

## CHARACTERIZATIONS IN A RANDOM RECORD MODEL WITH A NONIDENTICALLY DISTRIBUTED INITIAL RECORD

GADI BARLEVY,\* *Federal Reserve Bank of Chicago*

H. N. NAGARAJA,\*\* *The Ohio State University*

### Abstract

We consider a sequence, of random length  $M$ , of independent, continuous observations  $X_i$ ,  $1 \leq i \leq M$ , where  $M$  is geometric,  $X_1$  has cumulative distribution function (CDF)  $G$ , and  $X_i$ ,  $i \geq 2$ , have CDF  $F$ . Let  $N$  be the number of upper records and let  $R_n$ ,  $n \geq 1$ , be the  $n$ th record value. We show that  $N$  is independent of  $F$  if and only if  $G(x) = G_0(F(x))$  for some CDF  $G_0$ , and that if  $E(|X_2|)$  is finite then so is  $E(|R_n|)$ ,  $n \geq 2$ , whenever  $N \geq n$  or  $N = n$ . We prove that the distribution of  $N$ , along with appropriately chosen subsequences of  $E(R_n)$ , characterize  $F$  and  $G$  and, along with subsequences of  $E(R_n - R_{n-1})$ , characterize  $F$  and  $G$  up to a common location shift. We discuss some applications to the identification of the wage offer distribution in job search models.

*Keywords:* Moment sequence; number of records; record spacing; geometric distribution; Müntz–Szász theorem; Titchmarsh convolution theorem; job search model

2000 Mathematics Subject Classification: Primary 62E10

Secondary 60G70; 62G32

### 1. Introduction

Let  $\{X_i, i \geq 1\}$  be an infinite sequence of independent random variables. Suppose that  $M$  is a positive, integer-valued random variable independent of the  $X_i$ , and assume that only  $X_i$ ,  $1 \leq i \leq M$ , are observed. Define  $L(1) = 1$  and  $L(n) = \min\{k: X_k > X_{L(n-1)}\}$  for  $n > 1$  and  $R_n = X_{L(n)}$  for  $n \geq 1$ . Then  $R_1$  is the initial record (sometimes called the trivial record) and the  $R_n$ ,  $n \geq 2$ , represent the upper record values from the sequence  $\{X_i, i \geq 1\}$ . The total number of records we observe is given by  $N = \max\{j: L(j) \leq M\}$  and is itself a random variable.

When the  $X_i$  are identically distributed, this model is called a *random record model* (see, e.g. Arnold *et al.* (1998, p. 224)). When we further assume that  $M$  has a geometric distribution, i.e.  $\Pr(M = m) = q^{m-1}p$  for  $m \geq 1$ , where  $0 < p < 1$  and  $q = 1 - p$ , we have a geometric random record (GRR) model. Nagaraja and Barlevy (2003) derived several characterization results for the GRR model using record moments. In this paper, we consider a variation of the GRR model in which the initial observation,  $X_1$ , has a potentially different distribution from the sequence of remaining observations  $\{X_i, i \geq 2\}$ . We refer to this as a *GRR model with a nonidentically distributed initial record*, or a *modified GRR model*. Our purpose in this paper is to determine whether there exist analogous characterization results for this alternative

---

Received 28 September 2005; revision received 30 August 2006.

\* Postal address: Economic Research Department, Federal Reserve Bank of Chicago, 230 South La Salle Street, Chicago, IL 60604-1413, USA. Email address: gbarlevy@frbchi.org

\*\* Postal address: Department of Statistics, The Ohio State University, 1958 Neil Avenue, Columbus, OH 43210-1247, USA. Email address: hnn@stat.ohio-state.edu

formulation and to discuss some applications of this variation concerning identification of job search models.

Formally, let  $X_1$  be distributed with continuous cumulative distribution function (CDF)  $G$ , where  $G(x) = \Pr(X_1 \leq x)$ , and let the  $X_i, i \geq 2$ , be independent and identically distributed with continuous CDF  $F$ , where  $F(x) = \Pr(X_i \leq x)$ .

Define a mapping  $\Gamma$  from the set of continuous CDFs into itself such that  $G = \Gamma(F)$ . This notation allows us to view the model as being parameterized by a single CDF,  $F$ . We make the following assumption of  $\Gamma$ .

**Assumption 1.** *The probability measure implied by  $G = \Gamma(F)$  is absolutely continuous with respect to the probability measure implied by  $F$ .*

This assumption implies that the support of  $G$  forms a subset of the support of  $F$ . An implication of this assumption is that we can express the CDF  $G$  at the  $u$ th percentile in the distribution  $F$  as an absolutely continuous function of  $u$ . Formally, define  $F^{-1}(u) = \inf\{x : F(x) \geq u\}$ . We then have the following result.

**Lemma 1.**  *$G(F^{-1}(u))$  is absolutely continuous in  $u \in (0, 1)$ .*

*Proof.* By the Radon–Nikodým theorem, Assumption 1 implies the existence of a function  $\delta$  such that  $G(y) = \int_{-\infty}^y \delta(x) dF(x)$  for all  $y$ . Moreover, since  $G$  is a CDF,  $\int_{-\infty}^{\infty} \delta(x) dF(x) = 1$ . Define  $u = F(x)$ ; then we can write  $G(y) = \int_0^{F(y)} \delta(F^{-1}(u)) du$ . Since  $F$  is continuous,  $F(F^{-1}(u)) = u$  and

$$G(F^{-1}(v)) = \int_0^{F(F^{-1}(v))} \delta(F^{-1}(u)) du = \int_0^v \delta(F^{-1}(u)) du,$$

confirming that  $G(F^{-1}(v))$  is absolutely continuous in  $v$ .

**Remark 1.** The identity mapping  $\Gamma : F \mapsto F$  trivially satisfies Assumption 1. Our model thus includes the GRR model as a special case.

**Remark 2.** Our formulation is itself a special case of the Pfeifer (1982) model, in which the distribution of the underlying observations changes after each record is set. Here the distribution changes only after the first record, and the distribution of the first record,  $G = \Gamma(F)$ , is assumed to satisfy Assumption 1. Although Pfeifer assumed that  $M = \infty$ , Bunge and Nagaraja (1991) subsequently generalized his model to allow the number of observations to be random.

**Remark 3.** Our assumptions do not require  $\Gamma$  to be one-to-one, as illustrated in Example 2, below. However, Assumption 1 implies that  $\Gamma$  cannot assign a single  $G$  to all CDFs  $F$ . Thus,  $G$  cannot be independent of  $F$ .

Here are some examples of functions  $G(x) = \Gamma(F)(x)$  that satisfy Assumption 1. The motivation for these examples will become clear in Section 6, when we discuss how the model can be applied to estimate job search models.

1.  $\Gamma(F)(x) = \frac{F(x)}{1 + \kappa(1 - F(x))}$  for some constant  $\kappa$ .
2.  $\Gamma(F)(x) = F(x)/z$  if  $F(x) \leq z$  and  $\Gamma(F)(x) = 1$  if  $F(x) > z$ , for some constant  $z \in (0, 1)$ .

$$3. \Gamma(F)(x) = \frac{\int_{-\infty}^x H(w) dF(w)}{\int_{-\infty}^{\infty} H(w) dF(w)} = \frac{\int_0^{F(x)} H(F^{-1}(u)) du}{\int_0^1 H(F^{-1}(u)) du}, \text{ where } H(\cdot) \text{ is a CDF.}$$

In Example 1  $\Gamma(F) = \sum_{i=1}^{\infty} pq^{i-1}F^i$  with  $p = (1 + \kappa)^{-1}$ , implying that  $X_1$  has the same distribution as the maximum of a random (geometric) number of independent, identically distributed random variables distributed in the same way as  $X_2$ . In Example 2  $G$  arises from  $F$  by the truncation of the latter’s upper tail, and in Example 3  $G$  has the form of a weighted distribution. Note that, in the first two examples,  $\Gamma(F)(x)$  can be represented as  $G_0(F(x))$  for some function  $G_0$ , i.e. the CDF  $G = \Gamma(F)$  evaluated at  $x$  depends only on  $F(x)$  and not on the value of  $F$  at any point other than  $x$ . This is not true in the last example.

