Loan-Level ABS Estimation: Supplemental Material

This is an online companion supplement to the manuscript, *Estimating the time-to-event distribution for loan-level data within a consumer auto loan asset-backed security*. Please attribute any citations to the original manuscript. This companion includes incomplete data details for the ABS setting, proofs of all major results, statements and proofs of major asymptotic results, a numeric illustration of the likelihood function under right-censoring, a reference of derivative calculations for implementation, an outline of an approach to simulate data from h_*, additional simulation study details, and additional application details for ABS bonds AART (2017, 2019). For reference, all data and replication code are publicly available at https://github.com/jackson-lautier/consumer-auto-abs-parametric.

A Incomplete Data Details

Figure A1 provides details on the parameters and random variables that define the incomplete data setting relevant to ABS data. A version has also appeared in Lautier et al. (2023).

In Figure A1, we can see three possible ABS individual asset origination lifetime data outcomes at current time ε originated at random time T. Left-truncated (○): an origination at time T does not survive until $Y = m + \Delta + 1 - T$. Such an outcome would not be observable to an investor. Complete (□): a loan originated at time T survives longer than $Y = m+\Delta+1-T$ and terminates prior to time $C = Y+\varepsilon-(m+\Delta+1) \equiv Y+\tau$. The complete lifetime, X, (the length of the line segment from □ to •) is observable to the investor (though still conditional on surviving at least Y months). Right-censored (△): an origination at time T is still active as of time ε. The investor observes $X \geq C = Y + \varepsilon - (m + \Delta + 1) \equiv Y + \tau$ but does not observe the exact termination time, X.

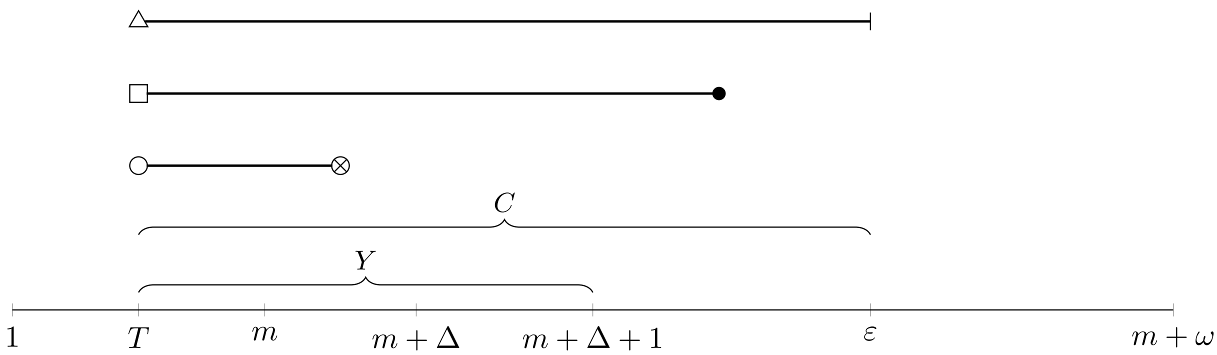

Figure A1: **Preliminary Incomplete Data Details**. For a description, see Section A.

B Proofs

Please see Sections 3.2 and 3.3 for complete statements.

B.1 Proof of Theorem 3.1

Proof. Without loss of generality, let $\Delta = 0$. It is equivalent to find the stationary points of the loglikelihood, $\log \mathcal{L}(\boldsymbol{\Theta} \mid \mathcal{S}_n)$. To handle the linear restrictions imposed by $\mathcal{C}$, we will proceed with the technique of Lagrange multipliers (e.g., Ravishanker and Dey, 2002, §2.9, pg. 69). Hence, the Lagrangian function is

$$\log \mathcal{L}(\boldsymbol{g}, p, \pi \mid \mathcal{S}_n)/n = -\log \alpha + \sum_{v=1}^{m} \sum_{u=v}^{\omega} \hat{h}_{uv} \{\log f(u; p) + \log g_v\} + \pi \left(1 - \sum_{v=1}^{m} g_v\right).$$

We now show $\hat{\pi} = 0$. Observe first from (2),

$$\frac{\partial \alpha}{\partial g_v} = \sum_{u=v}^{\omega} f(u; p), \quad v \in \mathcal{V},$$

and

$$\frac{\partial \alpha}{\partial p} = \sum_{u=1}^{\omega} \frac{\partial}{\partial p} f(u; p) \left(\sum_{v=1}^{\min(u,m)} g_v \right) \equiv \sum_{v=1}^{m} g_v \left(\sum_{u=v}^{\omega} \frac{\partial}{\partial p} f(u; p) \right).$$

For convenience of notation, define $\ell \coloneqq \log \mathcal{L}(\boldsymbol{g}, p, \pi \mid \mathcal{S}_n)/n$. Therefore, for $v \in \mathcal{V}$,

$$\begin{aligned}\frac{\partial \ell}{\partial g_v} &= -\frac{1}{\alpha}\frac{\partial \alpha}{\partial g_v} + \frac{\partial}{\partial g_v}\sum_{v=1}^{m}\sum_{u=v}^{\omega}\hat{h}_{uv}\log g_v - \pi \\ &= -\frac{1}{\alpha}\sum_{u=v}^{\omega} f(u;p) + \frac{\hat{h}_{\bullet v}}{g_v} - \pi.\end{aligned}$$

Observe,

$$g_v\left(\frac{\partial \ell}{\partial g_v}\right) = 0 \implies -\frac{1}{\alpha} g_v \sum_{u=v}^{\omega} f(u;p) + \hat{h}_{\bullet v} - \pi g_v = 0.$$

That is,

$$\sum_{v=1}^{m} g_v\left(\frac{\partial \ell}{\partial g_v}\right) = 0 \implies -\frac{1}{\alpha}\sum_{v=1}^{m} g_v\left(\sum_{u=v}^{\omega} f(u;p)\right) + \sum_{v=1}^{m}\hat{h}_{\bullet v} - \pi\sum_{v=1}^{m} g_v = 0.$$

