Online Resource A (ESM_A.pdf): Supplementary Derivations for Section 3

September 2023

by

Moritz Tarach

Technical University of Darmstadt, Department of Law and Economics
Hochschulstraße 1, D-64289 Darmstadt, Germany
e-mail: tarach@vwl.tu-darmstadt.de, tel.: +49 6151 1657 284

A Supplementary Derivations for Section 3 of "Bottom-up Aggregation of Field-Level Oil Production Profiles via a Successive Sampling Discovery Model and a Birth Process: An Application to the Gulf of Mexico and Norway"

A.1 Derivation of the Posterior Distribution from a Poisson prior

Under the Poisson prior assumption of $N \sim \text{Po}(\nu)$, Bayes' rule implies for the posterior of $N|\mathcal{F}_{\underline{t}}$ that

$$\mathbb{P}[N = n|\mathcal{F}_{\underline{t}}] = \frac{\mathrm{e}^{-\nu}\frac{\nu^n}{n!}\ell(\boldsymbol{x}|n)}{\sum_{n=k}^{\infty}\mathrm{e}^{-\nu}\frac{\nu^n}{n!}\ell(\boldsymbol{x}|n)}, \tag{A.1}$$

where $\ell(\boldsymbol{x}|n)$ is given in equation (3.13). Substituting $l = n - k$, the numerator can be rearranged as follows

$$\begin{aligned}\mathbb{P}[N - k = l|\mathcal{F}_{\underline{t}}] &\propto \mathrm{e}^{-\nu}\frac{\nu^{(k+l)}}{(k+l)!}\frac{(k+l)!}{l!}\prod_{i=1}^{k}\frac{f_{\boldsymbol{\theta}}(x_i)}{p_{\text{tr}}(\boldsymbol{\theta})}\frac{x_i^{\beta}}{b_i}\Big(\int_0^{\infty}\phi_{\boldsymbol{\theta}}(\gamma_k)^l g_{\boldsymbol{b}}(\gamma_k)\mathrm{d}\gamma_k\Big)\\ &= \nu^k\prod_{i=1}^{k}\frac{f_{\boldsymbol{\theta}}(x_i)}{p_{\text{tr}}(\boldsymbol{\theta})}\frac{x_i^{\beta}}{b_i}\Big(\int_0^{\infty}\mathrm{e}^{-\nu}\frac{(\nu\phi(\gamma_k))^l}{l!}g_{\boldsymbol{b}}(\gamma_k)\mathrm{d}\gamma_k\Big).\end{aligned}$$

Hence, the posterior of the number of undiscovered fields becomes

$$\mathbb{P}[N - k = l|\mathcal{F}_{\underline{t}}] = \frac{\int_0^{\infty}\mathrm{e}^{-\nu}\frac{(\nu\phi(\gamma_k))^l}{l!}g_{\boldsymbol{b}}(\gamma_k)\mathrm{d}\gamma_k}{\sum_{l=0}^{\infty}\int_0^{\infty}\mathrm{e}^{-\nu}\frac{(\nu\phi(\gamma_k))^l}{l!}g_{\boldsymbol{b}}(\gamma_k)\mathrm{d}\gamma_k}.$$

By combining this distribution with the conditional distribution of $\Gamma_k|(\boldsymbol{X}, N-k)$, we can derive a specific refactoring of the joint distribution of $(\Gamma_k, N-k)|\mathcal{F}_{\underline{t}}$, which shows that $(N-k)|\mathcal{F}_{\underline{t}}$ actually has a mixed Poisson posterior. The formulas we derive in Lemma 3 for $\mathbb{E}[(N-k)^1|\mathcal{F}_{\underline{t}}]$ and $g^{(1)}(\gamma_k|\mathcal{F}_{\underline{t}})$ have already been derived by Lee (2008) (see equations (A.72)-(A.74) on pp. 195f.), while the refactoring in Lemma 2 has not been stated explicitly in the literature, at least to our knowledge.

Lemma 1. *It holds that*

$$\sum_{l=0}^{\infty}\int_0^{\infty}\mathrm{e}^{-\nu}\frac{(\nu\phi(\gamma_k))^l}{l!}g_{\boldsymbol{b}}(\gamma_k)\mathrm{d}\gamma_k = \int_0^{\infty}\mathrm{e}^{\nu(\phi(\gamma_k)-1)}g_{\boldsymbol{b}}(\gamma_k)\mathrm{d}\gamma_k. \tag{A.2}$$

Proof of Lemma 1: This follows directly from the definition of the exponential function.□

Lemma 2. *The joint distribution $g^{(0)}(\gamma_k|\boldsymbol{x}, l)\cdot\mathbb{P}[N-k=l|\mathcal{F}_{\underline{t}}]$ can be refactored as a continuous mixture of Poisson distributions. In particular,*

$$\begin{aligned} g^{(0)}(\gamma_k|\boldsymbol{x}, l)\cdot\mathbb{P}[N-k=l|\mathcal{F}_{\underline{t}}] &= \mathbb{P}[N-k=l|\Gamma_k=\gamma_k]\cdot g^{(0)}(\gamma_k|\mathcal{F}_{\underline{t}}),\\ \text{where } \mathbb{P}[N-k=l|\Gamma_k=\gamma_k] &= \frac{(\nu\phi(\gamma_k))^l}{l!}\mathrm{e}^{-\nu\phi(\gamma_k)},\\ \text{and } g^{(0)}(\gamma_k|\mathcal{F}_{\underline{t}}) &= \frac{\mathrm{e}^{\nu(\phi(\gamma_k)-1)}g_{\boldsymbol{b}}(\gamma_k)}{\int_0^{\infty}\mathrm{e}^{\nu(\phi(\gamma_k)-1)}g_{\boldsymbol{b}}(\gamma_k)\mathrm{d}\gamma_k}.\end{aligned} \tag{A.3}$$

Proof of Lemma 2:

$$\begin{aligned} g^{(0)}(\gamma_k|\boldsymbol{x}, l)\mathbb{P}[N-k=l|\mathcal{F}_{\underline{t}}] &= \left(\frac{\phi(\gamma_k)^l g_{\boldsymbol{b}}(\gamma_k)}{\int_0^{\infty}\phi(\gamma_k)^l g_{\boldsymbol{b}}(\gamma_k)\mathrm{d}\gamma_k}\right)\cdot\left(\frac{\int_0^{\infty}\mathrm{e}^{-\nu}\frac{(\nu\phi(\gamma_k))^l}{l!}g_{\boldsymbol{b}}(\gamma_k)\mathrm{d}\gamma_k}{\int_0^{\infty}\mathrm{e}^{\nu(\phi(\gamma_k)-1)}g_{\boldsymbol{b}}(\gamma_k)\mathrm{d}\gamma_k}\right)\\ &= \phi(\gamma_k)^l g_{\boldsymbol{b}}(\gamma_k)\cdot\frac{\mathrm{e}^{-\nu}\frac{\nu^l}{l!}}{\int_0^{\infty}\mathrm{e}^{\nu(\phi(\gamma_k)-1)}g_{\boldsymbol{b}}(\gamma_k)\mathrm{d}\gamma_k}\\ &= \frac{(\nu\phi(\gamma_k))^l}{l!}\cdot\frac{\mathrm{e}^{-\nu}g_{\boldsymbol{b}}(\gamma_k)}{\int_0^{\infty}\mathrm{e}^{\nu(\phi(\gamma_k)-1)}g_{\boldsymbol{b}}(\gamma_k)\mathrm{d}\gamma_k}\\ &= \left(\frac{(\nu\phi(\gamma_k))^l}{l!}\mathrm{e}^{-\nu\phi(\gamma_k)}\right)\cdot\left(\frac{\mathrm{e}^{\nu(\phi(\gamma_k)-1)}g_{\boldsymbol{b}}(\gamma_k)}{\int_0^{\infty}\mathrm{e}^{\nu(\phi(\gamma_k)-1)}g_{\boldsymbol{b}}(\gamma_k)\mathrm{d}\gamma_k}\right).\end{aligned} \tag{A.4}$$

The formula for $g^{(0)}(\gamma_k|\boldsymbol{x}, n-k)$ is given in equation (3.14), and the posterior for $N-k$ was derived above. We then apply Lemma 1 in the denominator, cancel the integrals $\int_0^\infty \phi(\gamma_k)^l g_{\boldsymbol{b}}(\gamma_k)\mathrm{d}\gamma_k$, reorder the terms, and finally expand by $\mathrm{e}^{-\nu\phi(\gamma_k)}$. □

Lemma 3. *For $m = 1, 2$, it holds for the posterior moments $\mathbb{E}[(N-k)^m|\mathcal{F}_{\underline{t}}]$ that*

$$\begin{aligned}\mathbb{E}[(N-k)^1|\mathcal{F}_{\underline{t}}] &= \frac{\int_0^\infty \nu\phi(\gamma_k)\mathrm{e}^{\nu(\phi(\gamma_k)-1)}g_{\boldsymbol{b}}(\gamma_k)\mathrm{d}\gamma_k}{\int_0^\infty \mathrm{e}^{\nu(\phi(\gamma_k)-1)}g_{\boldsymbol{b}}(\gamma_k)\mathrm{d}\gamma_k},\\ \mathbb{E}[(N-k)^2|\mathcal{F}_{\underline{t}}] &= \frac{\int_0^\infty \nu\phi(\gamma_k)\big(1+\nu\phi(\gamma_k)\big)\mathrm{e}^{\nu(\phi(\gamma_k)-1)}g_{\boldsymbol{b}}(\gamma_k)\mathrm{d}\gamma_k}{\int_0^\infty \mathrm{e}^{\nu(\phi(\gamma_k)-1)}g_{\boldsymbol{b}}(\gamma_k)\mathrm{d}\gamma_k}.\end{aligned} \tag{A.5}$$

Also, for $m = 1, 2$, it holds for the densities

$$g^{(m)}(\gamma_k|\mathcal{F}_{\underline{t}}) = \Big(\sum_{l=0}^{\infty} l^m \mathbb{P}[N-k=l|\Gamma_k=\gamma_k]g^{(0)}(\gamma_k|\mathcal{F}_{\underline{t}})\Big)\Big/\mathbb{E}[(N-k)^m|\mathcal{F}_{\underline{t}}]$$

that

$$\begin{aligned}g^{(1)}(\gamma_k|\mathcal{F}_{\underline{t}}) &= \frac{\nu\phi(\gamma_k)\mathrm{e}^{\nu(\phi(\gamma_k)-1)}g_{\boldsymbol{b}}(\gamma_k)}{\int_0^\infty \nu\phi(\gamma_k)\mathrm{e}^{\nu(\phi(\gamma_k)-1)}g_{\boldsymbol{b}}(\gamma_k)\mathrm{d}\gamma_k},\\ g^{(2)}(\gamma_k|\mathcal{F}_{\underline{t}}) &= \frac{\nu\phi(\gamma_k)\big(1+\nu\phi(\gamma_k)\big)\mathrm{e}^{\nu(\phi(\gamma_k)-1)}g_{\boldsymbol{b}}(\gamma_k)}{\int_0^\infty \nu\phi(\gamma_k)\big(1+\nu\phi(\gamma_k)\big)\mathrm{e}^{\nu(\phi(\gamma_k)-1)}g_{\boldsymbol{b}}(\gamma_k)\mathrm{d}\gamma_k}.\end{aligned} \tag{A.6}$$