We show in Section 2 that  $\Gamma(F)(x) = G_0(F(x))$  for some function  $G_0$  if and only if the distribution of the number of observed records,  $N$ , is independent of  $F$ . In Section 3 we focus on the case where  $\Gamma(F)(x) = G_0(F(x))$ , and show that, under an additional assumption on  $\Gamma$ , the distribution  $F$  is characterized by subsequences of the following moments:

- (a)  $E(R_n \mid N \geq n)$ ,
- (b)  $E(R_n \mid N = n)$ ,
- (c)  $E(R_n - R_{n-1} \mid N \geq n)$ ,
- (d)  $E(R_n - R_{n-1} \mid N = n)$ .

In demonstrating this result, we appeal not only to the Müntz–Szász theorem (Müntz (1914), Szász (1916)), which is often invoked in moment-based characterization theorems (see Kamps (1998)), but also to a relatively unutilized convolution theorem due to Titchmarsh (1926). In Section 4 we consider more general mappings  $\Gamma$  and provide characterizations of  $F$  and  $G$  using the same moment subsequences together with a subset of values for  $\Pr(N = n)$ . All of these results are premised on there being a known geometric distribution for  $M$ . In Section 5 we analyze the case where  $M$  is known to have a geometric distribution but its parameter is unknown. In Section 6 we then discuss how our results can be used to nonparametrically identify the wage offer distribution in job search models when wage data can only be measured with noise.

In our proofs, we assume that  $F(x)$  is absolutely continuous in  $x$  and denote its density by  $f$ . By Lemma 1, this assumption implies that  $G(x)$  is also absolutely continuous in  $x$  and has density function  $g(x) = \delta(x)f(x)$ , where  $\delta(x)$  is as defined in Lemma 1. While this simplifies our proofs, our results hold even when  $F$  is merely continuous, since we can always define a series of random variables  $Y_i = F(X_i)$  whose respective distribution functions will be absolutely continuous for all  $i$  under Assumption 1, and repeat the analysis using transformations of record values from the sequence  $\{Y_i, i \geq 1\}$ .

## 2. Characterization results for $\Gamma$ and $N$

We begin with results that characterize the mapping  $\Gamma$ . Our first result shows that the number of observed records,  $N$ , is independent of  $F$  if and only if  $\Gamma(F)(x)$  can be expressed solely as a function of  $F(x)$ , i.e.  $F$  evaluated at the point  $x$ .

**Proposition 1.** *In the modified GRR model, the number of observed records,  $N$ , is independent of  $F$  if and only if  $\Gamma(F)(x) = G_0(F(x))$  for some absolutely continuous CDF  $G_0$  with support  $[0, 1]$ .*

*Proof.* Building on Bunge and Nagaraja (1991) and Nagaraja and Barlevy (2003), we can express the likelihood of having exactly  $n$  observed records with values  $r_1$  through  $r_n$  as

$$h(\{r_1, \dots, r_n\} \cap \{N = n\}) = \frac{(1 - q)g(r_1)}{1 - qF(r_1)} \prod_{i=2}^n \frac{qf(r_i)}{1 - qF(r_i)}.$$

Next, we integrate out  $r_2, \dots, r_n$ , to obtain

$$h(\{r_1\} \cap \{N = n\}) = \frac{1}{(n - 1)!} \frac{1 - q}{1 - qF(r_1)} \left[ \ln \left( \frac{1 - qF(r_1)}{1 - q} \right) \right]^{n-1} g(r_1).$$

Hence,  $\Pr(N = n)$  can be expressed as

$$\Pr(N = n) = \frac{1}{(n - 1)!} \int_{-\infty}^{\infty} \frac{1 - q}{1 - qF(r_1)} \left[ \ln \left( \frac{1 - qF(r_1)}{1 - q} \right) \right]^{n-1} g(r_1) dr_1.$$

Suppose that  $\Gamma(F)(x) = G_0(F(x))$  for some absolutely continuous function  $G_0$ . We want to show that  $\Pr(N = n)$  is independent of  $F$ . Since  $G_0$  is absolutely continuous, it has a related density function  $g_0(x) = (d/dx)G_0(x)$ . This implies that

$$g(r_1) = \frac{d}{dx} G_0(F(x)) \Big|_{x=r_1} = g_0(F(r_1))f(r_1).$$

Substituting this into the expression for  $\Pr(N = n)$  and using the change of variable  $u = F(r_1)$  yields

$$\Pr(N = n) = \frac{1}{(n - 1)!} \int_0^1 \frac{1 - q}{1 - qu} \left[ \ln \left( \frac{1 - qu}{1 - q} \right) \right]^{n-1} g_0(u) du,$$

which is indeed independent of  $F$ .

Next, suppose that  $\Pr(N = n)$  is independent of  $F$ . We want to show that this implies  $\Gamma(F)(x) = G_0(F(x))$  for some absolutely continuous CDF  $G_0(\cdot)$ . From Lemma 1, we can rewrite  $\Pr(N = n)$  using the change of variable  $u = F(r_1)$ , to obtain

$$\Pr(N = n) = \frac{1}{(n - 1)!} \int_0^1 \frac{1 - q}{1 - qu} \left[ \ln \left( \frac{1 - qu}{1 - q} \right) \right]^{n-1} \frac{g(F^{-1}(u))}{f(F^{-1}(u))} du. \tag{1}$$

Since  $\{\Pr(N = n), n \geq 1\}$  does not depend on the distribution of  $F$ , for any two distributions  $F_1$  and  $F_2$  we thus have, for  $n = 1, 2, 3, \dots$ ,

$$\begin{aligned} & \int_0^1 \frac{1 - q}{1 - qu} \left[ \ln \left( \frac{1 - qu}{1 - q} \right) \right]^{n-1} \frac{g(F_1^{-1}(u))}{f_1(F_1^{-1}(u))} du \\ &= \int_0^1 \frac{1 - q}{1 - qu} \left[ \ln \left( \frac{1 - qu}{1 - q} \right) \right]^{n-1} \frac{g(F_2^{-1}(u))}{f_2(F_2^{-1}(u))} du \end{aligned}$$

or, more compactly,

$$\int_0^1 h(u) \left[ \ln \left( \frac{1 - qu}{1 - q} \right) \right]^{n-1} du = 0, \tag{2}$$

where

$$h(u) = \frac{1 - q}{1 - qu} \left[ \frac{g(F_1^{-1}(u))}{f_1(F_1^{-1}(u))} - \frac{g(F_2^{-1}(u))}{f_2(F_2^{-1}(u))} \right].$$

We now appeal to the Müntz–Szász theorem, which states that, for any increasing sequence of positive integers,  $\{n_j, j \geq 1\}$ , the sequence of polynomials  $\{x^{n_j}\}$  is complete on  $L(0, 1)$  if and only if  $\sum_{j=1}^{\infty} n_j^{-1} = \infty$ . For a proof and history of the theorem, see Borwein and Erdélyi (1995, Section 4.2). From this theorem, it follows that (2) implies  $h(u) = 0$  for almost all  $u \in (0, 1)$ . Let  $F_1$  be any continuous CDF, and set  $F_2$  to be the uniform CDF, i.e.  $F_2^{-1}(u) = u$ , so  $g(F_2^{-1}(u))/f_2(F_2^{-1}(u)) = g(u)$ . Hence,

$$\frac{g(F_1^{-1}(u))}{f_1(F_1^{-1}(u))} = g(u). \tag{3}$$

Defining  $y = F_1^{-1}(u)$  in (3), we obtain  $g(y) = g(F_1(y))f_1(y)$ . Hence, we have

$$\begin{aligned} \Gamma(F_1)(x) &= \int_{-\infty}^x g(y) \, dy \\ &= \int_{-\infty}^x g(F_1(y))f_1(y) \, dy \\ &= \int_0^{F_1(x)} g(z) \, dz \\ &\equiv G_0(F_1(x)). \end{aligned}$$

Since  $G$  is a CDF, it follows that  $G_0$  is nondecreasing,  $G_0(0) = 0$ , and  $G_0(1) = 1$ . Hence,  $G_0$  is a CDF with support  $[0, 1]$ . The absolute continuity of  $G_0$  is immediate.

