interval models referred to as *Generalized Interval Arithmetic Center and Range* (GIA-CR) models by using generalized interval arithmetic. Parameters in GIA-CR models satisfy linear inequality constraints. An example of such linear inequality constraints is nonnegative constraints imposed to ensure valid interval forecasts. It has long been recognized that incorporating inequality constraints in parameter estimation may yield efficiency gain, see, e.g., Liew (1976), Judge et al. (1984), and more recent works of Chernozhukov and Hong (2004) and Moon and Schorfheide (2009). Moreover, as noted in Andrews (2001): "in cases where the restrictions on the parameter space arise from prior information, tests that utilize this information have a considerable power advantage over tests that do not." Compared with the CR model, the GIA-CR model allows the center/range of the dependent interval to depend on both centers and ranges of the covariate intervals and as such significantly broadens the scope of applications of linear interval models.

Second, we propose a constrained estimator of the parameter and a coefficient of determination for the GIA-CR model. We establish asymptotic distributions of both the constrained estimator of the GIA-CR model and the constrained estimator of model M_G in Blanco-Fernández et al. (2015). Although the asymptotic distribution of the constrained estimator for the GIA-CR model can be derived from results in Andrews (1999), the approach in Andrews (1999) is not applicable to the constrained estimator of model M_G because of the increasing number of random inequality constraints. Instead, we exploit the powerful tools developed in Knight (2001, 2006) for linear programming estimators and M-estimators of boundaries, i.e., epi-convergence in distribution and point process convergence for extreme values, to derive the asymptotic distributions of the constrained estimators of model M_G and the GIA-CR model.[1]

Third, we construct asymptotically uniformly valid tests for a class of linear equality constraints in the GIA-CR model. Specifically, let θ^* denote the true parameter vector satisfying linear inequality constraints of the form: $R\theta^* \geq r$ for known matrix R and vector r under the maintained hypothesis. The considered null hypothesis specifies the value of a subvector of $R\theta^*$. An important and motivating example for this inference set-up is that of testing the correct specification of the CR model against the Interval Arithmetic Center and Range (IA-CR) model in which the parameters in the range regression satisfy nonnegativity constraints. Due to the presence of undetermined inequalities in $R\theta^* \geq r$ under the null hypothesis of this type, the null asymptotic distribution of the constrained estimator for the GIA-

[1]We note that Chernozhukov and Hong (2004) employ the same techniques in likelihood-based estimation and inference for two-sided and one-sided regression models and derives asymptotic properties of likelihood-based estimators as well as Bayes and Wald inference.

https://doi.org/10.1017/S0266466621000360 Published online by Cambridge University Press

CR model is discontinuous in some model parameters posing technical challenges in constructing asymptotically uniformly valid tests. In the case of *subvector inference* in other contexts (see the works cited in the literature review below), a two-step procedure for constructing critical values for asymptotically uniformly valid inference has been widely adopted. In this procedure, a confidence set for the nuisance parameter is constructed in the first step and a Bonferroni-type correction is applied in the second step.[2] For the GIA-CR model, although the null hypothesis specifies the value of a subvector of $R\theta^*$, the remaining components in $R\theta^*$ could be linearly dependent rendering direct application of the two-step approach with Bonferroni-type correction for subvector inference inapplicable. To address this issue, we propose to use Gauss-Jordan elimination to identify the nuisance parameter defined as an appropriate subvector of the remaining components in $R\theta^*$. Given the identified nuisance parameter and the inequalities it satisfies, we apply the two-step method based on confidence sets for the nuisance parameter with Bonferroni-type correction to constructing asymptotically uniformly valid tests for linear hypotheses in the GIA-CR model.

Lastly, to gauge the finite sample performance of the proposed estimator and test, we conduct a simulation study. The results confirm the superior performance of the proposed estimator and the asymptotically uniformly valid test for the correct specification of the CR model against the IA-CR model in finite samples. Since testing the correct specification of the CR model considered in the simulation belongs to subvector inference, we compare our test with the Conditional Likelihood Ratio (CLR) test introduced in Ketz (2018). We note, however, that the CLR test is not (directly) applicable to the more general hypotheses studied in Section 4 of this paper.

1.2. Related Works

This paper builds on works in two distinct literatures. *The first is the literature on interval models*. Broadly speaking, there are two approaches to modeling interval data. They are the interval arithmetic approach and the bivariate regression approach. The interval arithmetic approach adopts interval arithmetic to model directly relations between random intervals. The most general model based on this approach in the current literature is model M_G proposed in Blanco-Fernández et al. (2015). Model M_G makes use of the canonical decomposition of an interval in terms of its center and range to model the dependent interval directly via covariate intervals with an interval error term. To ensure that the interval error is well defined, parameters in model M_G must satisfy an increasing number of random inequality constraints which are much stronger than the nonnegativity constraints necessary to ensure that the predicted interval at any covariate interval is always well defined. We refer interested readers to Blanco-Fernández et al. (2015) for

[2]To simplify the exposition, we refer to a parameter in which the asymptotic distribution of an estimator or the null asymptotic distribution of a test statistic is discontinuous as the nuisance parameter.

the constrained estimation of model M_G and a review of the interval arithmetic approach to modeling interval data.

The bivariate regression approach models jointly either the left and right end points of the dependent interval or the center and range of the dependent interval. An important example based on this approach is the CR model in which one regression relates the center of the dependent interval to the centers of the covariate intervals, and the other relates the range of the dependent interval to the ranges of the covariate intervals. To ensure that the predicted range of the dependent interval at any covariate interval is always nonnegative, parameters in the range regression are restricted to be nonnegative. Neto and de Carvalho (2010) propose to estimate the range regression of the CR model by a constrained OLS estimator restricting the coefficients in the range regression to be nonnegative.[3] Applications of the CR model and methods based on modeling the left and right end points of the dependent interval include Han et al. (2008); Han, Hong, and Wang (2012), and González-Rivera and Lin (2013).

The second is the literature on uniform subvector inference. Methods for constructing asymptotically uniformly valid subvector inference in the presence of discontinuity have been proposed in different contexts. They include Bounds tests, the least favorable approach, and tests based on confidence sets for nuisance parameters. See Sect. 4.3.2 in Silvapulle and Sen (2005) for a brief discussion of all three approaches.[4] Among these proposals, the two-step approach based on confidence sets for nuisance parameters and a Bonferroni-type correction has proven to perform well. There are several works that adopt this approach. Berger and Boos (1994) and Silvapulle (1996) study some specific parametric testing problems. In a single-equation instrumental variables regression with possibly "weak" instrumental variables, Staiger and Stock (1997) construct a confidence region for the parameters based on such a method. Romano and Wolf (2000) construct a confidence interval for a univariate mean that has finite sample validity. For moment equality models with overidentifying inequality moment conditions, Moon and Schorfheide (2009) propose asymptotically uniformly valid tests and confidence sets for the parameters of interest. Chernozhukov et al. (2013) construct confidence intervals for marginal effects in nonlinear panel data models. For testing a finite number of moment inequalities, Romano, Shaikh, and Wolf (2014) construct asymptotically uniformly valid confidence sets for parameters characterized by the moment inequalities. Finally, McCloskey (2017) considers general nonstandard testing problems in which the asymptotic distribution of a test statistic is discontinuous in a nuisance parameter under the null hypothesis. We refer interested readers to Romano et al. (2014) and McCloskey (2017) for other related works using similar two-step approaches.

[3] Golan and Ullah (2017) propose an information theoretic approach to estimating linear interval models.

[4] Wolak (1987, 1989, 1991) develop tests for the null hypothesis of inequality constraints based on the least favorable approach. Silvapulle and Sen (2005) provide a comprehensive and systematic treatment of constrained inference via the least favorable approach.

1.3. Organization of the Rest of This Paper and Notation

The rest of this paper is organized as follows. Section 2 introduces the GIA-CR model and two constrained estimators: one for the GIA-CR model and one for model M_G. It also constructs a goodness-of-fit measure for the GIA-CR model. Section 3 establishes an asymptotic theory for the constrained estimators for the GIA-CR model and model M_G. In Section 4, we first provide a detailed construction and technical treatment of asymptotically uniformly valid tests for the correct specification of the CR model against the IA-CR model. The theoretical analysis in this part builds on Romano et al. (2014) and McCloskey (2017). Then we construct asymptotically uniformly valid tests in the GIA-CR model. Section 5 reports results from a simulation study. The last section offers some concluding remarks. Technical proofs are collected in Appendix A. Appendix B contains a review of generalized interval arithmetic and random generalized intervals.

All limits are taken as $n \to \infty$. Let $|\boldsymbol{v}| \equiv \left(|\boldsymbol{v}_1|, \ldots, \left|\boldsymbol{v}_p\right|\right)'$ for any p-dimensional vector $\boldsymbol{v} = \left(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p\right)'$; $\boldsymbol{v} \geq \boldsymbol{u}$ means that $\boldsymbol{v}_j \geq \boldsymbol{u}_j$ for $j = 1, \ldots, p$; and $\|\boldsymbol{v}\|$ denotes the euclidean norm of $\boldsymbol{v}$. $\boldsymbol{I}_p$ denotes the identity matrix of dimension p. The notation $a \sim b$ means that $a/b \to 1$ as appropriate limits are taken. Let $\mathbb{1}(\cdot)$ be the indicator function. The remaining notations concern operations on generalized intervals and generalized random intervals; see Appendix B for details. For a generalized interval $A = [a_1, a_2]$, where a_1, a_2 are real numbers, let $\text{mid}A \equiv (a_1 + a_2)/2$ and $\text{sprA} \equiv (a_2 - a_1)/2$. A can also be expressed as $A = \left[\text{mid}A \pm \text{spr}A\right]$. For two generalized intervals A and B, the addition is defined as

$$A + B \equiv \left[(\text{mid}A + \text{mid}B) \pm (\text{spr}A + \text{spr}B)\right].$$

A generalized interval $A = [a_1, a_2]$ is a proper or simply an interval when $a_1 \leq a_2$; otherwise it is an improper interval. For two intervals A and B with $\text{spr}A \geq \text{spr}B$, the Hukuhara difference is an interval defined as

$$A -_H B \equiv \left[(\text{mid}A - \text{mid}B) \pm (\text{spr}A - \text{spr}B)\right].$$

For two generalized intervals A and B, $A -_{GH} B$ is a generalized interval denoting the Generalized Hukuhara difference between A and B:

$$A -_{GH} B \equiv \left[(\text{mid}A - \text{mid}B) \pm (\text{spr}A - \text{spr}B)\right].$$

In contrast to the Hukuhara difference, the Generalized Hukuhara difference always exists. Let $d(A, B)$ denote the L_2-type metric between intervals A and B defined as

$$d(A, B) \equiv \left((\text{mid}A - \text{mid}B)^2 + (\text{spr}A - \text{spr}B)^2\right)^{\frac{1}{2}}.$$

More discussion on the L_2-type metric between two intervals can be found in Appendix B.2. Let $\mathbb{E}(\cdot)$ denote the expectation on a random variable. For a generalized random interval X, define the Aumann expectation as

$$\mathbb{E}_A(X) \equiv \left[\mathbb{E}(\text{mid}X) \pm \mathbb{E}(\text{spr}X)\right],$$

https://doi.org/10.1017/S0266466621000360 Published online by Cambridge University Press

whenever $\mathbb{E}(\text{mid}X)$ and $\mathbb{E}(\text{spr}X)$ exist. Define the Fréchet variance of a generalized random interval as

$$Var_F(X) \equiv \mathbb{E}\left(d^2(X, \mathbb{E}_A(X))\right)$$

and the conditional expectation as

$$\mathbb{E}_A(X \mid \cdot) \equiv \left[\mathbb{E}(\text{mid}X \mid \cdot) \pm \mathbb{E}(\text{spr}X \mid \cdot)\right].$$

2. THE MODEL, ESTIMATION, AND COEFFICIENT OF DETERMINATION

Let $Y, X_1, \ldots, X_k$ denote $(k+1)$ random intervals, where Y is the dependent interval and $X_1, \ldots, X_k$ are covariate intervals. We call random interval X a degenerate random interval if $\Pr(\text{spr}X = 0) = 1$. We allow some of the covariate intervals $X_1, \ldots, X_k$ be degenerate intervals, and let $d \geq 0$ denote the number of degenerate covariate intervals. Without loss of generality, let the last d covariates be degenerate. Further let $\{Y_i, X_{1i}, \ldots, X_{ki}\}_{i=1}^n$ denote a random sample on $(Y, X_1, \ldots, X_k)$. Define a $(2k-d)$ dimensional random vector as $\boldsymbol{X}_i \equiv \left(\text{mid}\boldsymbol{X}_i', \text{spr}\boldsymbol{X}_i'\right)'$, where $\text{mid}\boldsymbol{X}_i \equiv (\text{mid}X_{1i}, \ldots, \text{mid}X_{ki})'$ and $\text{spr}\boldsymbol{X}_i \equiv \left(\text{spr}X_{1i}, \ldots, \text{spr}X_{(k-d)i}\right)'$. We note that $\text{spr}\boldsymbol{X}_i$ only contains the ranges of $(k-d)$ nondegenerate covariate intervals.

2.1. Model M_G and the GIA-CR Model

To motivate the generalized interval arithmetic representation of our model, we first introduce model M_G proposed in Blanco-Fernández et al. (2015). It takes the following form:

$$Y_i = \left[\boldsymbol{X}_i'\boldsymbol{\alpha}^* \pm |\boldsymbol{X}_i|'\boldsymbol{\beta}^*\right] + \Delta_i, \tag{1}$$

where $\boldsymbol{\alpha}^* \in \mathbb{R}^{2k-d}$ and $\boldsymbol{\beta}^* \in \mathbb{R}_{\geq 0}^{2k-d}$ are the coefficient vectors and Δ_i is the *random interval error* defined as

$$\Delta_i \equiv Y_i -_H \left[\boldsymbol{X}_i'\boldsymbol{\alpha}^* \pm |\boldsymbol{X}_i|'\boldsymbol{\beta}^*\right]$$

satisfying $\mathbb{E}_A(\Delta_i \mid \boldsymbol{X}_i) = [\gamma^* \pm \delta^*]$. Define the parameter vector of interest as

$$\theta^* \equiv \left(\boldsymbol{\alpha}^{*\prime}, \gamma^*, \boldsymbol{\beta}^{*\prime}, \delta^*\right)' \in \mathbb{R}^{2l} \text{ for } l \equiv 2k-d+1,$$

where θ^* is identified in model M_G under the rank conditions in Assumption 3.2(i) in Section 3. Notice that model M_G does not include an intercept term as in the linear regression model. Conditioning on $\boldsymbol{X}_i$, the expectations of the midpoint and spread of Δ_i are γ^* and δ^*, which are both independent of $\boldsymbol{X}_i$. To ensure that Δ_i is an interval, $\boldsymbol{\beta}^*$ in model M_G must satisfy the additional random constraints that

$$\text{spr}Y_i - |\boldsymbol{X}_i|'\boldsymbol{\beta}^* \geq 0, \text{ for } i = 1, \ldots, n.$$

Blanco-Fernández et al. (2015) propose an inequality constrained estimator of $\left(\boldsymbol{\beta}^*, \delta^*\right)$ without establishing its asymptotic distribution. The difficulty lies in the

https://doi.org/10.1017/S0266466621000360 Published online by Cambridge University Press

presence of an increasing number of random inequality constraints. By exploiting tools used in Knight (2001, 2006) for linear programming estimators and M-estimators of boundaries, we derive the asymptotic distribution of the constrained estimator of model M_G in Section 3. However, its complex nature makes inference based on it extremely difficult if not impossible.

To facilitate inference and broaden the scope of applications of interval models, we propose the GIA-CR model below. It generalizes model M_G by dispensing the random inequality constraints and by allowing for general linear inequality constraints rather than $\boldsymbol{\beta}^* \geq \mathbf{0}$ in model M_G. We make use of the generalized interval arithmetic and the concept of Generalized Hukuhara difference introduced in Section 1 and discussed in Appendix B.[5]

Specifically, the GIA-CR model is composed of two parts: (i) the model in (1), where the *random generalized interval error* Δ_i is defined in terms of the Generalized Hukuhara difference:

$$\Delta_i \equiv Y_i -_{GH} \left[\boldsymbol{X}_i'\boldsymbol{\alpha}^* \pm |\boldsymbol{X}_i|'\,\boldsymbol{\beta}^*\right]$$

satisfying

$$\mathbb{E}_A\left(\Delta_i \mid \boldsymbol{X}_i\right) = \left[\gamma^* \pm \delta^*\right]; \tag{2}$$

and (ii) the parameter space Θ for θ^* defined as

$$\Theta = \left\{\theta \in \mathbb{R}^{2l} : R\theta \geq r\right\}, \tag{3}$$

where R is a known matrix of dimension $l_R \times 2l$ and r is a known vector of dimension l_R. The matrix R can be row rank deficient to incorporate constraints like $0 \leq \theta \leq 1$. Note that unlike model M_G, both $\left[\boldsymbol{X}_i'\boldsymbol{\alpha}^* \pm |\boldsymbol{X}_i|'\,\boldsymbol{\beta}^*\right]$ and Δ_i in the GIA-CR model can be generalized intervals although our observations $Y_i, X_{1i}, \ldots, X_{ki}$ are all intervals.[6]

Remark 2.1. In the GIA-CR model, we can replace X_i and $|X_i|$ with any known transformations of X_i. But for model M_G and the important IA-CR model to be introduced in Example 2.1 below, it is convenient to use $|X_i|$ for the purpose of forecasting. To avoid introducing too many notations, we use X_i and $|X_i|$ in the GIA-CR model as well.

The GIA-CR model has an alternative bivariate regression representation:

$$\text{mid}Y_i = \boldsymbol{X}_i'\boldsymbol{\alpha}^* + \text{mid}\Delta_i = \text{mid}\boldsymbol{X}_i'\alpha_m^* + \text{spr}\boldsymbol{X}_i'\alpha_s^* + \text{mid}\Delta_i \text{ and} \tag{4}$$

$$\text{spr}Y_i = |\boldsymbol{X}_i|'\,\boldsymbol{\beta}^* + \text{spr}\Delta_i = |\text{mid}\boldsymbol{X}_i|'\,\beta_m^* + \text{spr}\boldsymbol{X}_i'\beta_s^* + \text{spr}\Delta_i, \tag{5}$$

[5]In Han et al. (2008; 2012), generalized interval arithmetic is used to construct linear time series models for generalized random intervals, i.e., the observations are generalized intervals and no constraints are imposed on model parameters. Instead, this paper focuses on the case that the observations are intervals and generalized interval arithmetic is used to handle general linear inequality constraints.

[6]In fact, all the results in this paper remain valid when the observations are generalized intervals and the same inequality constraints are imposed on the model.

https://doi.org/10.1017/S0266466621000360 Published online by Cambridge University Press

where $\boldsymbol{\alpha}^* \equiv \left(\alpha_m^{*\prime}, \alpha_s^{*\prime}\right)'$ with $\alpha_m^* \equiv \left(\alpha_{m,1}^*, \ldots, \alpha_{m,k}^*\right)'$ and $\alpha_s^* \equiv \left(\alpha_{s,1}^*, \ldots, \alpha_{s,(k-d)}^*\right)'$; $\boldsymbol{\beta}^* \equiv \left(\beta_m^{*\prime}, \beta_s^{*\prime}\right)'$ with $\beta_m^* \equiv \left(\beta_{m,1}^*, \ldots, \beta_{m,k}^*\right)'$ and $\beta_s^* \equiv \left(\beta_{s,1}^*, \ldots, \beta_{s,(k-d)}^*\right)'$. The center and range of Δ_i satisfy that $\mathbb{E}(\text{mid}\Delta_i \mid \boldsymbol{X}_i) = \gamma^*$ and $\mathbb{E}(\text{spr}\Delta_i \mid \boldsymbol{X}_i) = \delta^*$.

Example 2.1 below presents the IA-CR model which is a special case of the GIA-CR model.

Example 2.1 (The IA-CR model)**.** Let $R = (\mathbf{0}_{l\times l}, \boldsymbol{I}_l)$ and $r = \mathbf{0}_{l\times 1}$. Then the parameter space Θ becomes

$$\Theta_F = \left\{\theta \in \mathbb{R}^{2l} : \boldsymbol{\beta} \geq \mathbf{0} \text{ and } \delta \geq 0\right\}.$$

Model (1) with parameter space Θ_F is referred to as the IA-CR model. To ensure that the predicted dependent interval at any covariate interval is an interval, the regressor in the second term on the right hand side of (1) is $|\boldsymbol{X}_i|$ and all elements in $\boldsymbol{\beta}^*$ and δ^* are required to be nonnegative. When there is no degenerate covariate, $\Theta = \Theta_F$, and $\alpha_s^* = \beta_m^* = 0$, (4) and (5) reduce to the CR model in Neto and de Carvalho (2010). More general than the CR model, the IA-CR model allows both the center and range of each covariate interval to affect the center and range of the dependent interval Y_i.