Because $\sum_v \hat{h}_{\bullet v} = 1$, $g_v > 0$ by assumption, and (2), we must have $\hat{\pi} = 0$. Thus, any stationary point of the unconstrained optimization of (3) will also be a stationary point of the constrained optimization of (3) with solutions restricted to the convex subset, $\mathcal{C}$. This proves the final sentence of Theorem 3.1. Proceeding,

$$\left.\frac{\partial \ell}{\partial g_v}\right|_{\hat{\pi}} = -\frac{1}{\alpha}\sum_{u=v}^{\omega} f(u;p) + \frac{\hat{h}_{\bullet v}}{g_v} = 0 \iff g_v = \frac{\alpha \hat{h}_{\bullet v}}{\sum_{u=v}^{\omega} f(u;p)}. \tag{S.1}$$

Further,

$$\begin{aligned}\frac{\partial \ell}{\partial p} &= -\frac{1}{\alpha}\frac{\partial \alpha}{\partial p} + \frac{\partial}{\partial p}\sum_{v=1}^{m}\sum_{u=v}^{\omega}\hat{h}_{uv}\log f(u;p) \\ &= -\frac{1}{\alpha}\sum_{v=1}^{m} g_v\left(\sum_{u=v}^{\omega}\frac{\partial}{\partial p} f(u;p)\right) + \sum_{v=1}^{m}\sum_{u=v}^{\omega}\frac{\hat{h}_{uv}}{f(u;p)}\frac{\partial}{\partial p} f(u;p).\end{aligned}$$

Hence, by (S.1),

$$\left.\frac{\partial \ell}{\partial p}\right|_{g_v} = -\frac{1}{\alpha}\left[\sum_{v=1}^{m}\left(\frac{\alpha \hat{h}_{\bullet v}}{\sum_{u=v}^{\omega} f(u;p)}\right)\left(\sum_{u=v}^{\omega}\frac{\partial}{\partial p} f(u;p)\right)\right] + \sum_{v=1}^{m}\sum_{u=v}^{\omega}\frac{\hat{h}_{uv}}{f(u;p)}\frac{\partial}{\partial p} f(u;p)$$

$$= -\sum_{v=1}^{m}\left(\frac{\hat{h}_{\bullet v}}{\sum_{u=v}^{\omega} f(u;p)}\right)\left(\sum_{u=v}^{\omega}\frac{\partial}{\partial p}f(u;p)\right) + \sum_{v=1}^{m}\sum_{u=v}^{\omega}\frac{\hat{h}_{uv}}{f(u;p)}\frac{\partial}{\partial p}f(u;p).$$

Thus,

$$\left.\frac{\partial \ell}{\partial p}\right|_{g_v} = 0 \iff \sum_{v=1}^{m}\left(\frac{\hat{h}_{\bullet v}}{\sum_{u=v}^{\omega} f(u;p)}\right)\left(\sum_{u=v}^{\omega}\frac{\partial}{\partial p}f(u;p)\right) = \sum_{v=1}^{m}\sum_{u=v}^{\omega}\frac{\hat{h}_{uv}}{f(u;p)}\frac{\partial}{\partial p}f(u;p).$$

This proves (7). Finally, recall the constraint $\sum_v g_v = 1$. Hence, returning to (S.1), we must have

$$1 = \sum_{v\in\mathcal{V}} g_v = \sum_{v\in\mathcal{V}} \frac{\alpha\hat{h}_{\bullet v}}{\sum_{u=v}^{\omega} f(u;p)} \implies \alpha = \left[\sum_{k=1}^{m}\frac{\hat{h}_{\bullet k}}{S(k;p)}\right]^{-1}.$$

Therefore, for any $\hat{p}$ (i.e., $p \in \mathcal{P}$ such that (7) is satisfied) and all $v \in \mathcal{V}$,

$$\hat{g}_v = \frac{\alpha(\hat{p})\hat{h}_{\bullet v}}{\sum_{u=v}^{\omega} f(u;\hat{p})} = \frac{\hat{h}_{\bullet v}}{S(v;\hat{p})}\left[\sum_{k=1}^{m}\frac{\hat{h}_{\bullet k}}{S(k;\hat{p})}\right]^{-1}.$$

This recovers (5) and completes the proof. □

B.2 Proof of Corollary 3.1.1

Proof. Without loss of generality, assume $\Delta = 0$. The proof closely follows the proof of Theorem 3.1, and so we omit repetitive details. Recall the form of the likelihood in (9) to define the equivalent Lagrangian function

$$\log\mathcal{L}(\boldsymbol{g},\boldsymbol{p}\mid\mathcal{S}'_n)/n = -\log\alpha + \sum_{v=1}^{m}\sum_{u=v}^{\omega}\hat{h}_{uv}\{\log f(u;\boldsymbol{p}) + \log g_v\} + \pi\left(1 - \sum_{v=1}^{m} g_v\right).$$

Because

$$\frac{\partial\alpha}{\partial g_v} = \sum_{u=v}^{\omega} f(u;\boldsymbol{p}), \quad v\in\mathcal{V},$$

$$\frac{\partial\alpha}{\partial p_j} = \sum_{u=1}^{\omega}\frac{\partial}{\partial p_j}f(u;\boldsymbol{p})\left(\sum_{v=1}^{\min(u,m)} g_v\right) \equiv \sum_{v=1}^{m} g_v\left(\sum_{u=v}^{\omega}\frac{\partial}{\partial p_j}f(u;\boldsymbol{p})\right),$$

for $j = 1, \ldots, r$, and

$$\frac{\partial \log \mathcal{L}(\boldsymbol{g}, \boldsymbol{p} \mid \mathcal{S}'_n)/n}{\partial g_v} = -\frac{1}{\alpha}\sum_{u=v}^{\omega} f(u; \boldsymbol{p}) + \frac{\hat{h}_{\bullet v}}{g_v} - \pi,$$

for all $v \in \mathcal{V}$, it follows that $\hat{\pi} = 0$. Further,

$$\left.\frac{\partial \log \mathcal{L}(\boldsymbol{g}, \boldsymbol{p} \mid \mathcal{S}'_n)/n}{\partial g_v}\right|_{\hat{\pi}} = 0 \iff g_v = \frac{\alpha \hat{h}_{\bullet v}}{\sum_{u=v}^{\omega} f(u; \boldsymbol{p})}. \tag{S.2}$$

