



HAL
open science

Generalized method of moments for an extended gamma process

Zeina Al Masry, Sophie Mercier, Ghislain Verdier

► **To cite this version:**

Zeina Al Masry, Sophie Mercier, Ghislain Verdier. Generalized method of moments for an extended gamma process. *Communications in Statistics - Theory and Methods*, 2018, 47 (15), pp.3687-3714. 10.1080/03610926.2017.1361988 . hal-01576969

HAL Id: hal-01576969

<https://hal.science/hal-01576969>

Submitted on 24 Aug 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Generalized method of moments for an extended gamma process

Z. Al Masry¹, S. Mercier² & G. Verdier²

¹Laboratoire Angevin de REcherche en MATHématiques (UMR CNRS 6093),
Université d'Angers, Angers, France

²Laboratoire de Mathématiques et de leurs Applications - Pau (UMR CNRS 5142),
Université de Pau et des Pays de l'Adour, Pau, France

Abstract: In reliability theory, a widely used process to model the phenomena of the cumulative deterioration of a system over time is the standard gamma process. Based on several restrictions, such as a constant variance-to-mean ratio, this process is not always a suitable choice to describe the deterioration. A way to overcome these restrictions is to use an extended version of the gamma process introduced by Cinlar (1980), which is characterized by shape and scale functions. In this paper, the aim is to propose statistical methods to estimate the unknown parameters of parametric forms of the shape and scale functions. We here develop two generalized methods of moments (Hansen, 1982), based either on the moments or on the Laplace transform of an extended gamma process. Asymptotic properties are provided and a Wald-type test is derived, which allows to test standard gamma processes against extended ones with a specific parametric shape function. Also, the performance of the proposed estimation methods is illustrated on simulated and real data.

Keywords: Process with non-stationary independent increments; Extended gamma process; Laplace transform; Parametric estimation; Generalized method of moments.

1 Introduction

Safety and dependability is a crucial issue in many industries, which has led to the development of a huge literature devoted to the so-called reliability theory. In the oldest literature, the lifetimes of industrial systems or components were usually directly modeled through random variables, see, e.g., Barlow and Proschan (1965) for a pioneer work on the subject. Based on the development of on-line monitoring which allows the effective measurement of a system deterioration, numerous papers nowadays model the degradation in itself, which is often considered to be accumulating over time. This can be done through the use of non decreasing stochastic processes, among which standard gamma processes are widely used (see van Noortwijk (2009) for an overview).

Setting $A : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ to be a measurable, increasing and right-continuous function with $A(0) = 0$ and $b_0 > 0$, let us recall that a standard (non homogeneous) gamma process $\mathbf{Y} = (Y_t)_{t \geq 0} \sim \Gamma_0(A(\cdot), b_0)$, with $A(\cdot)$ as shape function and b_0 as (constant) scale parameter, is a stochastic process with independent, non-negative and gamma distributed increments such that $Y_0 = 0$ almost surely. The probability density function of an increment $Y_t - Y_s$ (with $0 < s < t$) is given by

$$f(x) = \frac{b_0^{A(t)-A(s)}}{\Gamma(A(t) - A(s))} x^{A(t)-A(s)-1} \exp(-b_0 x), \forall x \geq 0 \quad (1)$$

(e.g. see Abdel-Hameed (1975)).

The mean, variance and Laplace transform of Y_t are given by

$$\mathbb{E}[Y_t] = \frac{A(t)}{b_0}; \quad \mathbb{V}[Y_t] = \frac{A(t)}{b_0^2}; \quad \mathcal{L}_{Y_t}(\lambda) = \mathbb{E}\left(e^{-\lambda Y_t}\right) = \left(\frac{b_0}{b_0 + \lambda}\right)^{A(t)}$$

for all $t \geq 0$ and $\lambda \geq 0$, so that the variance-to-mean ratio is constant over time.

As a consequence, a Standard Gamma Process (SGP) is not a suitable choice when there is some "empirical evidence" against a constant variance-to-mean ratio (Guida et al. (2012)). A way to overcome this restriction is to use an Extended Gamma Process (EGP), mostly introduced simultaneously by Çinlar (1980) and Dykstra and Laud (1981). An EGP is a non decreasing stochastic process with independent increments. It is characterized by a similar shape function as a SGP plus a scale function, which is not assumed to be constant any more. To be more specific, let $A : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ be a measurable, increasing and right-continuous function such that $A(0) = 0$ and let $b : \mathbb{R}_+ \rightarrow \mathbb{R}_+^*$ be a measurable positive function. The functions $A(\cdot)$ and $b(\cdot)$ are assumed to be such that

$$\int_{(0,t]} \frac{dA(s)}{b(s)} < \infty, \quad \int_{(0,t]} \frac{dA(s)}{b(s)^2} < \infty \quad (2)$$

for all $t \geq 0$. The process $\mathbf{X} = (X_t)_{t \in \mathbb{R}^+}$ is then said to be an EGP with shape function $A(\cdot)$ and scale function $b(\cdot)$ (written $\mathbf{X} \sim \Gamma(A(\cdot), b(\cdot))$) if it can be represented as the following stochastic integral with respect to a SGP $(Y_t)_{t \in \mathbb{R}^+} \sim \Gamma(A(\cdot), 1)$:

$$X_t = \int_{(0,t]} \frac{dY_s}{b(s)}, \quad \forall t \in \mathbb{R}^+ \quad (3)$$

and $X_0 = 0$. The probability density function (pdf) of an EGP increment is not available in full form. However, both its Laplace transform and moments are explicitly known.

The statistical inference of a stochastic model is an important task for its practical use in an industrial reliability context. As far as SGPs are concerned, this has been much studied in the literature, both from a parametric and semi-parametric point of view, see, e.g., Çinlar et al. (1977) and Ye et al. (2013). However, it seems that the statistical inference of EGPs has received much less attention, see however Guida et al. (2012) for a specific parametric procedure and a discrete-time version of an EGP.

In this paper, the aim is to propose estimation methods for parametric forms of the shape and scale functions of an EGP. Based on the fact that the pdf of an increment is not known in full form, standard maximum likelihood estimation is not possible. To overcome the non availability of the pdf, other methods based on some Empirical Likelihood (EL) function have been developed (Qin and Lawless (1994)) and used for processes with independent increments, see, e.g., Kunitomo and Owada (2004), Sueishi and Nishiyama (2005), Chan et al. (2009). However, in the present case of an EGP, EL method requires estimating too many parameters (due to the non-stationarity of the EGP increments) and it consequently does not seem to be adapted either. Based on the fact that both moments and Laplace transform are here known in full form, we suggest to use the Generalized Method of Moments (GMM) introduced by Hansen (1982), which can be based on moments, Laplace/Fourier transforms, or many other quantities. GMM based on the Fourier transform has been the subject of many papers in the literature (see, e.g., Carrasco and Florens

(2002), Feuerverger and McDunnough (1981), Sueishi and Nishiyama (2005)). In the present study, we develop GMM for an EGP based on either the moments or on the Laplace transform. We provide conditions under which the proposed estimators are proved to be consistent and asymptotically normal, using previous results from Hall (2005), Hansen (1982) and Newey and McFadden (1994).

This paper is organized as follows. Section 2 briefly introduces the general approach of GMM. Then, we elaborate about the two approaches of GMM for an EGP in Section 3 and provide asymptotic properties. In Section 4, the focus is on a particular parametric form for the shape and scale functions. Sufficient conditions are provided for the asymptotic properties to hold and a Wald-type test is derived, which allows to check whether the EGP adjusted to some given data is significantly different from a SGP. Numerical illustrations are presented in Section 5 along with an application to real data, and we finally conclude in Section 6.

2 General specification of GMM

2.1 GMM approach

Let \mathbf{W} be a random vector of dimension d and $\{\mathbf{W}_n, n = 1, \dots, N\}$ be a set of independent and identically distributed (i.i.d.) random vectors sharing the same distribution with \mathbf{W} . Let $\boldsymbol{\theta}_0$ be the true unknown parameter vector in a parameter space $\Theta \subseteq \mathbb{R}^p$ that indexes the distribution of \mathbf{W} . Let also $\mathbf{f} : \mathbb{R}^d \times \Theta \rightarrow \mathbb{R}^q$ ($q \geq p$) be a function such that

$$\mathbf{f}(\mathbf{w}, \boldsymbol{\theta}) = \left(\mathbf{f}^{(i)}(w^{(i)}, \boldsymbol{\theta}) \right)_{1 \leq i \leq d} = \begin{pmatrix} \mathbf{f}^{(1)}(w^{(1)}, \boldsymbol{\theta}) \\ \vdots \\ \mathbf{f}^{(d)}(w^{(d)}, \boldsymbol{\theta}) \end{pmatrix}, \quad (4)$$

where $\mathbf{w} = (w^{(1)}, \dots, w^{(d)})$, $\boldsymbol{\theta} = (\theta_1, \dots, \theta_p)$ and $\mathbf{f}^{(i)}(w^{(i)}, \boldsymbol{\theta}), i = 1, \dots, d$ is a column vector with dimension k .

The gradient matrix of $\mathbf{f}(\mathbf{w}, \boldsymbol{\theta})$ is given by

$$\frac{\partial \mathbf{f}(\mathbf{w}, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \begin{pmatrix} \frac{\partial \mathbf{f}^{(1)}(w^{(1)}, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \\ \vdots \\ \frac{\partial \mathbf{f}^{(d)}(w^{(d)}, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \end{pmatrix} = \begin{pmatrix} \frac{\partial \mathbf{f}^{(1)}(w^{(1)}, \boldsymbol{\theta})}{\partial \theta_1} & \dots & \frac{\partial \mathbf{f}^{(1)}(w^{(1)}, \boldsymbol{\theta})}{\partial \theta_p} \\ \vdots & & \vdots \\ \frac{\partial \mathbf{f}^{(d)}(w^{(d)}, \boldsymbol{\theta})}{\partial \theta_1} & \dots & \frac{\partial \mathbf{f}^{(d)}(w^{(d)}, \boldsymbol{\theta})}{\partial \theta_p} \end{pmatrix}$$

whenever it exists, for all $\mathbf{w} \in \mathbb{R}^d, \boldsymbol{\theta} \in \Theta$, where $\frac{\partial \mathbf{f}^{(i)}(w^{(i)}, \boldsymbol{\theta})}{\partial \theta_j}, i = 1, \dots, d, j = 1, \dots, p$ is a k -dimensional column vector.

The function \mathbf{f} is assumed to be such that the following equation

$$\mathbb{E}[\mathbf{f}(\mathbf{W}, \boldsymbol{\theta})] = \mathbf{0}, \boldsymbol{\theta} \in \Theta, \quad (5)$$

is solved by $\boldsymbol{\theta}_0$. Equation (5) is called population moment condition. It represents a set of $q = kd$ equations for the p -dimensional unknown $\boldsymbol{\theta}$. The population moment condition should provide sufficient information to identify $\boldsymbol{\theta}_0$ (identification condition). Since \mathbf{W} is issued from $\boldsymbol{\theta}_0$, we cannot compute the expectation and we should rely on sample averages. The corresponding sample moment condition is given by

$$\hat{\mathbf{g}}_N(\boldsymbol{\theta}) = \mathbf{0} \quad (6)$$

where $\hat{\mathbf{g}}_N(\boldsymbol{\theta})$ is the sample moment

$$\hat{\mathbf{g}}_N(\boldsymbol{\theta}) = \frac{1}{N} \sum_{n=1}^N \mathbf{f}(\mathbf{W}_n, \boldsymbol{\theta}).$$

Next, the GMM estimator is defined as follows:

Definition 2.1. (Hall, 2005, Definition 1.2, p.14) Let (\mathbf{P}_N) be a sequence of positive semi-definite weighting matrices that converges in probability to a constant positive definite matrix \mathbf{P} . Then, the GMM estimator based on the population moment condition (5) is given by

$$\hat{\boldsymbol{\theta}}_N = \arg \min_{\boldsymbol{\theta} \in \Theta} \mathbf{Q}_N(\boldsymbol{\theta}) \quad (7)$$

where $\mathbf{Q}_N(\boldsymbol{\theta}) = \hat{\mathbf{g}}_N(\boldsymbol{\theta})^T \mathbf{P}_N \hat{\mathbf{g}}_N(\boldsymbol{\theta})$.

In the following, the convergence in probability is denoted by "Prob", the almost sure convergence by "a.s." and the convergence in distribution by "D".

2.2 Asymptotic properties

Under a few technical assumptions, GMM estimator is consistent and asymptotically normal (see Hall (2005), Hansen (1982), Newey and McFadden (1994)). There are a variety of ways to prove such asymptotic properties. We here follow the results of Newey and McFadden (1994).

Let $\|\cdot\|_\infty$ design the infinity norm. The following theorem is a direct consequence of (Newey and McFadden, 1994, Theorem 2.6, p.2132).

Theorem 2.2. Assume that

(H₁) $\{\mathbf{W}_n, n = 1, \dots, N\}$ are i.i.d.,

(H₂) (\mathbf{P}_N) converges in probability to a constant positive definite matrix \mathbf{P} ,

(H₃) $\mathbb{E}[\mathbf{f}(\mathbf{W}, \boldsymbol{\theta})] = 0$ if and only if $\boldsymbol{\theta} = \boldsymbol{\theta}_0$ (identification condition),

(H₄) Θ is compact,

(H₅) $\mathbf{w} \mapsto \mathbf{f}(\mathbf{w}, \boldsymbol{\theta})$ is Borel measurable for each $\boldsymbol{\theta} \in \Theta$ and $\boldsymbol{\theta} \mapsto \mathbf{f}(\mathbf{w}, \boldsymbol{\theta})$ is continuous for each $\mathbf{w} \in \mathbb{R}^d$,

(H₆) $\mathbb{E} \left[\sup_{\boldsymbol{\theta} \in \Theta} \|\mathbf{f}(\mathbf{W}, \boldsymbol{\theta})\|_\infty \right] < \infty$.

Then $\hat{\boldsymbol{\theta}}_N \xrightarrow{\text{Prob}} \boldsymbol{\theta}_0$.

Remark 2.3. Theorem 2.2 is a weak consistency result. To get the strong consistency, it is necessary and sufficient to replace the convergence in probability by the almost sure convergence in assumption (H₂), see (Hansen, 1982, Theorem 2.1, p.1035).

The following theorem states an asymptotic normality result.