We next argue that the mapping  $\Gamma$  is further characterized by the dependence structure of the record indicators derived from  $\{X_i, i \geq 1\}$ . Define the  $m$ th record indicator,  $I_m$ , to be a random variable that takes the value 1 if  $X_m$  is a record, i.e. if  $X_m > \max\{X_1, \dots, X_{m-1}\}$ , and takes the value 0 otherwise. Note that  $I_m$  is defined even for  $m > M$ , since the infinite sequence of  $X_i$  is defined. Previous work by Nevzorov (1986) has shown that if the  $X_i$  are independent random variables and  $X_i$  has CDF

$$F_i(x) = \{F(x)\}^{\alpha(i)} \tag{4}$$

for some common  $F$ , then the  $I_m, m \geq 2$ , are independent and  $I_m$  is Bernoulli with success probability  $\alpha(m)/\sum_{i=1}^m \alpha(i)$ . Conversely, Nevzorov showed that if the supports of the  $X_i$  are not disjoint and  $I_2, \dots, I_{n-1}$  are independent of  $I_n$  for any  $F_n$ , then (4) holds for  $i = 1, \dots, n-1$  (see also Arnold *et al.* (1998, p. 219)). We now derive a related result in our model. In particular, we argue that

$$\Gamma(F)(x) = \{F(x)\}^{\alpha}, \tag{5}$$

for some  $\alpha > 0$ , if and only if the record indicator  $I_2$  is independent of the record indicators  $I_3, \dots, I_m$  for all  $m$ , conditional on at least  $m$  observations.

**Proposition 2.** *In a random record model with a nonidentically distributed initial record,*

$$\Pr(I_3 = 1, \dots, I_m = 1 \mid M \geq m) = \Pr(I_3 = 1, \dots, I_m = 1 \mid I_2 = 1, M \geq m), \tag{6}$$

for all  $m \equiv m_j \geq 3, j \geq 1$ , such that  $\sum_j m_j^{-1} = \infty$ , if and only if (5) holds.

*Proof.* Condition (6) can be expressed as

$$\Pr(I_2 = 1, \dots, I_m = 1) = \Pr(I_3 = 1, \dots, I_m = 1) \Pr(I_2 = 1).$$

Without loss of generality we take  $F(x) = x$ . Upon conditioning on  $X_1$ , the left-hand side (denoted LHS) can be simplified to

$$\frac{1}{(m - 1)!} \int_0^1 (1 - x)^{m-1} dG(x).$$

Applying integration by parts yields

$$\frac{1}{(m - 2)!} \int_0^1 (1 - x)^{m-2} G(x) dx.$$

Furthermore,  $\Pr(I_3 = 1, \dots, I_m = 1) = \text{LHS} + \Pr(I_2 = 0, I_3 = 1, \dots, I_m = 1)$ . The second term on the right-hand side here, upon conditioning on  $X_1$ , can be written as

$$\frac{1}{(m - 2)!} \int_0^1 (1 - x)^{m-2} xg(x) dx.$$

Let  $c = \Pr(I_2 = 1) = \int_0^1 G(x) dx$ . Then (6) reduces to the condition

$$\int_0^1 (1 - x)^{m-2} [(1 - c)G(x) - cxg(x)] dx = 0, \quad m = m_j, j \geq 1. \tag{7}$$

By the Müntz–Szász theorem, it follows that, for almost all  $x \in (0, 1)$ ,

$$\frac{g(x)}{G(x)} = \frac{1 - c}{c} \frac{1}{x}.$$

This differential equation shows that (5) holds with  $\alpha = (1 - c)/c$ .

Conversely, if (4) holds with  $F(x) = x$  and  $G(x) = x^\alpha$ , then set  $c = (1 + \alpha)^{-1}$ . It follows that  $(1 - c)G(x) = cxg(x)$  for all  $x \in (0, 1)$ . Thus, (7) holds for all  $m_j \geq 1$  and, so, (6) holds.

**Corollary 1.** *If (6) holds then (5) holds and, in turn, the record indicators  $I_m, m \geq 2$ , are all independent and the distribution of the number of observed records,  $N$ , is independent of  $F$ .*

### 3. Results in a special case

We now turn to characterization results for  $F$  in the modified GRR model. We begin by deriving expressions for the relevant moments, and provide conditions for these moments to exist. We then consider the special case where  $\Gamma(F)(x) = G_0(F(x))$ , and provide a characterization result.

Consider record moments that condition on the event  $\{N \geq n\}$ . Using the expression for the likelihood  $h(\{r_1, \dots, r_n\} \cap \{N \geq n\})$  from Bunge and Nagaraja (1991), and integrating out  $r_2, \dots, r_{n-2}$ , yields the following expression for the joint likelihood of  $r_1, r_{n-1}$ , and  $r_n$  when  $N \geq n, n \geq 3$ :

$$\begin{aligned} & h(\{r_1, r_{n-1}, r_n\} \cap \{N \geq n\}) \\ &= \frac{1}{(n - 3)!} \left[ -\ln \left( \frac{1 - qF(r_{n-1})}{1 - qF(r_1)} \right) \right]^{n-3} \frac{qg(r_1)}{1 - qF(r_1)} \frac{qf(r_{n-1})}{1 - qF(r_{n-1})} f(r_n). \end{aligned}$$

Limiting our attention to  $n \geq 3$ , we derive the following expression for  $E(R_n | N \geq n)$ :

$$\begin{aligned}
 E(R_n | N \geq n) &= \frac{1}{\Pr(N \geq n)} \int_{-\infty}^{\infty} \int_{-\infty}^{r_{n-1}} \int_{r_{n-1}}^{\infty} r_n h(\{r_1, r_{n-1}, r_n\} \cap \{N \geq n\}) dr_n dr_1 dr_{n-1} \\
 &= \int_0^1 \int_0^{u_{n-1}} \frac{[-\ln((1 - qu_{n-1})/(1 - qu_1))]^{n-3}}{(n - 3)! \Pr(N \geq n)} \frac{g(F^{-1}(u_1))}{f(F^{-1}(u_1))} \\
 &\quad \times \phi_F(u_{n-1}) \frac{qdu_1}{1 - qu_1} \frac{qdu_{n-1}}{1 - qu_{n-1}}.
 \end{aligned} \tag{8}$$

Here

$$\phi_F(u_{n-1}) = \int_{u_{n-1}}^1 F^{-1}(u_n) du_n. \tag{9}$$

Similarly, the expected record spacing  $E(R_n - R_{n-1} | N \geq n)$ ,  $n \geq 3$ , can be expressed as

$$\begin{aligned}
 E(R_n - R_{n-1} | N \geq n) &= \int_0^1 \int_0^{u_{n-1}} \frac{[-\ln((1 - qu_{n-1})/(1 - qu_1))]^{n-3}}{(n - 3)! \Pr(N \geq n)} \frac{g(F^{-1}(u_1))}{f(F^{-1}(u_1))} \\
 &\quad \times \phi_F(u_{n-1}) \frac{qdu_1}{1 - qu_1} \frac{qdu_{n-1}}{1 - qu_{n-1}},
 \end{aligned}$$

where

$$\phi_F(u_{n-1}) = \int_{u_{n-1}}^1 [F^{-1}(u_n) - F^{-1}(u_{n-1})] du_n. \tag{10}$$

In the same fashion, we can use the expression for  $h(\{r_1, \dots, r_n\} \cap \{N = n\})$  derived in Section 2, and integrate out  $r_2, \dots, r_{n-2}$ , to obtain the joint likelihood of  $r_1, r_{n-1}$ , and  $r_n$  when  $N = n$ ,  $n \geq 3$ :

$$\begin{aligned}
 &h(\{r_1, r_{n-1}, r_n\} \cap \{N = n\}) \\
 &= \frac{1}{(n - 3)!} \left[ -\ln\left(\frac{1 - qF(r_{n-1})}{1 - qF(r_1)}\right) \right]^{n-3} \frac{qg(r_1)}{1 - qF(r_1)} \frac{qf(r_{n-1})}{1 - qF(r_{n-1})} \frac{(1 - q)f(r_n)}{1 - qF(r_n)}.
 \end{aligned}$$

Using this, we deduce that

$$\begin{aligned}
 E(R_n | N = n) &= \int_0^1 \int_0^{u_{n-1}} \frac{[-\ln((1 - qu_{n-1})/(1 - qu_1))]^{n-3}}{(n - 3)! \Pr(N = n)} \frac{g(F^{-1}(u_1))}{f(F^{-1}(u_1))} \\
 &\quad \times \phi_F(u_{n-1}) \frac{qdu_1}{1 - qu_1} \frac{qdu_{n-1}}{1 - qu_{n-1}},
 \end{aligned}$$

where

$$\phi_F(u_{n-1}) = \int_{u_{n-1}}^1 F^{-1}(u_n) \frac{(1 - q)du_n}{1 - qu_n}, \tag{11}$$

and, likewise, that

$$\begin{aligned}
 E(R_n - R_{n-1} | N = n) &= \int_0^1 \int_0^{u_{n-1}} \frac{[-\ln((1 - qu_{n-1})/(1 - qu_1))]^{n-3}}{(n - 3)! \Pr(N = n)} \frac{g(F^{-1}(u_1))}{f(F^{-1}(u_1))} \\
 &\quad \times \phi_F(u_{n-1}) \frac{qdu_1}{1 - qu_1} \frac{qdu_{n-1}}{1 - qu_{n-1}},
 \end{aligned}$$

where

$$\phi_F(u_{n-1}) = \int_{u_{n-1}}^1 \frac{1-q}{1-qu_n} [F^{-1}(u_n) - F^{-1}(u_{n-1})] du_n. \tag{12}$$

All four moment sequences above can thus be expressed as an integral of a common term multiplying a function  $\phi_F(u_{n-1})$  that varies with the particular moment at hand.