Thus, from (S.2) and

$$\frac{\partial \log \mathcal{L}(\boldsymbol{g}, \boldsymbol{p} \mid \mathcal{S}'_n)/n}{\partial p_j} = -\frac{1}{\alpha}\sum_{v=1}^{m} g_v \left(\sum_{u=v}^{\omega} \frac{\partial}{\partial p_j} f(u; \boldsymbol{p})\right) + \sum_{v=1}^{m}\sum_{u=v}^{\omega} \frac{\hat{h}_{uv}}{f(u; \boldsymbol{p})} \frac{\partial}{\partial p_j} f(u; \boldsymbol{p}),$$

it follows that

$$\left.\frac{\partial \log \mathcal{L}(\boldsymbol{g}, \boldsymbol{p} \mid \mathcal{S}'_n)/n}{\partial p_j}\right|_{g_v} = -\eta_1(j) + \eta_2(j) = 0 \iff \eta_1(j) = \eta_2(j), \forall j = 1, \ldots, r,$$

where

$$\eta_1(j) = \sum_{v=\Delta+1}^{\Delta+m} \left(\frac{\hat{h}_{\bullet v}}{\sum_{u=v}^{\omega} f(u; \boldsymbol{p})}\right)\left(\sum_{u=v}^{\omega} \frac{\partial}{\partial p_j} f(u; \boldsymbol{p})\right),$$

and

$$\eta_2(j) = \sum_{v=\Delta+1}^{\Delta+m}\sum_{u=v}^{\omega} \frac{\hat{h}_{uv}}{f(u; \boldsymbol{p})} \frac{\partial}{\partial p_j} f(u; \boldsymbol{p}).$$

The set of simultaneous solutions, $\hat{\boldsymbol{p}}$, recovers the estimator (11). The proof is complete by replacing $\hat{\boldsymbol{p}}$ in (S.2) and using the constraint $\sum_{\mathcal{V}} g_v = 1$ to recover (10). □

B.3 Statement & Proof of Theorem B.1

We demonstrate (7) of Theorem 3.1 takes the form of an asymptotically consistent M-estimator (van der Vaart, 1998, §5.3, pg. 51). The conditions of Theorem B.1 are typically satisfied under the standard regularity conditions (e.g., van der Vaart, 1998, §5.3, pg. 51;

Mukhopadhyay, 2000, §12.2, pg. 539), including the existence of $\partial^2 f/\partial p^2$.

Theorem B.1. *Let $\hat{p}_n$ satisfy (7) and denote p_0 as the true parameter value. Define*

$$\Psi_n(p,\mathcal{S}_n) = \sum_{v=\Delta+1}^{\Delta+m}\left(\frac{\hat{h}_{\bullet v}}{\sum_{u=v}^{\omega} f(u;p)}\right)\left(\sum_{u=v}^{\omega}\frac{\partial}{\partial p}f(u;p)\right) - \sum_{v=\Delta+1}^{\Delta+m}\sum_{u=v}^{\omega}\frac{\hat{h}_{uv}}{f(u;p)}\frac{\partial}{\partial p}f(u;p).$$

Then $\Psi_n(p,\mathcal{S}_n)\equiv\Psi_n(p)$ is an asymptotically consistent M-estimator of $\mathbf{E}\psi(X_i,Y_i,p)$ for all $p\in\mathcal{P}$ (van der Vaart, 1998, §5.3, pg. 51), where

$$\begin{aligned}\psi(X_i,Y_i,p) = &\sum_{v=\Delta+1}^{\Delta+m}\left(\frac{\sum_{u=v}^{\omega} W_i}{\sum_{u=v}^{\omega} f(u;p)}\right)\left(\sum_{u=v}^{\omega}\frac{\partial}{\partial p}f(u;p)\right)\\ &-\sum_{v=\Delta+1}^{\Delta+m}\sum_{u=v}^{\omega}\frac{W_i}{f(u;p)}\frac{\partial}{\partial p}f(u;p),\end{aligned}$$

and $W_i(u,v) = \mathbf{1}((X_i,Y_i)=(u,v))$ for $1\le i\le n$. Further, $\Psi_n(\hat{p}_n)=0$. If we also assume

(i) $\hat{p}_n \xrightarrow{\mathbf{P}} p_0$,

(ii) $\mathbf{E}[\psi(X_i,Y_i,p_0)]^2<\infty$,

(iii) $\mathbf{E}[\partial\psi(X_i,Y_i,p_0)/\partial p]$ *exists, and*

(iv) $\partial^2\Psi_n(\tilde{p})/\partial p^2$ *is* $O_{\mathbf{E}\psi}(1)$, *where $\tilde{p}$ is a point between $\hat{p}_n$ and p_0,*

then,

$$\sqrt{n}(\hat{p}_n - p_0)\xrightarrow{\mathcal{L}} N\left(0,\frac{\mathbf{E}[\psi(X_i,Y_i,p_0)^2]}{(\mathbf{E}[\partial\psi(X_i,Y_i,p_0)/\partial p])^2}\right),$$

where

$$\begin{aligned}\frac{\partial}{\partial p}\psi(X_i,Y_i,p_0) = &\sum_{v=\Delta+1}^{\Delta+m}\left(\sum_{u=v}^{\omega}W_i\right)\left[\frac{(\sum_{u=v}^{\omega}f''(u))(\sum_{u=v}^{\omega}f(u)) - (\sum_{u=v}^{\omega}f'(u))^2}{(\sum_{u=v}^{\omega}f(u))^2}\right]\\ &-\sum_{v=\Delta+1}^{\Delta+m}\sum_{u=v}^{\omega}W_i\left[\frac{f''(u)f(u)-f'(u)^2}{f(u)^2}\right],\end{aligned}$$

and f' and f'' denote $\partial f/\partial p$ and $\partial^2 f/\partial p^2$, respectively.