Theorem 2.4. (Newey and McFadden, 1994, Theorem 3.2, p.2145) Suppose that $(H_1), (H_2), (H_3), (H_4), (H_5), (H_6)$ are satisfied and

(H_7) $\boldsymbol{\theta}_0$ is an interior point of Θ ,

(H_8) $\boldsymbol{\theta} \mapsto \frac{\partial \mathbf{f}(\mathbf{w}, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}}$ is continuous on some neighborhood $O \subset \Theta$ of $\boldsymbol{\theta}_0$ for each $\mathbf{w} \in \mathbb{R}^d$,

(H_9) $\sqrt{N} \hat{\mathbf{g}}_N(\boldsymbol{\theta}_0) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \mathbf{S})$ with $\mathbf{S} = \mathbb{E} [\mathbf{f}(\mathbf{W}, \boldsymbol{\theta}_0) \mathbf{f}(\mathbf{W}, \boldsymbol{\theta}_0)^T]$,

(H_{10}) $\boldsymbol{\theta} \mapsto \mathbb{E} \left[\frac{\partial \mathbf{f}(\mathbf{W}, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right]$ is continuous at $\boldsymbol{\theta}_0$ and

$$\sup_{\boldsymbol{\theta} \in O} \left\| \frac{1}{N} \sum_{n=1}^N \left[\frac{\partial \mathbf{f}(\mathbf{W}_n, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right] - \mathbb{E} \left[\frac{\partial \mathbf{f}(\mathbf{W}, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right] \right\|_{\infty} \xrightarrow{\text{Prob}} 0,$$

(H_{11}) $\mathbf{D}_0^T \mathbf{P} \mathbf{D}_0$ is non-singular with

$$\mathbf{D}_0 = \mathbb{E} \left[\frac{\partial \mathbf{f}(\mathbf{W}, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} \right]. \quad (8)$$

Then

$$\sqrt{N} (\hat{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_0) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \mathbf{H} \mathbf{S} \mathbf{H}^T) \quad (9)$$

with $\mathbf{H} = (\mathbf{D}_0^T \mathbf{P} \mathbf{D}_0)^{-1} \mathbf{D}_0^T \mathbf{P}$.

The optimal weighting matrix is given by the following theorem which is derived from the result of Hall for stationary and ergodic random vectors (see (Hall, 2005, Theorem 3.4, p.88)).

Theorem 2.5. Suppose that the hypotheses for the asymptotic normality are satisfied and the matrix \mathbf{S} given in (H_9) is non-singular. Then the smallest possible asymptotic variance is

$$\mathbf{V} = (\mathbf{D}_0^T \mathbf{S}^{-1} \mathbf{D}_0)^{-1} \quad (10)$$

and this can be obtained by setting $\mathbf{P} = \mathbf{S}^{-1}$.

2.3 Estimation

As the matrix \mathbf{S} depends on the unknown parameter vector $\boldsymbol{\theta}_0$, it is also unknown and it should be estimated. An easy way to proceed is to adopt a two-step procedure (see Hansen (1982)) as follows:

1. Set $\mathbf{P}_N = \mathbf{I}$, where \mathbf{I} is the identity matrix, and compute

$$\hat{\boldsymbol{\theta}}_N^{(1)} = \arg \min_{\boldsymbol{\theta} \in \Theta} \mathbf{Q}_N(\boldsymbol{\theta}). \quad (11)$$

2. Construct an estimator of \mathbf{S} based on the initial GMM estimator $\hat{\boldsymbol{\theta}}_N^{(1)}$

$$\hat{\mathbf{S}}_N^{(1)} = \frac{1}{N} \sum_{n=1}^N \mathbf{f}(\mathbf{W}_n, \hat{\boldsymbol{\theta}}_N^{(1)}) \mathbf{f}(\mathbf{W}_n, \hat{\boldsymbol{\theta}}_N^{(1)})^T. \quad (12)$$

Theorem 2.5 implies that a good choice for \mathbf{P}_N is $(\hat{\mathbf{S}}_N^{(1)})^{-1}$, under the condition that $\hat{\mathbf{S}}_N^{(1)}$ is non-singular. Then, the two-step GMM estimator of $\boldsymbol{\theta}_0$ is given by

$$\hat{\boldsymbol{\theta}}_N^{(2)} = \arg \min_{\boldsymbol{\theta} \in \Theta} \hat{\mathbf{g}}_N(\boldsymbol{\theta})^T (\hat{\mathbf{S}}_N^{(1)})^{-1} \hat{\mathbf{g}}_N(\boldsymbol{\theta}). \quad (13)$$

3 GMM for an EGP

Section 2 introduced the general framework of GMM. The focus in this section is on GMM for an EGP.

3.1 Assumptions

The deterioration is modeled by an EGP $\mathbf{X} = (X_t)_{t \in \mathbb{R}^+} \sim \Gamma(A(\cdot), b(\cdot))$ (see (3)), which is assumed to be observed on a compact set $[0, T]$. The shape and scale functions depend on a parameter $\boldsymbol{\theta}_0 \in \Theta$ and we set $A(\cdot, \boldsymbol{\theta}_0)$ and $b(\cdot, \boldsymbol{\theta}_0)$ to highlight the dependence with respect to $\boldsymbol{\theta}_0$. We also set $a(t, \boldsymbol{\theta}_0)$ to be the derivative of $A(t, \boldsymbol{\theta}_0)$ with respect to t (which exists almost surely). Based on the fact that the process is observed on $[0, T]$, it is enough to assume that (2) holds only on $[0, T]$, namely that

$$\int_{(0,t]} \frac{a(s, \boldsymbol{\theta}_0)}{b(s, \boldsymbol{\theta}_0)} ds < \infty, \quad \int_{(0,t]} \frac{a(s, \boldsymbol{\theta}_0)}{b(s, \boldsymbol{\theta}_0)^2} ds < \infty \quad (14)$$

for all $t \in [0, T]$.

The Laplace transform of an EGP increment is given by

$$\mathcal{L}_{X_{t+h}-X_t}(\lambda, \boldsymbol{\theta}_0) := \mathbb{E} \left(e^{-\lambda(X_{t+h}-X_t)} \right) = \exp \left(- \int_{(t,t+h]} \log \left(1 + \frac{\lambda}{b(s, \boldsymbol{\theta}_0)} \right) a(s, \boldsymbol{\theta}_0) ds \right), \quad (15)$$

for all $t \in [0, T]$, $\lambda \geq 0$ and $h > 0$ (see Çinlar (1980)).

The mean and variance are given by

$$\mathbb{E}(X_{t+h} - X_t) = \int_{(t,t+h]} \frac{a(s, \boldsymbol{\theta}_0) ds}{b(s, \boldsymbol{\theta}_0)}, \quad \mathbb{V}(X_{t+h} - X_t) = \int_{(t,t+h]} \frac{a(s, \boldsymbol{\theta}_0) ds}{b(s, \boldsymbol{\theta}_0)^2} \quad (16)$$

(see Çinlar (1980)).

We assume that N independent processes $\mathbf{X}_n, n = 1, \dots, N$ are observed at the same inspection times $t_0 = 0 < t_1 < \dots < t_d = T$, which are copies of the process \mathbf{X} . For the process \mathbf{X} , we set

$$W^{(i)} = X_{t_i} - X_{t_{i-1}}, i = 1, \dots, d \quad (17)$$

to be the i -th increment and

$$\mathbf{W} = \left(W^{(1)}, \dots, W^{(d)} \right)^T. \quad (18)$$

Defining \mathbf{W}_n in the same way as \mathbf{W} with \mathbf{X} substituted by \mathbf{X}_n , this leads to a sequence of i.i.d. d -dimensional random vectors $\{\mathbf{W}_n, n = 1, \dots, N\}$.

Now, we turn to discuss the choice of the population moment condition as well as the asymptotic properties for each of the following approaches: GMM based on the moments and GMM based on the Laplace transform.

3.2 GMM based on the moments

3.2.1 Approach

As in the classical method of moments, the population moment condition for GMM based on the moments is defined by matching the theoretical moments of an EGP with the appropriate empirical ones. This matching is done for each increment. Define

$$\mathbf{f}_{MM}(\mathbf{w}, \boldsymbol{\theta}) = \left(\begin{array}{c} w^{(i)} - m^{(i)}(\boldsymbol{\theta}) \\ (w^{(i)} - m^{(i)}(\boldsymbol{\theta}))^2 - v^{(i)}(\boldsymbol{\theta}) \end{array} \right)_{1 \leq i \leq d} \quad (19)$$

where

$$m^{(i)}(\boldsymbol{\theta}) = \int_{t_{i-1}}^{t_i} \frac{a(s, \boldsymbol{\theta})}{b(s, \boldsymbol{\theta})} ds, \quad v^{(i)}(\boldsymbol{\theta}) = \int_{t_{i-1}}^{t_i} \frac{a(s, \boldsymbol{\theta})}{b(s, \boldsymbol{\theta})^2} ds. \quad (20)$$

Based on (16), \mathbf{f}_{MM} satisfies $\mathbb{E}[\mathbf{f}_{MM}(\mathbf{W}, \boldsymbol{\theta}_0)] = 0$. Based on assumptions given in (14), we have $\mathbb{E}[\mathbf{f}_{MM}(\mathbf{W}, \boldsymbol{\theta})] < \infty$ for all $\boldsymbol{\theta} \in \Theta$. Also

$$\hat{\mathbf{g}}_N(\boldsymbol{\theta}) = \left(\begin{array}{c} \frac{1}{N} \sum_{n=1}^N [W_n^{(i)} - m^{(i)}(\boldsymbol{\theta})] \\ \frac{1}{N} \sum_{n=1}^N [(W_n^{(i)} - m^{(i)}(\boldsymbol{\theta}))^2 - v^{(i)}(\boldsymbol{\theta})] \end{array} \right)_{1 \leq i \leq d} = \left(\begin{array}{c} \hat{m}^{(i)} - m^{(i)}(\boldsymbol{\theta}) \\ \hat{v}^{(i)}(\boldsymbol{\theta}) - v^{(i)}(\boldsymbol{\theta}) \end{array} \right)_{1 \leq i \leq d}$$

with

$$\hat{m}^{(i)} = \frac{1}{N} \sum_{n=1}^N W_n^{(i)}; \quad \hat{v}^{(i)}(\boldsymbol{\theta}) = \frac{1}{N} \sum_{n=1}^N (W_n^{(i)} - m^{(i)}(\boldsymbol{\theta}))^2.$$

The gradient matrix is given by

$$\frac{\partial \mathbf{f}_{MM}(\mathbf{w}, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \left(\begin{array}{c} -\frac{\partial m^{(i)}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \\ -2 \frac{\partial m^{(i)}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} (w^{(i)} - m^{(i)}(\boldsymbol{\theta})) - \frac{\partial v^{(i)}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \end{array} \right)_{1 \leq i \leq d}. \quad (21)$$

3.2.2 Asymptotic properties

In this subsection, conditions under which consistency and asymptotic normality of GMM estimators for an EGP based on the moments hold are presented in a general setting. These conditions will be checked for a particular parametric form of the shape and scale functions in Section 4. The results are derived from Theorems 2.2, 2.4, 2.5.

Before presenting the results, we list some assumptions we shall use to show the asymptotic properties for GMM based on both moments and Laplace transform (in Section 3.3).

Assumptions.

(M_{1j}) For $j = 1, 2$: There exists $I_j : [0, T] \rightarrow (0, +\infty)$ such that

$$\left| \frac{a(s, \boldsymbol{\theta})}{b^j(s, \boldsymbol{\theta})} \right| \leq I_j(s), \quad \text{for all } s \in [0, T], \boldsymbol{\theta} \in \Theta$$

with $\int_0^T I_j(s) ds < \infty$.

(M_{2j}) For $j = 1, 2$: There exists $J_j : [0, T] \rightarrow (0, +\infty)$ such that

$$\left\| \frac{\partial}{\partial \boldsymbol{\theta}} \left(\frac{a(s, \boldsymbol{\theta})}{b^j(s, \boldsymbol{\theta})} \right) \right\|_{\infty} \leq J_j(s), \text{ for all } s \in [0, T], \boldsymbol{\theta} \in \Theta$$

with $\int_0^T J_j(s) ds < \infty$.

(M_{3j}) For $j = 1, 2$: $\boldsymbol{\theta} \mapsto \frac{\partial}{\partial \boldsymbol{\theta}} \left(\frac{a(s, \boldsymbol{\theta})}{b^j(s, \boldsymbol{\theta})} \right)$ are continuous for each $s \in [0, T]$.

(M_{4j}) For $j = 1, 2$: $m_{j+2}([0, T], \boldsymbol{\theta}) := \int_0^T \frac{a(s, \boldsymbol{\theta})}{b^{j+2}(s, \boldsymbol{\theta})} ds < \infty$.

Let us consider \mathbf{W} and \mathbf{f}_{MM} as defined in (18) and (19), respectively. Then, the following lemma is easy to check and stated without proof.

Lemma 3.1. *Suppose that*

(a) Θ is compact,

(b) $\boldsymbol{\theta} \mapsto m^{(i)}(\boldsymbol{\theta})$ and $\boldsymbol{\theta} \mapsto v^{(i)}(\boldsymbol{\theta})$ are continuous for $1 \leq i \leq d$.

Then (i) (H₆) holds; (ii) if (H₇) holds and $\boldsymbol{\theta} \mapsto \frac{\partial m^{(i)}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}$ and $\boldsymbol{\theta} \mapsto \frac{\partial v^{(i)}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}$ are continuous, then (H₁₀) holds.

Next, a consistency result for GMM based on the moments is formulated.

Corollary 3.2 (Consistency). *Assumptions (H₂), (H₃), (H₄), (M₁₁) and (M₁₂) imply*

$$\hat{\boldsymbol{\theta}}_N \xrightarrow[N \rightarrow \infty]{Prob} \boldsymbol{\theta}_0.$$

Proof. By hypothesis and Lemma 3.1 (part (i)), we can easily verify that the conditions of Theorem 2.2 are satisfied.

The following corollary states the asymptotic normality of GMM estimator for an EGP based on the moments.

Corollary 3.3 (Asymptotic Normality). *Suppose that Θ is compact and $\boldsymbol{\theta}_0$ is an interior point of Θ . If (H₂), (H₃), (H₁₁) and (M_{ij}), $i \in \{1, 2, 3, 4\}$, $j \in \{1, 2\}$ are satisfied, then $\sqrt{N}(\hat{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_0)$ is asymptotically normal with mean 0 and asymptotic variance $\mathbf{V} = \mathbf{HSH}^T$ where*

$$\mathbf{H} = (\mathbf{D}_0^T \mathbf{P} \mathbf{D}_0)^{-1} \mathbf{D}_0^T \mathbf{P},$$

$$\mathbf{D}_0 = - \left(\begin{array}{c} \frac{\partial m^{(i)}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} \\ \frac{\partial v^{(i)}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} \end{array} \right)_{1 \leq i \leq d} = - \left(\begin{array}{ccc} \frac{\partial m^{(i)}(\boldsymbol{\theta})}{\partial \theta_1} \Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} & \cdots & \frac{\partial m^{(i)}(\boldsymbol{\theta})}{\partial \theta_p} \Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} \\ \frac{\partial v^{(i)}(\boldsymbol{\theta})}{\partial \theta_1} \Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} & \cdots & \frac{\partial v^{(i)}(\boldsymbol{\theta})}{\partial \theta_p} \Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} \end{array} \right)_{1 \leq i \leq d} \quad (22)$$

(from (8) and (21)).

Proof. The point here is to check the conditions for applying Theorem 2.4. Based on Assumptions $(M_{ij}), i \in 2, 3, j \in 1, 2$, Lemma 3.1, Lemma 2.4 from Newey and McFadden (1994) and using the dominated convergence theorem, one can easily verify conditions (H_8) and (H_{10}) . It remains to show (H_9) .