We now provide a sufficient condition for the above moments to exist. Here we use the fact that, for any random variable  $Y$ ,  $E(Y)$  exists if and only if  $E(|Y|) < \infty$ .

**Proposition 3.** *If  $E(|X_2|) < \infty$  then  $E(|R_n| \mid N \geq n)$  and  $E(|R_n| \mid N = n)$  exist for all  $n \geq 2$ .*

*Proof.* Suppose first that  $n \geq 3$ . From (8), we have

$$\begin{aligned} E(|R_n| \mid N \geq n) &= \int_0^1 \int_0^{u_{n-1}} \int_{u_{n-1}}^1 \frac{[-\ln((1-qu_{n-1})/(1-qu_1))]^{n-3}}{(n-3)! \Pr(N \geq n)} \frac{g(F^{-1}(u_1))}{f(F^{-1}(u_1))} \\ &\quad \times |F^{-1}(u_n)| du_n \frac{q du_1}{1-qu_1} \frac{q du_{n-1}}{1-qu_{n-1}}. \end{aligned}$$

Since  $u_1$  and  $u_{n-1}$  both lie in  $[0, 1]$ , the above expression is less than or equal to

$$\begin{aligned} &\frac{[-\ln(1-q)]^{n-3}}{(n-3)! \Pr(N \geq n)} \left(\frac{q}{1-q}\right)^2 \int_0^1 \int_{u_{n-1}}^1 \left\{ \int_0^{u_{n-1}} \frac{g(F^{-1}(u_1))}{f(F^{-1}(u_1))} du_1 \right\} |F^{-1}(u_n)| du_n du_{n-1} \\ &= \frac{[-\ln(1-q)]^{n-3}}{(n-3)! \Pr(N \geq n)} \left(\frac{q}{1-q}\right)^2 \int_0^1 \int_{u_{n-1}}^1 G(F^{-1}(u_{n-1})) |F^{-1}(u_n)| du_n du_{n-1} \\ &\leq \frac{[-\ln(1-q)]^{n-3}}{(n-3)! \Pr(N \geq n)} \left(\frac{q}{1-q}\right)^2 \int_0^1 \int_0^1 G(F^{-1}(u_{n-1})) |F^{-1}(u_n)| du_n du_{n-1} \\ &\leq \frac{[-\ln(1-q)]^{n-3}}{(n-3)! \Pr(N \geq n)} \left(\frac{q}{1-q}\right)^2 E(|X_2|). \end{aligned}$$

As long as  $E(|X_2|) < \infty$ ,  $E(|R_n| \mid N \geq n)$  exists. A similar argument applies in the case where  $n = 2$ . Extending the argument for  $E(|R_n| \mid N = n)$  is straightforward.

**Remark 4.** For  $n \geq 2$ , the existence of the  $n$ th record moment does not depend on whether  $E(X_1)$  exists. This is because, for  $n \geq 2$ , the relevant moment is always conditioned on an event in which  $M \geq n$  and  $\max\{X_2, \dots, X_M\} > X_1$ . As demonstrated in Nagaraja and Barlevy (2003), if  $M$  is geometric with success probability  $p$ , then  $E(E(|\max\{X_2, \dots, X_M\}| \mid M)) < p^{-1} E(|X_2|)$ . Thus, we are conditioning on the event that  $X_1$  is exceeded by a random variable whose mean is finite. Even if  $X_1$  does not have a well-defined unconditional mean, conditioning on the event that its value is exceeded by a random variable with a finite mean suffices to ensure that  $E(X_1 \mid N \geq n)$  is finite.

Equipped with these preliminaries, we turn to characterizing  $F$  using moment sequences. However, we first need to make an additional assumption about the range of  $\Gamma$ .

**Assumption 2.** *For any continuous CDF  $F$ , there exists an  $\varepsilon > 0$  such that*

$$g(F^{-1}(u)) = \frac{d}{dx} \Gamma(F)(x) \Big|_{x=F^{-1}(u)}$$

*is positive for almost all  $u \in (0, \varepsilon)$ .*



This assumption implies that  $G^{-1}(0) \leq F^{-1}(0)$ . However, by Assumption 1,  $G^{-1}(0) \geq F^{-1}(0)$ . Hence,  $G^{-1}(0) = F^{-1}(0)$ . This restriction is clearly necessary: if two distributions  $F_1$  and  $F_2$ ,  $F_1 \neq F_2$ , differ only below  $G^{-1}(0)$ , they will necessarily yield identical record moment sequences.

We first focus on a special case for  $\Gamma$ , namely that in which  $\Gamma(F) = G_0(F(x))$  for some absolutely continuous CDF  $G_0: [0, 1] \rightarrow [0, 1]$ . Recall from Proposition 1 that this is true if and only if  $\Pr(N = n)$  does not depend on  $F$ . We return to the more general case in the next section.

**Proposition 4.** *Suppose that Assumptions 1 and 2 are satisfied, that  $E(|X_2|) < \infty$ , and that  $\Gamma(F)(x) = G_0(F(x))$  for some absolutely continuous CDF  $G_0: [0, 1] \rightarrow [0, 1]$ .*

- (a) *If two distributions  $F_1$  and  $F_2$  give rise to either the same sequence  $\{E(R_n \mid N \geq n), n = n_1, n_2, \dots\}$  or the same sequence  $\{E(R_n \mid N = n), n = n_1, n_2, \dots\}$ , where  $\{n_j, j \geq 1\}$  is an increasing sequence of positive integers and  $\sum_j n_j^{-1} = \infty$ , then  $F_1(x) = F_2(x)$ .*
- (b) *If two distributions  $F_1$  and  $F_2$  give rise to either the same sequence  $\{E(R_n - R_{n-1} \mid N \geq n), n = n_1, n_2, \dots\}$  or the same sequence  $\{E(R_n - R_{n-1} \mid N = n), n = n_1, n_2, \dots\}$ , where  $\{n_j, j \geq 1\}$  is an increasing sequence of positive integers and  $\sum_j n_j^{-1} = \infty$ , then there exists a  $c$  such that  $F_1(x) = F_2(x - c)$ .*

*Proof.* Since  $\Gamma(F)(x) = G_0(F(x))$ , it follows that  $g(x) = g_0(F(x))f(x)$ , which implies that  $g(F^{-1}(u_1))/f(F^{-1}(u_1)) = g_0(u_1)$ . We can therefore express the respective record moments in the statement of the proposition as

$$\frac{1}{(n-3)! a_n} \int_0^1 \int_{u_1}^1 g_0(u_1) \left[ -\ln\left(\frac{1 - qu_{n-1}}{1 - qu_1}\right) \right]^{n-3} \phi_F(u_{n-1}) \frac{q \, du_{n-1}}{1 - qu_{n-1}} \frac{q \, du_1}{1 - qu_1},$$

where  $a_n = \Pr(N \geq n)$  or  $a_n = \Pr(N = n)$ , and  $\phi_F$  depends on the particular moment and conditioning event, i.e. (9) or (11) in case (a) and (10) or (12) in case (b). Changing variables according to

$$\begin{aligned} t = -\ln(1 - qu_{n-1}) &\implies dt = \frac{q \, du_{n-1}}{1 - qu_{n-1}}, & u_{n-1} &= \frac{1 - e^{-t}}{q}, \\ s = -\ln(1 - qu_1) &\implies ds = \frac{q \, du_1}{1 - qu_1}, & u_1 &= \frac{1 - e^{-s}}{q}, \end{aligned}$$

and setting  $c = -\ln(1 - q)$  allows us to rewrite the above expression as

$$\frac{1}{(n-3)! a_n} \int_0^c \int_s^c g_0\left(\frac{1 - e^{-s}}{q}\right) \phi_F\left(\frac{1 - e^{-t}}{q}\right) (t - s)^{n-3} \, dt \, ds.$$