2.2. Constrained Estimation and the Coefficient of Determination

The constrained estimator of the parameters in model M_G in Blanco-Fernández, Corral, and González-Rodríguez (2011) is defined by minimizing the sum of the squared d-distances between Y_i and $\left(\left[\boldsymbol{X}_i'\boldsymbol{\alpha} \pm |\boldsymbol{X}_i|'\boldsymbol{\beta}\right] + [\gamma \pm \delta]\right)$ via the following constrained minimization problem:

$$\begin{aligned}\widehat{\theta} &= \arg\min_{\theta} \sum_{i=1}^{n} d^2\left(Y_i, \left[\boldsymbol{X}_i'\boldsymbol{\alpha} \pm |\boldsymbol{X}_i|'\boldsymbol{\beta}\right] + [\gamma \pm \delta]\right) \\ &= \arg\min_{\theta}\left[\sum_{i=1}^{n}\left(\text{mid}Y_i - \boldsymbol{X}_i'\boldsymbol{\alpha} - \gamma\right)^2 + \sum_{i=1}^{n}\left(\text{spr}Y_i - |\boldsymbol{X}_i|'\boldsymbol{\beta} - \delta\right)^2\right] \\ &\text{s.t. } \text{spr}Y_i - |\boldsymbol{X}_i|'\boldsymbol{\beta} \geq 0, \text{ for } i = 1, \ldots, n, \text{ and } \boldsymbol{\beta} \geq \mathbf{0}.\end{aligned} \tag{6}$$

Similarly, in the GIA-CR model, we construct the constrained estimator of θ^* by minimizing the sum of the squared d-distances between Y_i and $\left(\left[\boldsymbol{X}_i'\boldsymbol{\alpha} \pm |\boldsymbol{X}_i|'\boldsymbol{\beta}\right] + [\gamma \pm \delta]\right)$ in the parameter space Θ:

$$\begin{aligned}\widetilde{\theta} &= \arg\min_{\theta\in\Theta} \sum_{i=1}^{n} d^2\left(Y_i, \left[\boldsymbol{X}_i'\boldsymbol{\alpha} \pm |\boldsymbol{X}_i|'\boldsymbol{\beta}\right] + [\gamma \pm \delta]\right) \\ &= \arg\min_{\theta\in\Theta}\left[\sum_{i=1}^{n}\left(\text{mid}Y_i - \boldsymbol{X}_i'\boldsymbol{\alpha} - \gamma\right)^2 + \sum_{i=1}^{n}\left(\text{spr}Y_i - |\boldsymbol{X}_i|'\boldsymbol{\beta} - \delta\right)^2\right].\end{aligned} \tag{7}$$

https://doi.org/10.1017/S0266466621000360 Published online by Cambridge University Press

The estimators $\widehat{\theta}$ and $\widetilde{\theta}$ are solutions to quadratic minimization problems with linear inequality constraints and can be computed via built-in algorithms such as *quadprog* and *lsqlin* in Matlab.

With the obtained estimators, the interval residual $\widehat{\Delta}_i$ for model M_G is computed as

$$\widehat{\Delta}_i \equiv Y_i -_H \left[\boldsymbol{X}_i'\widehat{\boldsymbol{\alpha}} \pm |\boldsymbol{X}_i|'\widehat{\boldsymbol{\beta}}\right];$$

and the generalized interval residual $\widetilde{\Delta}_i$ for the GIA-CR model is defined as

$$\widetilde{\Delta}_i \equiv Y_i -_{GH} \left[\boldsymbol{X}_i'\widetilde{\boldsymbol{\alpha}} \pm |\boldsymbol{X}_i|'\widetilde{\boldsymbol{\beta}}\right]. \tag{8}$$

For any i, $\widehat{\Delta}_i$ is an interval because of the constraints that $\mathrm{spr}Y_i - |\boldsymbol{X}_i|'\boldsymbol{\beta} \geq 0$ for $i = 1, \ldots, n$; whereas $\widetilde{\Delta}_i$ is a generalized interval.

Remark 2.2. In model M_G, the constraints are only imposed on $\boldsymbol{\beta}$. Therefore, the minimization problem in (6) can be solved separately for the mid and range regressions. In fact, $\widehat{\boldsymbol{\alpha}}$ and $\widehat{\gamma}$ are OLS estimators of the slope coefficient and intercept term in the linear regression of $\mathrm{mid}Y_i$ on $\boldsymbol{X}_i$ and hence have closed-form expressions. In contrast, $\left(\widehat{\boldsymbol{\beta}}', \widehat{\delta}\right)$ solves the following constrained optimization problem and does not have closed-form expressions in general:

$$\left(\widehat{\boldsymbol{\beta}}', \widehat{\delta}\right) = \arg\min_{\boldsymbol{\beta}, \delta} \sum_{i=1}^{n} \left(\mathrm{spr}Y_i - |\boldsymbol{X}_i|'\boldsymbol{\beta} - \delta\right)^2$$
$$\text{s.t. } \mathrm{spr}Y_i - |\boldsymbol{X}_i|'\boldsymbol{\beta} \geq 0, \text{ for } i = 1, \ldots, n, \text{ and } \boldsymbol{\beta} \geq \boldsymbol{0}.$$

On the contrary, because the restrictions in the parameter space Θ for the GIA-CR model are imposed on θ, the minimization problem in (7) cannot be solved separately for the mid and range regressions.

Example 2.1 (continued)**.** Similar to model M_G, in the IA-CR model, the constraints are imposed on $\boldsymbol{\beta}$ and δ. Therefore, the minimization problem can be done separately: $\widetilde{\boldsymbol{\alpha}}$ and $\widetilde{\gamma}$ are OLS estimators; $\widetilde{\boldsymbol{\beta}}$ and $\widetilde{\delta}$ solve the following constrained optimization problem:

$$\left(\widetilde{\boldsymbol{\beta}}', \widetilde{\delta}\right) = \arg\min_{\boldsymbol{\beta}, \delta} \sum_{i=1}^{n} \left(\mathrm{spr}Y_i - |\boldsymbol{X}_i|'\boldsymbol{\beta} - \delta\right)^2$$
$$\text{s.t. } \boldsymbol{\beta} \geq \boldsymbol{0} \text{ and } \delta \geq 0.$$

The nonnegative constraints imposed in Θ_F ensure valid forecasts. (2) implies that

$$\mathbb{E}_A(Y_i \mid X_i) = \left[\boldsymbol{X}_i'\boldsymbol{\alpha}^* \pm |\boldsymbol{X}_i|'\boldsymbol{\beta}^*\right] + \left[\gamma^* \pm \delta^*\right].$$

Using the constrained estimator, the predictor of Y_0 at covariate interval $\boldsymbol{x}_0$ is given by

$$\widetilde{Y}_0 = \left[\boldsymbol{x}_0'\widetilde{\boldsymbol{\alpha}} \pm |\boldsymbol{x}_0|'\widetilde{\boldsymbol{\beta}}\right] + \left[\widetilde{\gamma} \pm \widetilde{\delta}\right].$$

https://doi.org/10.1017/S0266466621000360 Published online by Cambridge University Press

As pointed out in Neto and de Carvalho (2010), the conceptual difficulty in constructing measures of goodness-of-fit for the CR model is due to the presence of two multiple linear regressions. As a result, some ad hoc combinations of the coefficients of determination for the center and range regressions are adopted. For example, Neto and de Carvalho (2010) propose the following three goodness-of-fit measures for the CR model:

$$\min\left(\mathcal{R}_c^2, \mathcal{R}_r^2\right), \frac{\mathcal{R}_c^2 + \mathcal{R}_r^2}{2}, \text{ and } \max\left(\mathcal{R}_c^2, \mathcal{R}_r^2\right),$$

where $\mathcal{R}_c^2$ and $\mathcal{R}_r^2$ are the coefficients of determination for the center and range regressions, respectively.

Based on the d-metric and the fact that $\mathbb{E}_A(\Delta_i|X_i) = [\gamma^* \pm \delta^*]$, we now extend the coefficient of determination for linear regressions for random variables to the GIA-CR model for random intervals avoiding the ad hoc nature of existing goodness-of-fit measures for the CR model. Define the residual sum of squares for the GIA-CR model as

$$RSS_G \equiv \sum_{i=1}^{n} d^2\left(\widetilde{\Delta}_i, \left[\widetilde{\gamma} \pm \widetilde{\delta}\right]\right) = \sum_{i=1}^{n} d^2\left(Y_i, \widetilde{Y}_i\right),$$

where $\left[\widetilde{\gamma} \pm \widetilde{\delta}\right]$ is the estimated mean of Δ_i, $\widetilde{\Delta}_i$ is defined in (8), and

$$\widetilde{Y}_i = \left[\boldsymbol{X}_i'\widetilde{\boldsymbol{\alpha}} \pm |\boldsymbol{X}_i|'\widetilde{\boldsymbol{\beta}}\right] + \left[\widetilde{\gamma} \pm \widetilde{\delta}\right].$$

Together with the analogous definition of the total sum of squares as $TSS_G \equiv \sum_{i=1}^n d^2\left(Y_i, \overline{Y}\right)$, where $\overline{Y} = n^{-1}\sum_{i=1}^n Y_i$, we define the coefficient of determination for the GIA-CR model as

$$\mathcal{R}_G^2 \equiv 1 - \frac{RSS_G}{TSS_G} = 1 - \frac{\sum_{i=1}^n d^2\left(\widetilde{\Delta}_i, \left[\widetilde{\gamma} \pm \widetilde{\delta}\right]\right)}{\sum_{i=1}^n d^2\left(Y_i, \overline{Y}\right)}. \quad \textbf{(9)}$$

In the special case when Y is degenerate, i.e., a random variable, $\mathcal{R}_G^2$ reduces to the coefficient of determination for multiple regressions.

The coefficient of determination $\mathcal{R}_G^2$ inherits all the properties of that for multiple regression which are summarized in the following proposition. The value of $\mathcal{R}_G^2$ ranges from 0 to 1. Further, $\mathcal{R}_G^2 = 0$ indicates no linear relationship between the dependent interval and covariates; $\mathcal{R}_G^2 = 1$ indicates that the fitted model explains all variation of the dependent interval; and the value $\mathcal{R}_G^2$ is nondecreasing with inclusion of more covariates.

PROPOSITION 2.1. *(i)* $0 \le \mathcal{R}_G^2 \le 1$ *and (ii)* $\mathcal{R}_G^2$ *is nondeceasing in the number of covariates.*

3. ASYMPTOTIC PROPERTIES OF THE CONSTRAINED ESTIMATORS

The estimators $\widetilde{\theta}$ and $\widehat{\theta}$ defined in (7) and (6) for the GIA-CR model and model M_G are both inequality constrained estimators and their asymptotic properties are more difficult to establish than unconstrained estimators. This is particularly true for $\left(\widehat{\boldsymbol{\beta}}',\widehat{\delta}\right)$ due to the increasing number of random inequality constraints that it must satisfy, i.e., $\mathrm{spr}Y_i - |\boldsymbol{X}_i|'\boldsymbol{\beta} \geq 0$, for $i = 1,\ldots,n$. For a special case of model M_G referred to as model M with one covariate, $\left(\widehat{\boldsymbol{\beta}}',\widehat{\delta}\right)$ has a closed-form expression and its asymptotic properties are established in Blanco-Fernández, Colubi, and González-Rodríguez (2012). However, for general model M_G, there is no closed-form expression for $\left(\widehat{\boldsymbol{\beta}}',\widehat{\delta}\right)$, and the approach in Blanco-Fernández et al. (2012) breaks down.

Asymptotic properties of $\widetilde{\theta}$ can be established by applying the general results in Andrews (1999). To handle the increasing number of inequality constraints imposed on $\widehat{\theta}$, we exploit the powerful techniques used in Knight (2001, 2006) for linear programming estimators and M-estimators of boundaries, i.e., epi-convergence in distribution (Geyer (1994, 1996); Knight (1999); Pflug (1994, 1995)) and point process convergence for extreme values (Kallenberg (1983); Leadbetter, Lindgren, and Rootzén (1987); Resnick (1987)). Since the same techniques apply to $\widetilde{\theta}$ as well, we establish asymptotic properties of both $\widehat{\theta}$ and $\widetilde{\theta}$, including consistency and asymptotic distributions in this section.

Throughout the rest of this paper, we make the following assumption.[7]

Assumption 3.1. $\{Y_i, X_{1i},\ldots,X_{ki}\}_{i=1}^n$ denotes a random sample on $(Y, X_1,\ldots,X_k)$.

3.1. Consistency

Recall that $\theta \equiv \left(\boldsymbol{\alpha}',\gamma,\boldsymbol{\beta}',\delta\right)'$. Define $Z_{1n}(\cdot)$ as

$$Z_{1n}(\theta) = \frac{1}{n}\sum_{i=1}^{n}\left(\mathrm{mid}Y_i - \boldsymbol{X}_i'\boldsymbol{\alpha} - \gamma\right)^2 + \frac{1}{n}\sum_{i=1}^{n}\left(\mathrm{spr}Y_i - |\boldsymbol{X}_i|'\boldsymbol{\beta} - \delta\right)^2 + \varphi_{1n}(\theta),$$

where

$$\varphi_{1n}(\theta) = \begin{cases} 0, & \text{if } \mathrm{spr}Y_i - |\boldsymbol{X}_i|'\boldsymbol{\beta} \geq 0, \text{ for } i = 1,\ldots,n; \text{ and } \boldsymbol{\beta} \geq \boldsymbol{0} \\ \infty, & \text{otherwise} \end{cases}.$$

Using $Z_{1n}(\cdot)$, we can reformulate the constrained estimator of model M_G as an unconstrained estimator: $\widehat{\theta} = \arg\min_{\theta\in\mathbb{R}^{2l}} Z_{1n}(\theta)$. Similarly, let

$$Z_{2n}(\theta) = \frac{1}{n}\sum_{i=1}^{n}\left(\mathrm{mid}Y_i - \boldsymbol{X}_i'\boldsymbol{\alpha} - \gamma\right)^2 + \frac{1}{n}\sum_{i=1}^{n}\left(\mathrm{spr}Y_i - |\boldsymbol{X}_i|'\boldsymbol{\beta} - \delta\right)^2 + \varphi_{2}(\theta),$$

[7] The asymptotic results we establish for the GIA-CR model for i.i.d. data have straightforward extensions to strictly stationary short-memory time series such as various mixing processes. Due to space limitations, we will report details in future work.

https://doi.org/10.1017/S0266466621000360 Published online by Cambridge University Press

where

$$\varphi_2(\theta) = \begin{cases} 0, & \text{if } R\theta \geq r \\ \infty, & \text{otherwise} \end{cases}.$$

We obtain the alternative formulation for the estimator of the GIA-CR model as: $\widetilde{\theta} = \arg\min_{\theta \in \mathbb{R}^{2l}} Z_{2n}(\theta)$. Let $\dot{\boldsymbol{X}}_i \equiv (\boldsymbol{X}_i', 1)'$, $P_{xx} \equiv \mathbb{E}\left(\dot{\boldsymbol{X}}_i \dot{\boldsymbol{X}}_i'\right)$, and $Q_{xx} \equiv \mathbb{E}\left(\left|\dot{\boldsymbol{X}}_i\right| \left|\dot{\boldsymbol{X}}_i\right|'\right)$. We prove the consistency result using the notion of epi-convergence under the following assumption. The definition of $Var_F(\cdot)$ can be found in Section 1.

Assumption 3.2. (i) P_{xx} and Q_{xx} are nonsingular; (ii) $Var_F(\Delta_i) < \infty$.

Assumption 3.2 imposes standard regularity conditions on the random intervals in the model. Assumption 3.2 (i) is a rank condition ensuring that the true parameter θ^* is point identified. The following theorem states the consistency of $\widehat{\theta}$ and $\widetilde{\theta}$ for the GIA-CR model and Model M_G.

THEOREM 3.1. *Under Assumptions 3.1 and 3.2,* $\widehat{\theta} \xrightarrow{p} \theta^*$ *and* $\widetilde{\theta} \xrightarrow{p} \theta^*$.

3.2. Asymptotic Distribution

The asymptotic distribution of $\widehat{\theta}$ in model M_G remains unknown and is more involved than that of $\widetilde{\theta}$. We present it first.

3.2.1. *Model M_G.* The constraints: $\mathrm{spr}Y_i - |\boldsymbol{X}_i|' \boldsymbol{\beta} \geq 0$, for $i = 1, \ldots, n$, are equivalent to the constraint that

$$\min_{i=1,\ldots,n} \left[\mathrm{spr}Y_i - |\boldsymbol{X}_i|' \boldsymbol{\beta}\right] \geq 0. \quad \textbf{(10)}$$

Because $\min_{i=1,\ldots,n} \left[\mathrm{spr}Y_i - |\boldsymbol{X}_i|' \boldsymbol{\beta}\right]$ essentially describes the conditional distribution of $\mathrm{spr}\Delta_i$ near the endpoint of its support, the behavior of the conditional distribution function of $\mathrm{spr}\Delta_i$ near zero is the critical component determining the asymptotic distribution of the estimator $\widehat{\theta}$.

The following assumption imposes restrictions on the conditional distribution of $\mathrm{spr}\Delta_i$ commonly adopted in the extreme value literature. Let the conditional distribution function of $\mathrm{spr}\Delta_i$ given $|\boldsymbol{X}_i| = \boldsymbol{x}$ be $F_s(\cdot \mid \boldsymbol{x})$.

Assumption 3.3. Assume that

$$F_s(z \mid \boldsymbol{x}) \sim g(\boldsymbol{x}) F_s(z) \text{ as } z \searrow 0 \text{ uniformly in } \boldsymbol{x},$$

where $g(\cdot) > 0$ is a continuous function and $F_s(\cdot)$ is a distribution function.

Assumption 3.3 requires that for any $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$, the tail behaviors of $\mathrm{spr}\Delta_i$ conditional on $|\boldsymbol{X}_i| = \boldsymbol{x}_1$ or $|\boldsymbol{X}_i| = \boldsymbol{x}_2$ are equivalent up to a constant. If $|\boldsymbol{X}_i|$ and $\mathrm{spr}\Delta_i$ are independent, this assumption is trivially satisfied with $g(\cdot) \equiv 1$. Knight

(2001), Chernozhukov (2005), and Chernozhukov and Fernández-Val (2011) also impose similar assumptions. However, they require $F_s(\cdot)$ to have Pareto-type tail distribution due to their specific settings, while we impose no such restriction. Let $1/\kappa \equiv \lim_{t\to\infty} tF_s\left(1/\sqrt{t}\right) \in [0, +\infty]$. The value of κ characterizes the distribution of $\text{spr}\Delta_i$ near its left endpoint and determines the effect of the random constraints on the asymptotic distribution of $\widehat{\theta}$: (i) If $\text{spr}\Delta_i$ has a relatively high probability of being close to zero, e.g., when $F_s(0) > 0$, then $\lim_{t\to\infty} tF_s\left(1/\sqrt{t}\right) = +\infty$ and $\kappa = 0$; (ii) If $\text{spr}\Delta_i$ is very "unlikely" of being near zero, e.g., when $F_s(z) = 0$ for all z in a neighborhood of zero, then $\lim_{t\to\infty} tF_s\left(1/\sqrt{t}\right) = 0$ and $\kappa = +\infty$; and (iii) when $F_s(z)$ behaves like $h(z)z^2$ near zero, where $h(z)$ is a slowly varying function at zero,[8] then $\kappa \in (0, +\infty)$.

Assumption 3.4. $\mathbb{E}\left(|\text{mid}\Delta_i|^4\right) < \infty$, $\mathbb{E}\left(|\text{spr}\Delta_i|^4\right) < \infty$, and $\mathbb{E}\left(\|\boldsymbol{X}_i\|^4\right) < \infty$.

Denote

$$M(\psi) \equiv \psi' \begin{pmatrix} P_{xx} & \mathbf{0} \\ \mathbf{0} & Q_{xx} \end{pmatrix} \psi - 2\psi' \begin{pmatrix} \boldsymbol{I}_l & \mathbf{0} \\ \mathbf{0} & \boldsymbol{I}_l \end{pmatrix} W, \quad \textbf{(11)}$$

where $\psi \in \mathbb{R}^{2l}$ and $W \sim \mathcal{N}(0, \Lambda)$ in which $\mathcal{N}(0, \Lambda)$ denotes the normal distribution with mean zero and covariance matrix $\Lambda = Var\left(\dot{\boldsymbol{X}}_i'\text{mid}\Delta_i, \left|\dot{\boldsymbol{X}}_i\right|' \text{spr}\Delta_i\right)$. Furthermore, let $\psi \equiv \left(\boldsymbol{p}', q, \boldsymbol{u}', v\right)'$ with $\boldsymbol{p} \in \mathbb{R}^{2k-d}$, $q \in \mathbb{R}$, $\boldsymbol{u} \in \mathbb{R}^{2k-d}$, and $v \in \mathbb{R}$.