We first point out that \mathbf{S} is the covariance matrix of

$$\dot{\mathbf{W}} = \left(W^{(1)}, \left(W^{(1)} - m^{(1)}(\boldsymbol{\theta}_0) \right)^2, \dots, W^{(d)}, \left(W^{(d)} - m^{(d)}(\boldsymbol{\theta}_0) \right)^2 \right)^T. \quad (23)$$

Then, we compute \mathbf{S} .

$$\mathbf{S} = \mathbb{E} \left[\mathbf{f}_{MM}(\mathbf{W}, \boldsymbol{\theta}_0) \mathbf{f}_{MM}(\mathbf{W}, \boldsymbol{\theta}_0)^T \right] = \begin{pmatrix} \mathbf{B}^{(1)} & \mathbf{Z}^{(1,2)} & \dots & \mathbf{Z}^{(1,d-1)} & \mathbf{Z}^{(1,d)} \\ \mathbf{Z}^{(2,1)} & \mathbf{B}^{(2)} & \dots & \mathbf{Z}^{(2,d-1)} & \mathbf{Z}^{(2,d)} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{Z}^{(d-1,1)} & \mathbf{Z}^{(d-1,2)} & \dots & \mathbf{B}^{(d-1)} & \mathbf{Z}^{(d-1,d)} \\ \mathbf{Z}^{(d,1)} & \mathbf{Z}^{(d,2)} & \dots & \mathbf{Z}^{(d,d-1)} & \mathbf{B}^{(d)} \end{pmatrix}$$

with

$$\mathbf{B}^{(i)} = \begin{pmatrix} \mathbb{E} \left[\left(W^{(i)} - m^{(i)}(\boldsymbol{\theta}_0) \right)^2 \right] & \mathbb{E} \left[\left(W^{(i)} - m^{(i)}(\boldsymbol{\theta}_0) \right) \left(\left(W^{(i)} - m^{(i)}(\boldsymbol{\theta}_0) \right)^2 - v^{(i)}(\boldsymbol{\theta}_0) \right) \right] \\ \mathbb{E} \left[\left(W^{(i)} - m^{(i)}(\boldsymbol{\theta}_0) \right) \left(\left(W^{(i)} - m^{(i)}(\boldsymbol{\theta}_0) \right)^2 - v^{(i)}(\boldsymbol{\theta}_0) \right) \right] & \mathbb{E} \left[\left(\left(W^{(i)} - m^{(i)}(\boldsymbol{\theta}_0) \right)^2 - v^{(i)}(\boldsymbol{\theta}_0) \right)^2 \right] \end{pmatrix}, \quad (24)$$

$$\mathbf{Z}^{(i,j)} = \begin{pmatrix} \mathbb{E} \left[\left(W^{(i)} - m^{(i)}(\boldsymbol{\theta}_0) \right) \left(W^{(j)} - m^{(j)}(\boldsymbol{\theta}_0) \right) \right] & \mathbb{E} \left[\left(W^{(i)} - m^{(i)}(\boldsymbol{\theta}_0) \right) \left(\left(W^{(j)} - m^{(j)}(\boldsymbol{\theta}_0) \right)^2 - v^{(j)}(\boldsymbol{\theta}_0) \right) \right] \\ \mathbb{E} \left[\left(W^{(j)} - m^{(j)}(\boldsymbol{\theta}_0) \right) \left(\left(W^{(i)} - m^{(i)}(\boldsymbol{\theta}_0) \right)^2 - v^{(i)}(\boldsymbol{\theta}_0) \right) \right] & \mathbb{E} \left[\left(\left(W^{(i)} - m^{(i)}(\boldsymbol{\theta}_0) \right)^2 - v^{(i)}(\boldsymbol{\theta}_0) \right) \left(\left(W^{(j)} - m^{(j)}(\boldsymbol{\theta}_0) \right)^2 - v^{(j)}(\boldsymbol{\theta}_0) \right) \right] \end{pmatrix}$$

for $i \neq j$.

Based on the fact that $W^{(i)}$ and $W^{(j)}$ are independent ($i \neq j$), it is readily seen that $\mathbf{Z}^{(i,j)} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$. Furthermore, we have

$$\begin{aligned} \mathbb{E} \left[\left(W^{(i)} - m^{(i)}(\boldsymbol{\theta}_0) \right)^2 \right] &= v^{(i)}(\boldsymbol{\theta}_0) < \infty \text{ (by Assumption (14))}, \\ \mathbb{E} \left[\left(W^{(i)} - m^{(i)}(\boldsymbol{\theta}_0) \right) \left(\left(W^{(i)} - m^{(i)}(\boldsymbol{\theta}_0) \right)^2 - v^{(i)}(\boldsymbol{\theta}_0) \right) \right] &= 2m_3([t_i, t_{i+1}], \boldsymbol{\theta}_0) < \infty \text{ (by } (M_{41})), \\ \mathbb{E} \left[\left(\left(W^{(i)} - m^{(i)}(\boldsymbol{\theta}_0) \right)^2 - v^{(i)}(\boldsymbol{\theta}_0) \right)^2 \right] &= 6m_4([t_i, t_{i+1}], \boldsymbol{\theta}_0) + 2v^{(i)}(\boldsymbol{\theta}_0)^2 < \infty \text{ (by (14) and } (M_{42})). \end{aligned}$$

Consequently, $\mathbf{B}^{(i)}$ is finite for all i and then \mathbf{S} is finite too. Hence, by the central limit theorem for i.i.d. random variables, we obtain

$$\sqrt{N} \hat{\mathbf{g}}_N(\boldsymbol{\theta}_0) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \mathbf{S}).$$

This ends the proof. ■

Corollary 3.4 (Optimal weighting matrix). *Suppose that Θ is compact and $\boldsymbol{\theta}_0$ is an interior point of Θ . If $(H_3), (M_{ij}), i \in \{1, 2, 3, 4\}, j \in \{1, 2\}$ are satisfied and $(H_2), (H_{11})$ hold for $\mathbf{P} = \mathbf{S}^{-1}$, the minimal asymptotic variance of $\hat{\boldsymbol{\theta}}_N$ is $(\mathbf{D}_0^T \mathbf{S}^{-1} \mathbf{D}_0)^{-1}$ where*

$$\mathbf{S} = \begin{pmatrix} \mathbf{B}^{(1)} & \mathbf{0}_{2,2} & \dots & \mathbf{0}_{2,2} & \mathbf{0}_{2,2} \\ \mathbf{0}_{2,2} & \mathbf{B}^{(2)} & \dots & \mathbf{0}_{2,2} & \mathbf{0}_{2,2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0}_{2,2} & \mathbf{0}_{2,2} & \dots & \mathbf{B}^{(d-1)} & \mathbf{0}_{2,2} \\ \mathbf{0}_{2,2} & \mathbf{0}_{2,2} & \dots & \mathbf{0}_{2,2} & \mathbf{B}^{(d)} \end{pmatrix}$$

with

$$\mathbf{0}_{2,2} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \mathbf{B}^{(i)} = \begin{pmatrix} v^{(i)}(\boldsymbol{\theta}_0) & 2m_3([t_i, t_{i+1}], \boldsymbol{\theta}_0) \\ 2m_3([t_i, t_{i+1}], \boldsymbol{\theta}_0) & 6m_4([t_i, t_{i+1}], \boldsymbol{\theta}_0) + 2v^{(i)}(\boldsymbol{\theta}_0)^2 \end{pmatrix}.$$

Proof. We need to show that \mathbf{S} (the covariance matrix of $\mathring{\mathbf{W}}$ given in (23)) is a positive definite matrix (see Theorem 2.5). Let us start by recalling that a covariance matrix is a positive semi-definite matrix. Also, it is known to be positive definite if and only if there does not exist any almost sure linear relation between the components of the random vector $\mathring{\mathbf{W}}$ (see (Koch, 1999, Theorem (2.79), p.101)).

Assume that there exist a vector $\underline{\mathbf{h}} = (h_1, \check{h}_1, \dots, h_d, \check{h}_d)$ and a constant \bar{c} such that

$$\mathbb{P} \left[\underline{\mathbf{h}} \mathring{\mathbf{W}} = \bar{c} \right] = 1. \quad (25)$$

By setting

$$\mathcal{E}^{(i)}(w^{(i)}) = h_i w^{(i)} + \check{h}_i \left(w^{(i)} - m^{(i)}(\boldsymbol{\theta}_0) \right)^2, \quad (26)$$

Equation (25) can be rewritten as

$$\mathbb{P} \left[\sum_{i=1}^d \mathcal{E}^{(i)}(W^{(i)}) = \bar{c} \right] = 1. \quad (27)$$

Equation (27) implies that $\sum_{i=1}^d \mathcal{E}^{(i)}(W^{(i)})$ is almost surely constant and then

$\mathbb{V} \left[\sum_{i=1}^d \mathcal{E}^{(i)}(W^{(i)}) \right] = 0$. Since $W^{(1)}, \dots, W^{(d)}$ are independent, it follows that $\sum_{i=1}^d \mathbb{V}[\mathcal{E}^{(i)}(W^{(i)})] = 0$ and then $\mathbb{V}[\mathcal{E}^{(i)}(W^{(i)})] = 0$ for all $i = 1, \dots, d$. This means that, for $i = 1, \dots, d$, the random variable $\mathcal{E}(W^{(i)})$ is almost surely constant, namely that there exists $c^{(i)}$ such that

$$\mathcal{E}^{(i)}(W^{(i)}) - c^{(i)} = 0 \text{ almost surely.} \quad (28)$$

Assume now that $(h_i, \check{h}_i) \neq (0, 0)$. Then, for any $c^{(i)}$, the second degree polynomial $\mathcal{E}(w^{(i)}) - c^{(i)}$ (see (26)) has at most two zeros. Hence, (28) means that $W^{(i)}$ can take only two possible values, which is false, because $W^{(i)}$ is a continuous random variable. Therefore, $(h_i, \check{h}_i) = (0, 0)$ for all i and consequently $\underline{\mathbf{h}} = \mathbf{0}$. The matrix \mathbf{S} hence is positive definite. ■

In the sequel, \mathbf{S} is estimated by $\hat{\mathbf{S}}_N$, where

$$\hat{\mathbf{S}}_N = \frac{1}{N} \sum_{n=1}^N \mathbf{f}_{MM}(\mathbf{W}_n, \boldsymbol{\theta}) \mathbf{f}_{MM}(\mathbf{W}_n, \boldsymbol{\theta})^T \quad (29)$$

is the empirical version of \mathbf{S} .

3.2.3 Optimal two-step GMM estimator for an EGP

Let us recall that our GMM estimator is obtained through the two-step procedure given in Subsection 2.3. The point now is to show that the optimal two-step GMM estimator $\hat{\boldsymbol{\theta}}_N^{(2)}$ given in (13) converges almost surely to $\boldsymbol{\theta}_0$. The result is provided below for GMM based on the moments.

Proposition 3.5. *Consider \mathbf{W} and \mathbf{f}_{MM} as defined in (18) and (19), respectively. Suppose that Θ is compact and $\boldsymbol{\theta}_0$ is an interior point of Θ . If (H_3) , (M_{ij}) , $i \in \{1, \dots, 4\}$, $j \in \{1, 2\}$ hold, (H_{11}) is satisfied for $\mathbf{P} = \mathbf{S}^{-1}$, then*

$$\hat{\boldsymbol{\theta}}_N^{(2)} \xrightarrow[N \rightarrow \infty]{a.s.} \boldsymbol{\theta}_0.$$

Proof. Considering $\mathbf{P}_N = \mathbf{I}$, we know that $\hat{\boldsymbol{\theta}}_N^{(1)}$ (given by (11)) converges almost surely to $\boldsymbol{\theta}_0$ (by Remark 2.3 and Corollary 3.2). We first show that $\hat{\mathbf{S}}_N^{(1)} \xrightarrow[N \rightarrow \infty]{a.s.} \mathbf{S}$ where

$$\hat{\mathbf{S}}_N^{(1)} = \frac{1}{N} \sum_{n=1}^N \mathbf{f}_{MM}(\mathbf{w}_n, \hat{\boldsymbol{\theta}}_N^{(1)}) \mathbf{f}_{MM}(\mathbf{w}_n, \hat{\boldsymbol{\theta}}_N^{(1)})^T.$$

Set $\phi(\mathbf{w}, \boldsymbol{\theta}) = \mathbf{f}_{MM}(\mathbf{w}, \boldsymbol{\theta}) \mathbf{f}_{MM}(\mathbf{w}, \boldsymbol{\theta})^T$, $\boldsymbol{\theta} \in \Theta$. By the triangular inequality, we get

$$\begin{aligned} \left| \frac{1}{N} \sum_{n=1}^N \phi(\mathbf{w}_n, \hat{\boldsymbol{\theta}}_N^{(1)}) - \mathbb{E}[\phi(\mathbf{W}, \boldsymbol{\theta}_0)] \right| &\leq \left| \frac{1}{N} \sum_{n=1}^N \phi(\mathbf{w}_n, \hat{\boldsymbol{\theta}}_N^{(1)}) - \mathbb{E}[\phi(\mathbf{W}, \hat{\boldsymbol{\theta}}_N^{(1)})] \right| \\ &\quad + \left| \mathbb{E}[\phi(\mathbf{W}, \hat{\boldsymbol{\theta}}_N^{(1)})] - \mathbb{E}[\phi(\mathbf{W}, \boldsymbol{\theta}_0)] \right|. \end{aligned}$$

It can be recalled that $\mathbf{w} \mapsto \mathbf{f}_{MM}(\mathbf{w}, \boldsymbol{\theta})$ is measurable for each $\boldsymbol{\theta} \in \Theta$ and $\boldsymbol{\theta} \mapsto \mathbf{f}_{MM}(\mathbf{w}, \boldsymbol{\theta})$ is continuous for each $\mathbf{w} \in \mathbb{R}^d$. It follows that $\mathbf{w} \mapsto \phi(\mathbf{w}, \boldsymbol{\theta})$ is measurable for each $\boldsymbol{\theta}$ and $\boldsymbol{\theta} \mapsto \phi(\mathbf{w}, \boldsymbol{\theta})$ is continuous for each \mathbf{w} . Since Θ is compact, it can be easily verified that there exists a function $d(\mathbf{w})$ such that $\|\phi(\mathbf{w}, \boldsymbol{\theta})\|_\infty \leq d(\mathbf{w})$ for all $\mathbf{w} \in \mathbb{R}^d$, $\boldsymbol{\theta} \in \Theta$, where $d(\mathbf{w})$ is of the form

$$d(\mathbf{w}) = \sum_{i=1}^d \left(w^{(i)4} + C_1 w^{(i)3} + C_2 w^{(i)2} + C_3 w^{(i)} \right) + C_4,$$

where C_1, C_2, C_3 and C_4 are positive constants, and where

$$\mathbb{E}[d(\mathbf{W})] < \infty \text{ (by hypothesis).}$$

Then, by applying Theorem 2 of Jennerich (1969), we get

$$\sup_{\boldsymbol{\theta} \in \Theta} \left| \frac{1}{N} \sum_{n=1}^N \phi(\mathbf{W}_n, \boldsymbol{\theta}_1) - \mathbb{E}[\phi(\mathbf{W}, \boldsymbol{\theta}_1)] \right| \xrightarrow[N \rightarrow \infty]{a.s.} 0.$$

Besides, we have

$$\mathbb{E} \left[\phi \left(\mathbf{W}, \hat{\boldsymbol{\theta}}_N^{(1)} \right) \right] \xrightarrow{N \rightarrow \infty} \mathbb{E} [\phi(\mathbf{W}, \boldsymbol{\theta}_0)],$$

since $\hat{\boldsymbol{\theta}}_N^{(1)} \xrightarrow[N \rightarrow \infty]{a.s.} \boldsymbol{\theta}_0$ and $\boldsymbol{\theta} \mapsto \mathbb{E} [\phi(\mathbf{W}, \boldsymbol{\theta})]$ is continuous.