We change variable yet again by setting  $\omega = t - s$ , and rewrite the above as

$$\frac{1}{(n-3)! a_n} \int_0^c \int_\omega^c g_0\left(\frac{1 - e^{-(t-\omega)}}{q}\right) \phi_F\left(\frac{1 - e^{-t}}{q}\right) \omega^{n-3} \, dt \, d\omega.$$

Define

$$\eta_F(\omega) = \int_\omega^c g_0\left(\frac{1 - e^{-(t-\omega)}}{q}\right) \phi_F\left(\frac{1 - e^{-t}}{q}\right) \, dt.$$

Let  $F_1$  and  $F_2$  denote two continuous CDFs that give rise to the same subsequence of moments. Since  $\Gamma(F)(x) = G_0(F(x))$  implies that  $a_n$  is independent of  $F$ , it follows that, for all  $n \geq 3$ ,

$$\int_0^c \eta_{F_1}(\omega)\omega^{n-3} d\omega = \int_0^c \eta_{F_2}(\omega)\omega^{n-3} d\omega.$$

By the Müntz–Szász theorem, it follows that  $\eta_{F_1}(\omega) = \eta_{F_2}(\omega)$  for almost all  $\omega \in (0, c)$ , i.e. for almost all  $\omega \in (0, c)$ ,

$$\int_{\omega}^c g_0\left(\frac{1 - e^{-(t-\omega)}}{q}\right)\phi_{F_1}\left(\frac{1 - e^{-t}}{q}\right) dt = \int_{\omega}^c g_0\left(\frac{1 - e^{-(t-\omega)}}{q}\right)\phi_{F_2}\left(\frac{1 - e^{-t}}{q}\right) dt. \tag{13}$$

We next argue that (13) implies  $\phi_{F_1}((1 - e^{-t})/q) = \phi_{F_2}((1 - e^{-t})/q)$  for almost all  $t \in (0, c)$ . It will suffice to prove that if

$$\int_{\omega}^c g_0\left(\frac{1 - e^{-(t-\omega)}}{q}\right)\phi(t) dt = 0 \quad \text{for almost all } \omega \in (0, c), \tag{14}$$

then  $\phi(t) = 0$  for almost all  $t \in (0, c)$ . By making a change in variables to  $w = c - t$  and  $z = c - \omega$ , (14) can be transformed into the following integral equation, where  $a(x) = g_0((1 - e^{-x})/q)$  and  $b(x) = \phi(c - x)$ :

$$\int_0^z a(z - w)b(w) dw = 0 \quad \text{for almost all } z \in (0, c). \tag{15}$$

From Theorem VII of Titchmarsh (1926) (see also Boas (1954, p. 234) for a more accessible reference), there exists a  $c^*$  such that  $a(x) = 0$  for all  $x \in (0, c^*)$  and  $b(x) = 0$  for all  $x \in (0, c - c^*)$ . However, Assumption 2 implies that there exists an  $\varepsilon > 0$  such that  $g_0(z) > 0$  for almost all  $z \in (0, \varepsilon)$ , which in turn implies that  $a(z) > 0$  for almost all  $z \in (0, \varepsilon)$ . Hence,  $c^*$  must equal 0, implying that  $b(z) = 0$  for almost all  $z \in (0, c)$ . Then  $\phi(t) = b(c - t) = 0$  for almost all  $t \in (0, c)$ , as claimed.

Lastly, we need to show that the statement of the proposition follows from the fact that  $\phi_{F_1}((1 - e^{-t})/q) = \phi_{F_2}((1 - e^{-t})/q)$  for almost all  $t \in (0, c)$ . In case (a), with  $\phi_F$  given by (9) or (11),

$$\int_{(1-e^{-t})/q}^1 [F_1^{-1}(u) - F_2^{-1}(u)] du = 0$$

or

$$\int_{(1-e^{-t})/q}^1 \frac{F_1^{-1}(u) - F_2^{-1}(u)}{1 - qu} du = 0,$$

for almost all  $t \in (0, c)$ . However, from Taylor (1965, p. 415), this implies that the function inside the integral is equal to 0 almost everywhere. It follows that  $F_1(u) = F_2(u)$  almost everywhere. Since  $F_1$  and  $F_2$  are both right continuous (in fact, under our assumption both are continuous),  $F_1$  and  $F_2$  must be equal everywhere. In case (b), with  $\phi_F$  given by (10), the fact that  $\phi_{F_1}(t) = \phi_{F_2}(t)$  for almost all  $t \in (0, c)$  implies that, for almost all  $u \in (0, 1)$ ,

$$\int_u^1 [F_1^{-1}(u_n) - F_1^{-1}(u)] du_n = \int_u^1 [F_2^{-1}(u_n) - F_2^{-1}(u)] du_n.$$

With  $\phi_F$  as given in (12), we analogously obtain

$$\int_u^1 \frac{F_1^{-1}(u_n) - F_1^{-1}(u)}{1 - qu_n} du_n = \int_u^1 \frac{F_2^{-1}(u_n) - F_2^{-1}(u)}{1 - qu_n} du_n.$$

Nagaraja and Barlevy (2003) have shown that this implies the existence of a constant  $c$  such that  $F_1^{-1}(u) = F_2^{-1}(u) + c$ , or  $F_1(x) = F_2(x - c)$ .

**Remark 5.** From the proof above, we can further deduce what happens when we relax Assumption 2, i.e. when we assume that  $G_0^{-1}(0) > 0$ . By the Titchmarsh (1926) convolution theorem, for any solution  $b(w)$  to (15) there exists a value  $c^*$  such that  $a(x) = 0$  for all  $x \in (0, c^*)$  and  $b(x) = 0$  for all  $x \in (0, c - c^*)$ . However, without making Assumption 2, we can conclude only that  $c^* \leq -\ln(1 - qG_0^{-1}(0))$ . Consequently, we can deduce that  $\phi(t) = b(c - t) = 0$  for almost all  $t \in (-\ln(1 - qG_0^{-1}(0)), c)$  and, hence, that  $\phi_{F_1}((1 - e^{-t})/q) = \phi_{F_2}((1 - e^{-t})/q)$  for almost all  $t \in (-\ln(1 - qG_0^{-1}(0)), c)$ . In case (a) it would therefore follow that  $F_1(x) = F_2(x)$  for  $x > F^{-1}(G_0^{-1}(0))$ , and in case (b) it would follow that  $F_1(x) = F_2(x - c)$  for  $x > F^{-1}(G_0^{-1}(0))$ . In other words, we can generalize Proposition 4 to imply that the moment sequences in the statement of the proposition uniquely characterize the distribution  $F$  over the range  $(F^{-1}(G_0^{-1}(0)), \infty)$ .

#### 4. Results in the general case

We now move to the general case of any arbitrary mapping  $\Gamma$  that satisfies Assumptions 1 and 2. In this case, it will no longer be true that record moments alone characterize the distribution  $F$ . However, record moments together with the distribution of the number of records do characterize  $F$ . We begin with the following lemma.

**Lemma 2.** *In the modified GRR model, the sequence  $\{\Pr(N = n_j), j \geq 1\}$ , for any increasing sequence of positive integers  $n_j$  such that  $\sum_{j=1}^\infty n_j^{-1} = \infty$ , uniquely determines the probability mass function of  $N$ .*

*Proof.* Consider two modified GRR models with respective CDFs  $F_1$  and  $G_1$  and  $F_2$  and  $G_2$ . If  $\Pr(N = n_j)$  remains the same for these two models, from (1) we have

$$\int_0^1 \frac{1 - q}{1 - qu} \left[ \ln\left(\frac{1 - qu}{1 - q}\right) \right]^{n_j - 1} \left[ \frac{g_1(F_1^{-1}(u))}{f_1(F_1^{-1}(u))} - \frac{g_2(F_2^{-1}(u))}{f_2(F_2^{-1}(u))} \right] du = 0.$$

Set  $t = \ln((1 - qu)/(1 - q))$  and rewrite the equation above as

$$\int_0^{-\ln(1 - q)} h_1(t) t^{n_j - 1} dt = 0.$$

Since  $-\ln(1 - q) < \infty$ , the Müntz–Szász theorem is applicable and it thus follows that  $h_1(t) = 0$  almost everywhere or

$$\frac{g_1(F_1^{-1}(u))}{f_1(F_1^{-1}(u))} = \frac{g_2(F_2^{-1}(u))}{f_2(F_2^{-1}(u))} \tag{16}$$

for almost all  $u \in (0, 1)$ . Again, upon appealing to (1) we conclude that  $\Pr(N = n)$  and, thus,  $\Pr(N \geq n)$  remain the same for all  $n \geq 1$  under the two models.