Proof of Lemma 3: (A.5) follows from the properties of the $\mathrm{Po}(\lambda)$-distribution which has mean λ and second moment $\lambda(1+\lambda)$, and then applying the law of iterated expectations. For (A.6), note that integrating $\sum_{l=0}^\infty l^m \mathbb{P}[N-k=l|\Gamma_k=\gamma_k]g^{(0)}(\gamma_k|\mathcal{F}_{\underline{t}})$ over γ_k would yield precisely the m^{th} posterior moment as in (A.5). Hence the numerators in (A.6) are just the numerators in (A.5) without the integral. Dividing by $\mathbb{E}[(N-k)^m|\mathcal{F}_{\underline{t}}]$ is the normalization required such that $g^{(m)}(\gamma_k|\mathcal{F}_{\underline{t}})$ becomes a density. □

Lemma 4. *Consider any continuous transformation of Γ_k, denoted by $t^*(\Gamma_k)$. For $m = 1, 2$, provided that $\int_0^\infty t^*(\gamma_k)g^{(m)}(\gamma_k|\mathcal{F}_{\underline{t}})\mathrm{d}\gamma_k$ exists, it holds that*

$$\mathbb{E}[(N-k)^m t^*(\Gamma_k)|\mathcal{F}_{\underline{t}}] = \mathbb{E}[(N-k)^m|\mathcal{F}_{\underline{t}}]\int_0^\infty t^*(\gamma_k)g^{(m)}(\gamma_k|\mathcal{F}_{\underline{t}})\mathrm{d}\gamma_k. \tag{A.7}$$

Proof of Lemma 4: We write out the expectation with respect to $(N-k)|\Gamma_k$ and $\Gamma_k|\mathcal{F}_{\underline{t}}$ as factored in (A.3), expand by $\mathbb{E}[(N-k)^m|\mathcal{F}_{\underline{t}}]$, and then recognize the formula of $g^{(m)}(\gamma_k|\mathcal{F}_{\underline{t}})$. □

A.2 Numerical Calculation of the General Gamma Density

To simplify the notation in this section, we drop the index k in γ_k. A draw from the general gamma distribution $g_{\boldsymbol{b}}(\gamma)$ with parameters $\boldsymbol{b} = (b_1, ..., b_k)$ and support $[0, \infty)$ as defined in section 3.2 is easily obtained by the sum $\sum_{j=1}^k \varepsilon_j/b_j$, where ε_j are iid standard exponential variates. Thus, one can approximate the density $g_{\boldsymbol{b}}(\gamma)$ by Monte Carlo methods, and one could try to approximate the integrals of the type $\int_0^\infty \mathrm{e}^{\nu(\phi(\gamma)-1)}g_{\boldsymbol{b}}(\gamma)\mathrm{d}\gamma$ or $\int_0^\infty \phi(\gamma)^{n-k}g_{\boldsymbol{b}}(\gamma)\mathrm{d}\gamma$ by Monte Carlo integration. However, the integrands $\mathrm{e}^{-\nu[1-\phi(\gamma)]}$ or $\phi(\gamma)^{n-k}$ exhibit a very sharp decline after zero due to the large exponents ν and $n-k$.[1]

[1]The Laplace transform $\phi(\gamma)$ fulfills $\phi(0) = 1$, is convex, and declines monotonously towards the asymptote $\lim_{\gamma\to\infty}\phi(\gamma) = 0$. This carries over to the integrand $\phi(\gamma)^{n-k}$. The integrand $\mathrm{e}^{-\nu[1-\phi(\gamma)]}$ also starts at 1, is convex, and declines monotonously towards the asymptote $\mathrm{e}^{-\nu}$. Both integrands exhibit an extremely sharp decline in the proximity of zero.

Thus, a highly precise evaluation of $g_{\boldsymbol{b}}(\gamma)$ is required at its left tail in the proximity of zero. Since Monte Carlo methods do not perform well at approximating the tails of a distribution, they have not been used in the literature here to our knowledge.

The pdf $g_{\boldsymbol{b}}(\gamma)$ does have an analytical expression, which can be obtained via partial fractions expansion from its real-valued Laplace transform as given by $\psi_{\boldsymbol{b}}(s) = \prod_{j=1}^{k} b_j/(b_j + s)$. The resulting expression is (see McGill and Gibbon, 1965, pp. 4f.; Barouch and Kaufman, 1976, p. 13)

$$g_{\boldsymbol{b}}(\gamma) = \sum_{j=1}^{k} C_j b_j \mathrm{e}^{-\gamma b_j}, \text{ where } C_j = \prod_{l=1, l\neq j}^{k} \frac{b_l}{b_l - b_j} \text{ for } j = 1, \ldots, k.$$

However, this expression is practically useless for accurately evaluating the density, in particular at the important values near zero (Nair and Wang, 1989, pp. 430f.). The problem with this formula is that already for modest values of k the coefficients C_j, $j = 1, ..., k$, become extremely large (and are always alternating in sign) so that each coefficient would need to be calculated with numerically infeasible precision.

The approach chosen in the size-biased sampling literature to calculate the general gamma density is the Fourier-series method for numerically inverting the complex-valued Laplace transform of $g_{\boldsymbol{b}}(\gamma)$. Let $i = \sqrt{-1}$ denote the imaginary number, and let $s = a + i\omega$ be a complex number. The inversion integral $\psi_{\boldsymbol{b}}(s) \to g_{\boldsymbol{b}}(\gamma)$ of the Laplace transform $g_{\boldsymbol{b}}(\gamma) \to \psi_{\boldsymbol{b}}(s)$ is given by the following formula (Durbin, 1974, p. 371; Abate and Whitt, 1995, p. 37):

$$\begin{aligned} g_{\boldsymbol{b}}(\gamma) &= \frac{\mathrm{e}^{a\gamma}}{\pi} \int_0^\infty \big[\mathrm{Re}\{\psi_{\boldsymbol{b}}(a + i\omega)\}\cos\omega\gamma - \mathrm{Im}\{\psi_{\boldsymbol{b}}(a + i\omega)\}\sin\omega\gamma\big]\mathrm{d}\omega \\ &= \frac{2\mathrm{e}^{a\gamma}}{\pi} \int_0^\infty \mathrm{Re}\{\psi_{\boldsymbol{b}}(a + i\omega)\}\cos(\omega\gamma)\mathrm{d}\omega, \ \gamma \geq 0, \end{aligned} \tag{A.8}$$

where a can be an arbitrary number greater than the real parts of all singularities of the complex-valued Laplace transform, i.e. in this case $a > -b_k$.

Inverting Laplace transforms via the Fourier-series method means essentially that a trapezoidal rule is applied to either of the integrals from (A.8). It can be shown that for these integrals, certain trapezoidal rules are equivalent to Fourier-series approximations of the density on a certain compact interval. This makes the method particular effective as a numerical technique here (Durbin, 1974; Crump, 1976; Abate and Whitt, 1995).

The version of the Fourier-series method we found most useful and practical to implement for the problem of evaluating the general gamma density is the "Euler method", which is described in the context of probability applications by Abate and Whitt (1995), and which builds on earlier work by Dubner and Abate (1968) and Simon et al. (1972). In particular, for the "Euler method" a trapezoidal rule with a step-size of $h = \pi/2\gamma$ is applied in the second line of (A.8) (or equivalently with a step-size of $h = \pi/\gamma$ in the first line). Denoting the Fourier-series approximation of $g_{\boldsymbol{b}}(\gamma)$ by $\widetilde{g}_{\boldsymbol{b}}(\gamma)$, and also setting $a = A/2\gamma$, this yields for $\widetilde{g}_{\boldsymbol{b}}(\gamma)$ at each γ the nearly alternating series (Abate and Whitt, 1995, pp. 37f.)

$$\widetilde{g}_{\boldsymbol{b}}(\gamma) = \frac{\mathrm{e}^{A/2}}{\gamma}\Big[\frac{1}{2}\mathrm{Re}\big\{\psi_{\boldsymbol{b}}\big(\frac{A}{2\gamma}\big)\big\} + \sum_{m=1}^{\infty}(-1)^m \mathrm{Re}\big\{\psi_{\boldsymbol{b}}\big(\frac{A + i(2m\pi)}{2\gamma}\big)\big\}\Big]. \tag{A.9}$$

In the case of the general gamma distribution, the real part of the complex-valued Laplace transform is easily derived as

$$\mathrm{Re}\{\psi_{\boldsymbol{b}}(a+i\omega)\} = \Big(\prod_{j=1}^{k} \frac{b_j}{\sqrt{(b_j+a)^2+\omega^2}}\Big) \cdot \cos\Big(\sum_{j=1}^{k} \mathrm{atan}\big(\frac{\omega}{b_j+a}\big)\Big). \tag{A.10}$$

It can be shown generally that the discretization error of the approximation $\widetilde{g}_{\boldsymbol{b}}(\gamma)$, i.e. the difference between the series in (A.9) and the exact integral in (A.8), equals (Abate and Whitt, 1995, p. 38):

$$\widetilde{g}_{\boldsymbol{b}}(\gamma) - g_{\boldsymbol{b}}(\gamma) = \sum_{m=1}^{\infty} \mathrm{e}^{-mA} g_{\boldsymbol{b}}(\gamma(2m+1)). \tag{A.11}$$

The crux of the method is that the free parameter A can always be chosen such that the discretization error becomes as small as desired. For example, if we wanted a discretization error of exactly 10^{-E} at some γ, we would need to choose A such that the right-hand side of (A.11) equals 10^{-E}. Obviously, since $g_{\boldsymbol{b}}(\gamma(2m+1))$ is not known, this cannot be done. What can be done is bounding (A.11) by using a function which envelopes $g_{\boldsymbol{b}}(\gamma)$, and calculate A based on this envelope so that the discretization error is forced to be less than 10^{-E}.

However, the delicacy of the Fourier-series method is that there is a risk of setting A too high, which can lead to large error. The reason is that the discretization error (A.11) is not the only error, there is also always an error from truncating the series in (A.9). The term $\mathrm{e}^{A/2}$ in (A.9) reveals that this truncation error grows exponentially with A. To mitigate this problem, Euler summation for accelerating the convergence of the series in (A.9) has been proposed in the literature (Simon et al., 1972), which allows to approximate the limit of the nearly alternating series much more efficiently (Abate and Whitt, 1995), and which is where the name "Euler method" derives from. According to Abate and Whitt (1992; 1995), Euler summation for an alternating series is equivalent to a simple weighted average of the last ΔM partial sums, where the weights are from the binomial distribution with parameters ΔM and $p = 1/2$. Put more clearly, Euler summation of (A.9) amounts to calculating the terms inside the sum of (A.9) for $m = M, M+1, \ldots, M+\Delta M$, then calculating the corresponding partial sums while also adding the initial M terms, and finally averaging with the binomial probabilities.[2]

Note that one cannot evaluate (A.9) at multiple gridpoints at once, but can only evaluate it iteratively. This may be seen as a disadvantage of the method, but there is also a clear advantage of such an iterative evaluation: if we know that $g_{\boldsymbol{b}}(\gamma - \Delta\gamma) < g_{\boldsymbol{b}}(\gamma)$, we can exploit this monotony in order to bound the relative discretization error using $\frac{\widetilde{g}_{\boldsymbol{b}}(\gamma) - g_{\boldsymbol{b}}(\gamma)}{g_{\boldsymbol{b}}(\gamma)} < \frac{\widetilde{g}_{\boldsymbol{b}}(\gamma) - g_{\boldsymbol{b}}(\gamma)}{g_{\boldsymbol{b}}(\gamma-\Delta\gamma)}$, and replace the denominator by the approximation at the previous gridpoint, $\widetilde{g}_{\boldsymbol{b}}(\gamma - \Delta\gamma)$, provided that it has been calculated without large error.