Proof. Observe

$$\begin{aligned}
\mathbf{E}[\psi(X_i, Y_i, p)] &= \sum_{v=\Delta+1}^{\Delta+m} \mathbf{E}\left[\left(\frac{\sum_{u=v}^{\omega} W_i}{\sum_{u=v}^{\omega} f(u;p)}\right)\left(\sum_{u=v}^{\omega} \frac{\partial}{\partial p} f(u;p)\right) - \sum_{u=v}^{\omega} \frac{W_i}{f(u;p)} \frac{\partial}{\partial p} f(u;p)\right] \\
&= \sum_{v=\Delta+1}^{\Delta+m} \left(\frac{\sum_{u=v}^{\omega} \mathbf{E}[W_i]}{\sum_{u=v}^{\omega} f(u;p)}\right)\left(\sum_{u=v}^{\omega} \frac{\partial}{\partial p} f(u;p)\right) - \sum_{u=v}^{\omega} \frac{\mathbf{E}[W_i]}{f(u;p)} \frac{\partial}{\partial p} f(u;p).
\end{aligned}$$

But $\mathbf{E}[W_i(u,v)] = h_*(u,v)$ and so for any $v \in \{\Delta+1, \ldots, \Delta+m\}$,

$$\begin{aligned}
&\left(\frac{\sum_{u=v}^{\omega} \mathbf{E}[W_i]}{\sum_{u=v}^{\omega} f(u;p)}\right)\left(\sum_{u=v}^{\omega} \frac{\partial}{\partial p} f(u;p)\right) - \sum_{u=v}^{\omega} \frac{\mathbf{E}[W_i]}{f(u;p)} \frac{\partial}{\partial p} f(u;p) \\
&\qquad = \left(\frac{\sum_{u=v}^{\omega} h_*(u,v)}{\sum_{u=v}^{\omega} f(u;p)}\right)\left(\sum_{u=v}^{\omega} \frac{\partial}{\partial p} f(u;p)\right) - \sum_{u=v}^{\omega} \frac{h_*(u,v)}{f(u;p)} \frac{\partial}{\partial p} f(u;p) \\
&\qquad = \left(\frac{g_v \sum_{u=v}^{\omega} f(u;p)}{\alpha \sum_{u=v}^{\omega} f(u;p)}\right)\left(\sum_{u=v}^{\omega} \frac{\partial}{\partial p} f(u;p)\right) - \sum_{u=v}^{\omega} \frac{f(u;p) g_v}{\alpha f(u;p)} \frac{\partial}{\partial p} f(u;p) \\
&\qquad = \frac{g_v}{\alpha} \sum_{u=v}^{\omega} \frac{\partial}{\partial p} f(u;p) - \frac{g_v}{\alpha} \sum_{u=v}^{\omega} \frac{\partial}{\partial p} f(u;p) \\
&\qquad = 0.
\end{aligned}$$

Hence, $\mathbf{E}[\psi(X_i, Y_i, p)] = 0$. Further,

$$\Psi_n(p) = \frac{1}{n} \sum_{i=1}^{n} \psi(X_i, Y_i, p),$$

and so $\Psi_n(p) \xrightarrow{\mathbf{P}} \mathbf{E}[\psi(X_i, Y_i, p)]$ by the Law of Large Numbers (Lehmann and Casella, 1998, Theorem 8.2, pg. 54-55). That $\Psi_n(\hat{p}_n) = 0$ is immediate by the conditions of (7). The remainder follows the standard Taylor series analysis (e.g., van der Vaart, 1998, §5.3, pg. 51-52), with $\partial/\partial p(\psi)$ following by the quotient rule (Rudin, 1976, Theorem 5.3, pg. 104). □

Remark. *The conditions (i) through (iv) in Theorem B.1 may be relaxed. See van der Vaart (1998, Theorems 5.21 and 5.23 pg. 51-53) for details. Further, these results may be extended to higher dimensions of parameters, such as those assumed in Corollary 3.1.1. See the discussion van der Vaart (1998, Equation (5.20) pg. 51-52) and Section G for details.*

B.4 Statement & Proof of Corollary B.1.1

In practical settings, the true parameter, $p \in \mathcal{P}$, will not be known. Thus, techniques to estimate the asymptotic variance we derive in Theorem B.1 are necessary. The results of Corollary B.1.1 provide one such approach. The conditions of Corollary B.1.1 are typically satisfied under the standard regularity conditions (e.g., van der Vaart, 1998, §5.3, pg. 51; Mukhopadhyay, 2000, §12.2, pg. 539).

Corollary B.1.1. *Assume the conditions of Theorem B.1 and define* $U = \mathbf{E}[\partial\psi(X_i, Y_i, p_0)/\partial p]$, $U_n = \partial\Psi_n(\hat{p}_n)/\partial p$, $V = \mathrm{Var}[\psi(X_i, Y_i, p_0)]$, *and* $V_n = \sum_i \psi(X_i, Y_i, \hat{p}_n)^2/n$. *If* $U_n \xrightarrow{\mathbf{P}} U$ *and* $V_n \xrightarrow{\mathbf{P}} V$, *then*

$$[V_n/U_n^2]^{-1/2}\sqrt{n}(\hat{p}_n - p_0) \xrightarrow{\mathcal{L}} N(0,1). \tag{S.3}$$

Additionally, if the second Bartlett identity (Ferguson, 1996, pg. 120) is also satisfied, then $U = V$ *with* $\mathbf{U}$ *symmetric, and so* $[V_n]^{1/2}\sqrt{n}(\hat{p}_n - p_0) \xrightarrow{\mathcal{L}} N(0,1)$.