THEOREM 3.2. *Under Assumptions 3.1–3.4, we obtain:*

$$\sqrt{n}\left(\widehat{\theta} - \theta^*\right) \xrightarrow{d} \arg\min_{\psi} [M(\psi) + \phi_1(\psi)],$$

where $M(\psi)$ is defined in (11) and

$$\phi_1(\psi) = \begin{cases} 0, & \text{if } \sqrt{\frac{\kappa}{g(\Upsilon_i)}}\Gamma_i \geq \Upsilon_i'\boldsymbol{u}, \text{ for } i = 1, 2, \ldots; \\ & \text{and } \mathbb{1}\left(\boldsymbol{\beta}_j^* = 0\right)\boldsymbol{u}_j \geq 0 \text{ for } j = 1, \ldots, 2k-d \cdot \\ \infty, & \text{otherwise} \end{cases}$$

For each i, $\Gamma_i \equiv (\mathcal{E}_1 + \cdots + \mathcal{E}_i)^{\frac{1}{2}}$ for unit mean i.i.d. exponential random variables $\mathcal{E}_1, \mathcal{E}_2, \ldots$; $\Upsilon_1, \Upsilon_2, \ldots$ are independent and identically distributed with the same distribution as $|\boldsymbol{X}_i|$; the Γ_i's are independent of Υ_i's, and they are both independent of W.

Several observations surface from Theorem 3.2. First, the estimators of the center and range regressions $\left(\widehat{\boldsymbol{\alpha}}', \widehat{\gamma}\right)$ and $\left(\widehat{\boldsymbol{\beta}}', \widehat{\delta}\right)$ are asymptotically dependent through the covariance matrix Λ, unless $\text{mid}\Delta_i$ and $\text{spr}\Delta_i$ are conditionally independent. Second, the asymptotic distribution of $\sqrt{n}\left(\widehat{\boldsymbol{\alpha}}' - \boldsymbol{\alpha}^{*\prime}, \widehat{\gamma} - \gamma^*\right)$ is nor-

[8] A function $h(z) : (0, +\infty) \longmapsto (0, +\infty)$ is said to be slowly varying at z_0 if for any $m > 0$, it holds that $\lim_{z \searrow z_0} [h(z)/h(mz)] = 1$.

mal. Lastly, because of the inequality constraints, the asymptotic distribution of $\sqrt{n}\left(\widehat{\boldsymbol{\beta}}' - \boldsymbol{\beta}^{*\prime}, \widehat{\delta} - \delta^*\right)$ takes a complicated form given by the distribution of the minimizer of an inequality-constrained optimization problem. Although in general, there is no closed-form expression for the asymptotic distribution of $\left(\widehat{\boldsymbol{\beta}}', \widehat{\delta}\right)$, it is clear from the minimization problem that it is nonstandard and discontinuous in model parameters including parameters characterizing the tail behavior of the conditional distribution of $\mathrm{spr}\Delta_i$ through $\phi_1(\cdot)$. For example, the value of κ determines whether the constraint in (10) binds or not leading to different asymptotic distributions of $\left(\widehat{\boldsymbol{\beta}}', \widehat{\delta}\right)$.

3.2.2. *The GIA-CR Model.* Employing the same techniques, we obtain the following asymptotic distribution for the estimator $\widetilde{\theta}$. Denote R_b as the $l_b \times 2l$ submatrix composed of the l_b rows of R corresponding to binding inequalities in $R\theta^* \geq r$.

THEOREM 3.3. *Under Assumptions 3.1, 3.2, and 3.4, it holds that*

$$\sqrt{n}\left(\widetilde{\theta} - \theta^*\right) \xrightarrow{d} \arg\min_{\psi}\left[M(\psi) + \phi_2(\psi)\right],$$

where $M(\psi)$ is defined in (11) and 4

$$\boldsymbol{\phi}_2(\psi) = \begin{cases} 0, & \text{if } R_b\psi \geq \mathbf{0} \\ \infty, & \text{otherwise} \end{cases}.$$

Alternatively one may apply the approach in Andrews (1999) to obtain the asymptotic distribution of $\widetilde{\theta}$. Instead of studying the unconstrained minimization problem, Andrews (1999) focuses on the constrained minimization problem and shows that the asymptotic distribution of the estimator can be represented as the minimizer of a quadratic function over a convex cone.

Theorem 3.3 shows that the asymptotic distribution of $\widetilde{\theta}$ is discontinuous in $R\theta^*$ at r; see also Example 2.1 below. The asymptotic distribution of the Wald-type statistic in general exhibits discontinuity as well. This motivates the construction of the asymptotically uniformly valid test in the GIA-CR Model in Section 4.

Example 2.1 (continued)**.** The asymptotic distribution of the estimators for the IA-CR model can be obtained directly from Theorem 3.3. The matrix R_b corresponds to elements in $\left(\boldsymbol{\beta}^{*\prime}, \delta^*\right)$ that are zeros. Specifically, if Assumptions 3.1, 3.2, and 3.4 hold, then

$$\sqrt{n}\begin{pmatrix} \widetilde{\boldsymbol{\alpha}} - \boldsymbol{\alpha}^* \\ \widetilde{\gamma} - \gamma^* \end{pmatrix} \xrightarrow{d} P_{xx}^{-1} W_m,$$

https://doi.org/10.1017/S0266466621000360 Published online by Cambridge University Press

where $W_m \sim \mathcal{N}(0, \Lambda_m)$ with Λ_m being the covariance matrix of $\dot{\boldsymbol{X}}_i' \mathrm{mid}\Delta_i$; and

$$\sqrt{n}\begin{pmatrix} \widetilde{\boldsymbol{\beta}} - \boldsymbol{\beta}^* \\ \widetilde{\delta} - \delta^* \end{pmatrix} \xrightarrow{d} \arg\min_{\boldsymbol{u}, v}\left[(\boldsymbol{u}', v)\, Q_{xx} \begin{pmatrix} \boldsymbol{u} \\ v \end{pmatrix} - 2(\boldsymbol{u}', v)\, W_s + \phi_s(\boldsymbol{u}', v) \right],$$

where $W_s \sim \mathcal{N}(0, \Lambda_s)$ with Λ_s being the covariance matrix of $\left|\dot{\boldsymbol{X}}_i\right|' \mathrm{spr}\Delta_i$, and

$$\boldsymbol{\phi}_s(\boldsymbol{u}', v) = \begin{cases} 0, & \text{if } \mathbb{1}\left(\boldsymbol{\beta}_j^* = 0\right) \boldsymbol{u}_j \geq 0 \text{ for } j = 1, \ldots, 2k - d \text{ and } \mathbb{1}(\delta^* = 0)\, v \geq 0 \\ \infty, & \text{otherwise} \end{cases}.$$

The two convergences in distribution hold jointly. Like $\sqrt{n}\left(\widehat{\boldsymbol{\beta}}' - \boldsymbol{\beta}^{*\prime}, \widehat{\delta} - \delta^*\right)$, the asymptotic distribution of $\sqrt{n}\left(\widetilde{\boldsymbol{\beta}}' - \boldsymbol{\beta}^{*\prime}, \widetilde{\delta} - \delta^*\right)$ is discontinuous in some model parameters.

For the CR model with only one covariate, we can solve the minimization problem in Example 2.1 by following Andrews (1999). The CR model with one covariate can be expressed as follows:

$$\mathrm{mid}Y_i = \alpha_m^* \mathrm{mid}X_i + \gamma^* + \epsilon_i^c \text{ and}$$
$$\mathrm{spr}Y_i = \beta_s^* \mathrm{spr}X_i + \delta^* + \epsilon_i^r,$$

where $\mathbb{E}\left(\epsilon_i^c \mid X_i\right) = 0$, $\mathbb{E}\left(\epsilon_i^r \mid X_i\right) = 0$, $\alpha_m^* \in \mathbb{R}$, $\gamma^* \in \mathbb{R}$, $\beta_s^* \geq 0$, and $\delta^* \geq 0$. Let $K = \mathrm{diag}^{1/2}\left(Q_{xx}^{-1}\right)$, $Z = (Z_1, Z_2)' = K^{-1} Q_{xx}^{-1} W_s$, and $\rho_{ij} = \left[K^{-1} Q_{xx}^{-1} K^{-1}\right]_{ij}$ for $i, j = 1, 2$, where $\mathrm{diag}(A)$ returns a diagonal matrix whose elements are equal to the diagonal elements of matrix A. Suppose Assumptions 3.1, 3.2, and 3.4 hold.

(i) When $\beta_s^* > 0$ and $\delta^* > 0$,

$$\sqrt{n}\begin{pmatrix} \widetilde{\beta}_s - \beta_s^* \\ \widetilde{\delta} - \delta^* \end{pmatrix} \xrightarrow{d} \mathcal{N}\left(0, Q_{xx}^{-1} \Lambda_s Q_{xx}^{-1}\right);$$

(ii) When $\beta_s^* = 0$ and $\delta^* > 0$,

$$\sqrt{n}\begin{pmatrix} \widetilde{\beta}_s - \beta_s^* \\ \widetilde{\delta} - \delta^* \end{pmatrix} \xrightarrow{d} \mathbb{1}(Z_1 \geq 0)\, Q_{xx}^{-1} W_s + \mathbb{1}(Z_1 < 0)\, K \begin{pmatrix} 0 \\ Z_2 - \rho_{12} Z_1 \end{pmatrix};$$

(iii) When $\beta_s^* > 0$ and $\delta^* = 0$,

$$\sqrt{n}\begin{pmatrix} \widetilde{\beta}_s - \beta_s^* \\ \widetilde{\delta} - \delta^* \end{pmatrix} \xrightarrow{d} \mathbb{1}(Z_2 \geq 0)\, Q_{xx}^{-1} W_s + \mathbb{1}(Z_2 < 0)\, K \begin{pmatrix} Z_1 - \rho_{21} Z_2 \\ 0 \end{pmatrix};$$

(iv) When $\beta_s^* = 0$ and $\delta^* = 0$,

$$\sqrt{n}\begin{pmatrix} \widetilde{\beta}_s - \beta_s^* \\ \widetilde{\delta} - \delta^* \end{pmatrix}$$
$$\xrightarrow{d} \mathbb{1}(Z_1 > 0, Z_2 > 0)\, Q_{xx}^{-1} W_s + \mathbb{1}(Z_1 - \rho_{21} Z_2 > 0, Z_2 \leq 0)\, K \begin{pmatrix} Z_1 - \rho_{21} Z_2 \\ 0 \end{pmatrix}$$
$$+ \mathbb{1}(Z_1 \leq 0, Z_2 - \rho_{12} Z_1 > 0)\, K \begin{pmatrix} 0 \\ Z_2 - \rho_{12} Z_1 \end{pmatrix}.$$

https://doi.org/10.1017/S0266466621000360 Published online by Cambridge University Press

The asymptotic distribution of $\left(\widetilde{\beta}_s, \widetilde{\delta}\right)$ takes a complicated form and is discontinuous in β_s^* and δ^*: it is normal when β_s^* and δ^* are both positive; nonnormal otherwise.

4. WALD-TYPE TESTS FOR LINEAR HYPOTHESIS IN THE GIA-CR MODEL

In this section, we construct asymptotically uniformly valid tests for the linear hypothesis in the GIA-CR model of the following form:

$$H_0 : R_0\theta^* = r_0 \text{ against } H_1 : R_0\theta^* \neq r_0 \tag{12}$$

under the maintained hypothesis that $\theta^* \in \Theta = \{\theta : R\theta \geq r\}$, where R_0 and r_0 are known matrices of dimensions $J \times 2l$ and $J \times 1$, respectively. Furthermore, $R_0 \equiv \left(R_{00}', R_{01}'\right)'$ is of full row rank, and R_{00} is a submatrix of R. Without loss of generality, let $R = \left(R_{00}', R_{\Gamma}'\right)'$ and accordingly $r = \left(r_{-\Gamma}', r_{\Gamma}'\right)'$.

One motivating example for our testing framework is that of testing the correct specification of the CR model against the IA-CR model:

$$H_0^{CR} : \left(\alpha_s^{*\prime}, \beta_m^{*\prime}\right)' = \mathbf{0} \text{ against } H_1^{CR} : \left(\alpha_s^{*\prime}, \beta_m^{*\prime}\right)' \neq \mathbf{0},$$

under the maintained hypothesis that $\theta^* \in \Theta_F = \{\theta : \boldsymbol{\beta} \geq \mathbf{0} \text{ and } \delta \geq 0\}$. Components of α_s^* and β_m^* measure, respectively, the marginal effects of spr$\boldsymbol{X}_i$ on midY_i and $|\text{mid}\boldsymbol{X}_i|$ on sprY_i in the IA-CR model. Specification testing for the CR model against the IA-CR model is thus equivalent to testing the joint null hypothesis that spr$\boldsymbol{X}_i$ has no effect on midY_i and $|\text{mid}\boldsymbol{X}_i|$ has no effect on sprY_i. Individual marginal effects can also be tested using the test proposed in this paper. Testing H_0^{CR} belongs to the standard subvector inference with $\left(\beta_s^{*\prime}, \delta^*\right)'$ being the vector of nuisance parameters. For such a subvector hypothesis, the approach in Ketz (2018) can be applied to construct a CLR test statistic such that asymptotically the conditional null distribution of the test statistic given some sufficient statistic is nuisance parameter free.

As another example, consider testing equality constraints on parameters in the range regression of the GIA-CR model. Let $R = \left(\mathbf{0}_{l_R \times l}, R^*\right)$ and $R_0 = \left(\mathbf{0}_{J \times l}, R_0^*\right)$ for some known matrices R^* and R_0^*. Our tests become tests for linear equality constraints in linear regression models with linear inequality constraints. They extend the Wald-type test in Gourieroux, Holly, and Monfort (1982) for the special case that $R_0^* = R^*$ and the test in Silvapulle and Sen (2005) for the case that R^* is a submatrix of R_0^*.[9] In both cases considered in Gourieroux et al. (1982) and Silvapulle and Sen (2005), the inequalities in the parameter space are known to bind under H_0 and as a result, the asymptotic distribution of the inequality constrained estimator of θ^* or the test statistic under the null hypothesis is continuous in model parameters and is thus nuisance parameter free. When R_{00} is a proper submatrix of R, the asymptotic distribution of inequality constrained

[9] Rogers (1986) studies a modified Lagrange Multiplier test for the case that $R_0^* = R^*$.

https://doi.org/10.1017/S0266466621000360 Published online by Cambridge University Press

estimators or the null asymptotic distribution of the test statistic is typically discontinuous in nuisance parameters. Asymptotically uniformly valid tests for H_0 are not available in the current literature.

In the following subsections, we first introduce our test statistic and the asymptotic size of the test. Then we apply the two-stage approach with Bonferroni-type correction in McCloskey (2017) to construct asymptotically uniformly valid tests for H_0^{CR} against H_1^{CR}. Lastly, we construct tests for H_0 in the GIA-CR model.

4.1. The Test Statistic and Asymptotic Size

The test statistic we adopt is of Wald-type:

$$T_n(R_0, r_0) = n\left(R_0\widetilde{\theta} - r_0\right)'\left(R_0\Sigma_n R_0'\right)^{-1}\left(R_0\widetilde{\theta} - r_0\right), \tag{13}$$

for some positive definite weighting matrix $\Sigma_n \to_p \Sigma$ with Σ being a deterministic positive definite matrix. Σ_n can be chosen as the identity matrix or a consistent estimator of

$$\begin{pmatrix} P_{xx}^{-1} & \mathbf{0} \\ \mathbf{0} & Q_{xx}^{-1} \end{pmatrix} \Lambda \begin{pmatrix} P_{xx}^{-1} & \mathbf{0} \\ \mathbf{0} & Q_{xx}^{-1} \end{pmatrix}.$$

When R_{00} is a proper submatrix of R, the null asymptotic distribution of $R_0\widetilde{\theta}$ or $T_n(R_0, r_0)$ is discontinuous in $R_\Gamma\theta^*$.

We now introduce the test for H_0 based on $T_n(R_0, r_0)$ and its asymptotic size. The GIA-CR model can be fully characterized by the finite dimensional parameter $\theta^* \in \Theta$ and the infinite dimensional parameter $\mu^* \in \mathrm{M}$ that characterizes the distribution of $\{(Y_i, \boldsymbol{X}_i) : 1 \le i \le n\}$ and is consistent with the value θ^*. The space M can be restricted to be some compact metric space with a metric that induces weak convergence; see Andrews, Cheng, and Guggenberger (2020). Let $\omega \equiv (\theta^*, \mu^*) \in \mathcal{W}$. Denote $\boldsymbol{P}_\omega$ as the probability model indexed by ω, $\mathbb{E}_\omega$ as the expectation, Var_ω as the variance, and $\Pr_\omega$ as the probability computed with respect to $\boldsymbol{P}_\omega$. Let $\mathcal{W}_0$ be the collection of elements $\omega \in \mathcal{W}$ consistent with the null hypothesis and CV_n be a (possibly) sample dependent critical value (CV) for the test based on the test statistic $T_n(R_0, r_0)$. The asymptotic size of the resulting test is defined by

$$AsySz(T_n(R_0, r_0), CV_n) \equiv \limsup_{n\to\infty} \sup_{\omega\in\mathcal{W}_0} \Pr_\omega(T_n(R_0, r_0) > CV_n). \tag{14}$$

We aim to construct CV_n that controls the asymptotic size of the test based on $T_n(R_0, r_0)$.

Following Andrews and Cheng (2012, 2014), Cheng (2015), and Andrews et al. (2020), we will establish the asymptotic distribution of T_n under drifting parameter sequences $\omega_n \in \mathcal{W}_0 \to \omega \in \overline{\mathcal{W}}_0$, where $\overline{\mathcal{W}}_0$ is the closure of $\mathcal{W}_0$. For brevity, throughout the rest of the paper the terminology "$\omega_n \in \mathcal{W}_0$" refers to "drifting parameter sequence $\omega_n \in \mathcal{W}_0$ with limit $\omega \in \overline{\mathcal{W}}_0$". Let $\lambda_{\min}(A)$ denote the smallest eigenvalue of a matrix A. We make the following assumptions.

Assumption 4.1. For any $\omega \in \mathcal{W}_0$, (i) $\lambda_{\min}\left(\mathbb{E}_\omega\left(\dot{\boldsymbol{X}}_i\dot{\boldsymbol{X}}_i'\right)\right) \geq \lambda_1$ for some $\lambda_1 > 0$ and $\lambda_{\min}\left(\mathbb{E}_\omega\left(\left|\dot{\boldsymbol{X}}_i\right|\left|\dot{\boldsymbol{X}}_i\right|'\right)\right) \geq \lambda_2$ for some $\lambda_2 > 0$; (ii) $\mathbb{E}_\omega\left(|\text{mid}\Delta_i|^{4+\nu}\right) < M$, $\mathbb{E}_\omega\left(|\text{spr}\Delta_i|^{4+\nu}\right) < M$ and $\mathbb{E}_\omega\left(\|\boldsymbol{X}_i\|^{4+\nu}\right) < M$ for some $\nu > 0$ and $M < \infty$.

Assumption 4.2. For any $\omega_n \in \mathcal{W}_0$, $\Sigma_n \xrightarrow{p} \Sigma_\omega$, where Σ_ω is positive definite.

We finish this section by introducing notations that will be used in the subsequent analysis. Let $P_{\omega,xx} \equiv \mathbb{E}_\omega\left(\dot{\boldsymbol{X}}_i\dot{\boldsymbol{X}}_i'\right)$, $Q_{\omega,xx} \equiv \mathbb{E}_\omega\left(\left|\dot{\boldsymbol{X}}_i\right|\left|\dot{\boldsymbol{X}}_i\right|'\right)$, and

$$M_\omega(\psi) \equiv \psi'\begin{pmatrix} P_{\omega,xx} & \mathbf{0} \\ \mathbf{0} & Q_{\omega,xx} \end{pmatrix}\psi - 2\psi'\begin{pmatrix} \boldsymbol{I}_l & \mathbf{0} \\ \mathbf{0} & \boldsymbol{I}_l \end{pmatrix} W_\omega, \tag{15}$$

where $\psi \equiv (\boldsymbol{p}', q, \boldsymbol{u}', v)$ with $\boldsymbol{p} \in \mathbb{R}^{2k-d}$, $q \in \mathbb{R}$, $\boldsymbol{u} \in \mathbb{R}^{2k-d}$, $v \in \mathbb{R}$, and $W_\omega \sim \mathcal{N}(0, \Lambda_\omega)$ in which $\Lambda_\omega = Var_\omega\left(\dot{\boldsymbol{X}}_i'\text{mid}\Delta_i, \left|\dot{\boldsymbol{X}}_i\right|'\text{spr}\Delta_i\right)$.

4.2. Testing Correct Specification of the CR Model Against the IA-CR Model

To simplify notation, we denote the test statistic $T_n(R_0, r_0)$ in (13) for testing H_0^{CR} as T_n:

$$T_n = n\left(\widetilde{\alpha}_s', \widetilde{\beta}_m'\right)\left(R_0\Sigma_n R_0'\right)^{-1}\left(\widetilde{\alpha}_s', \widetilde{\beta}_m'\right)', \tag{16}$$

where $\widetilde{\alpha}_s$ and $\widetilde{\beta}_m$ are the estimators defined in (7) and

$$R_0 = \begin{pmatrix} \mathbf{0}_{(k-d)\times k} & \boldsymbol{I}_{(k-d)} & \mathbf{0}_{(k-d)\times 1} & \mathbf{0}_{(k-d)\times k} & \mathbf{0}_{(k-d)\times(k-d)} & \mathbf{0}_{(k-d)\times 1} \\ \mathbf{0}_{k\times k} & \mathbf{0}_{k\times(k-d)} & \mathbf{0}_{k\times 1} & \boldsymbol{I}_k & \mathbf{0}_{k\times(k-d)} & \mathbf{0}_{k\times 1} \end{pmatrix}.$$

The asymptotic distribution of T_n is discontinuous in $\left(\beta_m^{*\prime}, \beta_s^{*\prime}, \delta^*\right)'$, because the asymptotic distribution of $\sqrt{n}\left(\widetilde{\theta} - \theta^*\right)$ is discontinuous in $\left(\beta_m^{*\prime}, \beta_s^{*\prime}, \delta^*\right)'$ by applying Theorem 3.3. For any $\omega \in \mathcal{W}_0$, it holds that $\left(\alpha_s^{*\prime}, \beta_m^{*\prime}\right)' = \mathbf{0}$. Therefore, $\left(\beta_m^{*\prime}, \beta_s^{*\prime}, \delta^*\right)' = \left(\mathbf{0}, \beta_s^{*\prime}, \delta^*\right)'$ under the null hypothesis.