To further mitigate the problem of choosing A too high, the remainder of this section is concerned with how to choose A as a function of the desired error E. The usual candidate for bounding $g_{\boldsymbol{b}}(\gamma)$ is its maximum. As discussed in McGill and Gibbon (1965) (pp. 6ff.), it is clear that $g_{\boldsymbol{b}}(\gamma)$ behaves globally similar as a usual gamma density, in particular that $g_{\boldsymbol{b}}(0) = 0$ (for $k \geq 2$), $g_{\boldsymbol{b}}(\infty) = 0$, and that it passes through a unique maximum. The

[2] Abate and Whitt (1995) (p. 38) propose to use $\Delta M = 11$. In our applications we always use $\Delta M = 11$, and use $M = 100$, increasing M to 500 for very small values of γ.

maximum is not derivable analytically, but can be easily approximated by a Monte Carlo-simulation.[3]

In the following, we propose our idea how to get an improved envelope for $g_{\boldsymbol{b}}(\gamma)$ at very small values of γ, so that at these values, the error parameter E need not be chosen with as much care as when using the constant bound only.[4] It is known about $g_{\boldsymbol{b}}(\gamma)$ that the first $k-2$ derivatives vanish at zero, and that the Taylor expansion around zero equals (Barouch and Kaufman, 1976, p. 13; Nair and Wang, 1989, p. 430)[5]

$$g_{\boldsymbol{b}}(\gamma) = \prod_{j=1}^{k} b_j \frac{\gamma^{k-1}}{(k-1)!} + O(\gamma^k). \tag{A.12}$$

Combining the Taylor approximation with the maximum of the pdf ($\overline{g}$) as an envelope for $g_{\boldsymbol{b}}(\gamma)$, we can bound the relative discretization error in (A.10) at γ as follows:

$$\frac{\widetilde{g}_{\boldsymbol{b}}(\gamma) - g_{\boldsymbol{b}}(\gamma)}{g_{\boldsymbol{b}}(\gamma)} = \sum_{m=1}^{\infty} \mathrm{e}^{-mA} \frac{g_{\boldsymbol{b}}(\gamma(2m+1))}{g_{\boldsymbol{b}}(\gamma)} \leq \sum_{m=1}^{\infty} \mathrm{e}^{-mA} \min\Big\{(2m+1)^{k-1}, \frac{\overline{g}}{\widetilde{g}_{\boldsymbol{b}}(\gamma-\Delta\gamma)}\Big\}.$$

If we let M^* be the last index where the Taylor approximation yields the smaller bound (and set M^* to zero if this does not happen at all), we get that

$$M^* = \max\Big\{0, \Big\lfloor -\frac{1}{2} + \frac{1}{2}\Big(\frac{\overline{g}}{\widetilde{g}_{\boldsymbol{b}}(\gamma-\Delta\gamma)}\Big)^{\frac{1}{k-1}} \Big\rfloor\Big\},$$

and

$$\frac{\widetilde{g}_{\boldsymbol{b}}(\gamma) - g_{\boldsymbol{b}}(\gamma)}{g_{\boldsymbol{b}}(\gamma)} \leq \begin{cases} \sum_{m=1}^{M^*} \mathrm{e}^{-mA}(2m+1)^{k-1} + \mathrm{e}^{-(M^*+1)A} \frac{\overline{g}}{\widetilde{g}_{\boldsymbol{b}}(\gamma-\Delta\gamma)}, & M^* \geq 1, \\ \mathrm{e}^{-A} \frac{\overline{g}}{\widetilde{g}_{\boldsymbol{b}}(\gamma-\Delta\gamma)}, & M^* = 0, \end{cases} \tag{A.13}$$

where we omit the factor $1/(1-\mathrm{e}^{-A})$ from summing the geometric series, which for all practical purposes can be replaced by one since e^{-A} will be small. For $M^* = 0$, the right-hand side of (A.13) is easily solved for A to achieve a relative error of 10^{-E}. For $M^* \geq 1$, we need a root-finding algorithm to determine A.[6] The issue we are left to discuss is which lower and upper bounds $A_{\min}$ and $A_{\max}$ to feed into the root-finding algorithm so that the right-hand side of (A.13) is above 10^{-E} at $A_{\min}$ and below 10^{-E} at $A_{\max}$. Starting with $A_{\min}$, the right-hand side of (A.13) is obviously larger than its first summand, $\mathrm{e}^{-A_{\min}+(k-1)\ln 3}$, which is greater or equal to 1 already if $A_{\min} \leq (k-1)\ln 3 \approx k-1$. For this reason we use $A_{\min} = k-1$ in our routine. Continuing with $A_{\max}$, note that $0 > -1 \cdot A_{\max} + (k-1)\ln(2\cdot 1+1) > -2 \cdot A_{\max} + (k-1)\ln(2\cdot 2+1) > \ldots$ holds already if $A_{\max} > (k-1)\ln 3 \approx A_{\min}$. This implies that the sum over $m = 1, \ldots, M^*$ in (A.13) is always less than $M^* \mathrm{e}^{-A_{\max}+(k-1)\ln 3}$, and that the right-hand side of (A.13) is always less than $2 \cdot \max\big\{M^* \mathrm{e}^{-A_{\max}+(k-1)\ln 3}, \mathrm{e}^{-(1+M^*)A_{\max}} \frac{\overline{g}}{\widetilde{g}_{\boldsymbol{b}}(\gamma-\Delta\gamma)}\big\}$. Hence, we know that the right-hand side of (A.13) remains below 10^{-E} if we choose $A_{\max}$ as

$$A_{\max} = \max\Big\{\ln(2M^*) + (k-1)\ln 3 + E\ln 10, \frac{1}{1+M^*}\big(\ln(2\overline{g}) - \ln\widetilde{g}_{\boldsymbol{b}}(\gamma-\Delta\gamma) + E\ln 10\big)\Big\}.$$

[3] For example, one may compute a kernel estimate of the density at its median, which usually is not far away from the mode since $g_{\boldsymbol{b}}(\gamma)$ is sufficiently symmetric already for modest values of k.

[4] In our calculations, we used a five-digit accuracy (i.e. $E = 5$) which worked well throughout all values of γ.

[5] Note that this can be shown by evaluating $k-1$ derivatives of equation (5c) in McGill and Gibbon (1965) at zero.

[6] We use the R-method "uniroot". We implement this by solving the natural logarithm of (A.13) for $-E\ln 10$.

A.3 Further Derivations for the Size-Biased Sampling Model

We now prepare a derivation of the induced order statistics representation of $(X_1, \ldots, X_n)$ with the following Lemma, whereby we also give a proof of the crucial property from equation (3.12), which is the key equation that allows to use the general gamma density in the size-biased sampling model. To give a concise notation for order statistics as in equation (3.20), we introduce the following sets: let $\Delta_n(c)$ be the unordered, and be $\Delta_n^*(a,b)$ the ordered, n-dimensional, open simplex over the intervals $(0,c)$ and (a,b), respectively:[7]

$$\begin{aligned}\Delta_n(c) =& \big\{(\epsilon_1, \ldots, \epsilon_n) \in \mathbb{R}^n \big| \epsilon_1, \ldots, \epsilon_n > 0,\ \sum_{i=1}^n \epsilon_i < c\big\},\\ \Delta_n^*(a,b) =& \big\{(\gamma_1, \ldots, \gamma_n) \in \mathbb{R}^n \big| a < \gamma_1 < \gamma_2 < \ldots < \gamma_n < b\big\}.\end{aligned}$$

Note that with $c = b-a$, the latter set is a coordinate transform of the former set resulting as the partial sums plus a shift by a.

Lemma 5. *Let $w_1, \ldots, w_k > 0$, and $s \geq 0$. Upon defining the partial sums $b_i = w_i + \ldots + w_k$ for each $i = 1, \ldots, k$, the term $\prod_{i=1}^k (b_i + s)^{-1}$ has the two integral representations*

$$\begin{aligned}\prod_{i=1}^k (b_i+s)^{-1} &= \int_0^\infty \int_{\gamma_1}^\infty \cdots \int_{\gamma_{k-1}}^\infty \mathrm{e}^{-\gamma_k s} \prod_{i=1}^k \mathrm{e}^{-\gamma_i w_i} \mathrm{d}\gamma_i,\\ \prod_{i=1}^k (b_i+s)^{-1} &= \prod_{i=1}^k b_i^{-1} \int_0^\infty \mathrm{e}^{-\gamma_k s} g_{\boldsymbol{b}}(\gamma_k) \mathrm{d}\gamma_k,\end{aligned} \tag{A.14}$$

where $g_{\boldsymbol{b}}(\gamma_k)$ is the general gamma density with parameters $\boldsymbol{b} = (b_1, \ldots, b_k)$.

Proof of Lemma 5:

First, recall the straight-forward property of the exponential function that $\int_0^\infty \mathrm{e}^{-\epsilon(b+s)} \mathrm{d}\epsilon = (b+s)^{-1}$. Upon applying this for all $i = 1, \ldots, k$, we get

$$\begin{aligned}\prod_{i=1}^k (b_i+s)^{-1} &= \prod_{i=1}^k \Big[\int_0^\infty \mathrm{e}^{-\epsilon_i (b_i+s)} \mathrm{d}\epsilon_i \Big]\\ &= \int_0^\infty \cdots \int_0^\infty \mathrm{e}^{-(\epsilon_1 + \ldots + \epsilon_k)s} \mathrm{e}^{-(\epsilon_1 b_1 + \ldots + \epsilon_k b_k)} \mathrm{d}\epsilon_1 \cdots \mathrm{d}\epsilon_k,\end{aligned} \tag{A.15}$$

which follows directly from the distributive property.

To proceed, we rearrange the second exponent from (A.15) by applying a version of Abel's summation by parts, which is the discrete analogue to integration by parts as applicable to sequences.[8] Upon defining $\gamma_i = \epsilon_1 + \ldots + \epsilon_i$ for each $i = 1, \ldots, k$, the summation by parts formula is expressible as (see, e.g., Königsberger (2003), p. 305)

$$\begin{aligned}&\epsilon_1(w_1 + \ldots + w_k) + \epsilon_2(w_2 + \ldots + w_k) + \ldots + \epsilon_k w_k\\ &= \gamma_1 b_1 + (\gamma_2 - \gamma_1) b_2 + \ldots + (\gamma_k - \gamma_{k-1}) b_k\\ &= \gamma_1 (b_1 - b_2) + \ldots + \gamma_{k-1}(b_{k-1} - b_k) + \gamma_k b_k\\ &= \epsilon_1 w_1 + \ldots + (\epsilon_1 + \ldots + \epsilon_{k-1}) w_{k-1} + (\epsilon_1 + \ldots + \epsilon_k) w_k.\end{aligned} \tag{A.16}$$

[7] Since in this paper all rvs are continuous, we restrict without loss of generality the sets to be open and not closed, i.e. the inequalities to be strict instead of weak inequalities.