Proof. The result (S.3) follows from Theorem B.1 and Slutsky's Theorem (Lehmann and Casella, 1998, Theorem 8.10, pg. 58). The final sentence is a classical result of maximum likelihood theory (e.g., van der Vaart, 1998, §5.5). □

B.5 Proof of Theorem 3.2

Proof. From the definition of the survival function in (6), the left-hand side of (7) becomes

$$\begin{aligned}
\sum_{v=\Delta+1}^{\Delta+m}\left(\frac{\hat{h}_{\bullet v}}{\sum_{u=v}^{\omega} f(u;p)}\right)\left(\sum_{u=v}^{\omega}\frac{\partial}{\partial p}f(u;p)\right) &= \sum_{v=\Delta+1}^{\Delta+m}\left(\frac{\hat{h}_{\bullet v}}{\sum_{u=v}^{\omega} f(u;p)}\right)\frac{\partial}{\partial p}\left(\sum_{u=v}^{\omega} f(u;p)\right) \\
&= \sum_{v=\Delta+1}^{\Delta+m}\left(\frac{\hat{h}_{\bullet v}}{S(v;p)}\right)\frac{\partial}{\partial p}S(v;p) \\
&= \sum_{v=\Delta+1}^{\Delta+m}\hat{h}_{\bullet v}\frac{\partial}{\partial p}\ln S(v;p)
\end{aligned}$$

$$= \frac{\partial}{\partial p} \sum_{v=\Delta+1}^{\Delta+m} \hat{h}_{\bullet v} \ln S(v;p).$$

Similarly, on the right-hand side of (7),

$$\sum_{v=\Delta+1}^{\Delta+m} \sum_{u=v}^{\omega} \frac{\hat{h}_{uv}}{f(u;p)} \frac{\partial}{\partial p} f(u;p) = \frac{\partial}{\partial p} \sum_{u=\Delta+1}^{\omega} \hat{h}_{u\bullet} \ln f(u;p).$$

Thus, (7) may be equivalently stated as any $p \in \mathcal{P}$ such that

$$\frac{\partial}{\partial p} \sum_{v=\Delta+1}^{\Delta+m} \hat{h}_{\bullet v} \ln S(v;p) = \frac{\partial}{\partial p} \sum_{u=\Delta+1}^{\omega} \hat{h}_{u\bullet} \ln f(u;p),$$

or

$$\frac{\partial}{\partial p} \left(\sum_{v=\Delta+1}^{\Delta+m} \hat{h}_{\bullet v} \ln S(v;p) - \sum_{u=\Delta+1}^{\omega} \hat{h}_{u\bullet} \ln f(u;p) \right) = 0. \tag{S.4}$$

But,

$$\begin{aligned} \sum_{v=\Delta+1}^{\Delta+m} \hat{h}_{\bullet v} \ln S(v;p) - \sum_{u=\Delta+1}^{\omega} \hat{h}_{u\bullet} \ln f(u;p) &= \sum_{v=\Delta+1}^{\Delta+m} \ln S(v;p)^{\hat{h}_{\bullet v}} - \sum_{u=\Delta+1}^{\omega} \ln f(u;p)^{\hat{h}_{u\bullet}} \\ &= \ln \left(\frac{\prod_{v=\Delta+m}^{\Delta+m} S(v;p)^{\hat{h}_{\bullet v}}}{\prod_{u=\Delta+1}^{\omega} f(u;p)^{\hat{h}_{u\bullet}}} \right). \end{aligned}$$

Therefore, (S.4) may equivalently be written as

$$\frac{\partial}{\partial p} \ln \left(\frac{\prod_{v=\Delta+m}^{\Delta+m} S(v;p)^{\hat{h}_{\bullet v}}}{\prod_{u=\Delta+1}^{\omega} f(u;p)^{\hat{h}_{u\bullet}}} \right) = 0. \tag{S.5}$$

Because $f(u;p) > 0$ for all $u \in \mathcal{U}$, $p \in \mathcal{P}$ by assumption (and, by extension, $S(u;p) > 0$ for all $u \in \mathcal{U}$, $p \in \mathcal{P}$), (S.5) is true if and only if,

$$\frac{\partial}{\partial p} \frac{\prod_{v=\Delta+1}^{\Delta+m} S(v;p)^{\hat{h}_{\bullet v}}}{\prod_{u=\Delta+1}^{\omega} f(u;p)^{\hat{h}_{u\bullet}}} = 0.$$

This recovers (12) and completes the proof. □

B.6 Proof of Theorem 3.3

Proof. Without loss of generality, let $\Delta = 0$. Given (13), the survival function becomes

$$S_G(u; p) = (1-p)^{u-1}, \quad u \in \{1, \ldots, \omega\}.$$

Hence, (12) reduces to

$$\frac{\partial}{\partial p} \frac{\prod_{v=1}^{m} S(v;p)^{\hat{h}_{\bullet v}}}{\prod_{u=1}^{\omega} f(u;p)^{\hat{h}_{u\bullet}}} = \frac{\partial}{\partial p} \frac{(1-p)^a}{p^b} = \frac{(1-p)^a}{p^b}\left[\frac{a}{1-p} - \frac{b}{p}\right].$$

Because $0 < p < 1$,

$$\frac{(1-p)^a}{p^b}\left[\frac{a}{1-p} - \frac{b}{p}\right] = 0 \iff \frac{a}{1-p} - \frac{b}{p} = 0 \implies \hat{p} = \frac{b}{b-a},$$

which is unique. Trivially, $\hat{p} \in \mathcal{C}$. To find $\hat{\boldsymbol{g}}$, observe

$$S_G(u; \hat{p}) = \left(\frac{a}{a-b}\right)^{u-1},$$

for $u \in \mathcal{U}$. Hence, replace $S_G(u;\hat{p})$ in (5). That $\hat{\boldsymbol{g}}$ is unique follows from the uniqueness of $\hat{p}$. Further, by Theorem 3.1, $\hat{\boldsymbol{g}} \in \mathcal{C}$.