We decompose the model parameter $\omega \in \mathcal{W}_0$ into three groups: (η, π, ξ) based on their roles in the asymptotic distribution, where $\eta \equiv \left(\beta_s^{*\prime}, \delta^*\right)' \in \mathbb{R}_{\geq 0}^{k-d+1}$, $\pi \equiv (\text{vec}(P_{xx}), \text{vec}(Q_{xx}), \text{vec}(\Lambda), \text{vec}(\Sigma))' \in \Pi$, and $\xi \in \Xi$ consists of all other parameters and is infinite-dimensional. The space Ξ is consistent with the null hypothesis H_0. From the previous discussion, the null asymptotic distribution of T_n is discontinuous in η; π affects the limiting distribution of T_n but not its continuity; ξ does not affect the limiting distribution of T_n.

We consider the parameter sequence $\left\{(\eta_n, \pi_n, \xi_n) \in \mathbb{R}_{\geq 0}^{k-d+1} \times \Pi \times \Xi : n \geq 1\right\}$ and the localization parameter c and π_ω as the limit of $\sqrt{n}\eta_n$ and π_n:

$$\sqrt{n}\eta_n \to c = (c_1, \ldots, c_{k-d}, c_{k-d+1})' \in \overline{\mathbb{R}}_{\geq 0}^{k-d+1} \text{ and}$$
$$\pi_n \to \pi_\omega = \left(\text{vec}\left(P_{\omega,xx}\right), \text{vec}\left(Q_{\omega,xx}\right), \text{vec}(\Lambda_\omega), \text{vec}(\Sigma_\omega)\right)' \in \Pi,$$

https://doi.org/10.1017/S0266466621000360 Published online by Cambridge University Press

where $\overline{\mathbb{R}}_{\geq 0} \equiv \mathbb{R}_{\geq 0} \cup \{+\infty\}$. As shown in the lemma below, the asymptotic distribution of T_n under the null hypothesis and the drifting parameter sequence (η_n, π_n, ξ_n) depends on c and π_ω; whereas ξ_n (or the limiting value ξ_ω of ξ_n) does not affect the limiting distribution under any parameter sequence η_n and π_n.

Let

$$\Psi = \Psi(c, \pi_\omega) \equiv \arg\min_{\psi} [M_\omega(\psi) + \phi_\omega(\psi)],$$

where $M_\omega(\psi)$ is defined in (15) and

$$\phi_\omega(\psi) = \begin{cases} 0, & \text{if } \boldsymbol{u}_j \geq 0 \text{ for } j = 1, \ldots, k, \ \boldsymbol{u}_{j+k} + c_j \geq 0 \text{ for } j = 1, \ldots, k-d \\ & \text{and } v + c_{k-d+1} \geq 0 \\ \infty, & \text{otherwise} \end{cases}.$$

Then $R_0\Psi$ characterizes the limiting distribution of $\sqrt{n}\left(\widetilde{\alpha}_s', \widetilde{\beta}_m'\right)'$ under H_0^{CR} and the drifting parameter sequence (η_n, π_n, ξ_n).

LEMMA 4.1. *Under H_0^{CR} and the parameter sequence $(\eta_n, \pi_n, \xi_n) \in \mathbb{R}_{\geq 0}^{k-d+1} \times \Pi \times \Xi$ such that $\sqrt{n}\eta_n \to c$, $\pi_n \to \pi_\omega$ and $\xi_n \to \xi_\omega$, if Assumptions 3.1, 4.1, and 4.2 hold, then the asymptotic distribution of T_n is given by $(R_0\Psi)'\left(R_0\Sigma_\omega R_0'\right)^{-1}(R_0\Psi)$.*

As shown in Lemma 4.1, the null asymptotic distribution of T_n under the drifting sequences of distributions depends on the value of (c, π_ω). Let $\mathcal{C}_{c,\pi_\omega}(1-\vartheta)$ denote the $(1-\vartheta)$ quantile of the distribution of $(R_0\Psi)'\left(R_0\Sigma_\omega R_0'\right)^{-1}(R_0\Psi)$ given c and π_ω which can be simulated. Building on existing work, especially McCloskey (2017), we adopt the two-stage approach with Bonferroni-type correction to construct an asymptotically uniformly valid test for H_0^{CR}.

The detailed process consists of the following steps.

Step 1. (i) Construct the estimator $\widehat{\pi}$. Consistent estimator $\widehat{\pi}$ can be decomposed into two parts: $\left(\text{vec}\left(P_{\omega,xx}\right), \text{vec}\left(Q_{\omega,xx}\right)\right)$ is estimable from the sample $\{\boldsymbol{X}_i\}_{i=1}^n$; Λ_ω and Σ_ω can be estimated by using the residuals computed with $\widetilde{\theta}$; (ii) Construct confidence sets for $\sqrt{n}\eta_n$. Let $\widehat{\beta}_{s,OLS}$ and $\widehat{\delta}_{OLS}$ be the OLS estimators of β_s^* and δ^*. Simple calculation shows that

$$\sqrt{n}\left(\widehat{\beta}_{s,OLS}', \widehat{\delta}_{OLS}\right)' - \sqrt{n}\eta_n \xrightarrow{d} Z(\Lambda_{OLS}),$$

where $Z(\Lambda_{OLS})$ follows a multivariate normal distribution with zero mean and covariance matrix Λ_{OLS}. Denote $ES_\Lambda(\tau)$ as a set such that $\Pr(Z(\Lambda) \in ES_\Lambda(\tau)) = 1-\tau$. The confidence set I_τ for $\sqrt{n}\eta_n$ is defined as

$$I_\tau \equiv \left\{\varsigma \in \mathbb{R}_{\geq 0}^{k-d+1} : \varsigma \in \sqrt{n}\left(\widehat{\beta}_{s,OLS}', \widehat{\delta}_{OLS}\right)' - ES_{\widehat{\Lambda}_{OLS}}(\tau)\right\}.$$

The value $\widehat{\Lambda}_{OLS}$ can be computed using the standard approach in least squares estimation and is a consistent estimator for Λ_{OLS}. In cases where I_τ is empty, let $I_\tau = \{\mathbf{0}\}$.

https://doi.org/10.1017/S0266466621000360 Published online by Cambridge University Press

Step 2. We construct the ϑ level Bonferroni critical value CV_n^b for some $0 < \tau \leq \vartheta$ as

$$CV_n^b(\vartheta, \tau) \equiv \sup_{c \in I_{\vartheta - \tau}} \mathcal{C}_{c,\widehat{\pi}}(1-\tau). \tag{17}$$

The following proposition establishes the asymptotic validity of the test.

PROPOSITION 4.1. *Under Assumptions 3.1, 4.1, and 4.2, $AsySz(T_n, CV_n^b(\vartheta, \tau)) \leq \vartheta$.*

Remark 4.1. Adjusted Bonferroni critical value and Minimum Bonferroni critical value discussed in McCloskey (2017) are also applicable. The former considers the joint distribution of T_n and $\left(\widehat{\beta}'_{s,OLS}, \widehat{\delta}_{OLS}\right)$, and computes the optimal pair of significant levels for c and T_n; while the latter combines Bonferroni critical value and Adjusted Bonferroni critical value. The computational burden for the two critical values is significant when the dimension of β_s is large. Interested readers could refer to McCloskey (2017) for detailed implementation of such methods.

4.3. Testing H_0 in the GIA-CR Model

In this subsection, we extend the test for the correct specification of the CR model against the IA-CR model developed in the previous subsection to the problem of testing H_0 in (12). Since $R_{00}\theta^*$ is a subvector of $R\theta^*$, the null asymptotic distribution of the test statistic defined in (13) is discontinuous in $R_\Gamma\theta^*$, where θ^* satisfies the inequalities: $R_\Gamma\theta^* \geq r_\Gamma$. In contrast to the standard subvector inference such as that considered in the previous subsection, components of $R_\Gamma\theta^*$ could be linearly dependent rendering direct application of the first stage of the two-stage approach with Bonferroni-type correction in the previous subsection problematic. To address this potential issue, we suggest a three-stage approach for constructing asymptotically uniformly valid tests for H_0.

In the first stage, we identify binding and nonbinding inequalities in $R_{00}\theta^* \geq r_{-\Gamma}$; and employ the Gauss-Jordan elimination to identify a row basis of R_Γ based on which we define nuisance parameters and express the inequalities $R_\Gamma\theta^* \geq r_\Gamma$ in terms of the nuisance parameters. In the second stage, we construct confidence sets for the nuisance parameters. Lastly we construct the CV for our test using Bonferroni-type correction.

4.3.1. *Identification of Nuisance Parameters.* Under H_0, the inequalities in $R_{00}\theta^* \geq r_{-\Gamma}$ are known to bind or not to bind. Let R_{0b} denote the submatrix of R_{00} composed of rows corresponding to binding inequalities in $R_{00}\theta^* \geq r_{-\Gamma}$. R_{0b} can be identified directly from the null hypothesis.

Let $\eta \equiv R_\Gamma\theta^* \in \mathbb{R}^{l_\Gamma}$. The vector of nuisance parameters is defined in the following.

Definition 4.1. The vector of nuisance parameters, denoted as η^u, is defined as a subvector of η corresponding to a row basis of R_Γ.

https://doi.org/10.1017/S0266466621000360 Published online by Cambridge University Press

The nuisance parameters may not be unique, because row basis of R_Γ is not unique. On the other hand, the dimension of η^u is uniquely determined. By definition, the nuisance parameters are $\eta^u = R_\Gamma^u \theta^*$, where R_Γ^u is a submatrix of R_Γ with rows forming a row basis of R_Γ. When R_Γ is of *full row rank*, $R_\Gamma^u = R_\Gamma$ and the nuisance parameters are $\eta^u = \eta$. When R_Γ is *not of full row rank*, we compute R_Γ^u and Γ such that $R_\Gamma = \Gamma R_\Gamma^u$ by Gauss-Jordan elimination on the transpose of R_Γ. In terms of the nuisance parameters, the inequalities: $\eta = R_\Gamma \theta^* \geq r_\Gamma$ become:

$$\Gamma \eta^u \geq r_\Gamma. \tag{18}$$

This step does not depend on η^u being unique, because any η^u would allow us to express the inequalities $R_\Gamma \theta^* \geq r_\Gamma$ as (18). Below we present two examples to illustrate this step. For notational compactness, we let $\theta = (\theta_1, ..., \theta_{l^*})$ for an integer l^*.

Example 4.1. Suppose $l^* = 4$ and $H_0 : \theta_1^* = \theta_2^*$. Then $R_0 = (1, -1, 0, 0)$.

(i) Let $\Theta = \{\theta : \theta_1 - \theta_2 \geq 0, \theta_2 - \theta_3 \geq 0, \theta_3 - \theta_4 \geq 0\}$. Then

$$R = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix} \text{ and } r = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Under H_0, the first inequality in Θ binds resulting in $R_{0b} = (1, -1, 0, 0)$ and

$$R_\Gamma = \begin{pmatrix} 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix}.$$

Since R_Γ is of full row rank, the nuisance parameters are given by

$$\eta^u = \eta = R_\Gamma \theta^* = \begin{pmatrix} \theta_2^* - \theta_3^* \\ \theta_3^* - \theta_4^* \end{pmatrix}.$$

(ii) Let $\Theta = \{\theta : \theta_1 - \theta_2 \geq 0, 1 \geq \theta_2 - \theta_3 \geq 0, \theta_3 - \theta_4 \geq 0\}$. Then

$$R = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix} \text{ and } r = \begin{pmatrix} 0 \\ 0 \\ -1 \\ 0 \end{pmatrix}.$$

Under H_0, the first inequality in Θ binds resulting in $R_{0b} = (1, -1, 0, 0)$ and

$$R_\Gamma = \begin{pmatrix} 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix}.$$

In this case, R_Γ is not of full row rank. Applying Gauss-Jordan elimination to R_Γ' yields

$$R_\Gamma^u = \begin{pmatrix} 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix} \text{ and } \Gamma = \begin{pmatrix} 1 & 0 \\ -1 & 0 \\ 0 & 1 \end{pmatrix}.$$

https://doi.org/10.1017/S0266466621000360 Published online by Cambridge University Press

By definition, the nuisance parameters are given by

$$\eta^u = R_\Gamma^u \theta^* = \begin{pmatrix} \theta_2^* - \theta_3^* \\ \theta_3^* - \theta_4^* \end{pmatrix}$$

and the inequalities are

$$\Gamma \eta^u = \begin{pmatrix} \theta_2^* - \theta_3^* \\ -\left(\theta_2^* - \theta_3^*\right) \\ \theta_3^* - \theta_4^* \end{pmatrix} \geq \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix}.$$

Example 4.2. Let $l^* = 8$ and $H_0 : \theta_1^* = \theta_2^* = \theta_3^* = 0$. Then

$$R_0 = \begin{pmatrix} 1 & 0 & 0 & \\ 0 & 1 & 0 & \mathbf{0}_{3\times 5} \\ 0 & 0 & 1 & \end{pmatrix} \text{ and } r_0 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Let $\Theta = \{\theta : R\theta \geq r\}$, where $R' = \left(R_0', R_\Gamma'\right)$ and $r = (0, -1, 0, 0, -1, 0, 0, 0)'$ in which

$$R_\Gamma = \begin{pmatrix} & 1 & 0 & 0 & 0 & 0 \\ & -1 & 0 & 0 & 0 & 0 \\ \mathbf{0}_{5\times 3} & 0 & 1 & 0 & 1 & -1 \\ & 0 & 2 & 1 & 1 & 0 \\ & 0 & 1 & -1 & 2 & -3 \end{pmatrix}.$$

Under H_0, the first and third inequalities in Θ bind and the second inequality does not bind resulting in

$$R_{0b} = \begin{pmatrix} 1 & 0 & 0 & \mathbf{0}_{2\times 5} \\ 0 & 0 & 1 & \end{pmatrix}.$$

Since the rows of R_Γ are linearly dependent, we apply the Gauss-Jordan elimination to the transpose of R_Γ:

$$R_\Gamma' = \begin{pmatrix} & & \mathbf{0}_{3\times 5} & & \\ 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & 1 & 2 \\ 0 & 0 & -1 & 0 & -3 \end{pmatrix} \xrightarrow{\text{Gauss-Jordan elimination}} \begin{pmatrix} 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & 1 & -1 \\ & & \mathbf{0}_{5\times 5} & & \end{pmatrix},$$

https://doi.org/10.1017/S0266466621000360 Published online by Cambridge University Press

and conclude that the first, third, and fourth rows of R_Γ constitute a row basis of R_Γ. Finally, we get that

$$\Gamma = \begin{pmatrix} 1 & 0 & 0 & \\ -1 & 0 & 0 & \\ 0 & 1 & 0 & \mathbf{0}_{5\times 5} \\ 0 & 0 & 1 & \\ 0 & 3 & -1 & \end{pmatrix} \text{ and } \eta^u = \begin{pmatrix} \theta_4^* \\ \theta_5^* + \theta_7^* - \theta_8^* \\ 2\theta_5^* + \theta_6^* + \theta_7^* \end{pmatrix}.$$

4.3.2. *The Null Asymptotic Distribution of $T_n(R_0, r_0)$ Under Drifting Sequences and the Testing Procedure.* We consider the drifting model parameters $\left(\eta_n^u, \pi_n, \xi_n\right)$, where π_n and ξ_n are defined in the same way as in Section 4.2. Since the nuisance parameters η^u satisfy inequalities in (18), we consider local sequences η_n^u such that

$$c = \lim_{n\to\infty} \sqrt{n}\left(\Gamma\eta_n^u - r_\Gamma\right) \in \overline{\mathbb{R}}_{\geq 0}^{l_\Gamma}.$$

The value of c is unique given $\omega_n \in \mathcal{W}_0$. Let

$$\Psi = \Psi(c, \pi_\omega) \equiv \arg\min_{\psi} \left[M_\omega(\psi) + \phi_\omega(\psi)\right],$$

where $M_\omega(\psi)$ is defined in (15) and

$$\phi_\omega(\psi) = \begin{cases} 0, & \text{if } R_{0b}\psi \geq 0 \text{ and } R_\Gamma\psi + c \geq 0 \\ \infty, & \text{otherwise} \end{cases}.$$

LEMMA 4.2. *Under $H_0: R_0\theta^* = r_0$ and $\left(\eta_n^u, \pi_n, \xi_n\right)$ defined above, if Assumptions 3.1, 4.1, and 4.2 hold, then the asymptotic distribution of $T_n(R_0, r_0)$ is given by* $(R_0\Psi)'\left(R_0\Sigma_\omega R_0'\right)^{-1}(R_0\Psi)$.

The null asymptotic distribution of $T_n(R_0, r_0)$ stated in Lemma 4.2 suggests the following procedure for computing the critical value of our test.

Step 1. (i) Consistently estimate π_ω using the constrained estimator $\widetilde{\theta}$, denoted as $\widehat{\pi}$;

(ii) Construct confidence sets for $\sqrt{n}\left(\Gamma\eta_n^u - r_\Gamma\right)$. By definition, $\eta_n^u = R_\Gamma^u\theta_n$. Denote $\widehat{\theta}_u$ as the unconstrained estimator defined as the minimizer of (6) without any constraint. Since $\sqrt{n}\left(\widehat{\theta}_u - \theta_n\right) \xrightarrow{d} \mathcal{N}(0, \Lambda)$ for some covariance matrix Λ, it holds that

$$\sqrt{n}\left(R_\Gamma^u\widehat{\theta}_u - \eta_n^u\right) \xrightarrow{d} Z(\Lambda) \sim \mathcal{N}\left(0, R_\Gamma^u\Lambda R_\Gamma^{u\prime}\right).$$

The confidence set $I_\tau\left(\eta_n^u\right)$ for η_n^u is obtained as $R_\Gamma^u\widehat{\theta}_u - \frac{1}{\sqrt{n}}ES_{\widehat{\Lambda}}(\tau)$, where $ES_\Lambda(\tau)$ is the set such that $\Pr(Z(\Lambda) \in ES_\Lambda(\tau)) = 1 - \tau$ and $\widehat{\Lambda}$ is a consistent estimator

https://doi.org/10.1017/S0266466621000360 Published online by Cambridge University Press

of Λ. The confidence set I_τ for $\sqrt{n}\left(\Gamma\eta_n^u - r_\Gamma\right)$ is calculated as

$$I_\tau = \left\{\varsigma \in \mathbb{R}_{\geq 0}^{l_\Gamma} : \varsigma = \sqrt{n}\left(\Gamma\iota - r_\Gamma\right),\ \iota \in I_\tau\left(\eta_n^u\right)\right\}.$$

In cases where I_τ is an empty set, let $I_\tau = \{\mathbf{0}\}$.

Step 2. Compute the ϑ level Bonferroni critical value CV_n^b for some $0 < \tau \leq \vartheta$ as

$$CV_n^b(\vartheta,\tau) \equiv \sup_{c \in I_{\vartheta-\tau}} \mathcal{C}_{c,\widehat{\pi}}(1-\tau),$$

where $\mathcal{C}_{c,\pi_\omega}(1-\tau)$ is the $(1-\tau)$th quantile of $(R_0\Psi)'\left(R_0\Sigma_\omega R_0'\right)^{-1}(R_0\Psi)$ given (c,π_ω). For any given (c,π_ω), the distribution $(R_0\Psi)'\left(R_0\Sigma_\omega R_0'\right)^{-1}(R_0\Psi)$ may not have a closed form expression but can be simulated.

The following theorem shows that the test is asymptotically uniformly valid.

THEOREM 4.1. *Assume that* $(R_0\Psi)'\left(R_0\Sigma_\omega R_0'\right)^{-1}(R_0\Psi)$ *is continuous at* $\mathcal{C}_{c,\pi_\omega}(1-\tau)$ *for all* $(c,\pi_\omega) \in \overline{\mathbb{R}}_{\geq 0}^{l_\Gamma} \times \Pi$. *Under Assumptions 3.1, 4.1, and 4.2, it holds that* $AsySz\left(T_n(R_0,r_0), CV_n^b(\vartheta,\tau)\right) \leq \vartheta$.

The continuity assumption in Theorem 4.1 may restrict the range of τ. Take $\Theta = \left\{\theta \in \mathbb{R}^{2l} : \beta_{m,1} \geq 0 \text{ and } \beta_{s,1} \geq 0\right\}$ as an example: the assumption is satisfied for all τ when $H_0 : \beta_{m,1}^* = 1$; and is satisfied for $\tau < 0.25$ when $H_0 : \beta_{m,1}^* = 0$. In the special case that $R_0\theta^*$ involves parameters that are not subject to any inequality constraint under Θ, the assumption is automatically satisfied.