[8] For this case of finite-dimensional vectors, the summation by parts formula is also easily derived by simple matrix algebra: First, write the vector $(b_1, \ldots, b_k)^\top$ as the product of the upper triangular matrix

Then, the multivariate substitution $(\epsilon_1, \ldots, \epsilon_k) \to (\gamma_1, \ldots, \gamma_k)$ yields via the substitution rule for multivariate mappings[9]

$$\begin{aligned}\prod_{i=1}^{k}(b_i+s)^{-1} &= \int_0^\infty \cdots \int_0^\infty \mathrm{e}^{-(\epsilon_1+\ldots+\epsilon_k)s}\mathrm{e}^{-\{\epsilon_1 w_1+\ldots+(\epsilon_1+\ldots+\epsilon_k)w_k\}}\mathrm{d}\epsilon_1\cdots\mathrm{d}\epsilon_k \\ &= \int_0^\infty \int_{\gamma_1}^\infty \cdots \int_{\gamma_{k-1}}^\infty \mathrm{e}^{-\gamma_k s}\mathrm{e}^{-(\gamma_1 w_1+\ldots+\gamma_k w_k)}\mathrm{d}\gamma_1\cdots\mathrm{d}\gamma_k.\end{aligned}$$

To prove the second part, we proceed from (A.15) directly by substituting $(\epsilon_1, \ldots, \epsilon_{k-1}, \epsilon_k) \to (\epsilon_1, \ldots, \epsilon_{k-1}, \gamma_k)$, i.e. we substitute only $\gamma_k = \epsilon_1 + \ldots + \epsilon_k$:

$$\begin{aligned}\prod_{i=1}^{k}(b_i+s)^{-1} &= \int_0^\infty \int_0^\infty \cdots \int_0^\infty \mathbb{1}(\epsilon_1+\ldots+\epsilon_{k-1}<\gamma_k)\mathrm{e}^{-\gamma_k s}\mathrm{e}^{-\{\epsilon_1 b_1+\ldots+(\gamma_k-\epsilon_1-\ldots-\epsilon_{k-1})b_k\}}\mathrm{d}\epsilon_1\cdots\mathrm{d}\epsilon_{k-1}\mathrm{d}\gamma_k \\ &= \prod_{i=1}^{k} b_i^{-1}\int_0^\infty \mathrm{e}^{-\gamma_k s}\Big\{\int_{\Delta_{k-1}(\gamma_k)}\Big(\prod_{i=1}^{k-1} b_i\mathrm{e}^{-\epsilon_i b_i}\Big)b_k\mathrm{e}^{-(\gamma_k-\epsilon_1-\ldots-\epsilon_{k-1})b_k}\mathrm{d}\epsilon_1\cdots\mathrm{d}\epsilon_{k-1}\Big\}\mathrm{d}\gamma_k.\end{aligned}$$

The proof is completed by noting that the integral in the curly brackets is precisely the convolution of k exponential densities with parameters $b_1, \ldots, b_k$. □

Both representations from Lemma 5 have probabilistic interpretations which are known in the size-biased sampling literature. Starting with the second representation, multiplying by $\prod_{i=1}^k b_i^{-1}$ and setting $s = \tilde{x}_{k+1}^\beta + \ldots + \tilde{x}_n^\beta$ proves the first line of equation (3.12). The probabilistic derivation of equation (3.12) as found, for example, in Nair and Wang (1989) (p. 427) and Lee (2008) (p. 179) is as follows: $b_i/(b_i+s)$ is the Laplace transform of an $\exp(b_i)$-distributed rv, expressible as ε_i/b_i, where ε_i is standard exponential. The product $\prod_{i=1}^k b_i/(b_i+s)$, then, is the Laplace transform of $\varepsilon_1/b_1 + \ldots + \varepsilon_k/b_k$ under independence, i.e. of a rv which has the general gamma density with parameters $b_1, \ldots, b_k$. We state the probabilistic interpretation of the first representation in the next subsection after applying it to rearrange the joint pdf of $(X_1, ..., X_N)|N$.

---

filled with 1s and $(w_1, \ldots, w_k)^\top$. Then, rearrange the sum $\epsilon_1 b_1 + \ldots + \epsilon_k b_k$ as

$$\begin{aligned}&(\epsilon_1, \ldots, \epsilon_k)\cdot(w_1+\ldots+w_k, \ldots, w_k)^\top \\ &= (\epsilon_1, \ldots, \epsilon_k)\cdot\left(\begin{pmatrix}1 & \cdots & 1\\ \vdots & \ddots & \vdots\\ 0 & \cdots & 1\end{pmatrix}\cdot\begin{pmatrix}w_1\\ \vdots\\ w_k\end{pmatrix}\right) = \left(\begin{pmatrix}1 & \cdots & 0\\ \vdots & \ddots & \vdots\\ 1 & \cdots & 1\end{pmatrix}\cdot\begin{pmatrix}\epsilon_1\\ \vdots\\ \epsilon_k\end{pmatrix}\right)^\top\cdot\begin{pmatrix}w_1\\ \vdots\\ w_k\end{pmatrix} \\ &= (\epsilon_1, \ldots, \epsilon_1+\ldots+\epsilon_k)\cdot(w_1, \ldots, w_k)^\top.\end{aligned}$$

[9]Note from the last footnote that the Jacobian of this mapping is the lower triangular matrix of 1s, so that the Jacobian and its inverse have a constant determinant of 1.

A priori: the joint distribution of all field sizes in the order of discovery

We now use the first integral representation from Lemma 5 with $w_i = x_i^\beta$, $i = 1, \ldots, n$, and $s = 0$ to rearrange equation (3.7). Factoring all terms inside the integrals, we get

$$\begin{aligned}\mathbb{P}[X_1 \in \mathrm{d}x_1, \ldots, X_n \in \mathrm{d}x_n | N = n] &= n! \prod_{i=1}^{n} \frac{x_i^\beta}{x_i^\beta + \ldots + x_n^\beta} f(x_i)\mathrm{d}x_i \\ &= \int_0^\infty \int_{\gamma_1}^\infty \cdots \int_{\gamma_{n-1}}^\infty n! \prod_{i=1}^{n} x_i^\beta \mathrm{e}^{-\gamma_i x_i^\beta} f(x_i)\mathrm{d}x_i \mathrm{d}\gamma_i.\end{aligned} \tag{A.17}$$

Following Pitman and Tran (2015) (pp. 2489ff.), (A.17) can be summarized as follows: let $\tilde{X}_1, \ldots, \tilde{X}_n$ be iid as in assumption (1), $\varepsilon_1, \ldots, \varepsilon_n$ be iid standard exponential rvs, where the ε_i are also independent of the $\tilde{X}_i$, and define $\tilde{\Gamma}_i = \varepsilon_i / \tilde{X}_i^\beta$. By the fact that $\big((\tilde{X}_i, \tilde{\Gamma}_i), i = 1, \ldots, n\big)$ form n iid pairs from the joint pdf $x^\beta \mathrm{e}^{-\gamma x^\beta} f(x)$, and noting that $n! = |\mathcal{I}_n|$, we can restate (A.17) in probabilistic terms as

$$\begin{aligned}&\mathbb{P}[X_1 \in \mathrm{d}x_1, \ldots, X_n \in \mathrm{d}x_n | N = n] \\ &= \int_0^\infty \int_{\gamma_1}^\infty \cdots \int_{\gamma_{n-1}}^\infty \sum_{(i(1),\ldots,i(n)) \in \mathcal{I}_n} \mathbb{P}\big[\tilde{\Gamma}_{i(1)} \in \mathrm{d}\gamma_1, \tilde{X}_{i(1)} \in \mathrm{d}x_1, \ldots, \tilde{\Gamma}_{i(n)} \in \mathrm{d}\gamma_n, \tilde{X}_{i(n)} \in \mathrm{d}x_n\big] \\ &= \mathbb{P}\Bigg[\bigcup_{(i(1),\ldots,i(n)) \in \mathcal{I}_n} \{\tilde{\Gamma}_{i(1)} < \ldots < \tilde{\Gamma}_{i(n)}, \tilde{X}_{i(1)} \in \mathrm{d}x_1, \ldots, \tilde{X}_{i(n)} \in \mathrm{d}x_n\}\Bigg].\end{aligned} \tag{A.18}$$

For a fixed set of size-measures, $\{x_1, \ldots, x_n\}$, this interpretation was first proved by Gordon (1983). (A.18) expresses the probability that the $\tilde{X}_i$ which becomes the first size-biased pick has the smallest $\tilde{\Gamma}_i$, the one which becomes the second size-biased pick has the second smallest $\tilde{\Gamma}_i$, etc., which means precisely that $(X_1, \ldots, X_n)$ are the induced order statistics or concomitants of $(\tilde{\Gamma}_1, \ldots, \tilde{\Gamma}_n)$ (Pitman and Tran, 2015, p. 2489).

We proceed by expanding in (A.17) each term by (minus one times) the derivative of the Laplace transform.. Due to $-\phi'(\gamma) = \int_{\underline{x}}^{\overline{x}} x^\beta \mathrm{e}^{-\gamma x^\beta} f(x)\mathrm{d}x$, this expansion is mathematically equivalent to refactoring the joint density of $(\tilde{X}_i, \tilde{\Gamma}_i)$ according to Bayes' rule. Then, the joint pdf of $(X_1, \ldots, X_n)$ given $N = n$ can be restated as

$$\int_0^\infty \int_{\gamma_1}^\infty \cdots \int_{\gamma_{n-1}}^\infty \left\{ \prod_{i=1}^{n} \frac{x_i^\beta \mathrm{e}^{-\gamma_i x_i^\beta} f(x_i)\mathrm{d}x_i}{-\phi'(\gamma_i)} \right\} \cdot \left\{ n! \prod_{i=1}^{n} (-\phi'(\gamma_i))\mathrm{d}\gamma_i \right\}. \tag{A.19}$$

We denote the pdf from the first bracket by $h(x_i|\gamma_i) \propto x_i^\beta \mathrm{e}^{-\gamma_i x_i^\beta} f(x_i)$. The second bracket, together with the integral boundaries, shows that $(\Gamma_1, \ldots, \Gamma_n)$ are jointly distributed as (increasing) order statistics from the pdf $-\phi'(\gamma)$ (with cdf $1 - \phi(\gamma)$). Pitman and Tran (2015) (p. 2490) go further to express (A.19) in terms of decreasing uniform order statistics, which immediately follows since $\phi(\gamma)$ is a complementary cdf. Making the corresponding substitution $u_i = \phi(\gamma_i)$, (A.19) becomes

$$\int_0^1 \int_0^{u_1} \cdots \int_0^{u_{n-1}} \left\{ \prod_{i=1}^{n} \frac{x_i^\beta \mathrm{e}^{-\phi^{-1}(u_i) x_i^\beta} f(x_i)\mathrm{d}x_i}{-\phi'(\phi^{-1}(u_i))} \right\} \cdot \left\{ n! \prod_{i=1}^{n} \mathrm{d}u_i \right\}. \tag{A.20}$$

Overall, (A.17)-(A.20) imply two further procedures with which Monte Carlo draws for $(X_1, \ldots . X_n)$ can be obtained. The first one is to sample $(\tilde{X}_1, \ldots, \tilde{X}_n)$ randomly from f, then to draw $\tilde{\Gamma}_i|\tilde{X}_i$ as $\tilde{\Gamma}_i = \varepsilon_i/\tilde{X}_i^{\beta}$, and then to order the sequence as $\tilde{\Gamma}_{i(1)} < \ldots < \tilde{\Gamma}_{i(n)}$, from which one obtains $X_1 = \tilde{X}_{i(1)}, \ldots, X_n = \tilde{X}_{i(n)}$. The second one is to invert decreasing uniform order statistics via $u \stackrel{!}{=} \phi(\gamma)$, yielding $(\Gamma_1, \ldots, \Gamma_n)$, and then draw each $X_i|\Gamma_i$ independently from the h-density. For the weight function $w(x) = x^{\beta}$ we have found a simple accept/reject procedure to draw from the h-density, which we will explain below eq. (A.23).