To see that $\hat{p}$, $\hat{\boldsymbol{g}}$ are together the global maximum of $\mathcal{L}$, it is sufficient to examine the behavior of $\ell(\boldsymbol{g}, p \mid \mathcal{S}_n)/n \equiv \ell(\boldsymbol{g}, p, \bar{\pi} \mid \mathcal{S}_n)/n$ for the boundaries of $\mathcal{C}$ (recall the convexity of $\mathcal{C}$). When $p = 0$, $f_G(u;p) = 0$ for all $u \in \{1, \ldots, \omega - 1\}$. Thus, for any $u \in \{1, \ldots, \omega-1\}$, $\log f_G(u;p) \downarrow -\infty$ and $\ell(\boldsymbol{g}, p \mid \mathcal{S}_n)/n$ cannot obtain a maximum. When $p = 1$, $f_G(u;p) = 0$ for all $u \in \{1, \ldots, \omega\}$. Thus, $\log f_G(u;p) \downarrow -\infty$ for all $u \in \{1, \ldots, \omega\}$, and $\ell(\boldsymbol{g}, p \mid \mathcal{S}_n)/n$ similarly cannot obtain a maximum. For the boundaries of $\mathcal{C}$ in terms of $\boldsymbol{\mathcal{G}}$, the constraint $\sum_v g_v = 1$ requires at least one $g_v = 0$ for any $g_v = 1$ (or there is a $g_v = 0$ directly). Hence, $\log g_v \downarrow -\infty$ and a maximum cannot be obtained. Therefore, $\hat{p}$, $\hat{\boldsymbol{g}}$ are the MLE for the parameters $p, \boldsymbol{g}$ of the conditional bivariate probability mass function, h_*, defined in (1). □

B.7 Statement & Proof of Corollary B.7.1

This section provides a restatement of Theorem 3.3 under an alternative parameterization. Aside from completeness, one advantage of Corollary B.7.1 is the difference in parameter space for p. Under the PL geometric distribution in (13), $p \in (0, 1)$, whereas $p > 0$ for (S.6) in the discretized, PL exponential distribution. Such differences may have utility in any generalized linear model (GLM) regression analysis build from the model of (1).

Corollary B.7.1 (MLE of $\boldsymbol{g}$, p, discretized, PL exponential)**.** *Define the* discretized*, policy limit exponential distribution with parameter,* $p > 0$*, as*

$$f_G(u; p) = \begin{cases} \exp\left(- \dfrac{\{u - (\Delta + 1)\}}{p} \right) \left[1 - \exp\left(- \dfrac{1}{p} \right) \right] & \Delta + 1 \leq u \leq \omega - 1, \\ \exp\left(- \dfrac{\{u - (\Delta + 1)\}}{p} \right) & u = \omega. \end{cases} \tag{S.6}$$

Then, for the conditional bivariate probability mass function, h_**, defined in* (1)*, under the sampling conditions of Theorem 3.1, the MLE of the parameter* p *is*

$$\hat{p}_{\text{MLE}} = - \left[\ln\left(\frac{a}{a - b} \right) \right]^{-1}, \tag{S.7}$$

where a *and* b *follow* (15) *and* (16) *of Theorem 3.3, respectively. Further,* $S_G(\cdot; \hat{p})$ *is equivalent for* (S.6) *with* (S.7) *to* (13) *with* (14)*. Therefore, the MLE of* $\boldsymbol{g}$ *is equivalent to* (17) *in Theorem 3.3.*

Proof. Given the similarity to the proof of Theorem 3.3, we proceed with repetitive details omitted. Without loss of generality, let $\Delta = 0$. Given (S.6), the survival function then becomes the continuous equivalent,

$$S_G(u; p) = \exp\left(\frac{-(u - 1)}{p} \right),$$

for $u \in \{1, \ldots, \omega\}$. Hence, (12) simplifies. To see this, let $q(z; p) \equiv q(z) = \exp(-z/p)$ for

$z \in \{1, \ldots, \omega\}$ to write

$$\frac{\partial}{\partial p} \frac{\prod_{v=1}^{m} S(v;p)^{\hat{h}_{\bullet v}}}{\prod_{u=1}^{\omega} f(u;p)^{\hat{h}_{u\bullet}}} = \frac{q(a)\{1-q(1)\}^{-b}}{p^2}\left(a + \frac{b \cdot q(1)}{1-q(1)}\right).$$

Because $p > 0$,

$$\frac{q(a)\{1-q(1)\}^{-b}}{p^2}\left(a + \frac{b \cdot q(1)}{1-q(1)}\right) = 0 \iff a + \frac{b \cdot q(1)}{1-q(1)} = 0.$$

That is,

$$\hat{p} = -\left[\ln\left(\frac{a}{a-b}\right)\right]^{-1},$$

which is unique. Trivially, $\hat{p} \in \mathcal{C}$. To find $\hat{\boldsymbol{g}}$, replace $S_G(u;\hat{p})$ in (5). That $\hat{\boldsymbol{g}}$ is unique follows from the uniqueness of $\hat{p}$. Further, by Theorem 3.1, $\hat{\boldsymbol{g}} \in \mathcal{C}$.

To see that $\hat{p}$, $\hat{\boldsymbol{g}}$ are together the global maximum of $\mathcal{L}$, it is sufficient to examine the behavior of $\ell(\boldsymbol{g}, p \mid \mathcal{S}_n)/n \equiv \ell(\boldsymbol{g}, p, \hat{\pi} \mid \mathcal{S}_n)/n$ for the boundaries of $\mathcal{C}$. The analysis proceeds as in the final steps of the proof of Theorem 3.3. Therefore, $\hat{p}$, $\hat{\boldsymbol{g}}$ are the MLE for the parameters $p, \boldsymbol{g}$ of the conditional bivariate probability mass function, h_*, defined in (1). □

B.8 Proof of Theorem 3.4

Proof. The proof is similar to the proof of Theorem 3.1, and so we proceed with less detail. Without loss of generality, let $\Delta = 0$. For convenience of notation, define $\ell_\tau := \log \mathcal{L}_\tau(\boldsymbol{g}, p \mid \mathcal{S}_{\tau,n})/n$. The Lagrangian function (e.g., Ravishanker and Dey, 2002, §2.9, pg. 69) becomes

$$\begin{aligned}\ell_\tau = {} & -\log\alpha + \sum_{v=1}^{m} \hat{\gamma}_n(v) \log g_v \\ & + \frac{1}{n}\sum_{i=1}^{n}\{D_i \log f(Z_i;p) + (1-D_i)\log S(Z_i+1;p)\} + \pi\left(1 - \sum_{v=1}^{m} g_v\right).\end{aligned}$$