As a consequence, $\hat{\mathbf{S}}_N^{(1)} \xrightarrow[N \rightarrow \infty]{a.s.} \mathbf{S}$. Also, as \mathbf{S} is non-singular, $\hat{\mathbf{S}}_N^{(1)}$ is non-singular too for large N , based on the continuity of the determinant function. It follows that $\left(\hat{\mathbf{S}}_N^{(1)} \right)^{-1} \xrightarrow[N \rightarrow \infty]{a.s.} \mathbf{S}^{-1}$. Then $\hat{\boldsymbol{\theta}}_N^{(2)}$ can be computed by

$$\hat{\boldsymbol{\theta}}_N^{(2)} = \arg \min_{\boldsymbol{\theta} \in \Theta} \hat{\mathbf{g}}_N(\boldsymbol{\theta})^T \left(\hat{\mathbf{S}}_N^{(1)} \right)^{-1} \hat{\mathbf{g}}_N(\boldsymbol{\theta})$$

and it converges almost surely to $\boldsymbol{\theta}_0$ by Remark 2.3. This completes the proof. ■

3.3 GMM based on the Laplace transform

The second approach of GMM for an EGP is here presented. Following similar steps to those in Section 3.2, we first elaborate on the approach, then we provide the asymptotic properties and present a result on the optimal two-step GMM estimator for an EGP based on the Laplace transform.

3.3.1 Approach

A similar procedure as for the previous method in Subsection 3.2.1 is followed. Instead of using moments, we rely on the Laplace transform at distinct points $\lambda_l, 1 \leq l \leq o$ (where $o \in \mathbb{N}^*$) and we match it with the empirical Laplace transform at the same points λ_l . Let

$$\mathbf{f}_{Lap}(\mathbf{w}, \boldsymbol{\theta}) = \left(\exp(-\lambda_l w^{(i)}) - \mathcal{L}^{(i)}(\lambda_l, \boldsymbol{\theta}) \right)_{\substack{1 \leq i \leq d \\ 1 \leq l \leq o}} \quad (30)$$

where

$$\mathcal{L}^{(i)}(\lambda_l, \boldsymbol{\theta}) = \exp \left(- \int_{t_{i-1}}^{t_i} \log \left(1 + \frac{\lambda_l}{b(s, \boldsymbol{\theta})} \right) a(s, \boldsymbol{\theta}) ds \right).$$

Based on (15), if the arbitrary $\boldsymbol{\theta}$ is replaced by the true value, we have $\mathbb{E}[\mathbf{f}_{Lap}(\mathbf{W}, \boldsymbol{\theta}_0)] = 0$. Also, $\mathbb{E}[\mathbf{f}_{Lap}(\mathbf{W}, \boldsymbol{\theta})] < \infty$ for all $\boldsymbol{\theta} \in \Theta$ and the sample moment is given by

$$\hat{\mathbf{g}}_N(\boldsymbol{\theta}) = \left(\hat{\mathcal{L}}^{(i)}(\lambda_l) - \mathcal{L}^{(i)}(\lambda_l, \boldsymbol{\theta}) \right)_{\substack{1 \leq i \leq d \\ 1 \leq l \leq o}} \quad (31)$$

with $\hat{\mathcal{L}}^{(i)}(\lambda_l) = \frac{1}{N} \sum_{n=1}^N \exp(-\lambda_l W_n^{(i)})$, $i = 1, \dots, d, l = 1, \dots, o$.

The gradient matrix is given by

$$\frac{\partial \mathbf{f}_{Lap}(\mathbf{w}, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = - \left(\frac{\partial \mathcal{L}^{(i)}(\lambda_l, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right)_{\substack{1 \leq i \leq d \\ 1 \leq l \leq o}} \quad (32)$$

(independent of \mathbf{w}).

3.3.2 Asymptotic properties

Here again, we provide technical assumptions to show asymptotic properties for GMM based on the Laplace transform. Let us now consider \mathbf{W} and \mathbf{f}_{Lap} as defined in (18) and (30), respectively.

Corollary 3.6 (Consistency). *Assumptions $(H_2), (H_3), (H_4)$ and (M_{11}) imply*

$$\hat{\boldsymbol{\theta}}_N \xrightarrow[N \rightarrow \infty]{Prob} \boldsymbol{\theta}_0.$$

Proof. One can immediately show that the conditions of Theorem 2.2 are satisfied (by hypothesis and (M_{11})). ■

Corollary 3.7 (Asymptotic Normality). *If assumptions $(H_2), (H_3), (H_4), (H_7), (H_{11}), (M_{11}), (M_{21}), (M_{31})$ hold then*

$$\sqrt{N}(\hat{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_0) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \mathbf{V})$$

with $\mathbf{V} = \mathbf{H}\mathbf{S}\mathbf{H}^T$, $\mathbf{H} = (\mathbf{D}_0^T \mathbf{P} \mathbf{D}_0)^{-1} \mathbf{D}_0^T \mathbf{P}$ and

$$\mathbf{D}_0 = - \left(\frac{\partial \mathcal{L}^{(i)}(\lambda_l, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \bigg|_{\boldsymbol{\theta} = \boldsymbol{\theta}_0} \right)_{\substack{1 \leq i \leq d \\ 1 \leq l \leq o}}. \quad (33)$$

Proof. We proceed by verifying the conditions of Theorem 2.4. Based on the assumptions and Lemma 2.4 from Newey and McFadden (1994), we can easily justify (H_8) and (H_{10}) . As for Assumption (H_9) , it suffices to prove that $\mathbf{S} = \mathbb{E} [\mathbf{f}_{Lap}(\mathbf{W}, \boldsymbol{\theta}_0) \mathbf{f}_{Lap}(\mathbf{W}, \boldsymbol{\theta}_0)^T]$ is finite (see the proof of Corollary 3.3).

We have

$$\mathbf{S} = \mathbb{E} [\mathbf{f}_{Lap}(\mathbf{W}, \boldsymbol{\theta}_0) \mathbf{f}_{Lap}(\mathbf{W}, \boldsymbol{\theta}_0)^T] = \begin{pmatrix} \mathbf{B}^{(1)} & \mathbf{Z}^{(1,2)} & \dots & \mathbf{Z}^{(1,d-1)} & \mathbf{Z}^{(1,d)} \\ \mathbf{Z}^{(2,1)} & \mathbf{B}^{(2)} & \dots & \mathbf{Z}^{(2,d-1)} & \mathbf{Z}^{(2,d)} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{Z}^{(d-1,1)} & \mathbf{Z}^{(d-1,2)} & \dots & \mathbf{B}^{(d-1)} & \mathbf{Z}^{(d-1,d)} \\ \mathbf{Z}^{(d,1)} & \mathbf{Z}^{(d,2)} & \dots & \mathbf{Z}^{(d,d-1)} & \mathbf{B}^{(d)} \end{pmatrix}$$

with

$$\mathbf{B}^{(i)} = \begin{pmatrix} \mathbb{E} [u(\lambda_1, W^{(i)})^2] & \mathbb{E} [u(\lambda_1, W^{(i)})u(\lambda_2, W^{(i)})] & \dots & \mathbb{E} [u(\lambda_1, W^{(i)})u(\lambda_o, W^{(i)})] \\ \mathbb{E} [u(\lambda_1, W^{(i)})u(\lambda_2, W^{(i)})] & \mathbb{E} [u(\lambda_2, W^{(i)})^2] & \dots & \mathbb{E} [u(\lambda_2, W^{(i)})u(\lambda_o, W^{(i)})] \\ \vdots & \vdots & \ddots & \vdots \\ \mathbb{E} [u(\lambda_1, W^{(i)})u(\lambda_o, W^{(i)})] & \mathbb{E} [u(\lambda_2, W^{(i)})u(\lambda_o, W^{(i)})] & \dots & \mathbb{E} [u(\lambda_o, W^{(i)})^2] \end{pmatrix}, \quad (34)$$

$$\mathbf{Z}^{(i,j)} = \begin{pmatrix} \mathbb{E} [u(\lambda_1, W^{(i)})u(\lambda_1, W^{(j)})] & \mathbb{E} [u(\lambda_1, W^{(i)})u(\lambda_2, W^{(j)})] & \dots & \mathbb{E} [u(\lambda_1, W^{(i)})u(\lambda_o, W^{(j)})] \\ \mathbb{E} [u(\lambda_2, W^{(i)})u(\lambda_1, W^{(j)})] & \mathbb{E} [u(\lambda_2, W^{(i)})u(\lambda_2, W^{(j)})] & \dots & \mathbb{E} [u(\lambda_2, W^{(i)})u(\lambda_o, W^{(j)})] \\ \vdots & \vdots & \ddots & \vdots \\ \mathbb{E} [u(\lambda_o, W^{(i)})u(\lambda_1, W^{(j)})] & \mathbb{E} [u(\lambda_o, W^{(i)})u(\lambda_2, W^{(j)})] & \dots & \mathbb{E} [u(\lambda_o, W^{(i)})u(\lambda_o, W^{(j)})] \end{pmatrix},$$

$$u(\lambda_l, W^{(i)}) = \exp(-\lambda_l W^{(i)}) - \mathcal{L}^{(i)}(\lambda_l, \boldsymbol{\theta}_0), l = 1, \dots, o..$$

In addition, it can be readily seen that $B^{(i)} < \infty$ for all i and $Z^{(i,j)} = \begin{pmatrix} 0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \end{pmatrix}$.

Consequently, we get $\mathbf{S} < \infty$. ■

Corollary 3.8 (Optimal weighting matrix). *Under assumptions $(H_3), (H_4), (H_7), (M_{11}), (M_{21}), (M_{31})$ and if $(H_2), (H_{11})$ are satisfied for $\mathbf{P} = \mathbf{S}^{-1}$, then the minimal asymptotic variance of $\hat{\boldsymbol{\theta}}_N$ is $(\mathbf{D}_0^T \mathbf{S}^{-1} \mathbf{D}_0)^{-1}$ with*

$$\mathbf{S} = \begin{pmatrix} \mathbf{B}^{(1)} & \mathbf{0}_{o,o} & \cdots & \mathbf{0}_{o,o} & \mathbf{0}_{o,o} \\ \mathbf{0}_{o,o} & \mathbf{B}^{(2)} & \cdots & \mathbf{0}_{o,o} & \mathbf{0}_{o,o} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0}_{o,o} & \mathbf{0}_{o,o} & \cdots & \mathbf{B}^{(d-1)} & \mathbf{0}_{o,o} \\ \mathbf{0}_{o,o} & \mathbf{0}_{o,o} & \cdots & \mathbf{0}_{o,o} & \mathbf{B}^{(d)} \end{pmatrix}$$

where $\mathbf{0}_{o,o}$ is a null matrix and $\mathbf{B}^{(i)}$ is given by (34).

Proof. The proof is presented for $o = 2$. Let

$$\dot{\mathbf{W}} = \left(\exp(-\lambda_1 W^{(1)}), \exp(-\lambda_2 W^{(1)}), \dots, \exp(-\lambda_1 W^{(d)}), \exp(-\lambda_2 W^{(d)}) \right)^T. \quad (35)$$

Assume that there exist a vector $\underline{h} = (h_1, \ddot{h}_1, \dots, h_d, \ddot{h}_d)$ and a constant \bar{c} such that

$$\mathbb{P} \left[\sum_{i=1}^d \mathcal{E}^{(i)}(W^{(i)}) = \bar{c} \right] = 1 \quad (36)$$

with

$$\mathcal{E}^{(i)}(w^{(i)}) = h_i \exp(-\lambda_1 w^{(i)}) + \ddot{h}_i \exp(-\lambda_2 w^{(i)}).$$

This implies that $\mathcal{E}^{(i)}(w^{(i)})$ is almost surely constant as in the proof of Corollary 3.4. In the same way as in the latter proof, it is enough to prove that $\mathcal{E}^{(i)}(w^{(i)})$ admits a finite number of zeros if $(h_i, \ddot{h}_i) \neq (0, 0)$ to conclude. Assume that $(h_i, \ddot{h}_i) \neq (0, 0)$. We have

$$\mathcal{E}^{(i)'}(w^{(i)}) = \exp(-\lambda_1 w^{(i)}) \left(-h_i \lambda_1 - \ddot{h}_i \lambda_2 \exp((\lambda_1 - \lambda_2) w^{(i)}) \right).$$

and $\mathcal{E}^{(i)'}$ changes sign at most once (because $\lambda_1 \neq \lambda_2$). Hence $\mathcal{E}^{(i)}$ admits at most two zeros which establishes the desired result. ■

Just as for GMM based on the moments, \mathbf{S} will be estimated by $\hat{\mathbf{S}}_N$, with

$$\hat{\mathbf{S}}_N = \frac{1}{N} \sum_{n=1}^N \mathbf{f}_{Lap}(\mathbf{W}_n, \boldsymbol{\theta}) \mathbf{f}_{Lap}(\mathbf{W}_n, \boldsymbol{\theta})^T. \quad (37)$$

Remark 3.9. *Notice that less assumptions are required to show the asymptotic properties of GMM based on the Laplace transform than for GMM based on the moments.*

3.3.3 Optimal two-step GMM estimator for an EGP

Finally, we provide the result on the optimal two-step GMM estimator based on the Laplace transform.

Proposition 3.10. Consider \mathbf{W} and \mathbf{f}_{Lap} as defined in (18) and (30), respectively. Suppose that Θ is compact and $\boldsymbol{\theta}_0$ is an interior point of Θ . If $(H_3), (M_{11}), (M_{21}), (M_{31})$ hold and if (H_{11}) is satisfied for $\mathbf{P} = \mathbf{S}^{-1}$, then

$$\hat{\boldsymbol{\theta}}_N^{(2)} \xrightarrow[N \rightarrow \infty]{a.s.} \boldsymbol{\theta}_0.$$

The proof of this proposition is similar to the proof of Proposition 3.5.