5. A SIMULATION STUDY

In this section, we report results from a simulation study. The objective is twofold. First, we compare the finite sample performance of three estimators $\widehat{\theta}$, $\widetilde{\theta}$, and $\widehat{\theta}_u$ measured by the mean squared error (MSE), where $\widehat{\theta}$ is defined in (6), $\widetilde{\theta}$ is defined in (7), and $\widehat{\theta}_u$ is the unconstrained estimator defined as the minimizer of (6) without any constraint. Second, we compare the finite sample size and power performance of the test introduced in Section 4.1 denoted as the UF test, the Wald test based upon the unconstrained estimator $\widehat{\theta}_u$ denoted as WU test, and the CLR test introduced in Ketz (2018). Since the CLR test is developed for subvector inference, we focus on testing the correct specification of the CR model against the IA-CR model in this section.

5.1. Simulation Design

We generate data from the following IA-CR model with one covariate:

$$\begin{aligned}\text{mid}Y &= \alpha_m^*\text{mid}X + \alpha_s^*\text{spr}X + \text{mid}\Delta \text{ and}\\ \text{spr}Y &= \beta_m^*\left|\text{mid}X\right| + \beta_s^*\text{spr}X + \text{spr}\Delta,\end{aligned}$$

https://doi.org/10.1017/S0266466621000360 Published online by Cambridge University Press

where the random generalized interval error Δ is independent of $\boldsymbol{X}$. We consider five different DGPs: DGP1–DGP3 belong to model M_G, where Δ is a random interval; and DGP4 and DGP5 do not belong to model M_G.

It follows from Theorems 3.2 and 3.3 that the asymptotic distributions of the scaled and centered estimators $\widehat{\theta}$ and $\widetilde{\theta}$ are discontinuous with respect to the value of $\boldsymbol{\beta}^*$. The asymptotic distribution of $\widehat{\theta}$ further depends on the tail behavior of the distribution of sprΔ through the value of κ. To cover all the possible situations displayed in the theorems, we consider three different specifications of κ in DGP1–DGP3. In DGP4 and DGP5, the support of sprΔ includes negative values, and thus Δ is a random generalized interval. In DGP1–DGP5 below, midΔ and sprΔ are independent of each other, and the variance of sprΔ is designed to be the same. Let (Z_1, Z_2, Z_3, Z_4) be the random variables that follow the distribution

$$\begin{pmatrix} Z_1 \\ Z_2 \\ Z_3 \\ Z_4 \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{bmatrix} 1 & 0.8 & 0 & 0 \\ 0.8 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 8 \end{bmatrix} \right).$$

Denote $\Gamma(\alpha, \beta)$ as the gamma distribution with shape parameter α and rate parameter β.

DGP1 mid$X = Z_1$, spr$X = Z_2^2$, mid$\Delta = Z_3$, and spr$\Delta \sim \Gamma(2, 1/2) + 1$. The distribution of sprΔ corresponds to the case when $\kappa = +\infty$.

DGP2 mid$X = Z_1$, spr$X = Z_2^2$, mid$\Delta = Z_3$, and spr$\Delta \sim \Gamma(2, 1/2)$. The distribution of sprΔ corresponds to the case when $\kappa \in (0, +\infty)$.

DGP3 mid$X = Z_1$, spr$X = Z_2^2$, mid$\Delta = Z_3$, and spr$\Delta \sim \Gamma\left(1, 1/2\sqrt{2}\right)$. The distribution of sprΔ corresponds to the case when $\kappa = 0$.

DGP4 mid$X = Z_1$, spr$X = Z_2^2$, mid$\Delta = Z_3$ and spr$\Delta = Z_4$.

DGP5 mid$X = Z_1$, spr$X = Z_2^2$, mid$\Delta = Z_3$, and $-$spr$\Delta \sim \Gamma(2, 1/2) - 4$.

Within each DGP, we study two submodels, where $\left(\alpha_m^*, \alpha_s^*, \beta_m^*, \beta_s^*\right) = (1, 1, 1, 1)$ in model A, and $\left(\alpha_m^*, \alpha_s^*, \beta_m^*, \beta_s^*\right) = (1, 0, 0, 1)$ in model B. The two submodels differ in the value of $\boldsymbol{\beta}^*$: $\boldsymbol{\beta}^*$ is in the interior in model A and is at the boundary in model B. DGP1–DGP3 differ in the value of κ. In DGP4 and DGP5, sprΔ has positive probability of being negative; and in DPG5, $\delta^* = 0$ is also at the boundary.

For DGPs 1–5, $\widetilde{\theta}$ and $\widehat{\theta}_u$ are consistent estimators; whereas $\widehat{\theta}$ is consistent only for DGPs 1–3.

5.2. MSE of the Estimators

We focus on the comparison between estimators for $\left(\boldsymbol{\beta}^*, \delta^*\right)$ in terms of MSE for the aforementioned data set configurations, because the major difference between the three estimators lies in the constraints imposed in the minimization problem

https://doi.org/10.1017/S0266466621000360 Published online by Cambridge University Press

and the estimator vector for $(\boldsymbol{\alpha}^*, \gamma^*)$ is not affected by the constraints. The number of repetitions is 5,000.

Table 1 compares the MSE between different models under different data configurations. For all cases considered, as the sample size increases, the MSE of each estimator decreases. For model M_G, which corresponds to DGP1–DPG3, $(\widehat{\boldsymbol{\beta}}, \widehat{\delta})$ is the most efficient estimator among the three at any given sample size. More importantly, the fatter the tail distribution of sprΔ is, the more efficient the estimator $(\widehat{\boldsymbol{\beta}}, \widehat{\delta})$ is compared to $(\widetilde{\boldsymbol{\beta}}, \widetilde{\delta})$ and $(\widehat{\boldsymbol{\beta}}_u, \widehat{\delta}_u)$. The distribution of sprΔ verifies $\kappa = +\infty$ in DGP1, which corresponds to a thin tailed distribution; the value $\kappa = 0$ in DGP3 implies a fat tailed distribution. The tail distribution of sprΔ in DGP2 is in the middle, such that $\kappa \in (0, +\infty)$. For DGP1, the estimator vector $(\widehat{\boldsymbol{\beta}}, \widehat{\delta})$ is slightly better than $(\widetilde{\boldsymbol{\beta}}, \widetilde{\delta})$ when the sample size is small and the two are the same when the sample size is large. On the other hand, the MSE for $(\widehat{\boldsymbol{\beta}}, \widehat{\delta})$ is smaller than $(\widetilde{\boldsymbol{\beta}}, \widetilde{\delta})$ in DGP2, especially when $\beta_m^* = 0$. This phenomenon is more pronounced in DGP3 when the tail distribution of sprΔ is fat. The MSE for $(\widehat{\boldsymbol{\beta}}, \widehat{\delta})$ is half of that for $(\widetilde{\boldsymbol{\beta}}, \widetilde{\delta})$ in DGP3 A, and is less than a quarter in DGP3 B. The constraints that spr$Y_i - |\boldsymbol{X}_i|' \boldsymbol{\beta} \geq 0$ for $i = 1, \ldots, n$ provide little information on the parameter $\boldsymbol{\beta}$ when sprΔ has a relatively small probability of being close to zero. During the computation of the minimization problem, the constraints almost

Table 1. MSE comparison

	n	$(\widehat{\boldsymbol{\beta}}, \widehat{\delta})$	$(\widetilde{\boldsymbol{\beta}}, \widetilde{\delta})$	$(\widehat{\boldsymbol{\beta}}_u, \widehat{\delta}_u)$		n	$(\widehat{\boldsymbol{\beta}}, \widehat{\delta})$	$(\widetilde{\boldsymbol{\beta}}, \widetilde{\delta})$	$(\widehat{\boldsymbol{\beta}}_u, \widehat{\delta}_u)$
DGP1 A	100	0.5847	0.6145	0.6504	DGP1 B	100	0.3891	0.4200	0.6620
	500	0.1207	0.1234	0.1234		500	0.0743	0.0764	0.1226
	1,000	0.0626	0.0626	0.0626		1,000	0.0397	0.0397	0.0628
DGP2 A	100	0.5628	0.6302	0.6707	DGP2 B	100	0.0216	0.4003	0.6437
	500	0.0947	0.1291	0.1294		500	0.0422	0.0801	0.1253
	1,000	0.0410	0.0607	0.0607		1,000	0.0218	0.0404	0.0639
DGP3 A	100	0.3453	0.6324	0.6779	DGP3 B	100	0.2011	0.4194	0.6887
	500	0.0629	0.1283	0.1283		500	0.0182	0.0784	0.1247
	1,000	0.0318	0.0635	0.0635		1,000	0.0097	0.0396	0.0643
DGP4 A	100	N/A	0.6218	0.6723	DGP4 B	100	N/A	0.4072	0.6629
	500	N/A	0.1244	0.1247		500	N/A	0.0825	0.1290
	1,000	N/A	0.0637	0.0637		1,000	N/A	0.0395	0.0621
DGP5 A	100	N/A	0.4163	0.6637	DGP5 B	100	N/A	0.1650	0.6903
	500	N/A	0.0913	0.1297		500	N/A	0.0319	0.1240
	1,000	N/A	0.0434	0.0638		1,000	N/A	0.0154	0.0636

never bind in this case. On the other hand, if sprΔ is close to zero with a relatively high probability, the constraints bind for some $i \in 1,\ldots,n$. This improves the accuracy of the estimator vector $\left(\widehat{\boldsymbol{\beta}},\widehat{\delta}\right)$, because the true parameter vector satisfies the constraint. Section 3 provides more detailed discussion about the effect of the tail of the distribution of sprΔ on the asymptotic property of $\left(\widehat{\boldsymbol{\beta}},\widehat{\delta}\right)$. For the same reason, the MSEs for estimator vectors $\left(\widetilde{\boldsymbol{\beta}},\widetilde{\delta}\right)$ and $\left(\widehat{\boldsymbol{\beta}}_u,\widehat{\delta}_u\right)$ in DGP1 A–DGP4 A are close to each other, whereas $\left(\widetilde{\boldsymbol{\beta}},\widetilde{\delta}\right)$ is significantly more efficient than $\left(\widehat{\boldsymbol{\beta}}_u,\widehat{\delta}_u\right)$ when some of the parameters are at the boundary, as in DGP1 B–DGP4 B, DGP5 A, and DGP5 B. When the correct constraints bind during the estimation, the accuracy of the constrained estimator improves.

5.3. Size and Power of the Tests

In this section, we study and compare the finite sample size and power of the UF test, the WU test, and the CLR test when the null hypothesis involves range regression and the nuisance parameters are either at the discontinuity point or not.

For each DGP, we consider testing the correct specification of the CR model against the IA-CR, i.e., testing $H_0:\left(\alpha_s^*,\beta_m^*\right)=\mathbf{0}$ against H_1: $\left(\alpha_s^*,\beta_m^*\right)\neq\mathbf{0}$ under the maintained hypothesis that $\left(\boldsymbol{\beta}^{*\prime},\delta^*\right)\geq\mathbf{0}$ using one of the three tests. The test statistic for the WU test takes the same form as (16) with $\left(\widetilde{\alpha}_s',\widetilde{\beta}_m'\right)$ replaced by $\left(\widehat{\alpha}_{s,u}',\widehat{\beta}_{m,u}'\right)$. The CLR statistic is constructed using the unconstrained estimator $\widehat{\theta}_u$. For the UF test, Σ_n equals to an estimator of

$$\begin{pmatrix} P_{xx}^{-1} & \mathbf{0} \\ \mathbf{0} & Q_{xx}^{-1} \end{pmatrix}\Lambda\begin{pmatrix} P_{xx}^{-1} & \mathbf{0} \\ \mathbf{0} & Q_{xx}^{-1} \end{pmatrix}$$

using $\widetilde{\theta}$, while Σ_n for the WU test and CLR test is calculated using $\widehat{\theta}_u$.

The critical value for the WU test is the 95% quantile of the chi-squared distribution with 2 degrees of freedom, and the one for the CLR test is simulated based on the procedure in Ketz (2018). When implementing the UF test, the confidence set I_τ for c is constructed as a Wald ellipsoid using $\widehat{\beta}_{s,OLS}$ and $\widehat{\delta}_{OLS}$. The Bonferroni $CV_n^b(\vartheta,\tau)$ for some $0<\tau\leq\vartheta$ is defined in (17). Following Romano et al. (2014) and McCloskey (2017), the tuning parameter τ is set as $\vartheta-\vartheta/10$. We refer interested readers to Romano et al. (2014) and McCloskey (2017) for a general discussion of the choice of τ. A special choice of τ is to set it as ϑ. This corresponds to the least favorable critical value. The test based upon the least favorable critical value controls the asymptotic size, but its finite sample performance is poor comparing to the UF, WU, and CLR tests. To save space, we focus on the UF, WU, and CLR tests.

There are three inequality constraints in the maintained hypothesis: $\left(\boldsymbol{\beta}_m^*,\boldsymbol{\beta}_s^*,\delta^*\right)\geq\mathbf{0}$. In DGP1–DGP4, only one constraint is binding under the null hypothesis; whereas in DGP5, both $\boldsymbol{\beta}_m^*=0$ and $\delta^*=0$ hold under H_0. In the simulation, the nominal size is 5%, and the number of repetitions is 5,000. Table 2 reports the finite sample size and power performance of each test. We let $\alpha_s^*=0$

https://doi.org/10.1017/S0266466621000360 Published online by Cambridge University Press

Table 2. Size and power performance-reject percentage

DGP1	$\beta_m^*=0$			$\beta_m^*=0.25$			$\beta_m^*=0.5$			$\beta_m^*=0.75$			$\beta_m^*=1$		
n	UF	WU	CLR	UF	WU	CLR	UF	WU	CLR	UF	WU	CLR	UF	WU	CLR
100	5.14	5.94	5.68	7.26	7.54	7.68	13.26	12.88	13.04	22.16	20.88	21.08	34.02	32.78	33.24
200	5.02	5.54	5.46	9.04	9.18	9.30	19.68	18.50	20.12	38.14	36.08	38.04	58.26	56.62	59.66
500	4.92	5.36	5.14	12.52	12.78	13.00	39.22	37.38	40.96	72.60	71.48	73.80	94.68	94.22	95.08
DGP2	$\beta_m^*=0$			$\beta_m^*=0.25$			$\beta_m^*=0.5$			$\beta_m^*=0.75$			$\beta_m^*=1$		
n	UF	WU	CLR	UF	WU	CLR	UF	WU	CLR	UF	WU	CLR	UF	WU	CLR
100	5.22	5.84	5.68	7.66	7.10	7.74	13.42	12.56	13.88	23.28	20.38	22.40	34.12	32.24	33.28
200	4.94	5.24	5.06	8.68	8.32	8.70	18.56	17.74	18.44	37.82	36.00	38.02	58.46	57.08	58.22
500	4.88	5.18	4.94	13.56	13.40	13.62	38.48	37.66	38.88	75.18	74.32	75.68	94.06	93.78	94.22
DGP3	$\beta_m^*=0$			$\beta_m^*=0.25$			$\beta_m^*=0.5$			$\beta_m^*=0.75$			$\beta_m^*=1$		
n	UF	WU	CLR	UF	WU	CLR	UF	WU	CLR	UF	WU	CLR	UF	WU	CLR
100	5.28	5.74	5.62	7.28	7.14	7.46	13.26	11.94	13.80	22.18	20.76	22.72	36.32	34.32	36.88
200	5.04	5.46	5.22	9.88	9.42	9.84	20.22	18.86	20.96	36.14	35.04	37.26	59.22	57.64	59.42
500	4.94	5.46	5.16	12.46	12.34	12.80	39.20	39.08	39.96	72.40	72.36	73.02	93.84	93.52	93.98
DGP4	$\beta_m^*=0$			$\beta_m^*=0.25$			$\beta_m^*=0.5$			$\beta_m^*=0.75$			$\beta_m^*=1$		
n	UF	WU	CLR	UF	WU	CLR	UF	WU	CLR	UF	WU	CLR	UF	WU	CLR
100	4.68	4.32	4.26	6.24	6.64	6.80	10.12	10.98	12.68	17.26	17.68	19.48	30.12	30.54	33.68
200	4.76	4.70	4.68	7.82	8.12	8.48	16.14	16.50	17.90	34.02	34.20	36.22	60.98	61.30	63.58
500	4.62	4.88	4.84	12.44	12.70	12.48	36.28	36.60	36.64	71.64	71.80	72.48	92.62	92.70	94.08
DGP5	$\beta_m^*=0$			$\beta_m^*=0.25$			$\beta_m^*=0.5$			$\beta_m^*=0.75$			$\beta_m^*=1$		
n	UF	WU	CLR	UF	WU	CLR	UF	WU	CLR	UF	WU	CLR	UF	WU	CLR
100	4.82	4.66	4.74	7.42	6.86	6.92	13.80	11.74	12.06	24.20	21.44	21.48	35.88	33.58	33.96
200	4.86	4.94	4.88	8.92	8.26	8.30	20.66	18.80	18.92	39.04	36.20	36.42	60.08	58.18	58.06
500	4.92	5.10	5.06	12.20	12.08	12.24	40.06	38.98	39.02	73.78	72.72	72.80	94.64	93.88	93.94

throughout the simulation and change the value of $\boldsymbol{\beta}_m^*$ to obtain the size and power of the tests. $\beta_m^*=0$ corresponds the case where H_0 holds; whereas when β_m^* equals 0.25, 0.5, 0.75, and 1, the alternative hypothesis is true, and we deviate more and more from the null hypothesis.

First, by comparing the rejection probabilities in the case $\beta_m^*=0$, we see that the finite sample size of the UF test introduced in the paper is the closest to the nominal size 5%, especially when n is small. The finite sample null rejection probability of the WU test tends to depend heavily on the distribution of sprΔ, and the size distortion of the WU test is the most severe among the three tests when the sample size is small. The CLR test outperforms the WU test, but is slightly less accurate than the UF test when n is small. In DGP4 where the distributions of midΔ and sprΔ are indeed Gaussian, the Wald test based upon $\widehat{\theta_u}$ has the best finite sample size. However, even in this case, the UF and CLR tests have comparable size performance.

Second, the UF test has better finite sample power than the WU test in cases where the distribution of sprΔ is skewed. Except in DGP4 where the WU test is destined to perform well, the UF test dominates the WU test. In particular, when n is small and when there are more binding inequalities under H_0 like DGP5, the power of the UF test can be much higher than that of the WU test. The CLR test

https://doi.org/10.1017/S0266466621000360 Published online by Cambridge University Press

consistently beats the WU test. Between the UF and the CLR tests, the UF test tends to have better power performance than the CLR test in small samples and when many inequalities bind such as in DGP5.

In summary, both the UF test and the CLR test have advantages over the WU test. The CLR test has better finite sample performance over the WU test in almost all cases, and the UF test outperforms the WU test when the distribution of the error is skewed. The UF test and CLR test have similar finite sample performance in cases for which both are applicable, and the UF test can be applied to many other hypotheses to which the CLR test does not apply.

6. CONCLUDING REMARKS

We have made several contributions in this paper. First, we have proposed a flexible model, i.e., the GIA-CR model for random intervals via the generalized interval arithmetic approach and constructed a constrained estimator of parameter vector in the GIA-CR model. As a special member of the generalized model, the IA-CR model extends and overcomes the drawbacks of both model M_G and the CR model. Second, as a measure of goodness-of-fit, we have extended the coefficient of determination for linear regressions for random variables to our GIA-CR model for random intervals. Third, we have developed asymptotically uniformly valid tests for linear hypotheses in the GIA-CR model including a test for the correct specification of the CR model against the IA-CR model. Fourth, we have conducted a simulation study to examine the finite sample performance of our estimator and test. As a separate contribution to the current literature on interval arithmetic approach to modeling interval data, we have established the asymptotic distribution of the constrained estimator of model M_G.