A posteriori: the joint distribution of the remaining field sizes in the order of discovery

We now derive a posterior version of these equations for the distribution of the remaining field sizes $(X_{k+1}, \ldots, X_N)$ given $(X_1, \ldots, X_k, N)$. In (A.17-A.19), we interchange the order of integration so that the integral over γ_k is the outermost integral, and the integrals over $\gamma_1, \ldots, \gamma_k$ and $\gamma_{k+1}, \ldots, \gamma_n$ are each pooled inside. Then, (A.17-A.19) becomes

$$n! \int_0^{\infty} \Big\{ p_{(<)}(\boldsymbol{x}; \gamma_k) \mathrm{d}x_1 \cdots \mathrm{d}x_k \Big\} \cdot \int\limits_{\Delta^*_{n-k}(\gamma_k, \infty)} \Big\{ \prod_{i=k+1}^{n} \frac{x_i^{\beta} \mathrm{e}^{-\gamma_i x_i^{\beta}} f(x_i) \mathrm{d}x_i}{-\phi'(\gamma_i)} \Big\} \cdot \Big\{ \prod_{i=k+1}^{n} (-\phi'(\gamma_i)) \mathrm{d}\gamma_i \Big\} \cdot \mathrm{d}\gamma_k,$$

where we define $p_{(<)}(\boldsymbol{x}; \gamma_k)$ as the density of $(X_1, \ldots, X_k)$ jointly with the probability that the mixing rvs with index below k are smaller than the value γ_k, i.e.

$$p_{(<)}(\boldsymbol{x}; \gamma_k) = \int\limits_{\Delta^*_{k-1}(0, \gamma_k)} \prod_{i=1}^{k} x_i^{\beta} \mathrm{e}^{-\gamma_i x_i^{\beta}} f(x_i) \mathrm{d}\gamma_i.$$

We rearrange the integral over the ordered set $\Delta^*_{k-1}(0, \gamma_k)$ by applying Abel's formula (A.16) with $b_i = x_i^{\beta} + \ldots + x_k^{\beta}$, $i = 1, \ldots, k$. That is, we substitute back $(\gamma_1, \ldots, \gamma_{k-1}, \gamma_k) \to (\epsilon_1, \ldots, \epsilon_{k-1}, \gamma_k)$ to rearrange the $p_{(<)}$-term as in the second integral representation of Lemma 5 (with $s = 0$ and without the integral over γ_k which is kept fixed here):

$$\begin{aligned} p_{(<)}(\boldsymbol{x}; \gamma_k) &= \Big\{ \prod_{i=1}^{k} x_i^{\beta} f(x_i) \Big\} \cdot \Big\{ \int\limits_{\Delta^*_{k-1}(0, \gamma_k)} \mathrm{e}^{-(\gamma_1 x_1^{\beta} + \ldots + \gamma_k x_k^{\beta})} \mathrm{d}\gamma_1 \ldots \mathrm{d}\gamma_{k-1} \Big\} \\ &= \Big\{ \prod_{i=1}^{k} \frac{x_i^{\beta}}{b_i} f(x_i) \Big\} \cdot \Big\{ \int\limits_{\Delta_{k-1}(\gamma_k)} \Big(\prod_{i=1}^{k-1} b_i \mathrm{e}^{-\epsilon_i b_i} \Big) b_k \mathrm{e}^{-(\gamma_k - \epsilon_1 - \ldots - \epsilon_{k-1}) b_k} \mathrm{d}\epsilon_1 \cdots \mathrm{d}\epsilon_{k-1} \Big\} \\ &= \Big\{ \prod_{i=1}^{k} \frac{x_i^{\beta}}{b_i} f(x_i) \Big\} \cdot g_{\boldsymbol{b}}(\gamma_k). \end{aligned}$$

This shows that the general gamma density expresses how the $p_{(<)}$-term varies as a function of γ_k. Plugging this into the equation from above, and expanding by $(n-k)!$ and $\phi(\gamma_k)^{n-k}$, yields

$$\begin{aligned} &\frac{n!}{(n-k)!} \prod_{i=1}^{k} \frac{x_i^{\beta}}{b_i} f(x_i) \mathrm{d}x_i \cdot \int_0^{\infty} \phi(\gamma_k)^{n-k} g_{\boldsymbol{b}}(\gamma_k) \\ &\cdot \Big(\int\limits_{\Delta^*_{n-k}(\gamma_k, \infty)} \Big\{ \prod_{i=k+1}^{n} \frac{x_i^{\beta} \mathrm{e}^{-\gamma_i x_i^{\beta}} f(x_i) \mathrm{d}x_i}{-\phi'(\gamma_i)} \Big\} \cdot \Big\{ (n-k)! \prod_{i=k+1}^{n} \frac{-\phi'(\gamma_i)}{\phi(\gamma_k)} \mathrm{d}\gamma_i \Big\} \Big) \cdot \mathrm{d}\gamma_k \end{aligned} \tag{A.21}$$

Note that (A.21) expresses still the joint pdf of $(X_1, \ldots, X_N)$ given $N = n$. Dividing by $\ell(\boldsymbol{x}|n)\mathrm{d}x_1 \cdots \mathrm{d}x_k$ to get the conditional pdf given $(X_1, \ldots, X_k) = \boldsymbol{x}$, and dropping the integral over γ_k so that this becomes a joint pdf for $(X_{k+1}, \ldots, X_N, \Gamma_k)$, we get

$$\begin{aligned}
&\mathbb{P}[X_{k+1} \in \mathrm{d}x_{k+1}, \ldots, X_n \in \mathrm{d}x_n, \Gamma_k \in \mathrm{d}\gamma_k | (X_1, \ldots, X_k) = \boldsymbol{x}, N = n] \\
&= \left(\frac{\phi(\gamma_k)^{n-k} g_{\boldsymbol{b}}(\gamma_k) \mathrm{d}\gamma_k}{\int_0^\infty \phi(\gamma_k)^{n-k} g_{\boldsymbol{b}}(\gamma_k) \mathrm{d}\gamma_k} \right) \\
&\cdot \int_{\gamma_k}^{\infty} \int_{\gamma_{k+1}}^{\infty} \cdots \int_{\gamma_{n-1}}^{\infty} \left\{ \prod_{i=k+1}^{n} \frac{x_i^\beta \mathrm{e}^{-\gamma_i x_i^\beta} f(x_i) \mathrm{d}x_i}{-\phi'(\gamma_i)} \right\} \cdot \left\{ (n-k)! \prod_{i=k+1}^{n} \frac{-\phi'(\gamma_i)}{\phi(\gamma_k)} \mathrm{d}\gamma_i \right\}.
\end{aligned} \tag{A.22}$$

This shows that $(X_{k+1}, \ldots, X_n)$ depends on the observed data $(X_1, \ldots, X_k) = \boldsymbol{x}$ only via its dependence on the mixing rv Γ_k, which has the data-dependent pdf $g^{(0)}(\gamma_k|\boldsymbol{x}, n) \propto \phi(\gamma_k)^{n-k} g_{\boldsymbol{b}}(\gamma_k)$, as stated in equation (3.14). We also note from comparing (A.19) and (A.22) that the h-density remains unchanged, while the $n - k$ order statistics are now from the truncated cdf $1 - \phi(\gamma)/\phi(\gamma_k)$ defined on $[\gamma_k, \infty)$.

This shows that a Monte Carlo procedure based on inverting decreasing uniform order statistics via the Laplace transform is almost the same as in the "a-priori case", and the h-density for drawing $X_i|\Gamma_i$ remains unchanged. For the other procedure based on exponential order statistics, it is now required that the $(\tilde{X}_{k+1}, \ldots, \tilde{X}_n)$ are drawn from the remaining data pdf $\rho(x|\gamma_k)$ (see equations (3.14 or 3.17) for a definition), before proceeding as in the "a-priori case" to obtain $X_{k+1} = \tilde{X}_{i(k+1)}, \ldots, X_n = \tilde{X}_{i(n)}$.

An accept / reject procedure for drawing from the h-density

For the case of $w(x) = x^\beta$ (with $\beta \neq 0$), it is possible to make a rearrangement which relates the h-density to the pdf of the Weibull distribution. Let us denote the sign of β by $d = \mathrm{sign}\beta$, so that $w(x) = x^{d|\beta|}$. Then, the h-density is proportional to

$$\begin{aligned}
h(x|\gamma) &\propto x^{d|\beta|} \mathrm{e}^{-\gamma x^{d|\beta|}} f(x) \\
&\propto \gamma|\beta| x^{d(|\beta|-1)} \mathrm{e}^{-\gamma x^{d|\beta|}} \cdot x^d f(x), \ \forall x \in [\underline{x}, \overline{x}].
\end{aligned} \tag{A.23}$$

Letting $h^{(\mathrm{Wei})}_{(|\beta|,\gamma)}(s) = \gamma|\beta| s^{|\beta|-1} \mathrm{e}^{-\gamma s^{|\beta|}}$ denote the pdf of the Weibull distribution with shape parameter $|\beta|$ and scale parameter γ, (A.23) shows that $h(x|\gamma) \propto h^{(\mathrm{Wei})}_{(|\beta|,\gamma)}(x^d) \cdot x^d f(x)$. The parameterization adopted here with the scale parameter γ is the one often encountered in econometrics (e.g., Wooldridge, 2002, p. 689).[10]

Draws from the h-density for $X_i|\Gamma_{ii}$ are easily obtained by an accept / reject procedure (see, e.g., Robert and Casella, 2010). Consider first the case of $\beta > 0$. Let X denote repeated drawings from the truncated Wei$(|\beta|, \gamma_i)$-distribution on $[\underline{x}, \overline{x}]$, and U from the U$(0, 1)$-distribution. Then, since the function $h^{(\mathrm{Wei})}_{(|\beta|,\gamma_i)}(x) \cdot \max_{r \in [\underline{x}, \overline{x}]} \{rf(r)\}$ envelopes $h^{(\mathrm{Wei})}_{(|\beta|,\gamma_i)}(x) \cdot xf(x)$, the rule for accepting becomes as follows: accept X as a draw for X_i for the first pair (X, U) which fulfills the inequality $U \cdot \max_{r \in [\underline{x}, \overline{x}]} \{rf(r)\} < Xf(X)$.[11]

[10] An alternative parameterization encountered in statistics is based on substituting $\zeta = \gamma^{-1/|\beta|}$ as the scale parameter, which is then directly proportional to the mean of the Weibull distribution. This substitution might be impractical if $|\beta|$ is very small.