4 Study of a specific parametric form for the scale and shape functions

We have discussed how to show the asymptotic properties of GMM estimators for an EGP in a general setting. In this section, a parametric form for the shape and scale functions of an EGP is investigated:

$$A(t, \boldsymbol{\theta}) = \frac{a}{\alpha + 1} t^{\alpha+1}, a(t, \boldsymbol{\theta}) = at^\alpha, b(t, \boldsymbol{\theta}) = b(t + c)^\beta.$$

Throughout we let $\boldsymbol{\theta} = (a, \alpha, b, \beta, c)$, $\mathbb{R}_+ = [0, +\infty)$ and $\mathbb{R}_+^* = (0, +\infty)$. The parameter space is defined as follows:

$$\Theta = P_1 \cup P_2$$

where

$$P_1 = \mathbb{R}_+^* \times (-1, +\infty) \times \mathbb{R}_+^* \times \mathbb{R} \times \mathbb{R}_+^*, \quad (38)$$

$$P_2 = \{(a, \alpha, b, \beta, c) \in \mathbb{R}_+^* \times (-1, +\infty) \times \mathbb{R}_+^* \times \mathbb{R} \times \{0\} \text{ such that } \alpha > 2\beta - 1\}. \quad (39)$$

Remark 4.1. Note that if $\beta = 0$, c can be any value and in this case the EGP is reduced to a SGP. By convention, we take $c = 0$ when $\beta = 0$. The parameter vector is hence denoted by $\boldsymbol{\theta} = (a, b, \alpha, \beta, c)$ for all cases.

4.1 GMM based on the moments

Before showing that the model introduced in Subsection 3.2.1 is identified for this parametric form, it should be noted that

$$\mathbb{E}[\mathbf{f}_{MM}(\mathbf{W}, \boldsymbol{\theta})] = 0 \Leftrightarrow \begin{cases} m^{(i)}(\boldsymbol{\theta}_0) = m^{(i)}(\boldsymbol{\theta}) \\ v^{(i)}(\boldsymbol{\theta}_0) = v^{(i)}(\boldsymbol{\theta}) \end{cases}$$

for $i \in \{1, \dots, d\}$, where \mathbf{W} and \mathbf{f}_{MM} are given in (18) and (19), respectively.

We begin with a technical lemma, that we shall use repeatedly in the sequel. It is easy to prove and consequently stated without proof.

Lemma 4.2. Let $\boldsymbol{\theta}_1 = (a_1, \alpha_1, b_1, \beta_1, c_1)$ and $\boldsymbol{\theta}_2 = (a_2, \alpha_2, b_2, \beta_2, c_2)$ be two parameter vectors in Θ . If

$$r(s) = b_2 (c_2 + s)^{\beta_2} - b_1 (c_1 + s)^{\beta_1}, \quad (40)$$

then r has at most two strict sign changes on \mathbb{R}_+ and at most two zeros unless $(b_1, c_1, \beta_1) = (b_2, c_2, \beta_2)$.

Theorem 4.3. Let $\theta_1 = (a_1, \alpha_1, b_1, \beta_1, c_1)$ and $\theta_2 = (a_2, \alpha_2, b_2, \beta_2, c_2)$ be two parameter vectors in Θ . Assume that

$$m^{(i)}(\theta_1) = m^{(i)}(\theta_2) \quad (41)$$

$$v^{(i)}(\theta_1) = v^{(i)}(\theta_2) \quad (42)$$

for $i \in \{1, \dots, d\}$ and $d \geq 5$.

Then $\theta_1 = \theta_2$.

This shows that assumption (H_3) is satisfied, namely that the model is identifiable.

Proof. Equations (41, 42) yields

$$\frac{a_1}{b_1^j} \int_{t_{i-1}}^{t_i} \frac{s^{\alpha_1}}{(s+c_1)^{j\beta_1}} ds = \frac{a_2}{b_2^j} \int_{t_{i-1}}^{t_i} \frac{s^{\alpha_2}}{(s+c_2)^{j\beta_2}} ds, \quad (43)$$

for $j = 1, 2$, $1 \leq i \leq d$ and $d \geq 5$.

From (43), it is clear that if $(\alpha_1, \beta_1, c_1) = (\alpha_2, \beta_2, c_2)$, then $\frac{a_1}{b_1^j} = \frac{a_2}{b_2^j}$ for $j = 1, 2$ and consequently $a_1 = a_2$ and $b_1 = b_2$. Hence, it is sufficient to prove that $(\alpha_1, \beta_1, c_1) = (\alpha_2, \beta_2, c_2)$.

Set

$$G_j(t) = \int_0^t g_j(s) ds$$

with

$$g_j(s) = \tilde{K}_j \frac{s^{\alpha_1}}{(s+c_1)^{j\beta_1}} - \frac{s^{\alpha_2}}{(s+c_2)^{j\beta_2}}, \quad (44)$$

$$\tilde{K}_j = \frac{a_1 b_2^j}{a_2 b_1^j}, \tilde{K}_j \neq 0$$

for $j = 1, 2$.

By hypothesis, we have

$$G_j(0) = G_j(t_1) = \dots = G_j(t_d) \quad (45)$$

with $d \geq 5$ and $j = 1, 2$. This implies that $g_j = G_j'$ admits at least 5 zeros (by Rolle's theorem).

Now, assume that $\alpha_1 \neq \alpha_2$. Then, we have

$$g_1(s) = 0 \Leftrightarrow \bar{g}(s) = 0$$

where

$$\bar{g}(s) = \tilde{K}_1^\nu s - (s+c_1)^{\mu_1} (s+c_2)^{\mu_2}, \quad (46)$$

with $\nu = \frac{1}{\alpha_1 - \alpha_2}$, $\mu_1 = \frac{\beta_1}{\alpha_1 - \alpha_2}$, $\mu_2 = -\frac{\beta_2}{\alpha_1 - \alpha_2}$.

In addition,

$$\bar{g}''(s) = (s+c_1)^{\mu_1-2} (s+c_2)^{\mu_2-2} \bar{g}(s)$$

with

$$\begin{aligned} \bar{g}(s) &= (\mu_1 + \mu_2)(1 - \mu_1 - \mu_2)s^2 + 2 [-(\mu_1^2 - \mu_1)c_2 - \mu_1\mu_2(c_1 + c_2) - (\mu_2^2 - \mu_2)c_1] s \\ &\quad - (\mu_1c_2 + \mu_2c_1)^2 + \mu_1c_2^2 + \mu_2c_1^2. \end{aligned} \quad (47)$$

Assume that the polynomial $\bar{g}(s)$ is not identically equal to zero. Then, it has at most 2 zeros. This implies that \bar{g}' admits at most three zeros on $[0, T]$ and then \bar{g} (or g_1) admits at most 4 zeros. This is not possible, since g_1 admits at least 5 zeros on $[0, T]$ (by hypothesis). Hence, $\bar{g}(s) = 0$ for all $s \geq 0$, which means that:

$$\begin{cases} (\mu_1 + \mu_2)(1 - \mu_1 - \mu_2) = 0 & (L_1) \\ 2 [-(\mu_1^2 - \mu_1)c_2 - \mu_1\mu_2(c_1 + c_2) - (\mu_2^2 - \mu_2)c_1] = 0 & (L_2) \\ -(\mu_1c_2 + \mu_2c_1)^2 + \mu_1c_2^2 + \mu_2c_1^2 = 0 & (L_3) \end{cases}$$

Based on (L_1) , several cases are to be considered:

- **Case 1:** $\mu_1 = 1 - \mu_2$ (or equivalently: $\alpha_1 - \alpha_2 = \beta_1 - \beta_2$).
Line (L_3) can be rewritten as $\mu_1(1 - \mu_1)(c_1 - c_2)^2 = 0$. Based on the latter equation, we distinguish the following cases:
 - If $\mu_1 = 0$, it follows that $\beta_1 = 0$ (and, consequently, $c_1 = 0$ by Remark 4.1) and $\mu_2 = 1$. Then, we get $\bar{g}(s) = (\tilde{K}_1^\nu - 1)s - c_2$ (see (46)).
 - * If $c_2 \neq 0$, then \bar{g} (or g_1) admits at most one zero. This case is not possible because g_1 admits at least 5 zeros.
 - * If $c_2 = 0$, then $g_2(s) = \tilde{K}_2s^{\alpha_1} - s^{\alpha_2 - 2\beta_2}$ (by using Equation (44) for $j = 2$). Since $\mu_2 = 1$, it follows that $\alpha_1 = \alpha_2 - \beta_2$ and hence $g_2(s) = \tilde{K}_2s^{\alpha_1} - s^{\alpha_1 - \beta_2} = 0 \Leftrightarrow \tilde{K}_2s^{-\beta_2} = 1$. This implies that g_2 admits at most one zero (as $\beta_2 \neq 0$, because $\mu_2 = 1$) which contradicts the hypothesis (g_2 admits at least 5 zeros).
- Hence the case $\mu_1 = 0$ is impossible. The case $\mu_1 = 1$ (namely $\mu_2 = 0$) is similar and not possible either.
- If $c_1 = c_2$, then we obtain $\bar{g}(s) = (\tilde{K}_1^\nu - 1)s - c_1$.
 - * If $c_1 \neq 0$, then g_1 admits at most one zero which is not possible.
 - * If $c_1 = 0 = c_2$, then $g_2(s) = \tilde{K}_2s^{\alpha_1 - 2\beta_1} - s^{\alpha_2 - 2\beta_2} = 0 \Leftrightarrow \tilde{K}_2s^{\beta_2 - \beta_1} = 1$ (since $\alpha_1 - \alpha_2 = \beta_1 - \beta_2$). As $\alpha_1 - \alpha_2 = \beta_1 - \beta_2 \neq 0$ and $\tilde{K}_2 \neq 0$, g_2 admits at most one zero. This case is not possible either.

The case $\mu_1 = 1 - \mu_2$ hence is impossible.

- **Case 2:** $\mu_1 = -\mu_2$ (or equivalently: $\beta_1 = \beta_2$).
In this case line (L_2) can be rewritten as

$$\mu_2(c_1 - c_2) = 0. \quad (48)$$

- If $\mu_2 = 0$, it follows that $\beta_2 = \beta_1 = 0$. Consequently, we have $c_1 = c_2 = 0$ (see Remark 4.1).
- If $\mu_2 \neq 0$, then we get $c_1 = c_2$ by (48).

As a conclusion, under the assumption $\alpha_1 \neq \alpha_2$, we have $\mu_1 = -\mu_2$, namely $\beta_1 = \beta_2$, in which case, we have also proved that $c_1 = c_2$. Hence, $(\beta_1, c_1) = (\beta_2, c_2)$. Substituting β_1 and c_1 by β_1 and c_2 , respectively, into (44) yields $g_1(s) = \tilde{K}_1 s^{\alpha_1 - \alpha_2}$. Hence g_1 admits at most one zero (as $\alpha_1 - \alpha_2 \neq 0$), which contradicts the hypothesis. As a consequence, assuming $\alpha_1 \neq \alpha_2$ leads to a contradiction and we have proved that $\alpha_1 = \alpha_2$.

It remains to show that $(\beta_1, c_1) = (\beta_2, c_2)$. By substituting α_2 by α_1 into (44), we obtain

$$g_1(s) = s^{\alpha_1} \left(\frac{\tilde{K}_1}{(s + c_1)^{\beta_1}} - \frac{1}{(s + c_2)^{\beta_2}} \right) = 0 \Leftrightarrow \frac{\tilde{K}_1}{(s + c_1)^{\beta_1}} = \frac{1}{(s + c_2)^{\beta_2}} \Leftrightarrow \ddot{u}(s) = 0$$

with $\ddot{u}(s) = \tilde{K}_1 (s + c_2)^{\beta_2} - (s + c_1)^{\beta_1}$.

The function \ddot{u} takes a similar form to (40) given in Lemma 4.2. Thus, the only possibility for \ddot{u} (or g_1) to admit more than two zeros is that $\tilde{K}_1 = 1$ and $(\beta_1, c_1) = (\beta_2, c_2)$ (see Lemma 4.2). As a conclusion, $(\alpha_1, \beta_1, c_1) = (\alpha_2, \beta_2, c_2)$ and this completes the proof. ■

Now that the identification condition is checked, we can easily derive the consistency of the GMM estimator.

Corollary 4.4. *Let us consider \mathbf{f}_{MM} as defined in (19) with $d \geq 5$ and a compact set $\bar{\Theta}$ such that $\theta_0 \in \bar{\Theta} \subset \Theta$. Then assumption (H_2) implies*

$$\hat{\theta}_N \xrightarrow{Prob} \theta_0.$$

Proof. We apply Corollary 3.2. The conclusion then follows by hypothesis and Theorem 4.3. ■

Now we define

$$\Theta^{\mathcal{N}} = (\mathbb{R}_+^* \times (-1, +\infty) \times \mathbb{R}_+^* \times \mathbb{R} \times \mathbb{R}_+^*) \\ \cup \{(a, \alpha, b, \beta, c) \in \mathbb{R}_+^* \times (-1, +\infty) \times \mathbb{R}_+^* \times \mathbb{R} \times \{0\} \text{ such that } \alpha > 4\beta - 1\}.$$

Corollary 4.5. *Consider \mathbf{f}_{MM} as defined in (19) with $d \geq 5$. Also, consider a compact parameter space $\bar{\Theta}^{\mathcal{N}} \subset \Theta^{\mathcal{N}}$. If θ_0 is an interior point of $\bar{\Theta}^{\mathcal{N}}$ and (H_2) is satisfied for $\mathbf{P} = \mathbf{S}^{-1}$, then*

$$\sqrt{N} (\hat{\theta}_N - \theta_0) \xrightarrow{\mathcal{D}} \mathcal{N}(0, (\mathbf{D}_0^T \mathbf{S}^{-1} \mathbf{D}_0)^{-1}).$$

Proof. The proof proceeds by verifying the conditions of Corollary 3.3. It suffices to check (M_{kj}) , $k = 2, 3, 4$, $j = 1, 2$, and (H_{11}) . Here, we have

$$\frac{\partial}{\partial \theta} \left(\frac{a(s, \theta)}{b^j(s, \theta)} \right) = \left(\frac{1}{b^j} \frac{s^\alpha}{(s+c)^{j\beta}}, \frac{a \ln(s)}{b^j} \frac{s^\alpha}{(s+c)^{j\beta}}, -\frac{a}{b^j} \frac{s^\alpha}{(s+c)^{j\beta}}, -\frac{a}{b^j} \ln(s+c) \frac{s^\alpha}{(s+c)^{j\beta}}, -\frac{a\beta}{b^j} \frac{s^\alpha}{(s+c)^{j\beta}} \right)$$

for $j = 1, 2$.

It is clear that $\frac{\partial}{\partial \theta} \left(\frac{a(s, \theta)}{b^j(s, \theta)} \right)$ given above satisfies (M_{21}) , (M_{22}) , (M_{31}) and (M_{32}) . Also, since $\alpha > 4\beta - 1$ (by hypothesis), it follows that

$$\int_0^T \frac{a(s, \theta)}{b^k(s, \theta)} ds = \int_0^T \frac{as^\alpha}{b^k(s+c)^k} ds, k = 3, 4. \quad (49)$$

and so conditions (M_{41}) and (M_{42}) are satisfied. It remains to show (H_{11}) .