APPENDICES

A. Technical Proofs

A.1. Proof of Proposition 2.1. Using the alternative expression of RSS_G, $\mathcal{R}_G^2$ can be computed as

$$\mathcal{R}_G^2 = 1 - \frac{\sum_{i=1}^n d^2\left(Y_i, \widetilde{Y}_i\right)}{\sum_{i=1}^n d^2\left(Y_i, \overline{Y}\right)}.$$

The nonnegativity of $\frac{\sum_i d^2(Y_i, \widetilde{Y}_i)}{\sum_i d^2(Y_i, \overline{Y})}$ implies that $\mathcal{R}_G^2 \leq 1$; and for $\{Y_i, X_{1i}, \ldots, X_{ki}\}_{i=1}^n$ and $\widetilde{\theta}$ such that $Y_i = \widetilde{Y}_i$ for all i, it holds that $d^2\left(Y_i, \widetilde{Y}_i\right) = 0$ and thus $\mathcal{R}_G^2 = 1$. By (7) and the fact that

$$\sum_{i=1}^n d^2\left(Y_i, \overline{Y}\right) = \sum_{i=1}^n d^2\left(Y_i, \left[\check{\gamma} \pm \check{\delta}\right]\right), \text{ where}$$

https://doi.org/10.1017/S0266466621000360 Published online by Cambridge University Press

$$\left(\check{\gamma}, \check{\delta}\right)' = \arg\min_{\gamma,\delta} \sum_{i=1}^{n} d^2\left(Y_i, [\gamma \pm \delta]\right),$$

it holds that

$$\sum_{i=1}^{n} d^2\left(Y_i, \widetilde{Y}_i\right) = \sum_{i=1}^{n} d^2\left(Y_i, \left[\boldsymbol{X}_i'\widetilde{\boldsymbol{\alpha}} \pm |\boldsymbol{X}_i|'\widetilde{\boldsymbol{\beta}}\right] + \left[\widetilde{\gamma} \pm \widetilde{\delta}\right]\right)$$
$$\leq \sum_{i=1}^{n} d^2\left(Y_i, \left[\check{\gamma} \pm \check{\delta}\right]\right) = \sum_{i=1}^{n} d^2\left(Y_i, \overline{Y}\right).$$

The inequality comes from property of the minimization operation: $\widetilde{Y}_i$'s are obtained by a constrained minimization problem over $\boldsymbol{\alpha}, \gamma, \boldsymbol{\beta}$, and δ, with $\boldsymbol{\alpha} = \mathbf{0}$ and $\boldsymbol{\beta} = \mathbf{0}$ satisfying the constraint. Therefore, $\sum_{i=1}^{n} d^2\left(Y_i, \overline{Y}\right) \geq \sum_{i=1}^{n} d^2\left(Y_i, \widetilde{Y}_i\right)$ and $\mathcal{R}_G^2 \geq 0$, where the equality holds if $\widetilde{\boldsymbol{\alpha}}$ and $\widetilde{\boldsymbol{\beta}}$ are zero vectors. The first part of the proposition holds. The second part follows directly from the characteristics of the minimization problem that defines $\widetilde{\theta}$ in (7). ■

A.2. Proof of Theorem 3.1. We first show that $\widehat{\theta} \to_p \theta^*$. Define

$$Z_1(\theta) = \left(\boldsymbol{\alpha}^{*\prime} - \boldsymbol{\alpha}', \gamma^* - \gamma\right) P_{xx} \begin{pmatrix} \boldsymbol{\alpha}^* - \boldsymbol{\alpha} \\ \gamma^* - \gamma \end{pmatrix} + \sigma^2_{\mathrm{mid}\Delta}$$
$$+ \left(\boldsymbol{\beta}^{*\prime} - \boldsymbol{\beta}', \delta^* - \delta\right) Q_{xx} \begin{pmatrix} \boldsymbol{\beta}^* - \boldsymbol{\beta} \\ \delta^* - \delta \end{pmatrix} + \sigma^2_{\mathrm{spr}\Delta} + \varphi_1(\theta),$$

where $\sigma^2_{\mathrm{mid}\Delta} = Var(\mathrm{mid}\Delta)$, $\sigma^2_{\mathrm{spr}\Delta} = Var(\mathrm{spr}\Delta)$, and

$$\varphi_1(\theta) = \begin{cases} 0, & \text{if } \Pr\left(\mathrm{spr}Y - |\boldsymbol{X}|'\boldsymbol{\beta} \geq 0\right) = 1; \text{ and } \boldsymbol{\beta} \geq 0 \\ \infty, & \text{otherwise} \end{cases}.$$

It holds that $\theta^* = \arg\min_\theta Z_1(\theta)$. First, notice that $\Pr\left(\mathrm{spr}Y - |\boldsymbol{X}|'\boldsymbol{\beta}^* \geq 0\right) = 1$ and $\boldsymbol{\beta}^* \geq 0$ by the model specification. Second, by Assumption 3.2, both $\sigma^2_{\mathrm{mid}\Delta}$ and $\sigma^2_{\mathrm{spr}\Delta}$ are finite. Therefore, $Z_1(\theta)$ reaches its minimal value $\sigma^2_{\mathrm{mid}\Delta} + \sigma^2_{\mathrm{spr}\Delta}$ at θ^*. At last, since P_{xx} and Q_{xx} are both nonsingular by Assumption 3.2, they are positive definite and θ^* is the unique solution to the minimization problem. Therefore, $\theta^* = \arg\min_\theta Z_1(\theta)$.

Next we aim to show that $\arg\min_\theta Z_{1n}(\theta) \xrightarrow{p} \arg\min_\theta Z_1(\theta)$. $Z_{1n}(\theta)$ is a convex function because its epigraph is a convex set. Moreover, since $\arg\min_\theta Z_1(\theta)$ is unique, it suffices to show that $Z_{1n}(\theta)$ epi-converges to $Z_1(\theta)$ by the Convexity Lemma of Geyer (1996) and Knight (1999).

The finite dimensional convergence and finiteness of $Z_1(\theta)$ on an open set provide the epi-convergence, given that $Z_{1n}(\theta)$ is convex. For

$$s_0 \equiv \sup\left\{\max_{1\leq j\leq 2k-d} \boldsymbol{b}_j : \Pr\left(\mathrm{spr}Y - |\boldsymbol{X}|'\boldsymbol{b} \geq 0\right) = 1, \boldsymbol{b} \in \mathbb{R}^{2k-d}_{\geq 0}\right\},$$

if s_0 is strictly positive, one can always find an open set $O_{\boldsymbol{b}} \subset \mathbb{R}^{2k}$ such that $\Pr\left(\mathrm{spr}Y - |\boldsymbol{X}|'\boldsymbol{b} \geq 0\right) = 1$ for any $\boldsymbol{b} \in O_{\boldsymbol{b}}$. Then on the set $O_{\boldsymbol{\alpha}} \times O_\gamma \times O_{\boldsymbol{b}} \times O_\delta$, where

https://doi.org/10.1017/S0266466621000360 Published online by Cambridge University Press

$O_{\boldsymbol{\alpha}} \subset \mathbb{R}^{2k-d}$, $O_\gamma \subset \mathbb{R}$, and $O_\delta \subset \mathbb{R}$ are any open sets, $Z(\theta)$ is finite. We now show the finite dimensional convergence to complete the proof.

By the weak law of large numbers and the model specification, we have that

$$\frac{1}{n}\sum\left(\text{mid}Y_i - \boldsymbol{X}_i'\boldsymbol{\alpha} - \gamma\right)^2 \xrightarrow{p} \left(\boldsymbol{\alpha}^{*\prime} - \boldsymbol{\alpha}', \gamma^* - \gamma\right) P_{xx}\left(\boldsymbol{\alpha}^{*\prime} - \boldsymbol{\alpha}', \gamma^* - \gamma\right)' + \sigma^2_{\text{mid}\Delta} \text{ and}$$
$$\frac{1}{n}\sum\left(\text{spr}Y_i - |\boldsymbol{X}_i|'\boldsymbol{\beta} - \delta\right)^2 \xrightarrow{p} \left(\boldsymbol{\beta}^{*\prime} - \boldsymbol{\beta}', \delta^* - \delta\right) Q_{xx}\left(\boldsymbol{\beta}^{*\prime} - \boldsymbol{\beta}', \delta^* - \delta\right)' + \sigma^2_{\text{spr}\Delta},$$

for any pair $\left(\boldsymbol{\alpha}', \gamma, \boldsymbol{\beta}', \delta\right)$. Thus, according to Knight (2001), it suffices to show that for given $\theta^1, \ldots, \theta^m$,

$$\Pr\left(\varphi_{1n}\left(\theta^1\right) = 0, \ldots, \varphi_{1n}\left(\theta^m\right) = 0\right) \longrightarrow \Pr\left(\varphi_1\left(\theta^1\right) = 0, \ldots, \varphi_1\left(\theta^m\right) = 0\right)$$

when $n \to \infty$. The former probability equals to

$$\begin{aligned}
&\Pr\left(\text{spr}Y_i \geq |\boldsymbol{X}_i|'\boldsymbol{\beta}^j, \text{ for } i = 1, \ldots, n \text{ and } j = 1, \ldots, m\right)\\
&= \Pr{}^n\left(\text{spr}Y_i \geq \max_{1\leq j\leq m} |\boldsymbol{X}_i|'\boldsymbol{\beta}^j\right) \longrightarrow \begin{cases} 1, & \text{if } \Pr\left(\text{spr}Y \geq \max_{1\leq j\leq m} |\boldsymbol{X}|'\boldsymbol{\beta}^j\right) = 1\\ 0, & \text{if } \Pr\left(\text{spr}Y \geq \max_{1\leq j\leq m} |\boldsymbol{X}|'\boldsymbol{\beta}^j\right) < 1\end{cases}\\
&= \Pr\left(\varphi_1\left(\theta^1\right) = 0, \ldots, \varphi_1\left(\theta^m\right) = 0\right).
\end{aligned}$$

Therefore, we can conclude that $Z_{1n}(\theta)$ epi-converges to $Z_1(\theta)$ and $\widehat{\theta} \xrightarrow{p} \theta^*$ when $s_0 > 0$.

On the other hand, if $s_0 = 0$, then $\boldsymbol{\beta}_j^* = 0$ for all $j = 1, \ldots, 2k-d$. The minimization problem can be separated into two parts with one part containing only $\widehat{\boldsymbol{\alpha}}$ and $\widehat{\gamma}$ and the other contains only $\widehat{\boldsymbol{\beta}}$ and $\widehat{\delta}$. Since no constraints are imposed on $\left(\widehat{\boldsymbol{\alpha}}', \widehat{\gamma}\right)$, the consistency of $\left(\widehat{\boldsymbol{\alpha}}', \widehat{\gamma}\right)$ follows from standard arguments in the least squares estimation. Because $Z_{1n}\left(\boldsymbol{\alpha}', \gamma, \boldsymbol{0}', \delta\right)$ is finite, to prove that $\left(\widehat{\boldsymbol{\beta}}', \widehat{\delta}\right) \xrightarrow{p} \left(\boldsymbol{\beta}^{*\prime}, \delta^*\right) = \left(\boldsymbol{0}', \delta^*\right)$, it suffices to show that for any $\boldsymbol{b} \neq \boldsymbol{0}$, $\Pr\left(Z_{1n}(\theta) < \infty\right) \to 0$ when $n \to \infty$. This follows from the fact that for $\boldsymbol{b} \neq \boldsymbol{0}$,

$$\Pr\left(\text{spr}Y_i \geq |\boldsymbol{X}_i|'\boldsymbol{b}, \text{ for } i = 1, \ldots, n\right) = \Pr{}^n\left(\text{spr}Y_i - |\boldsymbol{X}_i|'\boldsymbol{b} \geq 0\right) = 0,$$

where the last equality is implied by $s_0 = 0$. Therefore, $\widehat{\boldsymbol{\beta}} \xrightarrow{p} \boldsymbol{0}$. The convergence of $\widehat{\delta}$ follows from the law of large numbers since $\delta^* = \mathbb{E}(\text{spr}Y)$ when $\boldsymbol{\beta}^* = \boldsymbol{0}$.

The first part of the theorem then follows by combining the two cases of $s_0 > 0$ and $s_0 = 0$.

https://doi.org/10.1017/S0266466621000360 Published online by Cambridge University Press

The proof for $\widetilde{\theta} \xrightarrow{p} \theta^*$ is essentially the same. Define

$$Z_2(\theta) = \left(\boldsymbol{\alpha}^{*\prime} - \boldsymbol{\alpha}', \gamma^* - \gamma\right) P_{xx} \begin{pmatrix} \boldsymbol{\alpha}^* - \boldsymbol{\alpha} \\ \gamma^* - \gamma \end{pmatrix} + \sigma^2_{\mathrm{mid}\Delta} + \left(\boldsymbol{\beta}^{*\prime} - \boldsymbol{\beta}', \delta^* - \delta\right) Q_{xx} \begin{pmatrix} \boldsymbol{\beta}^* - \boldsymbol{\beta} \\ \delta^* - \delta \end{pmatrix} + \sigma^2_{\mathrm{spr}\Delta} + \varphi_2(\theta),$$

where $\sigma^2_{\mathrm{mid}\Delta} = Var(\mathrm{mid}\Delta)$ and $\sigma^2_{\mathrm{spr}\Delta} = Var(\mathrm{spr}\Delta)$. With Assumption 3.2, we have that $\theta^* = \arg\min_\theta Z_2(\theta)$, because $R\theta^* \geq r$ and P_{xx} and Q_{xx} are nonsingular. The convexity of $Z_{2n}(\theta)$ is implied by its quadratic component and the geometry of the feasible set. Since $\varphi_2(\theta)$ is not random, the epi-convergence of $Z_{2n}(\theta)$ to $Z_2(\theta)$ follows from the finite dimensional convergence of the quadratic component. At last, since $Z_2(\theta)$ is finite on any open set that is contained in $\mathbb{R}^{2k-d} \times \mathbb{R} \times \mathbb{R}^{2k-d}_{\geq 0} \times \mathbb{R}_{\geq 0}$, we obtain the consistency of $\widetilde{\theta}$. ■

A.3. Proof of Theorem 3.2. We will prove the theorem for the different cases: (i) $\kappa \in (0, +\infty)$; (ii) $\kappa = +\infty$; and (iii) $\kappa = 0$. Let $M_1(\psi) = M(\psi) + \phi_1(\psi)$.

Note that $\sqrt{n}\left(\widehat{\theta} - \theta^*\right)$ is the solution to the minimization problem:

$$\min_{\psi} M_{1n}(\psi) \equiv \min_{\boldsymbol{p}, q, \boldsymbol{u}, v} \begin{bmatrix} \sum_{i=1}^n \left(\mathrm{mid}\Delta_i - \gamma^* - \frac{1}{\sqrt{n}} \boldsymbol{X}_i' \boldsymbol{p} - \frac{q}{\sqrt{n}}\right)^2 \\ + \sum_{i=1}^n \left(\mathrm{spr}\Delta_i - \delta^* - \frac{1}{\sqrt{n}} |\boldsymbol{X}_i|' \boldsymbol{u} - \frac{v}{\sqrt{n}}\right)^2 \\ - \sum_{i=1}^n \left(\mathrm{mid}\Delta_i - \gamma^*\right)^2 - \sum_{i=1}^n \left(\mathrm{spr}\Delta_i - \delta^*\right)^2 + \phi_{1n}(\psi) \end{bmatrix},$$

where

$$\phi_{1n}(\psi) = \begin{cases} 0, & \text{if } \sqrt{n}\mathrm{spr}\Delta_i \geq |\boldsymbol{X}_i|' \boldsymbol{u}, \text{ for } i = 1, 2, \ldots n; \\ & \text{and } \boldsymbol{u}_j + \sqrt{n}\boldsymbol{\beta}_j^* \geq 0, \text{ for } j = 1, \ldots, 2k - d\,. \\ \infty, & \text{otherwise} \end{cases}$$

The goal is to show that $\arg\min_\psi M_{1n}(\psi) \xrightarrow{d} \arg\min_\psi M_1(\psi)$. Since the set of ψ for $\phi_{1n}(\psi)$ being finite is convex, the convexity of $M_{1n}(\psi)$ is straightforward due to its quadratic component. Recall that by the Convexity Lemma of Geyer (1996) and Knight (1999), the following three conditions are sufficient for $\arg\min_\psi M_{1n}(\psi) \xrightarrow{d} \arg\min_\psi M_1(\psi)$ provided that $M_{1n}(\psi)$ is convex: (a) $M_{1n}(\psi)$ converges to $M_1(\psi)$ in the finite-dimensional sense ($\xrightarrow{f.d.}$), (b) $M_1(\psi)$ is finite on an open set, and (c) $M_1(\psi)$ is uniquely minimized with probability 1.

We now prove that these three conditions are satisfied when $\lim_{t\to\infty} tF_s\left(1/\sqrt{t}\right) \in (0, +\infty)$.

For any Borel subsets D of $\mathbb{D} := [0, +\infty) \times \mathcal{X}$, where $\mathcal{X}$ is the support of $|\boldsymbol{X}_i|$, define the point process (random measure): $\upsilon_n(D) := \sum_{i=1}^n \mathbb{1}\left\{\left(\sqrt{n}\mathrm{spr}\Delta_i, |\boldsymbol{X}_i|\right) \in D\right\}$. The point process $\upsilon_n(\cdot)$ tends in distribution with respect to the vague topology to a Poisson point process (random measure) $\upsilon(\cdot)$ in the metric space of point measure $\mathcal{M}_p(\mathbb{D})$. The limit Poisson process has the mean measure: $\mathbb{E}[\upsilon(D)] = \int_D \frac{2}{\kappa} wg(\boldsymbol{x})\, d\mu(\boldsymbol{x})\, dw$ and can be represented by $\upsilon(D) := \sum_{i=1}^\infty \mathbb{1}\left\{\left(\sqrt{\kappa} g^{-\frac{1}{2}}(\Upsilon_i)\Gamma_i, \Upsilon_i\right) \in D\right\}$ for all Borel subsets D of $\mathbb{D} := [0, +\infty) \times \mathcal{X}$, where $g(\cdot)$ is defined in Assumption 3.3, $\Gamma_i = (\mathcal{E}_1 + \cdots + \mathcal{E}_i)^{\frac{1}{2}}$

https://doi.org/10.1017/S0266466621000360 Published online by Cambridge University Press

for unit mean i.i.d. exponential random variables $\mathcal{E}_1, \mathcal{E}_2, \ldots$, and $\Upsilon_1, \Upsilon_2, \ldots$ are i.i.d. with distribution $\Pr(\Upsilon_i \in A) = \mu(A)$, where $\mu(\cdot)$ is the probability measure of $|\boldsymbol{X}_i|$. The Γ_i 's are independent of Υ_i's. By Assumption 3.3 and the fact that $\lim_{t\to\infty} tF_s\left(1/\sqrt{t}\right) = 1/\kappa$, $\lim_{n\to\infty} \mathbb{E}[\upsilon_n(D)] = \mathbb{E}[\upsilon(D)]$ and $\lim_{n\to\infty} \Pr\{\upsilon_n(D) = 0\} = e^{-\mathbb{E}[\upsilon(D)]}$. The claimed weak convergence result follows from Kallenberg's theorem (Resnick, 1987).

Using the above convergence result of the point process, we are now ready to show the finite dimensional weak convergence of $M_{1n}(\psi)$. The following convergence result is straightforward:

$$\begin{bmatrix} \sum_{i=1}^n \left(\text{mid}\Delta_i - \gamma^* - \frac{1}{\sqrt{n}}\boldsymbol{X}_i'\boldsymbol{p} - \frac{q}{\sqrt{n}}\right)^2 - \sum_{i=1}^n \left(\text{mid}\Delta_i - \gamma^*\right)^2 \\ + \sum_{i=1}^n \left(\text{spr}\Delta_i - \delta^* - \frac{1}{\sqrt{n}}|\boldsymbol{X}_i|'\boldsymbol{u} - \frac{v}{\sqrt{n}}\right)^2 - \sum_{i=1}^n \left(\text{spr}\Delta_i - \delta^*\right)^2 \end{bmatrix}$$
$$\xrightarrow{d} \psi' \begin{pmatrix} P_{xx} & \boldsymbol{0} \\ \boldsymbol{0} & Q_{xx} \end{pmatrix} \psi - 2\psi' \begin{pmatrix} \boldsymbol{I}_l & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}_l \end{pmatrix} W,$$

with $W \sim \mathcal{N}(0, \Lambda)$ and Λ being the covariance matrix of $\left(\dot{\boldsymbol{X}}_i'\text{mid}\Delta_i, \left|\dot{\boldsymbol{X}}_i\right|'\text{spr}\Delta_i\right)$. The asymptotic independence between W and the point process follows from the standard proof of asymptotic independence of sample average and sample minimal-order statistics, see e.g. Resnick (1987) and Lem. 21.19 in Van der Vaart (2000). A more detailed proof can be found in Chernozhukov and Hong (2002). Thus, it remains to show that for given $\psi^1, \ldots, \psi^m$,

$$\Pr\left[\phi_{1n}\left(\psi^1\right) = 0, \ldots, \phi_{1n}\left(\psi^m\right) = 0\right] \longrightarrow \Pr\left[\phi_1\left(\psi^1\right) = 0, \ldots, \phi_1\left(\psi^m\right) = 0\right],$$

as $n \to \infty$. Since no randomness is involved in the constraint $\boldsymbol{u}_j + \sqrt{n}\boldsymbol{\beta}_j^* \geq 0$ for $j = 1, \ldots, 2k-d$, its limit is straightforward. Thus, we only need to focus on the constraint $\sqrt{n}\text{spr}\Delta_i \geq |\boldsymbol{X}_i|'\boldsymbol{u}$ for $i = 1, 2, \ldots n$. Exploiting the convergence in distribution of $\upsilon_n(\cdot)$ to the Poisson random measure $\upsilon(\cdot)$, we have that

$$\Pr\left[\phi_{1n}\left(\psi^1\right) = 0, \ldots, \phi_{1n}\left(\psi^m\right) = 0\right] = \Pr\left[\sum_{i=1}^n \mathbb{1}\left(\sqrt{n}\text{spr}\Delta_i < \max_{1\leq j\leq m} |\boldsymbol{X}_i|'\boldsymbol{u}^j\right) = 0\right]$$
$$\to \exp\left(-\int_{\mathcal{X}} \max_{1\leq j\leq m} \frac{g(\boldsymbol{x})}{\kappa}\left(\boldsymbol{x}'\boldsymbol{u}^j\right)^2 d\mu(\boldsymbol{x})\right) = \Pr\left[\phi_1\left(\psi^1\right) = 0, \ldots, \phi_1\left(\psi^m\right) = 0\right].$$

Therefore, $M_{1n}(\psi) \overset{f.d.}{\to} M_1(\psi)$ as $n \to \infty$.