**Proposition 5.** *Suppose that Assumptions 1 and 2 are satisfied and that  $E(|X_2|) < \infty$ .*

- (a) *If two distributions  $F_1$  and  $F_2$  give rise to the same sequences  $\{E(R_n \mid N \geq n), n = n_1, n_2, \dots\}$  and  $\{\Pr(N = n), n = n_1, n_2, \dots\}$ , where  $\{n_j, j \geq 1\}$  is an increasing sequence of positive integers and  $\sum_j n_j^{-1} = \infty$ , then  $F_1 = F_2$  and  $G_1 = G_2$ .*
- (b) *If two distributions  $F_1$  and  $F_2$  give rise to the same sequences  $\{E(R_n - R_{n-1} \mid N \geq n), n = n_1, n_2, \dots\}$  and  $\{\Pr(N = n), n = n_1, n_2, \dots\}$ , where  $\{n_j, j \geq 1\}$  is an increasing sequence of positive integers and  $\sum_j n_j^{-1} = \infty$ , then  $F_1(x) = F_2(x - c)$  and  $G_1(x) = G_2(x - c)$  for some constant  $c$ .*

*Proof.* Since both  $F_1$  and  $F_2$  give rise to the same sequence  $\{\Pr(N = n_j), j \geq 1\}$  where  $\sum_j n_j^{-1} = \infty$ , from Lemma 2 and (16) we conclude that the  $\Pr(N \geq n)$  also match for  $n \geq 1$ , and that, for almost all  $u \in (0, 1)$ ,

$$\frac{g_1(F_1^{-1}(u))}{f_1(F_1^{-1}(u))} = \frac{g_2(F_2^{-1}(u))}{f_2(F_2^{-1}(u))}.$$

Let us write this common function as  $g_0(u)$ . In contrast to in the previous section,  $g_0(u)$  now depends on the sequence  $\{\Pr(N = n_j)\}$  as opposed to being a stand-alone function.

Recall that all four moment sequences above can respectively be written as

$$\int_0^1 \int_0^{u_{n-1}} \frac{[-\ln((1 - qu_{n-1})/(1 - qu_1))]^{n-3}}{(n - 3)! \Pr(N \geq n)} g_0(u_1) \phi_F(u_{n-1}) \frac{q du_1}{1 - qu_1} \frac{q du_{n-1}}{1 - qu_{n-1}}$$

for an appropriately defined  $\phi_F$ . Using the change of variables

$$\begin{aligned} t = -\ln(1 - qu_{n-1}) &\implies ds = \frac{q du_{n-1}}{1 - qu_{n-1}}, & u_{n-1} &= \frac{1 - e^{-t}}{q}, \\ s = -\ln(1 - qu_1) &\implies dt = \frac{q du_1}{1 - qu_1}, & u_1 &= \frac{1 - e^{-s}}{q}, \\ c &= -\ln(1 - q), \end{aligned}$$

we can rewrite this expression as

$$\int_0^c \int_s^c g_0\left(\frac{1 - e^{-s}}{q}\right) \phi_F(t) \frac{(t - s)^{n-3}}{(n - 3)! \Pr(N \geq n)} dt ds.$$

Setting  $\omega = t - s$ , we can further rewrite this expression as

$$\int_0^c \int_\omega^c g_0\left(\frac{1 - e^{-(t-\omega)}}{q}\right) \phi_F(t) \frac{\omega^{n-3}}{(n - 3)! \Pr(N \geq n)} dt d\omega.$$

Let us define

$$\eta_F(\omega) = \int_\omega^c g_0\left(\frac{1 - e^{-(t-\omega)}}{q}\right) \phi_F(t) dt,$$

and let  $F_1$  and  $F_2$  denote two continuous CDFs that give rise to the same sequences. Define  $N_1$  to be the number of records when  $X_2$  is distributed according to  $F_1$ , and  $N_2$  to be the number

of records when  $X_2$  is distributed according to  $F_2$ . If  $F_1$  and  $F_2$  give rise to the same moment sequences, then

$$\frac{1}{(n - 3)! \Pr(N_1 \geq n)} \int_0^c \eta_{F_1}(\omega)\omega^{n-3} d\omega = \frac{1}{(n - 3)! \Pr(N_2 \geq n)} \int_0^c \eta_{F_2}(\omega)\omega^{n-3} d\omega.$$

Since  $\Pr(N_1 \geq n) = \Pr(N_2 \geq n)$ , it follows that

$$\int_0^c \eta_{F_1}(\omega)\omega^{n-3} d\omega = \int_0^c \eta_{F_2}(\omega)\omega^{n-3} d\omega$$

and, thus, by the Müntz–Szász theorem, that  $\eta_{F_1}(\omega) = \eta_{F_2}(\omega)$  almost everywhere, i.e.

$$\int_\omega^c g_0\left(\frac{1 - e^{-(t-\omega)}}{q}\right)\phi_{F_1}(t) dt = \int_\omega^c g_0\left(\frac{1 - e^{-(t-\omega)}}{q}\right)\phi_{F_2}(t) dt.$$

As in the proof of Proposition 4, we rely on the Titchmarsh (1926) convolution theorem to establish that  $\phi_{F_1}(t) = \phi_{F_2}(t)$  almost everywhere, from which we conclude that  $F_1 = F_2$  in case (a) and  $F_1^{-1}(u) = F_2^{-1}(u) + c$  in case (b).

Next, in case (a), we use the fact that  $F_1(x) = F_2(x)$  and the fact (from Lemma 2) that

$$\frac{g_1(F_1^{-1}(u))}{f_1(F_1^{-1}(u))} = \frac{g_2(F_2^{-1}(u))}{f_2(F_2^{-1}(u))}$$

for almost all  $u$  to conclude that

$$g_1(F_1^{-1}(u)) = g_2(F_2^{-1}(u))$$

for almost all  $u \in (0, 1)$ . Hence,

$$G_1(x) = \int_{F_1^{-1}(0)}^x g_1(x) dx = \int_{F_2^{-1}(0)}^x g_2(x) dx = G_2(x),$$

as claimed.

In case (b), we use the fact that  $F_1^{-1}(u) = F_2^{-1}(u) + c$  for all  $u$  to conclude that

$$\frac{g_1(F_2^{-1}(u) + c)}{f_1(F_2^{-1}(u) + c)} = \frac{g_2(F_2^{-1}(u))}{f_2(F_2^{-1}(u))}$$

for almost all  $u \in (0, 1)$ . Also, since  $F_2(x) = F_1(x + c)$  implies that  $f_2(x) = f_1(x + c)$  almost everywhere, it follows that  $f_1(F_2^{-1}(u) + c) = f_2(F_2^{-1}(u))$  for almost all  $u \in (0, 1)$ . Hence, we have

$$\frac{g_1(F_2^{-1}(u) + c)}{f_2(F_2^{-1}(u))} = \frac{g_2(F_2^{-1}(u))}{f_2(F_2^{-1}(u))}.$$

This implies that, for almost every  $u \in (0, 1)$ ,

$$g_1(F_2^{-1}(u) + c) = g_2(F_2^{-1}(u))$$

and, hence,

$$G_1(x + c) = \int_{F_2^{-1}(0)+c}^{x+c} g_1(x) dx = \int_{F_2^{-1}(0)}^x g_2(x) dx = G_2(x),$$

which implies that  $G_1^{-1}(u) = G_2^{-1}(u) + c$ . This completes the proof.

By a similar argument, we can show that the proposition above remains true if we condition on the event that  $N = n$  rather than on the event that  $N \geq n$ .

**Remark 6.** If  $\Gamma$  is known then, once we identify  $F$ , we can also recover  $G = \Gamma(F)$ . However, Proposition 5 implies that  $G$  is itself characterized by the sequences  $\{E(R_n \mid N \geq n)\}$  and  $\{\Pr(N = n)\}$ . Hence, we can independently verify whether the distribution  $\Gamma(F)$  at the  $F$  we identify is the same as the distribution  $G$  directly implied by the moment sequences and the distribution of  $N$ .