By Theorem (1.171) in Koch (1999), it suffices to show that $\text{rank}(\mathbf{D}_0) = p = 5$. We proceed to show that the column vectors $v_i, i = 1, \dots, 5$, of \mathbf{D}_0 are linearly independent where

$$\mathbf{D}_0 = \begin{pmatrix} \int_{t_{i-1}}^{t_i} \frac{s^{\alpha_0}}{b_0(s+c_0)^{\beta_0}} ds & \int_{t_{i-1}}^{t_i} \frac{a_0 s^{\alpha_0} \ln(s)}{b_0(s+c_0)^{\beta_0}} ds & - \int_{t_{i-1}}^{t_i} \frac{a_0 s^{\alpha_0}}{b_0^2(s+c_0)^{\beta_0}} ds & - \int_{t_{i-1}}^{t_i} \frac{a_0 s^{\alpha_0} \ln(s+c_0)}{b_0(s+c_0)^{\beta_0}} ds & - \int_{t_{i-1}}^{t_i} \frac{a_0 \beta_0 s^{\alpha_0}}{b_0(s+c_0)^{\beta_0+1}} ds \\ \int_{t_{i-1}}^{t_i} \frac{s^{\alpha_0}}{b_0^2(s+c_0)^{2\beta_0}} ds & \int_{t_{i-1}}^{t_i} \frac{a_0 s^{\alpha_0} \ln(s)}{b_0^2(s+c_0)^{2\beta_0}} ds & - \int_{t_{i-1}}^{t_i} \frac{2a_0 s^{\alpha_0}}{b_0^3(s+c_0)^{2\beta_0}} ds & - \int_{t_{i-1}}^{t_i} \frac{2a_0 s^{\alpha_0} \ln(s+c_0)}{b_0^3(s+c_0)^{2\beta_0}} ds & - \int_{t_{i-1}}^{t_i} \frac{2a_0 \beta_0 s^{\alpha_0}}{b_0^3(s+c_0)^{2\beta_0+1}} ds \end{pmatrix}_{1 \leq i \leq d}$$

(from (22)).

Suppose that p_1, p_2, p_3, p_4, p_5 are real numbers such that

$$p_1 v_1 + p_2 v_2 + p_3 v_3 + p_4 v_4 + p_5 v_5 = 0. \quad (50)$$

Set

$$H(t) = \int_0^t h(s) ds, \quad \bar{H}(t) = \int_0^t \bar{h}(s) ds$$

with

$$h(s) = \frac{s^{\alpha_0}}{b_0(s+c_0)^{\beta_0}} \left[p_1 + p_2 a_0 \ln(s) - p_3 \frac{a_0}{b_0} - p_4 a_0 \ln(s+c_0) - p_5 \frac{a_0 \beta_0}{s+c_0} \right] \quad (51)$$

$$\bar{h}(s) = \frac{s^{\alpha_0}}{b_0^2(s+c_0)^{2\beta_0}} \left[p_1 + p_2 a_0 \ln(s) - 2p_3 \frac{a_0}{b_0} - 2p_4 a_0 \ln(s+c_0) - 2p_5 \frac{a_0 \beta_0}{s+c_0} \right]. \quad (52)$$

Equation (50) means that $H(0) = H(t_1) = \dots = H(t_d)$ and $\bar{H}(0) = \bar{H}(t_1) = \dots = \bar{H}(t_d)$ for $d \geq 5$. Thus, by Rolle's theorem, $h = H'$ and $\bar{h} = \bar{H}'$ admit at least 5 zeros on $[0, T]$. We have

$$h(s) = 0 \Leftrightarrow \dot{v}(s) = 0$$

where $\dot{v}(s) = p_1 + p_2 a_0 \ln(s) - p_3 \frac{a_0}{b_0} - p_4 a_0 \ln(s+c_0) - p_5 \frac{a_0 \beta_0}{s+c_0}$.

Then, we differentiate \dot{v}

$$\dot{v}'(s) = \frac{a_0}{s(s+c_0)^2} [(p_2 - p_4)s^2 + (2p_2 c_0 - p_4 c_0 + p_5 \beta_0)s + p_2 c_0^2].$$

If $p(s) = (p_2 - p_4)s^2 + (2p_2 c_0 - p_4 c_0 + p_5 \beta_0)s + p_2 c_0^2$ is a non identically zero polynomial, then \dot{v}' admits at most two zeros and hence h admits at most 3 zeros which contradicts the hypothesis (h admits at least 4 zeros). Thus, $p(s) = 0$ for all $s \geq 0$, which means that

$$\begin{cases} p_2 - p_4 = 0 \\ 2p_2 c_0 - p_4 c_0 + p_5 \beta_0 = 0 \\ p_2 = 0 \end{cases}$$

Consequently, we get $p_2 = p_4 = p_5 = 0$. It remains to show that $p_1 = p_3 = 0$. By substituting p_2, p_4 and p_5 by their values into (51,52), we get

$$h(s) = 0 \Leftrightarrow p_1 = p_3 \frac{a_0}{b_0}; \quad \bar{h}(s) = 0 \Leftrightarrow p_1 = 2p_3 \frac{a_0}{b_0}.$$

As h and \bar{h} have at least 5 zeros, this implies that $p_1 = p_3 = 0$ and achieves the proof of condition (H_{11}) .

Finally, by Corollary 3.4, we get

$$\sqrt{N} \left(\hat{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_0 \right) \xrightarrow{\mathcal{D}} \mathcal{N}(0, (\mathbf{D}_0^T \mathbf{S}^{-1} \mathbf{D}_0)^{-1}).$$

■

4.2 GMM based on the Laplace transform

We first check the identification condition using this approach. With that aim, we shall use the following technical Lemma, which it is easy to prove and hence stated without proof.

Lemma 4.6. *Let (a_1, α_1) and (a_2, α_2) be two parameter vectors in $\mathbb{R}^* \times \mathbb{R}$. If*

$$h(s) = a_1 s^{\alpha_1} - a_2 s^{\alpha_2}, \quad (53)$$

then h has at most one strict sign change on \mathbb{R}_+ and at most one zero unless $(a_1, \alpha_1) = (a_2, \alpha_2)$.

Here, we can remark that

$$\mathbb{E}[\mathbf{f}_{Lap}(\mathbf{W}, \boldsymbol{\theta})] = 0 \Leftrightarrow \mathcal{L}^{(i)}(\lambda_l, \boldsymbol{\theta}_0) = \mathcal{L}^{(i)}(\lambda_l, \boldsymbol{\theta})$$

for $l \in \{1, \dots, o\}$ and $i \in \{1, \dots, d\}$, where \mathbf{W} and \mathbf{f}_{Lap} are given in (18) and (30), respectively.

We next provide the identifiability result.

Theorem 4.7. *Let $\boldsymbol{\theta}_1 = (a_1, \alpha_1, b_1, \beta_1, c_1)$ and $\boldsymbol{\theta}_2 = (a_2, \alpha_2, b_2, \beta_2, c_2)$ be two parameter vectors in Θ . Assume that*

$$\mathcal{L}^{(i)}(\lambda_l, \boldsymbol{\theta}_1) = \mathcal{L}^{(i)}(\lambda_l, \boldsymbol{\theta}_2) \quad (54)$$

for $l \in \{1, \dots, o\}$, $o \geq 3$ and $i \in \{1, \dots, d\}$, $d \geq 4$.

Then $\boldsymbol{\theta}_1 = \boldsymbol{\theta}_2$.

This means that the model is identifiable (Assumption (H_3)).

Proof. Set

$$G_i(\lambda) = \bar{G}_{t_i}(\lambda) - \bar{G}_{t_{i-1}}(\lambda) = \int_{t_{i-1}}^{t_i} u_s(\lambda) ds$$

where

$$\begin{aligned} \bar{G}_t(\lambda) &= \int_0^t u_s(\lambda) ds, \\ u_s(\lambda) &= a_1 s^{\alpha_1} g_{1,s}(\lambda) - a_2 s^{\alpha_2} g_{2,s}(\lambda), \\ g_{j,s}(\lambda) &= \ln \left(1 + \frac{\lambda}{b_j (c_j + s)^{\beta_j}} \right), j = 1, 2. \end{aligned} \quad (55)$$

Equation (54) implies that $G_i(\lambda_l) = 0$ for $l \in \{1, \dots, o\}$, $o \geq 3$ and $i \in \{1, \dots, d\}$, $d \geq 4$. The derivatives of $g_{j,s}(\lambda)$ with respect to λ are given by

$$g'_{j,s}(\lambda) = \frac{1}{\lambda + b_j(c_j + s)^{\beta_j}}, \quad g''_{j,s}(\lambda) = -\frac{1}{\left(\lambda + b_j(c_j + s)^{\beta_j}\right)^2}.$$

In addition, we have $\left|s^{\alpha_j} g'_{j,s}(\lambda)\right| \leq \frac{s^{\alpha_j}}{b_j(c_j+s)^{\beta_j}}$ and $\left|s^{\alpha_j} g''_{j,s}(\lambda)\right| \leq \frac{s^{\alpha_j}}{b_j^2(c_j+s)^{2\beta_j}}$ with $\int_0^T \frac{s^{\alpha_j}}{b_j(c_j+s)^{\beta_j}} ds < \infty$ and $\int_0^T \frac{s^{\alpha_j}}{b_j^2(c_j+s)^{2\beta_j}} ds < \infty$ (from (14)). Then, based on the differentiation under the integral sign theorem, the derivatives of $G_i(\lambda)$ with respect to λ are given by

$$G'_i(\lambda) = \int_{t_{i-1}}^{t_i} (a_1 s^{\alpha_1} g'_{1,s}(\lambda) - a_2 s^{\alpha_2} g'_{2,s}(\lambda)) ds, \quad (56)$$

$$G''_i(\lambda) = \int_{t_{i-1}}^{t_i} (a_1 s^{\alpha_1} g''_{1,s}(\lambda) - a_2 s^{\alpha_2} g''_{2,s}(\lambda)) ds. \quad (57)$$

One can point out that

$$g_{1,s}(\lambda) < g_{2,s}(\lambda) \text{ if and only if } r(s) < 0 \quad (58)$$

$$g''_{1,s}(\lambda) < g''_{2,s}(\lambda) \text{ if and only if } r(s) > 0 \quad (59)$$

where $r(s)$ is given by (40) in Lemma 4.2.

The point is to show that $(a_1, \alpha_1) = (a_2, \alpha_2)$ and $(b_1, \beta_1, c_1) = (b_2, \beta_2, c_2)$. With that aim, we show that $(a_1, \alpha_1) \neq (a_2, \alpha_2)$ or $(b_1, \beta_1, c_1) \neq (b_2, \beta_2, c_2)$ is not possible. The proof is split into three different cases.

- Suppose first that $(a_1, \alpha_1) \neq (a_2, \alpha_2)$ and $(b_1, \beta_1, c_1) = (b_2, \beta_2, c_2)$. Then

$$u_s(\lambda) = (a_1 s^{\alpha_1} - a_2 s^{\alpha_2}) g_{1,s}(\lambda) = h(s) g_{1,s}(\lambda).$$

By Lemma 4.6, the function h changes sign at most one time on \mathbb{R}_+ . This implies that there exist at least three intervals among the $[t_{i-1}, t_i]$ on which h (and consequently $s \mapsto u_s(\lambda)$) keeps a constant sign (see Figure 1). We now work on one of these intervals, say $[t_{i_0-1}, t_{i_0}]$.

Let $l \in \{1, \dots, o\}$ be fixed. Since $G_{i_0}(\lambda_l) = \int_{t_{i_0-1}}^{t_{i_0}} u_s(\lambda_l) ds = 0$ (by hypothesis), it follows that

$$u_s(\lambda_l) = h(s) g_{1,s}(\lambda_l) = 0, \text{ for all } s \in [t_{i_0-1}, t_{i_0}], \text{ for each } \lambda_l. \quad (60)$$

As $g_{1,s}(\lambda_l) > 0$, we derive that $h(s) = 0$ for all $s \in [t_{i_0-1}, t_{i_0}]$ and hence $(a_1, \alpha_1) = (a_2, \alpha_2)$ (by Lemma 4.6) which is not possible.

- Assume now that $(a_1, \alpha_1) = (a_2, \alpha_2)$ and $(b_1, \beta_1, c_1) \neq (b_2, \beta_2, c_2)$. Then

$$u_s(\lambda) = a_1 s^{\alpha_1} (g_{1,s}(\lambda) - g_{2,s}(\lambda))$$

with sign provided by $r(s)$, see (58).

By Lemma 4.2, the function r changes sign at most two times on \mathbb{R}_+ . This implies that there exist at least two intervals (say $[t_{i_0-1}, t_{i_0}]$ and $[t_{i_1-1}, t_{i_1}]$) on which r keeps a constant sign

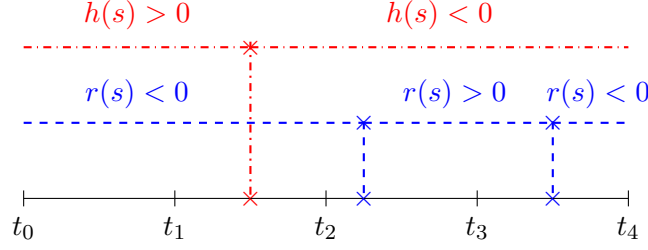


Figure 1: Schematic illustration of the sign change for h and r

(see Figure 1). It follows that $s \mapsto u_s(\lambda)$ keeps a constant sign on $[t_{i_0-1}, t_{i_0}]$ (independently of λ , see (58)) and based on $G_{i_0}(\lambda_l) = 0$, we get

$$u_s(\lambda_l) = 0, \text{ for all } s \in [t_{i_0-1}, t_{i_0}], \text{ for each } \lambda_l. \quad (61)$$

Hence $g_{1,s}(\lambda_l) = g_{2,s}(\lambda_l)$ for all $s \in (t_{i_0-1}, t_{i_0}]$, so that $r(s) = 0$ for all $s \in (t_{i_0-1}, t_{i_0}]$ and $(b_1, \beta_1, c_1) = (b_2, \beta_2, c_2)$ (see Lemma 4.2). This case is not possible either.

- Finally we suppose that $(a_1, \alpha_1) \neq (a_2, \alpha_2)$ and $(b_1, \beta_1, c_1) \neq (b_2, \beta_2, c_2)$. The function $h(s) = a_1 s^{\alpha_1} - a_2 s^{\alpha_2}$ changes sign at most one time on \mathbb{R}_+ and $r(s) = b_2 (c_2 + s)^{\beta_2} - b_1 (c_1 + s)^{\beta_1}$ changes sign at most two times on \mathbb{R}_+ (by Lemmas 4.2 and 4.6). Then, there exist at least two intervals $[t_{i-1}, t_i]$ on which r keeps a constant sign and at least three intervals on which h keeps a constant sign. Hence there is at least one common interval $[t_{i_0-1}, t_{i_0}]$ on which the functions r and h do not change sign (see Figure 1). We now work on this fixed interval $[t_{i_0-1}, t_{i_0}]$ on which both r and h keep a constant sign and we distinguish two cases accordingly whether the sign is the same or not.

- **Case 1:** Assume that $r(s)$ and $h(s)$ have the same sign on $[t_{i_0-1}, t_{i_0}]$. It follows, by (58), that $g_{1,s}(\lambda) - g_{2,s}(\lambda)$ and $h(s)$ have the same sign. Then, for each λ , the function $s \mapsto u_s(\lambda)$ (given by (55)) keeps a constant sign on $[t_{i_0-1}, t_{i_0}]$. Let $l \in \{1, \dots, o\}$ be fixed.