The other two conditions can be easily verified. On the set $O_p \times O_q \times (-\infty, 0)^{2k-d} \times O_v$, where $O_p \subset \mathbb{R}^{2k-d}$, $O_q \subset \mathbb{R}$ and $O_v \subset \mathbb{R}$ are any open sets, $M_1(\psi)$ is finite by its definition. And, for any realization of W, $\{\Gamma_i,\ i \geq 1\}$ and $\{\Upsilon_i,\ i \geq 1\}$, $M_1(\psi)$ will be uniquely minimized due to the quadratic form of $M(\psi)$ and the geometry of the constraint. If the minimal of $M(\psi)$ is in the constraint set, the uniqueness is trivially satisfied. If the minimal of $M(\psi)$ lies outside the constraint set, the solution to the minimization problem will be the intersect of the level set of $M(\psi)$ with the boundary of the constraint set. The level set of the quadratic component of $M(\psi)$ takes the shape of an ellipse in high dimension, while the constraint set is convex with boundary consisting of high dimensional planes. They can only intersect at one point. Thus, the latter two conditions in the Convexity Lemma are satisfied. Hence, when $\lim_{t\to\infty} tF_s\left(1/\sqrt{t}\right) \in (0, +\infty)$, $\arg\min_\psi M_{1n}(\psi) \xrightarrow{d} \arg\min_\psi M_1(\psi)$ as $n \to \infty$.

https://doi.org/10.1017/S0266466621000360 Published online by Cambridge University Press

The above result also holds when $\lim_{t\to\infty} tF_s\left(1/\sqrt{t}\right)=0$. Rewrite $M_1(\psi)$ by substituting $\kappa=+\infty$: $M_1(\psi)=M(\psi)+\phi_1(\psi)$, with

$$\phi_1(\psi)=\begin{cases}0, & \text{if } \Upsilon_i'\boldsymbol{u}\le+\infty, \text{ for } i=1,2,\ldots;\\ & \text{and } \mathbb{1}\left(\boldsymbol{\beta}_j^*=0\right)\boldsymbol{u}_j\ge 0 \text{ for } j=1,\ldots,2k-d\,.\\ \infty, & \text{otherwise}\end{cases}$$

Since Υ_i follows a tight probability measure, we can further simplify

$$\phi_1(\psi)=\begin{cases}0, & \text{if } \mathbb{1}\left(\boldsymbol{\beta}_j^*=0\right)\boldsymbol{u}_j\ge 0 \text{ for } j=1,\ldots,2k\\ \infty, & \text{otherwise}\end{cases}.$$

The finite dimensional convergence of the quadratic component of $M_{1n}(\psi)$ follows the same argument. We now show that for any given $\psi^1,\ldots,\psi^m$, as $n\to\infty$,

$$\Pr\left[\phi_{1n}\left(\psi^1\right)=0,\ldots,\phi_{1n}\left(\psi^m\right)=0\right]\longrightarrow\Pr\left[\phi_1\left(\psi^1\right)=0,\ldots,\phi_1\left(\psi^m\right)=0\right].$$

By writing the probability as an expectation of a conditional probability, we have that

$$\begin{aligned}\Pr\left[\phi_{1n}\left(\psi^1\right)=0,\ldots,\phi_{1n}\left(\psi^m\right)=0\right]&=\Pr\left[\sum_{i=1}^{n}\mathbb{1}\left(\sqrt{n}\operatorname{spr}\Delta_i<\max_{1\le j\le m}|\boldsymbol{X}_i|'\boldsymbol{u}^j\right)=0\right]\\ &=\Pr{}^n\left[\sqrt{n}\operatorname{spr}\Delta\ge\max_{1\le j\le m}|\boldsymbol{X}|'\boldsymbol{u}^j\right]=\mathbb{E}^n\left[\Pr\left[\operatorname{spr}\Delta\ge\frac{1}{\sqrt{n}}\max_{1\le j\le m}|\boldsymbol{X}|'\boldsymbol{u}^j\mid\boldsymbol{X}\right]\right]\\ &=\mathbb{E}^n\left[1-g(|\boldsymbol{X}|)F_s^-\left(\frac{1}{\sqrt{n}}\max_{1\le j\le m}|\boldsymbol{X}|'\boldsymbol{u}^j\right)\right]\\ &=\left[1-\mathbb{E}\left[g(|\boldsymbol{X}|)F_s^-\left(\frac{1}{\sqrt{n}}\max_{1\le j\le m}|\boldsymbol{X}|'\boldsymbol{u}^j\right)\right]\right]^n,\end{aligned}$$

where $F_s^-(a)\equiv\lim_{z\uparrow a}F_s(z)$. The indeterminate form has the limit of

$$\exp\left(-\lim_{n\to\infty}n\mathbb{E}\left[g(|\boldsymbol{X}|)F_s^-\left(\frac{1}{\sqrt{n}}\max_{1\le j\le m}|\boldsymbol{X}|'\boldsymbol{u}^j\right)\right]\right).$$

For $|\boldsymbol{X}|\in\mathcal{X}$ such that $\max_{1\le j\le m}|\boldsymbol{X}|'\boldsymbol{u}^j\le 0$, we have that $F_s^-\left(\frac{1}{\sqrt{n}}\max_{1\le j\le m}|\boldsymbol{X}|'\boldsymbol{u}^j\right)=0$; if $\max_{1\le j\le m}|\boldsymbol{X}|'\boldsymbol{u}^j>0$, the inequality $\lim_{n\to\infty}nF_s^-\left(\frac{1}{\sqrt{n}}\max_{1\le j\le m}|\boldsymbol{X}|'\boldsymbol{u}^j\right)\le\lim_{n\to\infty}nF_s\left(\frac{1}{\sqrt{n}}\max_{1\le j\le m}|\boldsymbol{X}|'\boldsymbol{u}^j\right)=0$ holds. By the dominated convergence theorem, we obtain that

$$\Pr\left[\phi_{1n}\left(\psi^1\right)=0,\ldots,\phi_{1n}\left(\psi^m\right)=0\right]\to\exp(0)=1=\Pr\left[\phi_1\left(\psi^1\right)=0,\ldots,\phi_1\left(\psi^m\right)=0\right].$$

Therefore, the finite dimensional convergence of $M_{1n}(\psi)$ is verified. By the same argument as in the first case where $\kappa\in(0,+\infty)$, we can show that $M_1(\psi)$ is finite on an open set and that its minimizer is unique with probability one. The Convexity Lemma then provides that $\arg\min_\psi M_{1n}(\psi)\xrightarrow{d}\arg\min_\psi M_1(\psi)$ when $\kappa=+\infty$.

The proof for the case when $\lim_{t\to\infty}tF_s\left(1/\sqrt{t}\right)=+\infty$ is essentially the same. Since the part $\mathbb{1}\left(\boldsymbol{\beta}_j^*=0\right)\boldsymbol{u}_j\ge 0$ for $j=1,\ldots,2k-d$ does not contain any randomness, we focus

https://doi.org/10.1017/S0266466621000360 Published online by Cambridge University Press

the proof on the constraint $\sqrt{n}\mathrm{spr}\Delta_i \geq |X_i|'\boldsymbol{u}$ for $i = 1,2,\ldots n$. For any given $\psi^1,\ldots,\psi^m$, it holds that

$$\Pr\left[\phi_{1n}\left(\psi^1\right) = 0,\ldots,\phi_{1n}\left(\psi^m\right) = 0\right] = \left[1 - \mathbb{E}\left[g\left(|\boldsymbol{X}|\right)F_s^-\left(\frac{1}{\sqrt{n}}\max_{1\leq j\leq m}|\boldsymbol{X}|'\boldsymbol{u}^j\right)\right]\right]^n$$

with the limit

$$\exp\left(-\lim_{n\to\infty} n\mathbb{E}\left[g\left(|\boldsymbol{X}|\right)F_s^-\left(\frac{1}{\sqrt{n}}\max_{1\leq j\leq m}|\boldsymbol{X}|'\boldsymbol{u}^j\right)\right]\right).$$

If $\max_{1\leq j\leq m}|\boldsymbol{X}|'\boldsymbol{u}^j \leq 0$, $F_s^-\left(\frac{1}{\sqrt{n}}\max_{1\leq j\leq m}|\boldsymbol{X}|'\boldsymbol{u}^j\right) \leq F_s^-(0) = 0$, because $F_s^-(0) = 0$. If $\max_{1\leq j\leq m}|\boldsymbol{X}|'\boldsymbol{u}^j > 0$, we have that $\lim_{n\to\infty} nF_s^-\left(\frac{1}{\sqrt{n}}\max_{1\leq j\leq m}|\boldsymbol{X}|'\boldsymbol{u}^j\right) \geq \lim_{n\to\infty} nF_s\left(\frac{1}{2\sqrt{n}}\max_{1\leq j\leq m}|\boldsymbol{X}|'\boldsymbol{u}^j\right) = \infty$ by condition $\lim_{t\to\infty} tF_s\left(1/\sqrt{t}\right) = +\infty$. Therefore for the given $\boldsymbol{u}$'s, when $\Pr\left(\max_{1\leq j\leq m}|\boldsymbol{X}|'\boldsymbol{u}^j \leq 0\right) = 1$,

$$\exp\left(-\lim_{n\to\infty} n\mathbb{E}\left[g\left(|\boldsymbol{X}|\right)F_s^-\left(\frac{1}{\sqrt{n}}\max_{1\leq j\leq m}|\boldsymbol{X}|'\boldsymbol{u}^j\right)\right]\right) \to 1;$$

and when $\Pr\left(\max_{1\leq j\leq m}|\boldsymbol{X}|'\boldsymbol{u}^j \leq 0\right) < 1$,

$$\exp\left(-\lim_{n\to\infty} n\mathbb{E}\left[g\left(|\boldsymbol{X}|\right)F_s^-\left(\frac{1}{\sqrt{n}}\max_{1\leq j\leq m}|\boldsymbol{X}|'\boldsymbol{u}^j\right)\right]\right) \to 0.$$

Now consider $\Pr\left[\phi_1\left(\psi^1\right) = 0,\ldots,\phi_1\left(\psi^m\right) = 0\right]$. The probability can be calculated as

$$\begin{aligned}\Pr\left[\phi_1\left(\psi^1\right) = 0,\ldots,\phi_1\left(\psi^m\right) = 0\right] &= \Pr\left[\max_{1\leq j\leq m}\Upsilon_i'\boldsymbol{u}^j \leq 0, : \text{for} : i = 1,2,\ldots\right] \\ &= \begin{cases} 1, & \text{if } \Pr\left(\max_{1\leq j\leq m}|\boldsymbol{X}|'\boldsymbol{u}^j \leq 0\right) = 1 \\ 0, & \text{if } \Pr\left(\max_{1\leq j\leq m}|\boldsymbol{X}|'\boldsymbol{u}^j \leq 0\right) < 1 \end{cases},\end{aligned}$$

where the last equality follows from the fact that the distribution of Υ_i is the same as $|\boldsymbol{X}|$ and i goes to infinity. Therefore, we have shown that for any given $\psi^1,\ldots,\psi^m$,

$$\Pr\left[\phi_{1n}\left(\psi^1\right) = 0,\ldots,\phi_{1n}\left(\psi^m\right) = 0\right] \longrightarrow \Pr\left[\phi_1\left(\psi^1\right) = 0,\ldots,\phi_1\left(\psi^m\right) = 0\right],$$

as $n\to\infty$. The rest is the same as in case $\lim_{t\to\infty} tF_s\left(1/\sqrt{t}\right) = 0$.

Hence, we have shown that $\arg\min_\psi M_{1n}(\psi)$ converges in law to $\arg\min_\psi M_1(\psi)$ for all different values of $\lim_{t\to\infty} tF_s\left(1/\sqrt{t}\right)$. The claimed theorem can be concluded. ■

A.4. Proof of Theorem 3.3. The proof is similar to the proof for Theorem 3.2. $\sqrt{n}\left(\widetilde{\theta} - \theta^*\right)$ is the solution to the minimization problem:

$$\min_{\boldsymbol{p},q,\boldsymbol{u},v}\left[\begin{array}{c}\sum_{i=1}^n\left(\mathrm{mid}\Delta_i - \gamma^* - \frac{1}{\sqrt{n}}\boldsymbol{X}_i'\boldsymbol{p} - \frac{q}{\sqrt{n}}\right)^2 - \sum_{i=1}^n\left(\mathrm{mid}\Delta_i - \gamma^*\right)^2 \\ +\sum_{i=1}^n\left(\mathrm{spr}\Delta_i - \delta^* - \frac{1}{\sqrt{n}}|\boldsymbol{X}_i|'\boldsymbol{u} - \frac{v}{\sqrt{n}}\right)^2 - \sum_{i=1}^n\left(\mathrm{spr}\Delta_i - \delta^*\right)^2 + \phi_{2n}(\psi)\end{array}\right],$$

where

$$\phi_{2n}(\psi) = \begin{cases} 0, & \text{if } R\psi \geq \sqrt{n}\left(r - R\theta^*\right) \\ \infty, & \text{otherwise} \end{cases}.$$

For the vector $\left(r - R\theta^*\right)$, some elements are zero and the rest are strictly negative. The zero elements correspond to the submatrix R_b of R by definition. This implies that $\phi_{2n}(\psi) \to \phi_2(\psi)$ pointwise as $n \to \infty$. The rest of the proof is the same for Theorem 3.2 by showing finite-dimensional convergence, $M(\psi) + \phi_2(\psi)$ being finite on an open set, and $M(\psi) + \phi_2(\psi)$ being uniquely minimized with probability 1. ■

A.5. Proof of Lemma 4.1. By Assumption 4.2, the claimed result follows if we show that $R_0\Psi$ represents the asymptotic distribution of $\sqrt{n}\left(\widetilde{\alpha}_s', \widetilde{\beta}_m'\right)'$ under the null hypothesis and the parameter sequence (η_n, π_n, ξ_n). Under H_0, it holds that $\beta_m^* = \mathbf{0}$. Simple manipulation of the proof for Theorem 3.3 would provide that $\sqrt{n}\left(\widetilde{\alpha}_s', \widetilde{\beta}_m'\right)' \xrightarrow{d} R_0\Psi$, with the Lindeberg-Lévy Central Limit Theorem replaced by Lyapunov Central Limit Theorem under Assumption 4.1, see Lem. A.3 in Cheng (2015). ■

A.6. Proof of Proposition 4.1. We prove the proposition by verifying assumptions in McCloskey (2017). Notice that the distribution $(R_0\Psi)'\left(R_0\Sigma_\omega R_0'\right)^{-1}(R_0\Psi)$ is finite with probability 1 for all c and π_ω in the localization parameter space. Assumption PS, Sel, and Inf in McCloskey (2017) are trivially satisfied. By the expression in Lemma 4.1, $\mathcal{C}_{c,\pi_\omega}(1-\vartheta)$ is continuous in c and π_ω. Moreover, $R_0\Psi$ follows a continuous distribution. Thus, Assumption Cont in McCloskey (2017) is satisfied. The requirement for the confidence set I_τ that $\lim_{n\to\infty}\Pr\left(\sqrt{n}\eta_n \in I_\tau\right) \geq 1-\tau$ fulfils Assumption CS in McCloskey (2017). It suffices to prove that Assumption DS in McCloskey (2017) is satisfied for the first claim of the proposition.

Lemma 4.1 provides that the asymptotic distribution of the test statistic T_n is $(R_0\Psi)'\left(R_0\Sigma_\omega R_0'\right)^{-1}(R_0\Psi)$ under the full parameter sequence (η_n, π_n, ξ_n); and the asymptotic convergence of $\sqrt{n}\left(\widehat{\beta}_{s,OLS}', \widehat{\delta}_{OLS}\right)' - \sqrt{n}\eta_n \xrightarrow{d} Z(\Lambda_{OLS})$ is straightforward. We follow Lem. 2.1 in Andrews et al. (2020) to establish the equivalence of results under full sequences and subsequences provided that Assump. B2 in Andrews et al. (2020) holds. Therefore, the goal is to show that for any subsequence there exists a full sequence that has the same limit (possibly infinity) and has its subsequence equal to the original one. Denote the subsequence as $\left\{\eta_{p_n}, \pi_{p_n} : n \geq 1\right\}$ such that $\left(\sqrt{p_n}\eta_{p_n}, \pi_{p_n}\right) \to (c, \pi_\omega)$. We aim to construct a full sequence $\left\{\eta_n^\star, \pi_n^\star : n \geq 1\right\}$ satisfying that $\left(\sqrt{n}\eta_n^\star, \pi_n^\star\right) \to (c, \pi_\omega)$ and $\left(\eta_n^\star, \pi_n^\star\right) = \left(\eta_{p_n}, \pi_{p_n}\right)$, $\forall n \geq 1$. To clarify the notation, let the full sequence be indexed by m: $\left\{\eta_m^\star, \pi_m^\star : m \geq 1\right\}$. For $\forall m = p_n$, define $\left(\eta_m^\star, \pi_m^\star\right) = \left(\eta_{p_n}, \pi_{p_n}\right)$; and for $\forall m \in \left(p_n, p_{n+1}\right)$, define

$$\delta_m^\star = \begin{cases} \frac{\sqrt{p_n}\delta_{p_n}}{\sqrt{m}}, & \text{if } \sqrt{p_n}\delta_{p_n} \to c_{k-d+1} \in \mathbb{R}_{\geq 0} \\ \delta_{p_n}, & \text{if } \sqrt{p_n}\delta_{p_n} \to +\infty \end{cases} \text{ and}$$

$$\beta_{s,j,m}^\star = \begin{cases} \frac{\sqrt{p_n}\beta_{s,j,p_n}}{\sqrt{m}}, & \text{if } \sqrt{p_n}\beta_{s,j,p_n} \to c_j \in \mathbb{R}_{\geq 0} \\ \beta_{s,j,p_n}, & \text{if } \sqrt{p_n}\beta_{s,j,p_n} \to +\infty \end{cases}$$

https://doi.org/10.1017/S0266466621000360 Published online by Cambridge University Press

for $j = 1, \ldots, k-d$ and $\pi^{\star}_{m} = \pi_{p_n}$. It is trivial that the constructed full sequence satisfies the second requirement that $\left(\eta^{\star}_{p_n}, \pi^{\star}_{p_n}\right) = \left(\eta_{p_n}, \pi_{p_n}\right)$ for $\forall n \geq 1$. To see that the first requirement is also satisfied, please refer to pp. 225–226 in Cheng (2015) for a detailed derivation. ■

A.7. Proof of Lemma 4.2. With Assumption 4.2, the lemma follows if we can show that $\sqrt{n}\left(\widetilde{\theta} - \theta_n\right) \xrightarrow{d} \Psi$ under the model parameters $\left(\eta^{u}_{n}, \pi_n, \xi_n\right)$. The estimator is defined as $\widetilde{\theta} = \arg\min_{\theta \in \mathbb{R}^{2l}} Z_{2n}(\theta)$. Note that $\sqrt{n}\left(\widetilde{\theta} - \theta_n\right)$ is the solution to the minimization problem:

$$\min_{\boldsymbol{p}, q, \boldsymbol{u}, v} \left[\begin{array}{c} \sum_{i=1}^{n} \left(\text{mid}\Delta_i - \gamma_n - \frac{1}{\sqrt{n}} \boldsymbol{X}_i' \boldsymbol{p} - \frac{q}{\sqrt{n}}\right)^2 - \sum_{i=1}^{n} \left(\text{mid}\Delta_i - \gamma_n\right)^2 \\ + \sum_{i=1}^{n} \left(\text{spr}\Delta_i - \delta_n - \frac{1}{\sqrt{n}} |\boldsymbol{X}_i|' \boldsymbol{u} - \frac{v}{\sqrt{n}}\right)^2 - \sum_{i=1}^{n} \left(\text{spr}\Delta_i - \delta_n\right)^2 + \phi_n(\psi) \end{array} \right],$$

where

$$\phi_{2n}(\psi) = \begin{cases} 0, & \text{if } R\psi + \sqrt{n}\left(R\theta_n - r\right) \geq 0 \\ \infty, & \text{otherwise} \end{cases}.$$

By Lyapunov Central Limit Theorem, we obtain that

$$\left[\begin{array}{c} \sum_{i=1}^{n} \left(\text{mid}\Delta_i - \gamma_n - \frac{1}{\sqrt{n}} \boldsymbol{X}_i' \boldsymbol{p} - \frac{q}{\sqrt{n}}\right)^2 - \sum_{i=1}^{n} \left(\text{mid}\Delta_i - \gamma_n\right)^2 \\ + \sum_{i=1}^{n} \left(\text{spr}\Delta_i - \delta_n - \frac{1}{\sqrt{n}} |\boldsymbol{X}_i|' \boldsymbol{u} - \frac{v}{\sqrt{n}}\right)^2 - \sum_{i=1}^{n} \left(\text{spr}\Delta_i - \delta_n\right)^2 \end{array} \right] \xrightarrow{d} M_{\omega}(\psi).$$

Decompose R into three submatrices: R_{Γ}, R_{0b} and R_{0s}, where R_{0s} denotes the nonbinding inequalities in $R_{00}\theta \geq r_{-\Gamma}$. By the definition of c, it holds that $\Gamma\eta^{u}_{n} - r_{\Gamma} = R_{\Gamma}\theta_n - r_{\Gamma} \to c$ when $n \to \infty$; under the null hypothesis, R_{0b} and R_{0s} represent the binding and nonbinding inequalities $R_{00}\theta \geq r_{-\Gamma}$. Therefore, we have $\phi_{2n}(\psi) \to \phi_{\omega}(\psi)$ pointwise as $n \to \infty$. We obtain that $\sqrt{n}\left(\widetilde{\theta} - \theta_n\right) \xrightarrow{d} \Psi$ and the claimed lemma. ■

A.8. Proof of Theorem 4.1. The theorem follows from the same proof for Proposition 4.1 with the continuity of $\left(R_0\Psi\right)'\left(R_0\Sigma_{\omega}R_0'\right)^{-1}\left(R_0\Psi\right)$ at $\mathcal{C}_{c,\pi_{\omega}}(1-\vartheta)$ for all $(c, \pi_{\omega}) \in \overline{\mathbb{R}}^{l_{\Gamma}}_{\geq 0} \times \Pi$ being assumed in the theorem. ■

B. A Review of Generalized Interval Arithmetic and Random Generalized Intervals

B.1. Interval Arithmetic. Given $a_1, a_2 \in \mathbb{R}$ and $a_1 \leq a_2$, an interval A is defined by its left and right end points: $A = [a_1, a_2] = \{x \in \mathbb{R} : a_1 \leq x \leq a_2\}$, or by its center and range: $A = \left[\text{mid}A \pm \text{spr}A\right]$, where $\text{mid}A = (a_1 + a_2)/2$ and $\text{spr}A = (a_2 - a_1)/2 \geq 0$. The set of all intervals is denoted by $I(\mathbb{R})$. For all $A, B \in I(\mathbb{R})$ and $\lambda \in \mathbb{R}$, it holds that

(i) $A + B \equiv \left[(\text{mid}A + \text{mid}B) \pm (\text{spr}A + \text{spr}B)\right]$ and
(ii) $\lambda A \equiv \left[\lambda \text{mid}A \pm |\lambda| \text{spr}A\right]$.