### 5. Characterization results across GRR models

The modified GRR model we study can be summarized by a triple  $\{\Gamma, F, q\}$ . So far, we have implicitly focused on results that characterize  $F$  within a given model. That is, for a given  $q$  and  $\Gamma$ , we showed that there is at most one  $F$  for which the model is consistent with a given sequence of record moments and a given distribution for  $N$ . In this section, we ask whether it is possible to characterize the model itself as opposed to the distribution  $F$  within a given model. We show that if two modified GRR models,  $\{\Gamma_1, F_1, q_1\}$  and  $\{\Gamma_2, F_2, q_2\}$ , yield the same record moments and the same distribution for the number of records, then  $q_1 = q_2$ ,  $F_1 = F_2$ , and  $\Gamma_1(F_1) = \Gamma_2(F_2)$ . In other words, the sequences considered in Proposition 5 characterize not only  $F$  and  $G$  but also the particular geometric distribution of  $M$ .

**Proposition 6.** *Suppose that two models,  $\{\Gamma_1, F_1, q_1\}$  and  $\{\Gamma_2, F_2, q_2\}$ , both satisfy Assumptions 1 and 2, and that  $E(|X_2|) < \infty$  in both models. Let  $G_1 = \Gamma_1(F_1)$  and  $G_2 = \Gamma_2(F_2)$ , and assume that  $\sum_j n_j^{-1} = \infty$ .*

- (a) *If  $\{\Gamma_1, F_1, q_1\}$  and  $\{\Gamma_2, F_2, q_2\}$  give rise to the same sequence  $\{\Pr(N = n_j), j \geq 1\}$ , then  $q_1 = q_2$  and (16) holds, i.e.*

$$\frac{g_1(F_1^{-1}(u))}{f_1(F_1^{-1}(u))} = \frac{g_2(F_2^{-1}(u))}{f_2(F_2^{-1}(u))}$$

for almost all  $u \in (0, 1)$ .

- (b) *If, in addition to the condition in (a), the two GRR models give rise to the same sequence  $\{E(R_n \mid N \geq n), n = n_1, n_2, \dots\}$ , then  $G_1 = G_2$  and  $F_1 = F_2$ .*
- (c) *If, in addition to the condition in (a), the two GRR models give rise to the same sequence  $\{E(R_n - R_{n-1} \mid N \geq n), n = n_1, n_2, \dots\}$ , then there exists a  $c$  such that  $G_1(x) = G_2(x - c)$  and  $F_1(x) = F_2(x - c)$ .*

*Proof.* We prove (a). The two remaining claims then follow from Proposition 5.

Since both  $\{G_1, F_1, q_1\}$  and  $\{G_2, F_2, q_2\}$  give rise to the same sequence  $\{\Pr(N = n)\}$ , for  $n = n_1, n_2, \dots$  we have

$$\begin{aligned} & \int_0^1 \frac{1 - q_1}{1 - q_1 u} \left[ \ln \left( \frac{1 - q_1 u}{1 - q_1} \right) \right]^{n-1} \frac{g_1(F_1^{-1}(u))}{f_1(F_1^{-1}(u))} du \\ &= \int_0^1 \frac{1 - q_2}{1 - q_2 u} \left[ \ln \left( \frac{1 - q_2 u}{1 - q_2} \right) \right]^{n-1} \frac{g_2(F_2^{-1}(u))}{f_2(F_2^{-1}(u))} du. \end{aligned}$$

Set  $t = \ln((1 - q_1u)/(1 - q_1))$  on the left-hand side and  $t = \ln((1 - q_2u)(1 - q_2))$  on the right-hand side. We can then rewrite the equation above as

$$\int_0^{-\ln(1-q_1)} \frac{1 - q_1}{q_1} \frac{g_1(F_1^{-1}([1 - (1 - q_1)e^t]/q_1))}{f_1(F_1^{-1}([1 - (1 - q_1)e^t]/q_1))} t^{n-1} dt$$

$$= \int_0^{-\ln(1-q_2)} \frac{1 - q_2}{q_2} \frac{g_2(F_2^{-1}([1 - (1 - q_2)e^t]/q_2))}{f_2(F_2^{-1}([1 - (1 - q_2)e^t]/q_2))} t^{n-1} dt.$$

Let us rewrite this equation as

$$\int_0^{-\ln(1-q_1)} h_1(t)t^{n-1} dt = \int_0^{-\ln(1-q_2)} \frac{q_1}{q_2} \frac{1 - q_2}{1 - q_1} h_2(t)t^{n-1} dt.$$

We now proceed to prove the claim by contradiction. Suppose, without loss of generality, that  $q_2 > q_1$ . Then

$$-\ln(1 - q_2) > -\ln(1 - q_1).$$

Define

$$\hat{h}_1(t) = \begin{cases} h_1(t) & \text{if } t \leq -\ln(1 - q_1), \\ 0 & \text{if } t \in (-\ln(1 - q_1), -\ln(1 - q_2)). \end{cases}$$

Then

$$\int_0^{-\ln(1-q_2)} \hat{h}_1(t)t^{n-1} dt = \int_0^{-\ln(1-q_2)} \frac{q_1}{q_2} \frac{1 - q_2}{1 - q_1} h_2(t)t^{n-1} dt$$

for all  $n = 1, 2, \dots$ . By the Müntz–Szász theorem, it follows that

$$\hat{h}_1(t) = \frac{q_1}{q_2} \frac{1 - q_2}{1 - q_1} h_2(t) \quad \text{for almost all } t \in (0, -\ln(1 - q_2)),$$

which implies that  $h_2(t) = 0$  for almost all  $t \in (-\ln(1 - q_1), -\ln(1 - q_2))$ . However, this implies that  $g_2(x) = 0$  for almost all  $x \in [F_2^{-1}(0), F_2^{-1}((q_2 - q_1)/[q_2(1 - q_1)])]$ , which violates Assumption 2.

Given  $q_1 = q_2 \equiv q$ , the fact that both  $\{G_1, F_1, q\}$  and  $\{G_2, F_2, q\}$  give rise to the same sequence  $\{\Pr(N = n_j)\}$  implies, in view of Lemma 2, that (16) holds.

### 6. Application

Finally, we turn to applications of our results, specifically estimation of job search models that have been successfully applied by economists to study labor markets. In particular, we show that the offer distribution is identified in a larger class of models than was previously demonstrated in work by Nagaraja and Barlevy (2003) and Barlevy (2005). For a more comprehensive review of the literature on the identification and estimation of these models, see Eckstein and van den Berg (2006).

Consider the following job search model, which is frequently used in labor economics. At any point in time, a worker can be either employed or unemployed. While unemployed, a worker receives a fixed dollar amount  $X^*$  per unit time (which may be 0). This amount reflects unemployment benefits, as well as the monetary value of the leisure she enjoys while not working. Employed and unemployed workers encounter employers at a constant rate  $\lambda$  per

unit time. Each time a worker encounters an employer, the latter offers her a wage of  $X$  that is drawn independently from a continuous offer distribution  $F$ . The worker must then choose whether to stay in her current job (or remain unemployed) or to accept the new offer and change employers (or leave unemployment). In addition, a worker can lose her job, an event that occurs at constant rate  $\delta$  whenever she is employed. When a worker loses her job, she cannot recall any of her past offers, and instead becomes unemployed. Workers are assumed to maximize their earnings. Hence, the optimal strategy for an employed worker is to only accept offers that surpass her current wage. Similarly, while unemployed, the worker should only accept offers that exceed  $X^*$ . We assume that  $F^{-1}(0) \geq X^*$ , i.e. all employers offer at least  $X^*$  (otherwise their offers would never be accepted).

Let  $M$  denote the number of job offers a worker receives between intervening spells of unemployment, and index the offers, according to the order in which they arrive, so that  $X_i$  denotes the  $i$ th offer since the worker was last unemployed. Barlevy (2005) showed that  $M$  will have a geometric distribution, i.e.  $\Pr(M = m) = q^{m-1}p$ , where  $p = \delta(\delta + \lambda)^{-1}$  and  $q = 1 - p$ . Given the worker's strategy, the wages for the jobs the worker accepts correspond to records from the sequence  $\{X_i, 1 \leq i \leq M\}$ . In the typical data-sets economists use, workers are only queried on the jobs they work on, not on job offers they have received but turned down. Thus, the only available data consists of  $R_n, 1 \leq n \leq N$ , the record offers the worker received since her last unemployment spell, not the original offers,  $X_i, 1 \leq i \leq M$ , or even the number of offers,  $M$ . A question of interest for economists is whether this data can be used to identify the offer distribution  $F$ .