[11] Note that for $\beta > 0$ and in the case of a log-uniform field-size distribution, which has density $f(x) \propto 1/x$ on $[\underline{x}, \overline{x}]$, the result is that $X_i|\Gamma_i$ has the truncated Wei$(|\beta|, \gamma_i)$-distribution on $[\underline{x}, \overline{x}]$.

In the case of $\beta < 0$ we have $d = -1$, so that we get $h^{(\mathrm{Wei})}_{(|\beta|,\gamma_i)}(x^{-1}) \cdot x^{-1} f(x) = h^{(\mathrm{Wei})}_{(|\beta|,\gamma_i)}(x^{-1}) x^{-2} \cdot x f(x)$. In fact, integrating gives $\int_{\underline{x}}^{x} h^{(\mathrm{Wei})}_{(|\beta|,\gamma_i)}(r^{-1}) r^{-2} \mathrm{d}r = \int_{x^{-1}}^{\underline{x}^{-1}} h^{(\mathrm{Wei})}_{(|\beta|,\gamma_i)}(s) \mathrm{d}s$, so that we can make the repeated drawings X^{-1} from the truncated $\mathrm{Wei}(|\beta|, \gamma_i)$-distribution on support $[\overline{x}^{-1}, \underline{x}^{-1}]$. We then accept the inverse of the draw, X, as a draw for X_i for the first pair (X^{-1}, U) which fulfills, as above, $U \cdot \max_{r \in [\underline{x}, \overline{x}]} \{r f(r)\} < X f(X)$.

A.4 Maximum Likelihood Estimation

In this section we again show the parameter dependence and separate the truncation factor in the field-size density as $f_{\boldsymbol{\theta}} / p_{\mathrm{tr}}(\boldsymbol{\theta})$.

Point estimation

At first, we state the important relations which connect the log-derivatives of the parameter dependent function $a(\boldsymbol{\eta})$ from a pdf from the exponential family to the mean vector and covariance matrix of the sufficient statistics. If X is distributed with pdf $c(x) \mathrm{e}^{\boldsymbol{\eta}^\top \boldsymbol{s}(x)} / a(\boldsymbol{\eta})$, then it holds that (see Dempster et al., 1977, p. 5)

$$\begin{aligned} a(\boldsymbol{\eta}) &= \int c(x) \mathrm{e}^{\boldsymbol{\eta}^\top \boldsymbol{s}(x)} \mathrm{d}x, \\ \Rightarrow \frac{\partial \ln a(\boldsymbol{\eta})}{\partial \boldsymbol{\eta}} &= \mathbb{E}_{\boldsymbol{\eta}}\big[\boldsymbol{s}(X)\big], \\ \Rightarrow \frac{\partial^2 \ln a(\boldsymbol{\eta})}{\partial \boldsymbol{\eta} \partial \boldsymbol{\eta}^\top} &= \mathbb{V}_{\boldsymbol{\eta}}\big[\boldsymbol{s}(X)\big]. \end{aligned} \tag{A.24}$$

By truncating a pdf from the exponential family, the only change which occurs in the exponential family form is that the parameter dependent function needs to be multiplied by the truncation factor. Thus, the pdf of $\tilde{Y}_1$ has the parameter dependent function $a(\boldsymbol{\eta})(1 - p_{\mathrm{tr}}(\boldsymbol{\eta}))$. Moreover, the density of the remaining field sizes, ρ, can be written in exponential family form as

$$\rho_{\boldsymbol{\eta}}(x|\gamma) = \big[\mathrm{e}^{-\gamma x^\beta} c(x)\big] \mathrm{e}^{\boldsymbol{\eta}^\top \boldsymbol{s}(x)} \big/ \big[a(\boldsymbol{\eta}) \phi_{\boldsymbol{\eta}}(\gamma) p_{\mathrm{tr}}(\boldsymbol{\eta})\big], \tag{A.25}$$

which makes it clear that ρ has the parameter dependent function $a(\boldsymbol{\eta}) \phi_{\boldsymbol{\eta}}(\gamma) p_{\mathrm{tr}}(\boldsymbol{\eta})$. Table A1 shows the exponential family form for the lognormal distribution.[12]

[12] The pdf of the Pareto distribution can be rearranged as $f_\theta(x) = \theta \underline{x}^\theta / x^{\theta+1} = (1/x) e^{-\theta(\ln x - \ln \underline{x})} (1/\theta^{-1})$, hence $\eta(\theta) = -\theta$, $s(x) = \ln x - \ln \underline{x}$, and $\ln a(\eta) = -\ln(-\eta)$. Taking derivatives with respect to η and plugging in $\eta(\theta)$ yields $\partial \ln a(\theta) / \partial \eta = 1/\theta$, $\partial^2 \ln a(\theta) / \partial \eta^2 = 1/\theta^2$.

Table A.1: Exponential family form for the lognormal distribution

$\big(\eta_1(\mu,\sigma),\eta_2(\mu,\sigma)\big) = \big(\mu\sigma^{-2},(-2\sigma^2)^{-1}\big)$		
$\big(\mu(\eta_1,\eta_2),\sigma(\eta_1,\eta_2)\big) = \big(\eta_1(-2\eta_2)^{-1},(-2\eta_2)^{-\frac{1}{2}}\big)$		
$c(x)=\frac{1}{\sqrt{2\pi x}},\ s_1(x)=\ln x,\ s_2(x)=(\ln x)^2,\ \ln a(\boldsymbol{\eta})=-\frac{1}{4}\frac{\eta_1^2}{\eta_2}-\frac{1}{2}\ln(-2\eta_2)$		
Jacobian of $\boldsymbol{\eta}\to\boldsymbol{\theta}$:	$\mu(\eta_1,\eta_2)$	$\sigma(\eta_1,\eta_2)$
Derivative w.r.t. η_1	σ^2	0
Derivative w.r.t. η_2	$2\mu\sigma^2$	σ^3
Gradient of $\ln a(\boldsymbol{\eta})$:	$\partial\ln a(\boldsymbol{\theta})/\partial\eta_1=\mu,$	$\partial\ln a(\boldsymbol{\theta})/\partial\eta_2=\mu^2+\sigma^2$
Hessian of $\ln a(\boldsymbol{\eta})$:	$\partial\ln a(\boldsymbol{\theta})/\partial\eta_1$	$\partial\ln a(\boldsymbol{\theta})/\partial\eta_2$
Derivative w.r.t. η_1	σ^2	$2\mu\sigma^2$
Derivative w.r.t. η_2	$2\mu\sigma^2$	$4\mu^2\sigma^2+2\sigma^4$

We continue from equation (3.16) by reparameterizing $\boldsymbol{\theta}\to\boldsymbol{\eta}$ so that the exponential family form of this equation becomes

$$\begin{aligned}F(\boldsymbol{\eta}',\nu'|\boldsymbol{\eta},\nu) = k\cdot\Big(\frac{1}{k}\sum_{i=1}^{k}\boldsymbol{s}(x_i)^\top\boldsymbol{\eta}'\Big) - k\ln a(\boldsymbol{\eta}') + \mathbb{E}_{(\boldsymbol{\eta},\nu)}\Big\{\Big(\sum_{i=1}^{M}\boldsymbol{s}(\tilde{Y}_i)^\top\boldsymbol{\eta}'\Big) - M\ln a(\boldsymbol{\eta}')\Big|\mathcal{F}_{\underline{t}}\Big\}\\ + \mathbb{E}_{(\boldsymbol{\eta},\nu)}\Big\{\Big(\sum_{i=k+1}^{N}\boldsymbol{s}(\tilde{X}_i)^\top\boldsymbol{\eta}'\Big) - (N-k)\ln a(\boldsymbol{\eta}')\Big|\mathcal{F}_{\underline{t}}\Big\}\\ - \nu' + \ln(\nu')\mathbb{E}_{(\boldsymbol{\eta},\nu)}\big[N\big|\mathcal{F}_{\underline{t}}\big] + \text{constant}.\end{aligned}$$

Thus, the necessary conditions for maximization with respect to $(\boldsymbol{\eta}',\nu')$ are obtained as

$$\begin{aligned}0 &\overset{!}{=} \frac{\partial F(\boldsymbol{\eta}',\nu'|\boldsymbol{\eta},\nu)}{\partial\boldsymbol{\eta}'}\\ \Leftrightarrow \mathbb{E}_{(\boldsymbol{\eta},\nu)}[N+M|\mathcal{F}_{\underline{t}}]\frac{\partial\ln a(\boldsymbol{\eta}')}{\partial\boldsymbol{\eta}'} &= k\cdot\Big(\frac{1}{k}\sum_{i=1}^{k}\boldsymbol{s}(x_i)\Big) + \mathbb{E}_{(\boldsymbol{\eta},\nu)}[M|\mathcal{F}_{\underline{t}}]\mathbb{E}_{\boldsymbol{\eta}}\big[\boldsymbol{s}(\tilde{Y}_1)\big]\\ &\quad + \mathbb{E}_{(\boldsymbol{\eta},\nu)}\Big\{\sum_{i=k+1}^{N}\mathbb{E}_{\boldsymbol{\eta}}\big[\boldsymbol{s}(\tilde{X}_i)\big|N,\Gamma_k,\mathcal{F}_{\underline{t}}\big]\Big|\mathcal{F}_{\underline{t}}\Big\}\\ \Leftrightarrow \frac{\mathbb{E}_{(\boldsymbol{\eta},\nu)}[N|\mathcal{F}_{\underline{t}}]}{p_{\text{tr}}(\boldsymbol{\eta})}\frac{\partial\ln a(\boldsymbol{\eta}')}{\partial\boldsymbol{\eta}'} &= k\cdot\Big(\frac{1}{k}\sum_{i=1}^{k}\boldsymbol{s}(x_i)\Big) + \mathbb{E}_{(\boldsymbol{\eta},\nu)}[N|\mathcal{F}_{\underline{t}}]\frac{1-p_{\text{tr}}(\boldsymbol{\eta})}{p_{\text{tr}}(\boldsymbol{\eta})}\cdot\mathbb{E}_{\boldsymbol{\eta}}\big[\boldsymbol{s}(\tilde{Y}_1)\big]\\ &\quad + \mathbb{E}_{(\boldsymbol{\eta},\nu)}\Big\{(N-k)\cdot\mathbb{E}_{\boldsymbol{\eta}}\big[\boldsymbol{s}(\tilde{X}_{k+1})\big|\Gamma_k,\mathcal{F}_{\underline{t}}\big]\Big|\mathcal{F}_{\underline{t}}\Big\}\\ \Leftrightarrow \frac{\partial\ln a(\boldsymbol{\eta}')}{\partial\boldsymbol{\eta}'} &= p_{\text{tr}}(\boldsymbol{\eta})\alpha(\boldsymbol{\eta},\nu)\Big(\frac{1}{k}\sum_{i=1}^{k}\boldsymbol{s}(x_i)\Big) + \big(1-p_{\text{tr}}(\boldsymbol{\eta})\big)\mathbb{E}_{\boldsymbol{\eta}}\big[\boldsymbol{s}(\tilde{Y}_1)\big]\\ &\quad + p_{\text{tr}}(\boldsymbol{\eta})\big(1-\alpha(\boldsymbol{\eta},\nu)\big)\int_0^\infty\mathbb{E}_{\boldsymbol{\eta}}\big[\boldsymbol{s}(\tilde{X}_{k+1})\big|\Gamma_k=\gamma,\mathcal{F}_{\underline{t}}\big]g^{(1)}_{(\boldsymbol{\eta},\nu)}(\gamma|\mathcal{F}_{\underline{t}})\mathrm{d}\gamma,\end{aligned}$$