Combining (i) and (ii), we obtain that for all $A, B \in I(\mathbb{R})$ and $\lambda \in \mathbb{R}$,

$$A + \lambda B = \left[(\text{mid}A + \lambda \text{mid}B) \pm (\text{spr}A + |\lambda| \text{spr}B)\right]. \quad \textbf{(B.1)}$$

https://doi.org/10.1017/S0266466621000360 Published online by Cambridge University Press

It follows from (ii) that $-A \equiv (-1)A = \left[-\text{mid}A \pm \text{spr}A\right] = [-a_2, -a_1]$. Subtraction between two intervals A and B is defined as

$$A - B = A + (-B) = \left[(\text{mid}A - \text{mid}B) \pm (\text{spr}A + \text{spr}B)\right].$$

As a result, we have:

$$A - A = \left[0 \pm (2\text{spr}A)\right] \neq [0,0] \text{ and}$$
$$A - B + B = \left[\text{mid}A \pm (\text{spr}A + 2\text{spr}B)\right] \neq A.$$

To partly remedy this situation, Hukuhara (1967) introduces an alternative difference operation on intervals referred to as Hukuhara difference and denoted as $(-_H)$. Specifically, for any $A, B \in I(\mathbb{R})$, $A -_H B = C$ if there exists $C \in I(\mathbb{R})$ such that $A = B + C$. It can be shown that Hukuhara difference $A -_H B$ exists if and only if $\text{spr}A \geq \text{spr}B$ and when it exists,

$$A -_H B \equiv \left[(\text{mid}A - \text{mid}B) \pm (\text{spr}A - \text{spr}B)\right].$$

In contrast to subtraction $(-)$, Hukuhara difference satisfies: $A -_H A = [0,0]$ and $A -_H B + B = A$. However, Hukuhara difference between two intervals may not exist which limits the scope of applications of the interval arithmetic approach to modeling interval data.

B.2. Generalized Interval Arithmetic and an L_2-type Metric for Generalized Intervals. In the paper, we make use of generalized intervals studied in the mathematics literature, see e.g., Kaucher (1980) and Markov (1996), to fully explore advantages of the interval arithmetic approach to modeling interval data. Specifically, for $a_1, a_2 \in \mathbb{R}$, a generalized interval is an ordered couple denoted as $A = [a_1, a_2]$: it is a proper or simply an interval when $a_1 \leq a_2$; otherwise it is an improper interval. A generalized interval can also be represented as $A = \left[\text{mid}A \pm \text{spr}A\right]$: it is proper if $\text{spr}A \geq 0$; improper if $\text{spr}A < 0$. Denote $K(\mathbb{R})$ as the space of generalized intervals. For $A, B \in K(\mathbb{R})$, it turns out that the addition and scalar product operations can be computed in the same way as in (B.1), see Kaucher (1980) and Markov (1996) for details.

With generalized intervals, we can extend Hukuhara difference to any two intervals. Let A and B be two intervals. Generalized Hukuhara difference[10] is defined as follows:

$$A -_{GH} B \equiv A + \left(-\overline{B}\right) = \left[(\text{mid}A - \text{mid}B) \pm (\text{spr}A - \text{spr}B)\right],$$

where $\overline{B} = \left[\text{mid}B \mp \text{spr}B\right]$ is the conjugation or dual of B. In contrast to Hukuhara difference, the Generalized Hukuhara difference between two intervals A and B always exists: when $\text{spr}A \geq \text{spr}B$, $A -_{GH} B$ is an interval and $A -_{GH} B = A -_H B$; otherwise it is an improper interval. It is easy to see that $A -_{GH} A = [0,0]$ and $A -_{GH} B + B = A$.

Let A and B be two generalized intervals. We define an L_2-type metric d_λ between A and B as

$$d_\lambda(A,B) \equiv \left((\text{mid}A - \text{mid}B)^2 + \lambda(\text{spr}A - \text{spr}B)^2\right)^{\frac{1}{2}}$$

for some $\lambda \in (0, \infty)$ and the norm of $A \in K(\mathbb{R})$ as $\|A\|^2 = (\text{mid}A)^2 + \lambda(\text{spr}A)^2$. The d-metric discussed in the paper corresponds to the d_λ-metric when $\lambda = 1$. It is easy to verify

[10]It is sufficient for our purpose to define Generalized Hukuhara difference for intervals only. It turns out that the Generalized Hukuhara difference for intervals is the $-_h$ operation defined in Markov (1996) for generalized intervals when applied to intervals.

that the d_λ-metric satisfies nonnegativity, identity of indiscernibles, symmetry and triangle inequality. By choosing different values of λ, one can assign different relative importance for the squared distance between the ranges with respect to the square distance between the midpoints. When A and B are both intervals, the d_λ-metric generalizes the well-known Bertoluzza metric denoted as d_W in Bertoluzza, Corral Blanco, and Salas (1995) if the metric is required to be invariant to rigid motion (Trutschnig et al. (2009)). One common choice for λ is $1/3$, as it corresponds to d_W when W is chosen as the Lebesgue measure. For more discussions on different metrics on $I(\mathbb{R})$, see Bertoluzza et al. (1995) and Trutschnig et al. (2009).

B.3. Random Generalized Intervals. Let (Ω, Σ, P) be an abstract probability space.

Definition B.1. (i) A random generalized interval $X : \Omega \to K(\mathbb{R})$ is a map from the sample space Ω to the space of generalized intervals such that $\text{mid}X : \Omega \to \mathbb{R}$ and $\text{spr}X : \Omega \to \mathbb{R}$ are random variables; (ii) The expected value of a random generalized interval X, denoted as $\mathbb{E}_A(X) \in K(\mathbb{R})$, is defined as

$$\mathbb{E}_A(X) = \mathbb{E}_A\left(\left[\text{mid}X \pm \text{spr}X\right]\right) \equiv \left[\mathbb{E}(\text{mid}X) \pm \mathbb{E}(\text{spr}X)\right], \qquad \textbf{(B.2)}$$

whenever $\mathbb{E}(\text{mid}X)$ and $\mathbb{E}(\text{spr}X)$ exist.

When $\text{spr}X \geq 0$ with probability one, the random generalized interval X becomes a random interval, which is a measurable map from Ω to $I(\mathbb{R})$. When X is a random interval, the expectation defined in (B.2) agrees with the well-known Aumann expectation (Aumann, 1965). Let $\mathfrak{A}$ be a sub-σ-algebra of Σ, the conditional expectation of a generalized random interval X given $\mathfrak{A}$ is defined accordingly as

$$\mathbb{E}_A(X \mid \mathfrak{A}) = \mathbb{E}_A\left(\left[\text{mid}X \pm \text{spr}X\right] \mid \mathfrak{A}\right) \equiv \left[\mathbb{E}(\text{mid}X \mid \mathfrak{A}) \pm \mathbb{E}(\text{spr}X \mid \mathfrak{A})\right].$$

We follow the approach of Fréchet (1948) to define the variance of a random generalized interval as:

$$Var_F(X) \equiv \inf_{A \in K(\mathbb{R})} \mathbb{E}\left(d_\lambda^2(X, A)\right) \text{ whenever } \mathbb{E}\left(\|X\|^2\right) < \infty.$$

Since the expectation defined in (B.2) agrees with Fré chet expectation with respect to the metric d_λ (Körner (1997); Körner and Näther (2002)), Fréchet variance of a random generated interval X in the metric space $K(\mathbb{R})$ endowed with the d_λ-metric is simplified as $Var_F(X) = \mathbb{E}\left(d_\lambda^2(X, \mathbb{E}_A(X))\right)$.

REFERENCES

Andrews, D. W. (1999) Estimation when a parameter is on a boundary. *Econometrica* 67(6), 1341–1383.

Andrews, D. W. (2001) Testing when a parameter is on the boundary of the maintained hypothesis. *Econometrica* 69(3), 683–734.

Andrews, D. W. & X. Cheng (2012) Estimation and inference with weak, semi-strong, and strong identification. *Econometrica* 80(5), 2153–2211.

Andrews, D. W. & X. Cheng (2014) GMM estimation and uniform subvector inference with possible identification failure. *Econometric Theory* 30(2), 287–333.

Andrews, D. W., X. Cheng & P. Guggenberger (2020) Generic results for establishing the asymptotic size of confidence sets and tests. *Journal of Econometrics* 218(2), 496–531.

Artstein, Z. & R. A. Vitale (1975) A strong law of large numbers for random compact sets. *The Annals of Probability*, 879–882.

Aumann, R. J. (1965) Integrals of set-valued functions. *Journal of Mathematical Analysis and Applications* 12(1), 1–12.

Beresteanu, A., I. Molchanov & F. Molinari (2011) Sharp identification regions in models with convex moment predictions. *Econometrica* 79(6), 1785–1821.

Beresteanu, A. & F. Molinari (2008) Asymptotic properties for a class of partially identified models. *Econometrica* 76(4), 763–814.

Berger, R. L. & D. D. Boos (1994) P values maximized over a confidence set for the nuisance parameter. *Journal of the American Statistical Association* 89(427), 1012–1016.

Bertoluzza, C., N. Corral Blanco & A. Salas (1995) On a new class of distances between fuzzy numbers. *Mathware & Soft Computing* 2(2), 71–84.

Blanco-Fernández, A., A. Colubi & G. González-Rodríguez (2012) Confidence sets in a linear regression model for interval data. *Journal of Statistical Planning and Inference* 142(6), 1320–1329.

Blanco-Fernández, A., N. Corral & G. González-Rodríguez (2011) Estimation of a flexible simple linear model for interval data based on set arithmetic. *Computational Statistics & Data Analysis* 55(9), 2568–2578.

Blanco-Fernández, A., M. García-Bárzana, A. Colubi, & E.J. Kontoghiorghes (2015) Multiple Set Arithmetic-Based Linear Regression Models for Interval-Valued Variables. Unpublished Working Paper, Department of Statistics and OR and Mathematics Didactics, University of Oviedo.

Bontemps, C., T. Magnac & E. Maurin (2012) Set identified linear models. *Econometrica* 80(3), 1129–1155.

Cheng, X. (2015) Robust inference in nonlinear models with mixed identification strength. *Journal of Econometrics* 189(1), 207–228.

Chernozhukov, V. (2005) Extremal quantile regression. *Annals of Statistics* 33, 806–839.

Chernozhukov, V. & I. Fernández-Val (2011) Inference for extremal conditional quantile models, with an application to market and birthweight risks. *The Review of Economic Studies* 78(2), 559–589.

Chernozhukov, V., I. Fernández-Val, J. Hahn & W. Newey (2013) Average and quantile effects in nonseparable panel models. *Econometrica* 81(2), 535–580.

Chernozhukov, V. & H. Hong (2002) Likelihood Inference in a Class of Nonregular Econometric Models. Working Paper No. 02-05, Department of Economics, MIT Press.

Chernozhukov, V. & H. Hong (2004) Likelihood estimation and inference in a class of nonregular econometric models. *Econometrica* 72(5), 1445–1480.

Cowling, B. J., V. J. Fang, S. Riley, J. M. Peiris & G. M. Leung (2009) Estimation of the serial interval of influenza. *Epidemiology* 20(3), 344.

Debreu, G. (1967) Integration of correspondences. In *Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability*, vol. 2, pp. 351–372. University of California Press, Berkeley.

Demsetz, H. (1968) The cost of transacting. *The Quarterly Journal of Economics* 82(1), 33–53.

Du, Z., X. Xu, Y. Wu, L. Wang, B. J. Cowling & L. A. Meyers (2020) Serial interval of COVID-19 among publicly reported confirmed cases. *Emerging Infectious Diseases* 26(6), 1341.

Fréchet, M. (1948) Les éléments aléatoires de nature quelconque dans un espace distancié. *Annales de l'institut Henri Poincaré* 10(3), 215–310.

Ganyani, T., C. Kremer, D. Chen, A. Torneri, C. Faes, J. Wallinga, & N. Hens (2020) Estimating the generation interval for coronavirus disease (COVID-19) based on symptom onset data, march 2020. *Eurosurveillance* 25(17), 2000257.

García-Ascanio, C. & C. Maté (2010) Electric power demand forecasting using interval time series: A comparison between var and imlp. *Energy Policy* 38(2), 715–725.

Geyer, C. J. (1994) On the asymptotics of constrained M-estimation. *Annals of Statistics* 22, No. 4, 1993–2010.

Geyer, C.J. (1996) On the asymptotics of convex stochastic optimization. Unpublished Manuscript, Department of Statistics, University of Minnesota.

Golan, A. & A. Ullah (2017) Interval estimation: An information theoretic approach. *Econometric Reviews* 36(6-9), 781–795.

González-Rivera, G. & W. Lin (2013) Constrained regression for interval-valued data. *Journal of Business & Economic Statistics* 31(4), 473–490.

Gourieroux, C., A. Holly & A. Monfort (1982) Likelihood ratio test, wald test, and kuhn-tucker test in linear models with inequality constraints on the regression parameters. *Econometrica* 50(1), 63–80.

Han, A., Y. Hong, & S. Wang (2012) Autoregressive Conditional Models for Interval-Valued Time Series Data. Unpublished Manuscript, Department of Economics, Cornell University.

Han, A., H. Yongmiao, K. K. Lai & W. Shouyang (2008) Interval time series analysis with an application to the sterling-dollar exchange rate. *Journal of Systems Science and Complexity* 21(4), 558–573.

Hukuhara, M. (1967) Integration des applications mesurables dont la valeur est un compact convexe. *Funkcialaj Ekvacioj* 10(3), 205–223.

Judge, G., T. Yancey, M. Bock & R. Bohrer (1984) The non-optimality of the inequality restricted estimator under squared error loss. *Journal of Econometrics* 25(1-2), 165–177.

Kallenberg, O. (1983) *Random Measures*. Akademie-Verlag.

Kaucher, E. (1980) Interval analysis in the extended interval space ir. *Computing Supplementum* 2, 33–49.

Kenah, E., M. Lipsitch & J. M. Robins (2008) Generation interval contraction and epidemic data analysis. *Mathematical Biosciences* 213(1), 71–79.

Ketz, P. (2018) Subvector inference when the true parameter vector may be near or at the boundary. *Journal of Econometrics* 207(2), 285–306.

Knight, K. (1999) Epi-convergence in distribution and stochastic equi-semicontinuity. Unpublished Manuscript, Department of Statistics, University of Toronto.

Knight, K. (2001) Limiting distributions of linear programming estimators. *Extremes* 4(2), 87–103.

Knight, K. (2006) Asymptotic theory for M-estimators of boundaries. In S. Sperlich, W. Härdle, & G. Aydinli (eds.), *The Art of Semiparametrics*, pp. 1–21. Physica-Verlag.

Körner, R. (1997) On the variance of fuzzy random variables. *Fuzzy Sets and Systems* 92(1), 83–93.

Körner, R. & W. Näther (2002) On the variance of random fuzzy variables. In C. Bertoluzza, M.Á. Gil, & D.A. Ralescu (eds.), *Statistical Modeling, Analysis and Management of Fuzzy Data*, pp. 25–42. Physica.

Leadbetter, M. R., G. Lindgren & H. Rootzén (1987) *Extremes and Related Properties of Random Sequences and Processes*. Springer-Verlag.

Lee, C. M. (1993) Market integration and price execution for nyse-listed securities. *The Journal of Finance* 48(3), 1009–1038.

Liew, C. K. (1976) Inequality constrained least-squares estimation. *Journal of the American Statistical Association* 71(355), 746–751.

Lin, W. & G. González-Rivera (2016) Interval-valued time series models: Estimation based on order statistics exploring the agriculture marketing service data. *Computational Statistics & Data Analysis* 100, 694–711.

Manski, C. F. & E. Tamer (2002) Inference on regressions with interval data on a regressor or outcome. *Econometrica* 70(2), 519–546.

Markov, S. (1996) On directed interval arithmetic and its applications. *The Journal of Universal Computer Science* 200(1), 514–526.

McCloskey, A. (2017) Bonferroni-based size-correction for nonstandard testing problems. *Journal of Econometrics* 200(1), 17–35.

McShane, E. J. (1969) *A Riemann-Type Integral that Includes Lebesgue-Stieltjes, Bochner and Stochastic Integrals*. American Mathematical Society.

Moon, H. R. & F. Schorfheide (2009) Estimation with overidentifying inequality moment conditions. *Journal of Econometrics* 153(2), 136–154.

Neto, E. D. A. L. & F. D. A. de Carvalho (2010) Constrained linear regression models for symbolic interval-valued variables. *Computational Statistics & Data Analysis* 54(2), 333–347.

Nishiura, H., N. M. Linton & A. R. Akhmetzhanov (2020) Serial interval of novel coronavirus (COVID-19) infections. *International Journal of Infectious Diseases* 93, 284–286.

Pflug, G.C. (1994) On an argmax-distribution connected to the Poisson process. In P. Mandl, M. Hušková (eds.), *Asymptotic Statistics*, pp. 123–129. Physica-Verlag.

Pflug, G. C. (1995) Asymptotic stochastic programs. *Mathematics of Operations Research* 20(4), 769–789.

Resnick, S. I. (1987) *Extreme Values, Regular Variation, and Point Processes* (Springer).

Rogers, A. J. (1986) Modified lagrange multiplier tests for problems with one-sided alternatives. *Journal of Econometrics* 31(3), 341–361.

Romano, J. P., A. M. Shaikh & M. Wolf (2014) A practical two-step method for testing moment inequalities. *Econometrica* 82(5), 1979–2002.

Romano, J. P. & M. Wolf (2000) Finite sample nonparametric inference and large sample efficiency. *Annals of Statistics* 28(3), 756–778.

Silvapulle, M. & P. Sen (2005) *Constrained Statistical Inference* (Wiley).

Silvapulle, M. J. (1996) A test in the presence of nuisance parameters. *Journal of the American Statistical Association* 91(436), 1690–1693.

Staiger, D. O. & J. H. Stock (1997) Instrumental variables regression with weak instruments. *Econometrica* 65, 557–586.

Trutschnig, W., G. González-Rodríguez, A. Colubi & M. Á. Gil (2009) A new family of metrics for compact, convex (fuzzy) sets based on a generalized concept of mid and spread. *Information Sciences* 179(23), 3964–3972.

Van der Vaart, A. W. (2000) *Asymptotic Statistics* (Cambridge University Press).

Vynnycky, E. & P. E. Fine (2000) Lifetime risks, incubation period, and serial interval of tuberculosis. *American Journal of Epidemiology* 152(3), 247–263.

Wolak, F. A. (1987) An exact test for multiple inequality and equality constraints in the linear regression model. *Journal of the American Statistical Association* 82(399), 782–793.

Wolak, F. A. (1989) Local and global testing of linear and nonlinear inequality constraints in nonlinear econometric models. *Econometric Theory* 5(1), 1–35.

Wolak, F. A. (1991) The local nature of hypothesis tests involving inequality constraints in nonlinear models. *Econometrica* 59, No. 4, 981–995.