Since  $R_1 = X_1$ , the distribution  $F$  is obviously identified from the empirical distribution of wages of workers on the first job. However, a key obstacle in applying the model to real-world data is that, empirically, a considerable number of workers voluntarily move into lower-wage jobs, in direct violation of the model. To resolve this discrepancy, economists have argued that wages in the data are a noisy version of wages in the model, i.e. we observe not  $R_n$  but  $R_n + \varepsilon_n$  for some random variable  $\varepsilon_n$  with  $E(\varepsilon_n) = 0$ . The  $\varepsilon_n$  can be viewed as measurement errors, but alternative interpretations for this term have been offered (see Barlevy (2005) for a discussion). Once we assume that we only observe  $R_n + \varepsilon_n$ , we can no longer identify  $F$  from the distribution of  $X_1$ . Previous work, as summarized in Eckstein and van den Berg (2006), resorted to parametric assumptions on  $F$  and the distribution of  $\varepsilon$  to proceed with estimation. By contrast, Nagaraja and Barlevy (2003) and Barlevy (2005) argued that characterization results for the GRR model imply that  $F$  is identified nonparametrically, since it is possible to recover  $E(R_n)$  from noise-ridden data.

However, in order to apply this identification result, we need to keep track of all jobs between spells of unemployment, so that we can determine which record number  $n$  each job represents. Unfortunately, this is not possible in many data-sets. In particular, many surveys collect data on workers who are already employed. For those workers, we have no way of classifying which record number to assign to the jobs we observe for them. Although we could wait until the worker is next unemployed, unemployment is often a sufficiently low-probability event that a large part of the data would have to be thrown out. The results of this paper suggest a way to incorporate data for workers who are already employed. In particular, we know from previous work on search models, e.g. Burdett and Mortensen (1998), that the economy described by this model converges in the limit to a steady state in which the fraction of all employed workers who earn a wage of  $x$  or less is equal to

$$G(x) = \frac{F(x)}{1 + \lambda/\delta(1 - F(x))}. \quad (17)$$



Moreover, the number of offers the worker receives, starting from any job, continues to have a geometric distribution. Thus, as long as the economy we consider is in its steady state, the wages on the jobs we observe for a randomly chosen employed worker will correspond to records from a sequence  $\{X_i, 1 \leq i \leq M\}$  where  $M$  has a geometric distribution,  $X_1$  is distributed according to  $G$  as defined in (17), and  $X_2, \dots, X_M$  are distributed according to  $F$ . Since  $G \equiv G_0(F(x))$  with a known  $G_0$ , we can appeal to Proposition 4 to argue that average wages or average wage changes identify the wage offer distribution  $F$ . For example, if the average wage gains of workers is constant regardless of how many jobs they have had since the first job we observe them in, the wage offer distribution must be exponential.

More generally, our results can be applied whenever the distribution of the wages in the first job we observe for a worker differs from the offer distribution  $F$ . For example, some surveys focus on low-wage workers, using the initial earnings of a worker as a criterion for selection into the survey. In this case, even if we could track workers from their very first job out of unemployment, the distribution of wages on the worker’s first job would correspond to

$$G(x) = \begin{cases} F(x)/z & \text{if } F(x) < z, \\ 1 & \text{otherwise,} \end{cases}$$

where  $z$  reflects the percentile of the threshold wage workers must earn within the wage offer distribution to qualify for the survey. Once again, we can appeal to Proposition 4 to argue that average wages or average wage changes identify the offer distribution  $F$ .

Discrepancies between the wages in the first job we observe for a worker and the offer distribution  $F$  are not confined to sampling issues. Suppose that we can track workers from the first job out of unemployment and that no wages are censored, but that the amount,  $X^*$ , that workers earn while unemployed varies across workers. For example, they might enjoy leisure differently, or they might earn different unemployment benefits (which is not unreasonable, given that these often depend on what the worker earned in her last job before becoming unemployed). Let  $H(x)$  denote the fraction of workers for whom  $X^*$  is  $x$  or less, and suppose that  $H^{-1}(0) \leq F^{-1}(0)$ . Workers for whom  $X^*$  is bigger than  $x$  will hold out for a higher wage before they accept a job offer. The wages in the first job out of unemployment for a worker chosen at random from  $H$  is now given by

$$G(x) = \frac{\int_{-\infty}^x H(w) dF(w)}{\int_{-\infty}^{\infty} H(w) dF(w)} = \frac{\int_0^{F(x)} H(F^{-1}(u)) du}{\int_0^1 H(F^{-1}(u)) du}. \tag{18}$$

Workers continue to draw offers from  $F$  at rate  $\lambda$ , so the wages of a worker chosen at random between two consecutive unemployment spells will correspond to records from a sequence  $\{X_i, 1 \leq i \leq M\}$  where  $M$  has a geometric distribution,  $X_1$  is distributed according to  $G$  as defined in (18), and  $X_2, \dots, X_M$  are distributed according to  $F$ . Since  $G(x)$  cannot be represented as  $G_0(F(x))$  for some function  $G_0$ , we must appeal to Proposition 5 to argue that average wages or average wage changes, together with the distribution of the number of jobs workers hold between unemployment spells, identify  $F$ . If the distribution of reservation wages,  $H$ , is itself unknown, then Proposition 5 implies that we can also identify the distribution of wages of workers on their first job  $G$ . It is easily demonstrated that, given  $F$  and  $G$ , we can recover  $H$ . Thus, when workers have different reservation wages, not only is the common offer distribution,  $F$ , they face still identified, but so is the distribution of  $X^*$  across workers. Thus, we could infer the distribution of how much workers value leisure from the extent of job mobility we observe for them once they become employed. Lastly, Proposition 6 tells us that

we do not need to know the ratio  $\lambda/\delta$  in advance to identify  $F$ , since we can recover it from data on  $N$ , i.e. the distribution of how many jobs workers hold between consecutive unemployment spells. For an empirical implementation of these ideas using panel data on young workers, see Barlevy and Nagaraja (2006).

### Acknowledgements

We wish to thank an anonymous referee for pointing out that the claim in Lemma 1 follows directly from Assumption 1, and for other helpful comments.

### References

- ARNOLD, B. C., BALAKRISHNAN, N. AND NAGARAJA, H. N. (1998). *Records*. John Wiley, New York.
- BARLEVY, G. (2005). Identification of job search models using record statistics. FRB Chicago mimeo.
- BARLEVY, G. AND NAGARAJA, H. N. (2006). Identification of job search models with initial condition problems. FRB Chicago mimeo.
- BOAS, R. P. (1954). *Entire Functions*. Academic Press, New York.
- BORWEIN, P. AND ERDÉLYI, T. (1995). *Polynomials and Polynomial Inequalities*. Springer, New York.
- BUNGE, J. AND NAGARAJA, H. N. (1991). The distribution of certain record statistics from a random number of observations. *Stoch. Proc. Appl.* **38**, 167–183.
- BURDETT, K. AND MORTENSEN, D. (1998). Wage differentials, employer size, and unemployment. *Internat. Econom. Rev.* **39**, 257–273.
- ECKSTEIN, Z. AND VAN DEN BERG, G. (2006). Empirical labor search: a survey. To appear in *J. Econometrics*.
- KAMPS, U. (1998). Characterizations of distributions by recurrence relations and identities for moments of order statistics. In *Handbook of Statistics*, Vol. 16, *Order Statistics: Theory and Methods*, eds N. Balakrishnan and C. R. Rao, Elsevier, Amsterdam, pp. 291–311.
- MÜNTZ, C. H. (1914). Über den Approximationssatz von Weierstrass. In *H. A. Schwarz Festschrift*, Springer, Berlin, pp. 303–312.
- NAGARAJA, H. N. AND BARLEVY, G. (2003). Characterizations using record moments in a random record model and applications. *J. Appl. Prob.* **40**, 826–833.
- NEVZOROV, V. B. (1986). Two characterizations using records. In *Stability Problems for Stochastic Models* (Lecture Notes Math. **1223**), eds V. V. Kalashnikov, B. Penkov and V. M. Zolotarev, Springer, Berlin, pp. 79–85.
- PFEIFER, D. (1982). Characterizations of exponential distributions by independent nonstationary record increments. *J. Appl. Prob.* **19**, 127–135. (Correction: **19** (1982), 906.)
- SZÁSZ, O. (1916). Über die Approximation stetiger Funktionen durch lineare Aggregate von Potenzen. *Math. Ann.* **77**, 482–496.
- TAYLOR, A. (1965). *General Theory of Functions and Integration*. Blaisdell, New York.
- TITCHMARSH, E. C. (1926). The zeros of certain integral functions. *Proc. London Math. Soc.* **2** **25**, 283–302.