Because

$$\frac{\partial \ell_\tau}{\partial g_v} = -\frac{1}{\alpha}\left(\sum_{u=v}^{\xi} f(u;p)\right) + \hat{\gamma}_n(v)\frac{1}{g_v} - \pi,$$

we have

$$\sum_{v=1}^{m} g_v \left(\frac{\partial \ell_\tau}{\partial g_v}\right) = 0 \iff \hat{\pi} = 0,$$

as $\sum_v \hat{\gamma}_n(v) = 1$. Thus, any stationary point of the unconstrained optimization of $\mathcal{L}_\tau$ will also be a stationary point of the constrained optimization of $\mathcal{L}_\tau$ with solutions restricted to the convex subset, $\mathcal{C}$. This proves the final sentence of Theorem 3.4. Further, for all $v \in \mathcal{V}$,

$$\left.\frac{\partial \ell_\tau}{\partial g_v}\right|_{\hat{\pi}} = 0 \iff g_v = \frac{\alpha \hat{\gamma}_n(v)}{\sum_{u=v}^{\xi} f(u;p)}. \tag{S.8}$$

Thus, via (S.8),

$$\begin{aligned}\left.\frac{\partial \ell_\tau}{\partial p}\right|_{g_v} = 0 \iff & \sum_{v=1}^{m}\left(\frac{\hat{\gamma}_n(v)}{\sum_{u=v}^{\xi} f(u;p)}\right)\left(\sum_{u=v}^{\xi}\frac{\partial}{\partial p} f(u;p)\right) \\ &= \frac{1}{n}\sum_{i=1}^{n}\left(\frac{D_i}{f(Z_i;p)}\frac{\partial}{\partial p} f(Z_i;p) + \frac{1-D_i}{S(Z_i+1;p)}\frac{\partial}{\partial p} S(Z_i+1;p)\right).\end{aligned}$$

Finally, because we require $\sum_{\mathcal{V}} g_v = 1$, we have, for any $\hat{p}_\tau$ (i.e., $p \in \mathcal{P}$ such that (20) is satisfied) and $v \in \mathcal{V}$,

$$\hat{g}_\tau(v) = \frac{\hat{\gamma}_n(v)}{S(v;\hat{p}_\tau)}\left[\sum_{k=\Delta+1}^{\Delta+m}\frac{\hat{\gamma}_n(v)}{S(k;\hat{p}_\tau)}\right]^{-1}.$$

□

B.9 Proof of Corollary 3.4.1

Proof. Without loss of generality, assume $\Delta = 0$. The proof closely follows the proof of Corollary 3.1.1 and Theorem 3.4, and so we omit repetitive details. Recall the form of

$\mathcal{L}_\tau(\boldsymbol{g}, \boldsymbol{p} \mid \mathcal{S}_{\tau,n})$ to define the equivalent Lagrangian function

$$\log \mathcal{L}_\tau(\boldsymbol{g}, \boldsymbol{p} \mid \mathcal{S}_{\tau,n})/n = -\log \alpha + \sum_{v=1}^{m} \hat{\gamma}_n(v) \log g_v \\ + \frac{1}{n} \sum_{i=1}^{n} \{D_i \log f(Z_i; \boldsymbol{p}) + (1 - D_i) \log S(Z_i + 1; \boldsymbol{p})\} \\ + \pi_\tau \left(1 - \sum_{v=1}^{m} g_v\right).$$

Because

$$\frac{\partial \log \mathcal{L}_\tau(\boldsymbol{g}, \boldsymbol{p} \mid \mathcal{S}_{\tau,n})/n}{\partial g_v} = -\frac{1}{\alpha} \sum_{u=v}^{\xi} f(u; \boldsymbol{p}) + \frac{\hat{\gamma}_n(v)}{g_v} - \pi_\tau,$$

for all $v \in \mathcal{V}$, it follows that $\hat{\pi}_\tau = 0$. Further,

$$\left.\frac{\partial \log \mathcal{L}_\tau(\boldsymbol{g}, \boldsymbol{p} \mid \mathcal{S}_{\tau,n})/n}{\partial g_v}\right|_{\hat{\pi}_\tau} = 0 \iff g_v = \frac{\alpha \hat{\gamma}_n(v)}{\sum_{u=v}^{\xi} f(u; \boldsymbol{p})}. \tag{S.9}$$

Thus, from (S.9) and

$$\frac{\partial \log \mathcal{L}_\tau(\boldsymbol{g}, \boldsymbol{p} \mid \mathcal{S}_{\tau,n})/n}{\partial p_j} = -\frac{1}{\alpha}\left(\frac{\partial \alpha}{\partial p_j}\right) \\ + \frac{1}{n} \sum_{i=1}^{n} \left(\frac{D_i}{f(Z_i; \boldsymbol{p})} \frac{\partial}{\partial p_j} f(Z_i; \boldsymbol{p}) + \frac{1 - D_i}{S(Z_i + 1; \boldsymbol{p})} \frac{\partial}{\partial p_j} S(Z_i + 1; \boldsymbol{p})\right),$$

it follows that

$$\left.\frac{\partial \log \mathcal{L}_\tau(\boldsymbol{g}, \boldsymbol{p} \mid \mathcal{S}_{\tau,n})/n}{\partial p_j}\right|_{g_v} = -\varphi_1(j) + \varphi_2(j) = 0 \iff \varphi_1(j) = \varphi_2(j), \forall j = 1, \ldots, r',$$

where

$$\varphi_1(j) = \sum_{v=\Delta+1}^{\Delta+m} \left(\frac{\hat{\gamma}_n(v)}{\sum_{u=v}^{\xi} f(u; \boldsymbol{p})}\right)\left(\sum_{u=v}^{\xi} \frac{\partial}{\partial p_j} f(u; \boldsymbol{p})\right),$$

and

$$\varphi_2(j) = \frac{1}{n} \sum_{i=1}^{n} \left(\frac{D_i}{f(Z_i; \boldsymbol{p})} \frac{\partial}{\partial p_j} f(Z_i; \boldsymbol{p}) + \frac{1 - D_i}{S(Z_i + 1; \boldsymbol{p})} \frac{\partial}{\partial p_j} S(Z_i + 1; \boldsymbol{p})\right)\Bigg\}.$$