Since $G_{i_0}(\lambda_l) = \int_{t_{i_0-1}}^{t_{i_0}} u_s(\lambda_l) ds = 0$, we derive that

$$u_s(\lambda_l) = 0 \text{ for all } s \in [t_{i_0-1}, t_{i_0}], \text{ for each } l. \quad (62)$$

Let $s \in [t_{i_0-1}, t_{i_0}]$ be fixed. We have

$$u'_s(\lambda) = \frac{(a_1 s^{\alpha_1} - a_2 s^{\alpha_2}) \lambda + a_1 s^{\alpha_1} b_2 (c_2 + s)^{\beta_2} - a_2 s^{\alpha_2} b_1 (c_1 + s)^{\beta_1}}{(\lambda + b_1 (c_1 + s)^{\beta_1}) (\lambda + b_2 (c_2 + s)^{\beta_2})}.$$

If $a_1 s^{\alpha_1} - a_2 s^{\alpha_2} \neq 0$,

$$u'_s(\lambda) = 0 \Leftrightarrow \lambda = \frac{a_1 s^{\alpha_1} b_2 (c_2 + s)^{\beta_2} - a_2 s^{\alpha_2} b_1 (c_1 + s)^{\beta_1}}{a_1 s^{\alpha_1} - a_2 s^{\alpha_2}}.$$

This implies that $u'_s(\lambda)$ admits at most one zero. It follows that $u_s(\lambda)$ admits at most two zeros, which is false based on (62). Hence,

$$a_1 s^{\alpha_1} = a_2 s^{\alpha_2} \quad (63)$$

for all $s \in [t_{i_0-1}, t_{i_0}]$. Based on (62) again, we also get $g_{1,s}(\lambda) = g_{2,s}(\lambda)$, because $u_s(\lambda) = 0$ and $a_1 s^{\alpha_1} = a_2 s^{\alpha_2}$. This yields

$$b_1 (c_1 + s)^{\beta_1} = b_2 (c_2 + s)^{\beta_2} \quad (64)$$

for all $s \in [t_{i_0-1}, t_{i_0}]$. Finally, by Lemmas 4.2 and 4.6, Equation (63) admits at most one zero unless we have $(a_1, \alpha_1) = (a_2, \alpha_2)$ and Equation (64) admits at most two zeros unless $(b_1, c_1, \beta_1) = (b_2, c_2, \beta_2)$. We consequently have $(a_1, \alpha_1) = (a_2, \alpha_2)$ and $(b_1, \beta_1, c_1) = (b_2, \beta_2, c_2)$, namely $\boldsymbol{\theta}_1 = \boldsymbol{\theta}_2$.

- **Case 2:** Assume that $r(s)$ and $h(s)$ have different signs on $[t_{i_0-1}, t_{i_0}]$, say $-$ and $+$, respectively. It follows that $g''_{1,s}(\lambda) - g''_{2,s}(\lambda) > 0$ for each λ (see (59)). Hence $G''_{i_0}(\lambda) \geq 0$ (see (57)) and G_{i_0} is convex. If there does not exist any open interval $\mathcal{I} \subset \mathbb{R}_+$ such that $G_{i_0}(\lambda) = 0$ for all $\lambda \in \mathcal{I}$, then G_{i_0} admits at most two zeros, which contradicts the hypothesis. Accordingly, there exists an open interval $\mathcal{I} \subset \mathbb{R}_+$ such that $G_{i_0}(\lambda) = 0$ for all $\lambda \in \mathcal{I}$. This yields

$$G''_{i_0}(\lambda) = \int_{t_{i_0-1}}^{t_{i_0}} v_s(\lambda) = 0, \text{ for all } \lambda \in \mathcal{I}, \quad (65)$$

where

$$v_s(\lambda) = \frac{a_1 s^{\alpha_1}}{\left(\lambda + b_1 (c_1 + s)^{\beta_1}\right)^2} - \frac{a_2 s^{\alpha_2}}{\left(\lambda + b_2 (c_2 + s)^{\beta_2}\right)^2}.$$

Since $h(s) > 0$ and $r(s) < 0$ (or $g''_{1,s}(\lambda) - g''_{2,s}(\lambda) > 0$), then $s \mapsto v_s(\lambda)$ keeps a constant sign on $[t_{i_0-1}, t_{i_0}]$ and hence $v_s(\lambda) = 0$ for all $s \in [t_{i_0-1}, t_{i_0}]$ and $\lambda \in \mathcal{I}$ (by (65)).

For each $s \in [t_{i_0-1}, t_{i_0}]$, a necessary condition for $\lambda \mapsto v_s(\lambda)$ to admit more than two zeros is that $a_1 s^{\alpha_1} = a_2 s^{\alpha_2}$. We consequently have $a_1 s^{\alpha_1} = a_2 s^{\alpha_2}$ for all $s \in [t_{i_0-1}, t_{i_0}]$ and then we get $b_1 (c_1 + s)^{\beta_1} = b_2 (c_2 + s)^{\beta_2}$ for all $s \in [t_{i_0-1}, t_{i_0}]$ (since $v_s(\lambda) = 0$), which allows to conclude using the same arguments as in the previous case. ■

Corollary 4.8. *Let us consider \mathbf{f}_{Lap} as defined in (30) with $d \geq 4$, $o \geq 3$ and $\boldsymbol{\theta}_0 \in \bar{\Theta} \subset \Theta$ where $\bar{\Theta}$ is a compact set. If (H_2) holds, then $\hat{\boldsymbol{\theta}}_N \xrightarrow{\text{Prob}} \boldsymbol{\theta}_0$.*

Proof. We apply Corollary 3.6. The conclusion then follows by hypothesis and Theorem 4.7. ■

Corollary 4.9. *Consider \mathbf{f}_{Lap} as defined in (30), with $d \geq 4$, $o \geq 3$ and a compact parameter space $\bar{\Theta} \subset \Theta$. If $\boldsymbol{\theta}_0$ is an interior point of $\bar{\Theta}$ and (H_2) is satisfied for $\mathbf{P} = \mathbf{S}^{-1}$. Then*

$$\sqrt{N} \left(\hat{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_0 \right) \xrightarrow{\mathcal{D}} \mathcal{N}(0, (\mathbf{D}_0^T \mathbf{S}^{-1} \mathbf{D}_0)^{-1}).$$

Proof. The proof of this result follows by Corollary 3.7, Theorem 4.7 and the proof of Corollary 4.5. ■

Remark 4.10. *As already told in Remark 3.9 for asymptotic properties in a general setting, the identifiability requires less assumptions for the GMM based on the Laplace transform than on the moments (e.g., observation of at least 4 increments per trajectory instead of 5).*

4.3 Test of a SGP

A standard gamma process is a particular case of an extended gamma process: the one for which the scale function is constant. By considering the parametric form studied in this section for the scale function ($b(t, \theta) = b(t + c)^\beta$), an EGP reduces to a SGP if and only if the value of the parameter β is equal to 0. Thanks to the asymptotic normality results obtained in the corollaries 4.5 and 4.9, it is then possible to construct a statistical test for the nullity of the coefficient β :

$$H_0 : \beta = 0 \quad \text{against} \quad H_1 : \beta \neq 0.$$

More precisely, the corollaries 4.5 and 4.9 leads to:

$$\sqrt{N} \left(\hat{\beta}_N - \beta_0 \right) \xrightarrow{\mathcal{D}} \mathcal{N}(0, V(4, 4)),$$

where $V(4, 4)$ is the $(4, 4)$ element of the matrix \mathbf{V} defined in (10). When the hypothesis H_0 is true, it follows:

$$N \hat{\beta}_N^2 V(4, 4)^{-1} \xrightarrow{\mathcal{D}} \chi_1^2.$$

We then propose the following test: Let $\xi \in (0, 1)$. The hypothesis H_0 is rejected (i.e. \mathbf{X} is not a SGP) as soon as:

$$\hat{J}_N := N \hat{\beta}_N^2 \hat{\mathbf{V}}_N(4, 4)^{-1} > \chi_{1, 1-\xi}^2, \quad (66)$$

where $\chi_{1, 1-\xi}^2$ is the $1 - \xi$ -quantile of a chi-squared distribution with 1 degree of freedom and $\hat{\mathbf{V}}_N$ denotes the estimator of the asymptotic variance \mathbf{V} and is defined by:

$$\hat{\mathbf{V}}_N = \left(\hat{\mathbf{D}}_N^T \left(\hat{\mathbf{S}}_N \right)^{-1} \hat{\mathbf{D}}_N \right)^{-1} \quad (67)$$

with $\hat{\mathbf{D}}_N = \frac{1}{N} \sum_{n=1}^N \frac{\partial \mathbf{f}}{\partial \boldsymbol{\theta}} \left(\mathbf{w}_n, \hat{\boldsymbol{\theta}}_N \right)$ and $\hat{\mathbf{S}}_N = \frac{1}{N} \sum_{n=1}^N \mathbf{f} \left(\mathbf{w}_n, \hat{\boldsymbol{\theta}}_N \right) \mathbf{f} \left(\mathbf{w}_n, \hat{\boldsymbol{\theta}}_N \right)^T$.

The previous test is approximately (when N is large) of level ξ .

5 Numerical experiments

In this section, the performances of the proposed estimation approaches are investigated on simulated data and also on real data. All numerical experimentations are made with MATLAB[®] and the "fminsearch" function is used to estimate the unknown parameters. This function requires initial parameters, which are obtained through a standard least square method (not provided here). Please refer to Al Masry (2016) for further information on this method.

5.1 Numerical comparison between GMM_{MM} and GMM_{Lap}

We first test GMM based on the moments (GMM_{MM}) and Laplace transform (GMM_{Lap}) on simulated data. N independent approximate sample paths are generated on $[0, T]$ through the Rejection method (see Al Masry et al. (2015)). As previously mentioned in Subsection 2.3, a two-step procedure is used to estimate $\boldsymbol{\theta}_0$. An initial estimator is first computed by taking $\mathbf{P} = \mathbf{I}$. Then, using this estimator, we construct $\hat{\mathbf{S}}_N^{(1)}$ and compute $\hat{\boldsymbol{\theta}}_N^{(2)}$ using (12) and (13), respectively. In

order to study the behavior of the estimator, we independently simulate R sets of N independent approximate sample paths. This provides R estimations $\hat{\boldsymbol{\theta}}_{N,i}^{(2)}$ for $1 \leq i \leq R$.

Example 1:

We first consider the specific case where $A(t, \boldsymbol{\theta}) = t^\alpha$ and $b(t, \boldsymbol{\theta}) = t^\beta$, and we take $T = 10$ and $R = 200$. The mean and standard deviation of $\hat{\boldsymbol{\theta}}_{N,i}^{(2)}$ are summarized in Table 1 for different values of (N, d) , where we recall that d is the number of observations per trajectory. It can be observed in this table that GMM_{Lap} provides good results, which are clearly better than the ones based on GMM_{MM} .

Table 1: Mean and standard deviation for $A(t) = t^\alpha$, $b(t) = t^\beta$, $T = 10$, $R = 200$. True values of the parameters are $\alpha = 2$, $\beta = 0.75$

(N, d)	(20, 15)	(30, 20)	(50, 20)	(75, 20)	(100, 20)
	Mean (std)	Mean (std)	Mean (std)	Mean (std)	Mean (std)
$\hat{\alpha}$ (GMM_{MM})	1.9499 (0.1965)	1.9579 (0.1795)	1.9673 (0.1337)	1.9686 (0.0938)	1.9707 (0.0810)
$\hat{\alpha}$ (GMM_{Lap})	2.0037 (0.1670)	2.0038 (0.1215)	2.0058 (0.0858)	2.0055 (0.0760)	2.0038 (0.0596)
$\hat{\beta}$ (GMM_{MM})	0.6861 (0.2489)	0.6978 (0.2263)	0.7103 (0.1691)	0.7113 (0.1188)	0.7137 (0.1014)
$\hat{\beta}$ (GMM_{Lap})	0.7521 (0.2200)	0.7534 (0.1611)	0.7572 (0.1135)	0.7566 (0.1006)	0.7546 (0.0782)

Another indicator of interest is the coverage probability, which can be numerically assessed through

$$CP = \frac{\#\{i : \boldsymbol{\theta}_0 \in \mathcal{CR}_{N,i}\}}{R}, \quad (68)$$

for $1 \leq i \leq R$, where

$$\mathcal{CR}_{N,i} = \{\boldsymbol{\theta} : N(\hat{\boldsymbol{\theta}}_{N,i}^{(2)} - \boldsymbol{\theta})\hat{\mathbf{V}}_N^{-1}(\hat{\boldsymbol{\theta}}_{N,i}^{(2)} - \boldsymbol{\theta})^T \leq \chi_{p,1-\xi}^2\} \quad (69)$$

is a confidence region obtained from the asymptotic normality of the GMM estimators (Corollaries 4.5 and 4.9) and which is approximatively (with N large) of level ξ . Here, the level ξ is fixed at 0.05. The CPs for the cases from Table 1 are displayed in Table 2. It can be seen that the CPs are much lower than the nominal value 95%, so that the asymptotic normality is far from being reached.

Table 2: Coverage probabilities for $A(t) = t^2$, $b(t) = t^{0.75}$, $T = 10$, $R = 200$.

(N, d)	(20, 15)	(30, 20)	(50, 20)	(75, 20)	(100, 20)
	CP	CP	CP		
GMM_{MM}	44%	46.5%	66.33%	75%	80.5%
GMM_{Lap}	43.5%	53%	78.57%	81.5%	83.5%

Keeping the same example as in Table 1, the number of increments is now fixed to $d = 10$ and the number of observations increased to $N = 200$. As can be seen in Table 3, both GMM_{MM}

and GMM_{Lap} provide good results. The CP s are now approximately equal to 95% for both approaches (94% for GMM_{MM} and 93.5% for GMM_{Lap}) so that as expected, enlarging the number of observations allows to approach the asymptotic normality.

Table 3: Mean, standard deviation and quantiles for $A(t) = t^\alpha$, $b(t) = t^\beta$, $d = 10$, $T = 10$, $N = 200$, $R = 200$

	$\hat{\alpha}$	$\hat{\beta}$
True value	2	0.75
	Mean (std)	Mean (std)
GMM_{MM}	1.9903 (0.0527)	0.7376 (0.0665)
GMM_{Lap}	1.9985 (0.0461)	0.7475 (0.0597)
	$[Q_{.025}, Q_{.975}]$	$[Q_{.025}, Q_{.975}]$
GMM_{MM}	[1.8803, 2.0840]	[0.5990, 0.8549]
GMM_{Lap}	[1.9145, 2.0744]	[0.6377, 0.8447]

Example 2: We next consider $A(t, \theta) = at^\alpha$, $b(t, \theta) = bt^\beta$ and compute the four estimates using both GMM approaches. In Table 4, we display the estimates for different values of (N, d) . Here again, we can observe that GMM_{Lap} behaves well for a few number of trajectories, and it is more performing than GMM_{MM} . For example, the mean bias for GMM_{Lap} for the case $(50, 25)$ is lower than for the GMM_{MM} .