where $\alpha(\boldsymbol{\eta},\nu)=k/\mathbb{E}_{(\boldsymbol{\eta},\nu)}[N\big|\mathcal{F}_{\underline{t}}]$, and

$$\begin{aligned}0 &\overset{!}{=} \frac{\partial F(\boldsymbol{\eta}',\nu'|\boldsymbol{\eta},\nu)}{\partial\nu'},\\ \Leftrightarrow \nu' &= \mathbb{E}_{(\boldsymbol{\eta},\nu)}[N|\mathcal{F}_{\underline{t}}].\end{aligned}$$

After taking derivatives, we evaluate the expectations stepwise by the law of iterated expectations. To arrive at the second equality of the first order conditions for $\boldsymbol{\eta}$, we pull the constant $\mathbb{E}_{\boldsymbol{\eta}}\big[\boldsymbol{s}(\tilde{Y}_i)|\mathcal{F}_{\underline{t}}\big] = \mathbb{E}_{\boldsymbol{\eta}}\big[\boldsymbol{s}(\tilde{Y}_1)\big]$ out of the expectation. For the third equality, we first use the fact that the NegBin(n,p)-distribution has the mean $n(1-p)/p$ to arrive at $\mathbb{E}_{(\boldsymbol{\eta},\nu)}[M|\mathcal{F}_{\underline{t}}] = \mathbb{E}_{(\boldsymbol{\eta},\nu)}[N|\mathcal{F}_{\underline{t}}](1-p_{\mathrm{tr}}(\boldsymbol{\eta}))/p_{\mathrm{tr}}(\boldsymbol{\eta})$. This also implies that the posterior mean of $N+M$ is the posterior mean of N scaled up by the inverse of the truncation probability. Further, we use in the third equality the fact that

$$\mathbb{E}_{\boldsymbol{\eta}}\big[\boldsymbol{s}(\tilde{X}_i)\big|N,\Gamma_k,\mathcal{F}_{\underline{t}}\big] = \mathbb{E}_{\boldsymbol{\eta}}\big[\boldsymbol{s}(\tilde{X}_i)\big|\Gamma_k,\mathcal{F}_{\underline{t}}\big] = \mathbb{E}_{\boldsymbol{\eta}}\big[\boldsymbol{s}(\tilde{X}_{k+1})\big|\Gamma_k,\mathcal{F}_{\underline{t}}\big]$$

for all $i = k+1,\dots,N$, as can be seen from equation (3.14). For the fourth equality, we apply Corollary 1.3 (see A.7) to the outer expected value over the joint distribution of $(N,\Gamma_k)|\mathcal{F}_{\underline{t}}$.

Finally, note that the necessary conditions are in fact sufficient for $(\boldsymbol{\eta}',\nu')$ to be a maximum given $(\boldsymbol{\eta},\nu)$. The function $F(\cdot|\cdot)$ is globally concave in $(\boldsymbol{\eta}',\nu')$ since its hessian with respect to $(\boldsymbol{\eta}',\nu')$ equals

$$\frac{\partial^2 F(\boldsymbol{\eta}',\nu'|\boldsymbol{\eta},\nu)}{\partial(\boldsymbol{\eta}',\nu')\partial(\boldsymbol{\eta}',\nu')^\top} = \begin{bmatrix} -\frac{\mathbb{E}_{(\boldsymbol{\eta},\nu)}[N|\mathcal{F}_{\underline{t}}]}{p_{\mathrm{tr}}(\boldsymbol{\eta})}\frac{\partial^2 \ln a(\boldsymbol{\eta}')}{\partial\boldsymbol{\eta}'\boldsymbol{\eta}'^\top} & 0 \\ 0 & -\frac{\mathbb{E}_{(\boldsymbol{\eta},\nu)}[N|\mathcal{F}_{\underline{t}}]}{(\nu')^2} \end{bmatrix},$$

where $\frac{\partial^2 \ln a(\boldsymbol{\eta}')}{\partial\boldsymbol{\eta}'\boldsymbol{\eta}'^\top}$ is positive definite for each $\boldsymbol{\eta}'$, as can be seen by its relation to a covariance matrix from (A.24). Thus, independent of $(\boldsymbol{\eta}',\nu')$ all eigenvalues of the hessian are negative, which implies that it is globally negative definite.

Interval estimation

By reparametrizing equation (3.13) into exponential family form, and then taking second-order derivatives, we arrive at

$$\begin{aligned}
\frac{\partial^2 L(\boldsymbol{\theta},\nu)}{\partial\boldsymbol{\eta}\partial\boldsymbol{\eta}^\top} =& -k\Big(\frac{\partial^2 \ln a(\boldsymbol{\theta})}{\partial\boldsymbol{\eta}\partial\boldsymbol{\eta}^\top} + \frac{\partial^2 \ln p_{\mathrm{tr}}(\boldsymbol{\theta})}{\partial\boldsymbol{\eta}\partial\boldsymbol{\eta}^\top}\Big) \\
&+ \bigg\{\int_0^\infty \frac{\partial^2\phi_{\boldsymbol{\theta}}(\gamma)}{\partial\boldsymbol{\eta}\partial\boldsymbol{\eta}^\top}\frac{1}{\phi_{\boldsymbol{\theta}}(\gamma)}\nu\phi_{\boldsymbol{\theta}}(\gamma)\mathrm{e}^{\nu(\phi_{\boldsymbol{\theta}}(\gamma)-1)}g_{\boldsymbol{b}}(\gamma)\mathrm{d}\gamma \\
&+ \int_0^\infty \frac{\partial \ln\phi_{\boldsymbol{\theta}}(\gamma)}{\partial\boldsymbol{\eta}}\frac{\partial \ln\phi_{\boldsymbol{\theta}}(\gamma)}{\partial\boldsymbol{\eta}^\top}\big(\nu\phi_{\boldsymbol{\theta}}(\gamma)\big)^2\mathrm{e}^{\nu(\phi_{\boldsymbol{\theta}}(\gamma)-1)}g_{\boldsymbol{b}}(\gamma)\mathrm{d}\gamma \\
&- \Big(\int_0^\infty \frac{\partial \ln\phi_{\boldsymbol{\theta}}(\gamma)}{\partial\boldsymbol{\eta}}\nu\phi_{\boldsymbol{\theta}}(\gamma)\mathrm{e}^{\nu(\phi_{\boldsymbol{\theta}}(\gamma)-1)}g_{\boldsymbol{b}}(\gamma)\mathrm{d}\gamma\Big) \\
&\cdot \Big(\int_0^\infty \frac{\partial \ln\phi_{\boldsymbol{\theta}}(\gamma)}{\partial\boldsymbol{\eta}^\top}\nu\phi_{\boldsymbol{\theta}}(\gamma)\mathrm{e}^{\nu(\phi_{\boldsymbol{\theta}}(\gamma)-1)}g_{\boldsymbol{b}}(\gamma)\mathrm{d}\gamma\Big)\bigg\}\Big/\Big(\int_0^\infty \mathrm{e}^{\nu(\phi_{\boldsymbol{\theta}}(\gamma)-1)}g_{\boldsymbol{b}}(\gamma)\mathrm{d}\gamma\Big).
\end{aligned}$$

Using $\frac{\partial^2\phi_{\boldsymbol{\theta}}(\gamma)}{\partial\boldsymbol{\eta}\partial\boldsymbol{\eta}^\top}\frac{1}{\phi_{\boldsymbol{\theta}}(\gamma)} = \frac{\partial^2 \ln\phi_{\boldsymbol{\theta}}(\gamma)}{\partial\boldsymbol{\eta}\partial\boldsymbol{\eta}^\top} + \frac{\partial \ln\phi_{\boldsymbol{\theta}}(\gamma)}{\partial\boldsymbol{\eta}}\frac{\partial \ln\phi_{\boldsymbol{\theta}}(\gamma)}{\partial\boldsymbol{\eta}^\top}$ to split the second term, then combining the resulting $\frac{\partial \ln\phi_{\boldsymbol{\theta}}(\gamma)}{\partial\boldsymbol{\eta}}\frac{\partial \ln\phi_{\boldsymbol{\theta}}(\gamma)}{\partial\boldsymbol{\eta}^\top}$-term with the third term which also involves $\frac{\partial \ln\phi_{\boldsymbol{\theta}}(\gamma)}{\partial\boldsymbol{\eta}}\frac{\partial \ln\phi_{\boldsymbol{\theta}}(\gamma)}{\partial\boldsymbol{\eta}^\top}$, and finally using the formulas from (A.5) and (A.6) yields

$$\begin{aligned}
\frac{\partial^2 L(\boldsymbol{\theta},\nu)}{\partial\boldsymbol{\eta}\partial\boldsymbol{\eta}^\top} =& -k\Big(\frac{\partial^2 \ln a(\boldsymbol{\theta})}{\partial\boldsymbol{\eta}\partial\boldsymbol{\eta}^\top} + \frac{\partial^2 \ln p_{\mathrm{tr}}(\boldsymbol{\theta})}{\partial\boldsymbol{\eta}\partial\boldsymbol{\eta}^\top}\Big) \\
&+ \mathbb{E}_{(\boldsymbol{\theta},\nu)}[N-k|\mathcal{F}_{\underline{t}}]\int_0^\infty \frac{\partial^2 \ln\phi_{\boldsymbol{\theta}}(\gamma)}{\partial\boldsymbol{\eta}\partial\boldsymbol{\eta}^\top}g^{(1)}_{(\boldsymbol{\theta},\nu)}(\gamma|\mathcal{F}_{\underline{t}})\mathrm{d}\gamma \\
&+ \mathbb{E}_{(\boldsymbol{\theta},\nu)}[(N-k)^2|\mathcal{F}_{\underline{t}}]\Big\{\int_0^\infty \frac{\partial \ln\phi_{\boldsymbol{\theta}}(\gamma)}{\partial\boldsymbol{\eta}}\frac{\partial \ln\phi_{\boldsymbol{\theta}}(\gamma)}{\partial\boldsymbol{\eta}^\top}g^{(2)}_{(\boldsymbol{\theta},\nu)}(\gamma|\mathcal{F}_{\underline{t}})\mathrm{d}\gamma\Big\} \\
&- \mathbb{E}_{(\boldsymbol{\theta},\nu)}[N-k|\mathcal{F}_{\underline{t}}]^2\Big\{\int_0^\infty \frac{\partial \ln\phi_{\boldsymbol{\theta}}(\gamma)}{\partial\boldsymbol{\eta}}g^{(1)}_{(\boldsymbol{\theta},\nu)}(\gamma|\mathcal{F}_{\underline{t}})\mathrm{d}\gamma\Big\}\Big\{\int_0^\infty \frac{\partial \ln\phi_{\boldsymbol{\theta}}(\gamma)}{\partial\boldsymbol{\eta}^\top}g^{(1)}_{(\boldsymbol{\theta},\nu)}(\gamma|\mathcal{F}_{\underline{t}})\mathrm{d}\gamma\Big\}.
\end{aligned}$$

Moreover, the relations from (A.24-A.25) imply

$$\begin{aligned}
\frac{\partial \ln \phi_{\boldsymbol{\theta}}(\gamma)}{\partial \boldsymbol{\eta}} =& \frac{\partial \ln[a(\boldsymbol{\theta}) p_{\mathrm{tr}}(\boldsymbol{\theta}) \phi_{\boldsymbol{\theta}}(\gamma)]}{\partial \boldsymbol{\eta}} - \frac{\partial \ln[a(\boldsymbol{\theta}) p_{\mathrm{tr}}(\boldsymbol{\theta})]}{\partial \boldsymbol{\eta}} = \mathbb{E}_{\boldsymbol{\theta}}\big[\boldsymbol{s}(\tilde{X}_{k+1}) \big| \Gamma_k = \gamma, \mathcal{F}_{\underline{t}}\big] - \mathbb{E}_{\boldsymbol{\theta}}\big[\boldsymbol{s}(\tilde{X}_1)\big], \\
\frac{\partial^2 \ln \phi_{\boldsymbol{\theta}}(\gamma)}{\partial \boldsymbol{\eta} \partial \boldsymbol{\eta}^\top} =& \frac{\partial^2 \ln[a(\boldsymbol{\theta}) p_{\mathrm{tr}}(\boldsymbol{\theta}) \phi_{\boldsymbol{\theta}}(\gamma)]}{\partial \boldsymbol{\eta} \partial \boldsymbol{\eta}^\top} - \frac{\partial^2 \ln[a(\boldsymbol{\theta}) p_{\mathrm{tr}}(\boldsymbol{\theta})]}{\partial \boldsymbol{\eta} \partial \boldsymbol{\eta}^\top} = \mathbb{V}_{\boldsymbol{\theta}}\big[\boldsymbol{s}(\tilde{X}_{k+1}) \big| \Gamma_k = \gamma, \mathcal{F}_{\underline{t}}\big] - \mathbb{V}_{\boldsymbol{\theta}}\big[\boldsymbol{s}(\tilde{X}_1)\big], \\
& \Big(\frac{\partial^2 \ln a(\boldsymbol{\theta})}{\partial \boldsymbol{\eta} \partial \boldsymbol{\eta}^\top} + \frac{\partial^2 \ln p_{\mathrm{tr}}(\boldsymbol{\theta})}{\partial \boldsymbol{\eta} \partial \boldsymbol{\eta}^\top}\Big) = \mathbb{V}_{\boldsymbol{\theta}}\big[\boldsymbol{s}(\tilde{X}_1)\big],
\end{aligned}$$

which finally results for the Hessian of the log-likelihood w.r.t. $\boldsymbol{\eta}$ in the following formula

$$\begin{aligned}
\frac{\partial^2 L(\boldsymbol{\theta}, \nu)}{\partial \boldsymbol{\eta} \partial \boldsymbol{\eta}^\top} =& - k \cdot \mathbb{V}_{\boldsymbol{\theta}}\big[\boldsymbol{s}(\tilde{X}_1)\big] \\
& + \mathbb{E}_{(\boldsymbol{\theta},\nu)}[N - k | \mathcal{F}_{\underline{t}}] \cdot \Big\{ \int_0^\infty \Big(\mathbb{V}_{\boldsymbol{\theta}}\big[\boldsymbol{s}(\tilde{X}_{k+1}) \big| \Gamma_k = \gamma, \mathcal{F}_{\underline{t}}\big] - \mathbb{V}_{\boldsymbol{\theta}}\big[\boldsymbol{s}(\tilde{X}_1)\big] \Big) g^{(1)}_{(\boldsymbol{\theta},\nu)}(\gamma | \mathcal{F}_{\underline{t}}) \mathrm{d}\gamma \Big\} \\
& + \mathbb{E}_{(\boldsymbol{\theta},\nu)}[(N - k)^2 | \mathcal{F}_{\underline{t}}] \cdot \Big\{ \int_0^\infty \Big(\mathbb{E}_{\boldsymbol{\theta}}\big[\boldsymbol{s}(\tilde{X}_{k+1}) \big| \Gamma_k = \gamma, \mathcal{F}_{\underline{t}}\big] - \mathbb{E}_{\boldsymbol{\theta}}\big[\boldsymbol{s}(\tilde{X}_1)\big] \Big) \\
& \qquad \cdot \Big(\mathbb{E}_{\boldsymbol{\theta}}\big[\boldsymbol{s}(\tilde{X}_{k+1}) \big| \Gamma_k = \gamma, \mathcal{F}_{\underline{t}}\big] - \mathbb{E}_{\boldsymbol{\theta}}\big[\boldsymbol{s}(\tilde{X}_1)\big] \Big)^\top g^{(2)}_{(\boldsymbol{\theta},\nu)}(\gamma | \mathcal{F}_{\underline{t}}) \mathrm{d}\gamma \Big\} \\
& - (\mathbb{E}_{(\boldsymbol{\theta},\nu)}[N - k | \mathcal{F}_{\underline{t}}])^2 \cdot \Big\{ \int_0^\infty \Big(\mathbb{E}_{\boldsymbol{\theta}}\big[\boldsymbol{s}(\tilde{X}_{k+1}) \big| \Gamma_k = \gamma, \mathcal{F}_{\underline{t}}\big] - \mathbb{E}_{\boldsymbol{\theta}}\big[\boldsymbol{s}(\tilde{X}_1)\big] \Big) g^{(1)}_{(\boldsymbol{\theta},\nu)}(\gamma | \mathcal{F}_{\underline{t}}) \mathrm{d}\gamma \Big\} \\
& \qquad \cdot \Big\{ \int_0^\infty \Big(\mathbb{E}_{\boldsymbol{\theta}}\big[\boldsymbol{s}(\tilde{X}_{k+1}) \big| \Gamma_k = \gamma, \mathcal{F}_{\underline{t}}\big] - \mathbb{E}_{\boldsymbol{\theta}}\big[\boldsymbol{s}(\tilde{X}_1)\big] \Big)^\top g^{(1)}_{(\boldsymbol{\theta},\nu)}(\gamma | \mathcal{F}_{\underline{t}}) \mathrm{d}\gamma \Big\}.
\end{aligned}$$

A.5 Estimation of the Parameters of R_i under Assumption 4b

Define $e_{k+1} = \mathbb{E}\big[X^\beta_{k+1} + \ldots + X^\beta_N \big| \mathcal{F}_{\underline{t}}\big]$, and $e_i = x^\beta_i + \ldots + x^\beta_k + e_{k+1}$ for each $i = 1, \ldots, k$. For the first k discovered fields, we can use information about the number of exploratory wells that had to be drilled to discover a field around the time of their discoveries. Let this number be denoted by r_i^{-1}, (such that r_i is the success rate for finding the i^{th} field). Extending equation (3.21) backwards to $i = 1, \ldots, k$ and replacing $X^\beta_i + \ldots + X^\beta_N$ with its expected value given the data, e_i, we can rearrange equation (3.21) into the linear regression specification

$$r_i^{-1} = a_0 + b_0 e_i^{-1} + u_i, \quad i = 1, \ldots, k, \tag{A.26}$$

where u_i is an error term, and from which a_0 and b_0 can be readily estimated by OLS.[13] If instead the parameter b_0 was replaced by a positive, non-linear function of the cumulative exploratory wells z, e.g. $b_0(z) = c_0 + c_1 \mathrm{e}^{-c_2 z}$, one could plug in the exploratory well data for z and estimate eq. (A.26) by non-linear least squares.

[13] Setting $a_0 = 1$, b_0 could also be determined by interpolation of only the most recent value, i.e. using only r_k and e_k. Solving eq. (A.26) with $u_k = 0$ would give $b_0 = e_k(r_k^{-1} - 1)$.

References

Abate, J. and W. Whitt (1992): "The Fourier-series method for inverting transforms of probability distributions," *Queueing Systems*, 10 (1), 5–87.

——— (1995): "Numerical inversion of Laplace transforms of probability distributions," *ORSA Journal on Computing*, 7 (1), 36–43.

Barouch, E. and G. Kaufman (1976): "Oil and gas discovery modelled as sampling proportional to random size," *Working Paper Alfred P. Sloan School of Management*.

Crump, K. S. (1976): "Numerical inversion of Laplace transforms using a Fourier-series approximation," *Journal of the ACM (JACM)*, 23 (1), 89–96.

Dempster, A. P., N. M. Laird, and D. B. Rubin (1977): "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society. Series B (Methodological)*, 39 (1), 1–38.

Dubner, H. and J. Abate (1968): "Numerical inversion of Laplace transforms by relating them to the finite Fourier cosine transform," *Journal of the ACM (JACM)*, 15 (1), 115–123.

Durbin, F. (1974): "Numerical inversion of Laplace transforms: an efficient improvement to Dubner and Abate's method," *The Computer Journal*, 17 (4), 371–376.

Gordon, L. (1983): "Successive sampling in large finite populations," *The Annals of Statistics*, 11 (2), 702–706.

Königsberger, K. (2003): *Analysis 1, 6. Auflage*, Springer-Verlag Berlin Heidelberg.

Lee, P. J. (2008): *Statistical methods for estimating petroleum resoures*, Oxford University Press.

McGill, W. J. and J. Gibbon (1965): "The general-gamma distribution and reaction times," *Journal of Mathematical Psychology*, 2 (1), 1–18.

Nair, V. N. and P. C. C. Wang (1989): "Maximum likelihood estimation under a successive sampling discovery model," *Technometrics*, 31 (4), 423–436.

Pitman, J. and N. M. Tran (2015): "Size-biased permutation of a finite sequence with independent and identically distributed terms," *Bernoulli*, 21 (4), 2484–2512.

Robert, C. P. and G. Casella (2010): *Monte Carlo statistical methods, 2. ed.*, New York, NY : Springer.

Simon, R. M., M. T. Stroot, and G. H. Weiss (1972): "Numerical inversion of Laplace transforms with application to percentage labeled mitoses experiments," *Computers and Biomedical Research*, 5 (6), 596–607.

Wooldridge, J. M. (2002): *Econometric analysis of cross section and panel data*, Cambridge, Mass.: MIT Press.