Table 4: Mean and standard deviation for $A(t) = at^\alpha$, $b(t) = bt^\beta$, $T = 10$, $R = 200$. True values of the parameters are $a = 1$, $\alpha = 2$, $b=1$, $\beta = 0.5$

(N, d)	(50, 25)	(100, 50)
	Mean (std)	Mean (std)
\hat{a} (GMM_{MM})	1.1892 (0.2303)	1.1099 (0.1806)
\hat{a} (GMM_{Lap})	1.0216 (0.1118)	1.0059 (0.0499)
$\hat{\alpha}$ (GMM_{MM})	1.9571 (0.1128)	1.9748 (0.0972)
$\hat{\alpha}$ (GMM_{Lap})	2.0094 (0.1118)	2.0046 (0.0458)
\hat{b} (GMM_{MM})	1.1011 (0.1599)	1.0633 (0.1362)
\hat{b} (GMM_{Lap})	1.0382 (0.1453)	1.0171 (0.0839)
$\hat{\beta}$ (GMM_{MM})	0.4753 (0.1165)	0.4842 (0.1018)
$\hat{\beta}$ (GMM_{Lap})	0.5034 (0.1271)	0.5017 (0.0597)

Then, we generate $R = 500$ sets of $N = 500$ approximate samples paths and present the results in Table 5 (in order to test the asymptotic normality). From Table 5, it is clear that GMM_{Lap} shows better performance against GMM_{MM} . In addition, in Table 6, we see that the CP for GMM_{Lap}

becomes closer to 95% by increasing N while for GMM_{MM} it tends to be less than the nominal value.

Table 5: Mean, standard deviation and quantiles for $A(t) = at^\alpha$, $b(t) = bt^\beta$, $d = 10$, $T = 10$, $N = 500$, $R = 500$

	\hat{a}	$\hat{\alpha}$	\hat{b}	$\hat{\beta}$
True value	1	2	1	0.5
	Mean (std)	Mean (std)	Mean (std)	Mean (std)
GMM_{MM}	1.0402 (0.0635)	1.9897 (0.0282)	1.0398 (0.0538)	0.4873 (0.0299)
GMM_{Lap}	1.0053 (0.0501)	2.0030 (0.0247)	1.0083 (0.0440)	0.5023 (0.0270)
	$[Q_{.025}, Q_{.975}]$	$[Q_{.025}, Q_{.975}]$	$[Q_{.025}, Q_{.975}]$	$[Q_{.025}, Q_{.975}]$
GMM_{MM}	[0.9138, 1.1725]	[1.9369, 2.0478]	[0.9387, 1.1457]	[0.4286, 0.5487]
GMM_{Lap}	[0.9159, 1.1081]	[1.9534, 2.0513]	[0.9301, 1.1025]	[0.4490, 0.5582]

Table 6: Coverage probability of hyper-ellipsoid confidence regions for $A(t) = at^\alpha$, $b(t) = bt^\beta$, $d = 10$, $T = 10$, $R = 500$

	CP (N=500)	CP (N=800)
GMM_{MM}	84.8%	87.6%
GMM_{Lap}	89.8%	93.4%

5.2 Test of a SGP

In this subsection, the "Wald-type" test presented in Section 4.3 is applied on simulated data. To this end, the following parametric form is considered: $A(t) = t^{1.5}$ and $b(t) = t^\beta$. From N ($= 100$ or 500) generated trajectories with $d = 20$ increments, an estimate $\hat{\beta}$ is computed by GMM_{Lap} (and also by GMM_{MM}) and the test for the nullity of the parameter β is then applied with $\xi = 0.05$. The previous procedure is repeated $R = 200$ times.

In Table 7 are reported the proportions of rejection of the $H_0 : \beta = 0$ hypothesis after the R repetitions. These proportions correspond to the observed type I error of the test for $\beta = 0$ and to the observed power of the test for $\beta \neq 0$. We can see that for GMM_{Lap} and $N = 500$ the observed level of the test is close to the theoretical value ξ , which is not the case for GMM_{MM} . Moreover, we remark that the power of the test for GMM_{MM} and GMM_{Lap} is relatively high for $\beta = 0.1$ and $\beta = 0.2$, which indicates that the proposed test can be used in order to test the nullity of the parameter β , that is to say in order to test if the EGP adjusted to a given set of data is significantly different from a SGP.

Table 7: Test of a SGP with $A(t) = t^{1.5}$, $b(t) = t^\beta$, $d = 20$, $T = 10$, $R = 200$

		$\beta = 0$	$\beta = 0.1$	$\beta = 0.2$
$N = 100$	GMM_{MM}	0.355	0.63	0.97
	GMM_{Lap}	0.355	0.64	0.985
$N = 500$	GMM_{MM}	0.145	0.96	1
	GMM_{Lap}	0.055	1	1

5.3 Application to real data

We finally consider the famous Hudak crack growth data from the literature (see Hudak et al. (1978)), which are based on the observation of 21 metallic specimens subject to degradation. The data are displayed in Figure 2. For estimation and testing purpose, we only consider the 20 trajectories for which 10 increments are observed in the following. The point here is to test whether SGPs are as well adapted as EGPs for matching the data. The shape function and the scale parameter of a SGP are first estimated using the semi-parametric estimation procedure developed by Ye et al. (2013). The parameters of an EGP are next estimated using GMM based on the Laplace transform and the following parametric forms $A(t) = at^\alpha$, $b(t) = bt^\beta$. In the present case, the least square method previously proposed for the initialization of the optimization procedure does not provide correct results. To overcome this problem, we have plotted the empirical mean and variance, and chosen initial parameters which provide "similar" curves. This gives us $(a, \alpha, b, \beta) = (14, 0.5, 4, -0.9)$ as a (possible) candidate for the initialization. Starting from this point, we obtain the following estimates:

$$(\hat{a}, \hat{\alpha}, \hat{b}, \hat{\beta}) = (22.6787, 0.3283, 3.0958, -1.1192).$$

For comparison purpose, the mean, variance and variance-to-mean ratio are next plotted in Figure 3 for the data as well as for the estimated EGP and SGP. As can be seen, EGP provides closer results to the data than SGP. In particular, conversely to the data, the SGP mean starts from 0. This can be due to the fact that the SGP mean and variance are necessary proportional and that the variance starts from 0. As for the EGP, its mean starts from a similar non zero point as for the data.

Finally, we apply the Wald-type test of Section 4.3 with $A(t) = at^\alpha$, $b(t) = bt^\beta$. The null assumption $\beta = 0$ (SGP) is here clearly rejected with $\hat{J}_N \approx 2.17 \times 10^4 > \chi_{1,1-\xi}^2 \approx 3.8415$.

Either considering a non-parametric shape function for the SGP or a similar parametric shape function as for the EGP, there hence are arguments for thinking that SGPs are less adapted than EGPs to model the Hudak data.

Figure 2: Hudak crack growth data

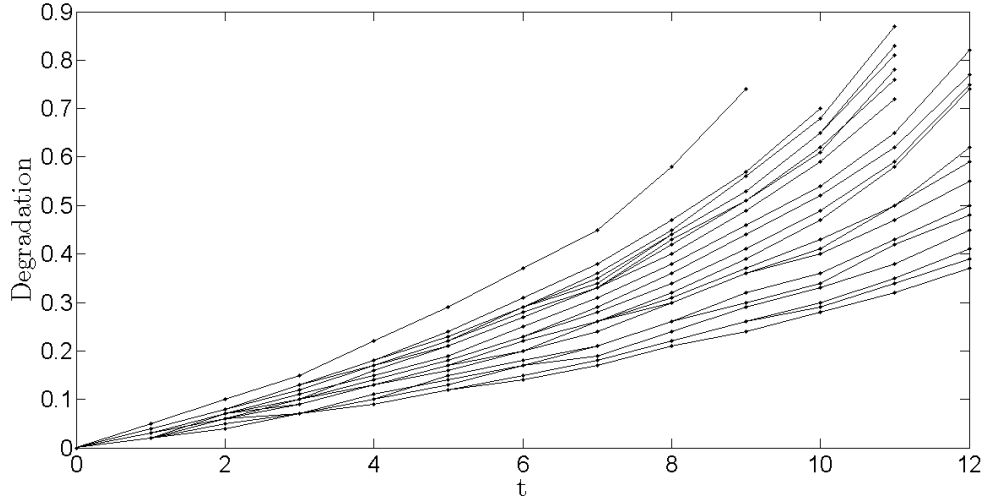
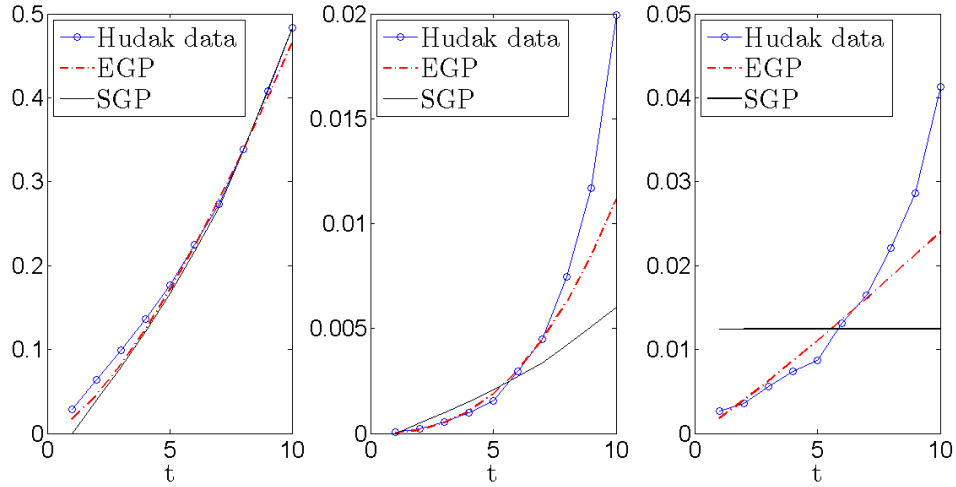


Figure 3: Mean, variance and variance-to-mean ratio for Hudak crack growth data, EGP and SGP



6 Conclusion

In this work, we have dealt with parametric estimation techniques to estimate the shape and scale functions of an EGP. Two approaches dedicated for an EGP have been proposed: GMM based on the moments and GMM based on the Laplace transform. Technical assumptions under which GMM estimators for an EGP are consistent and asymptotically normal have been provided. As a by-product, a Wald-type test has been proposed, which tests a SGP among EGPs with a specific parametric shape function.

It should be noted that, as often, the identifiability was the hardest condition to verify, which

has required the development of original techniques. These techniques can be used to check the identifiability of other parametric forms (see e.g. Al Masry (2016)). As for the practical effectiveness, it has been seen here and in other numerical experiments that GMM based on the Laplace transform is generally more performing than GMM based on the moments.

Finally, the Laplace-based GMM procedure has been applied to a real data set from the literature (the Hudak data). It has been seen that EGPs seem to be better adapted than SGPs to model this data set. This clearly motivates the study of EGPs for modeling purpose.

In regard to inspections, the proposed methods are applicable when the observation times are common to all trajectories. This is not always the case in an industrial context. One can then think about an approximate maximum likelihood method, based on the approximated pdf of an EGP proposed in Al Masry et al. (2015). This method requires a discretization of the scale function of an EGP and a numerical computation of infinite integral. It has been tested on a few simulated data sets. The results were correct but the computational times were rather long. Also, the theoretical basis of such a method remains to be developed. Indeed, no convergence result is provided in Al Masry et al. (2015) for the approximate pdf towards the exact pdf (only for the cdf). Hence, there is a large gap to be filled, before being able to develop asymptotic properties of the resulting estimator (convergence of the approximate pdf plus control of the error made through the approximation).

Acknowledgements

This work is part of Zeina Al Masry's PhD studies which was supported by a French Ministry grant.

References

- M. Abdel-Hameed. A gamma wear process. *IEEE Trans. Rel.*, 24(2):152–153, 1975.
- Z. Al Masry. *Extended gamma processes in view of application to reliability (English)*. PhD thesis, École doctorale des Sciences exactes et leurs applications (ED 211), Université de Pau et des Pays de l'Adour, 2016.
- Z. Al Masry, S. Mercier, and G. Verdier. Approximate simulation techniques and distribution of an extended gamma process. *Methodol. Comput. Appl. Probab.*, available online, 2015.
- R. E. Barlow and F. Proschan. *Mathematical theory of reliability*, volume 17 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1965. With contributions by Larry C. Hunter, 1996.
- M. Carrasco and J. Florens. Efficient GMM estimation using the empirical characteristic function. *Working Paper, Department of Economics. University of Rochester*, 2002.
- N. H. Chan, S. X. Chen, L. Peng, and C. L. Yu. Empirical likelihood methods based on characteristic functions with applications to Lévy processes. *J. Am. Stat. Assoc.*, 104(488):1621–1630, 2009.
- E. Çinlar. On a generalization of gamma processes. *J. Appl. Probab.*, 17:467–480, 1980.
- E. Çinlar, Z. P. Bažant, and E. Osman. Stochastic process for extrapolating concrete creep. *J. of the Engrg. Mech. Div., ASCE*, 103(EM6):1069–1088, 1977.

- R. L. Dykstra and P. Laud. A bayesian nonparametric approach to reliability. *Ann. Stat.*, 9(2): 356–367, 1981.
- A. Feuerverger and P. McDunnough. On some Fourier methods for inference. *J. Am. Stat. Assoc.*, 76(374):379–387, 1981.
- M. Guida, F. Postiglione, and G. Pulcini. A time-discrete extended gamma process for time-dependent degradation phenomena. *Reliab. Eng. Syst. Safe.*, 105:73–79, 2012.
- A. R. Hall. *Generalized method of moments*. Oxford University Press, Oxford, UK, 2005.
- L. P. Hansen. Large sample properties of generalized method of moments estimators. *Econometrica*, 50(4):1029–1054, 1982.
- S.J. Hudak, A. Saxena, R.J. Bucci, and R.C. Malcolm. Development of standard methods of testing and analyzing fatigue crack growth rate data. Technical report AFML-TR-78-40, Westinghouse R&D Center, Westinghouse Electric Corporation. Pittsburgh, PA, 1978.
- R. I. Jennerich. Asymptotic properties of non-linear least squares estimators. *Ann. Math. Stat.*, 40(2):633–643, 1969.
- K.-R. Koch. *Parameter estimation and hypothesis testing in linear models*. Springer, 2 edition, 1999.
- N. Kunitomo and T. Owada. Empirical likelihood estimation of Lévy processes. *Graduate School of Economics, University of Tokyo Discussion Paper*, CIRJE-F-272, 2004.
- W. K. Newey and D. L. McFadden. *Handbook of econometrics*, volume 4, chapter Large sample estimation and hypothesis testing, pages 2113–2245. Elsevier Science Publishers, Amsterdam, The Netherlands, 1994.
- J. Qin and J. Lawless. Empirical likelihood and general estimating equations. *Ann. Stat.*, 22: 300–325, 1994.
- N. Sueishi and Y. Nishiyama. Estimation of Lévy processes in mathematical finance: A comparative study. *MSSANZ*, pages 953–959, 2005.
- J. M. van Noortwijk. A survey of the application of gamma processes in maintenance. *Reliab. Eng. Syst. Safe.*, 94:2–21, 2009.
- Z. S. Ye, M. Xie, L. C. Tang, and N. Chen. Efficient semiparametric estimation of gamma processes for deteriorating products. *Technometrics*, 2013.