The set of simultaneous solutions, $\hat{\boldsymbol{p}}_\tau$, recovers the estimator (22). The proof is complete by replacing $\hat{\boldsymbol{p}}_\tau$ in (S.9) and using the constraint $\sum_{\mathcal{V}} g_v = 1$ to recover (21). □

B.10 Statement & Proof of Theorem B.2

As with (7), the estimator (20) takes the form of an asymptotically consistent M-estimator (van der Vaart, 1998, §5.3, pg. 51). We thus prove the equivalent of Theorem B.1 in Theorem B.2, from which the exact form of the asymptotically normal distribution of the estimator, (20), follows. The conditions of Theorem B.2 are typically satisfied under the standard regularity conditions (e.g., van der Vaart, 1998, §5.3, pg. 51; Mukhopadhyay, 2000, §12.2, pg. 539), including the existence of $\partial^2 f/\partial p^2$.

Theorem B.2. *Let $\hat{p}_{\tau,n}$ satisfy* (20) *and denote p_0 as the true parameter value. Define*

$$\begin{aligned}\Psi_{\tau,n}(p, \mathcal{S}_{\tau,n}) = &\sum_{v=\Delta+1}^{\Delta+m} \left(\frac{\hat{\gamma}_n(v)}{\sum_{u=v}^{\xi} f(u;p)}\right)\left(\sum_{u=v}^{\xi} \frac{\partial}{\partial p} f(u;p)\right)\\ &- \frac{1}{n}\sum_{i=1}^{n}\left(\frac{D_i}{f(Z_i;p)}\frac{\partial}{\partial p} f(Z_i;p) + \frac{1-D_i}{S(Z_i+1;p)}\frac{\partial}{\partial p} S(Z_i+1;p)\right).\end{aligned}$$

Then $\Psi_{\tau,n}(p,\mathcal{S}_{\tau,n}) \equiv \Psi_{\tau,n}(p)$ is an asymptotically consistent M-estimator of $\mathbf{E}\psi_\tau(Y_i, Z_i, D_i, p)$ for all $p \in \mathcal{P}$ (van der Vaart, 1998, §5.3, pg. 51), where

$$\begin{aligned}\psi_\tau(Y_i, Z_i, D_i, p) = &\sum_{v_*=\Delta+1}^{\Delta+m} \left(\frac{\mathbf{1}(Y_i = v_*)}{\sum_{u=v_*}^{\xi} f(u;p)}\right)\left(\sum_{u=v_*}^{\xi} \frac{\partial}{\partial p} f(u;p)\right)\\ &- \left(\frac{D_i}{f(Z_i;p)}\frac{\partial}{\partial p} f(Z_i;p) + \frac{1-D_i}{S(Z_i+1;p)}\frac{\partial}{\partial p} S(Z_i+1;p)\right),\end{aligned}$$

for $1 \le i \le n$. Further, $\Psi_{\tau,n}(\hat{p}_{\tau,n}) = 0$. If we also assume

(i) $\hat{p}_{\tau,n} \xrightarrow{\mathbf{P}} p_0$,

(ii) $\mathbf{E}[\psi_\tau(Y_i, Z_i, D_i, p_0)]^2 < \infty$,

(iii) $\mathbf{E}[\partial\psi_\tau(Y_i, Z_i, D_i, p_0)/\partial p]$ *exists, and*

(iv) $\partial^2\Psi_{\tau,n}(\tilde{p}_\tau)/\partial p^2$ *is* $O_{\mathbf{E}\psi_\tau}(1)$*, where* $\tilde{p}_\tau$ *is a point between* $\hat{p}_{\tau,n}$ *and* p_0*,*

then

$$\sqrt{n}(\hat{p}_{\tau,n} - p_0) \xrightarrow{\mathcal{L}} N\left(0, \frac{\mathbf{E}[\psi_\tau(Y_i, Z_i, D_i, p_0)^2]}{(\mathbf{E}[\partial\psi_\tau(Y_i, Z_i, D_i, p_0)/\partial p])^2}\right),$$

where

$$\frac{\partial}{\partial p}\psi_\tau(Y_i, Z_i, D_i, p_0) = \sum_{v_*=\Delta+1}^{\Delta+m} \mathbf{1}(Y_i = v_*)\left[\frac{(\sum_{u=v_*}^{\xi} f''(u))(\sum_{u=v_*}^{\xi} f(u)) - (\sum_{u=v_*}^{\xi} f'(u))^2}{(\sum_{u=v_*}^{\xi} f(u))^2}\right]$$
$$- \left(D_i\left[\frac{f''(Z_i)f(Z_i) - f'(Z_i)^2}{f(Z_i)^2}\right] + (1 - D_i)\left[\frac{S''(Z_i+1)S(Z_i+1) - S'(Z_i+1)^2}{S(Z_i+1)^2}\right]\right),$$

and f'*,* f''*,* S'*, and* S'' *denote* $\partial f/\partial p$*,* $\partial^2 f/\partial p^2$*,* $\partial S/\partial p$*, and* $\partial^2 S/\partial p^2$*, respectively.*

Proof. Recall $D = 0$ (which implies $\mathbf{1}(u = \tau + v)$) if an observation is right-censored and $D = 1$ (which implies $\mathbf{1}(u \leq \tau + v)$) otherwise (see Section 3.3 as needed). It is first instructive to show by (2) and (6),

$$\begin{aligned}
&\sum_{v=\Delta+1}^{m+\Delta}\sum_{u=v}^{\xi}\sum_{d=0}^{1}\{(d)h_*(u,v)\mathbf{1}(u \leq v+\tau) + (1-d)\bar{h}_*(u,v)\mathbf{1}(u = v+\tau)\}\\
&= \sum_{v=\Delta+1}^{m+\Delta}\frac{g_v}{\alpha}\sum_{u=v}^{\xi}\sum_{d=0}^{1}\{(d)f(u;p)\mathbf{1}(u \leq v+\tau) + (1-d)S(u+1;p)\mathbf{1}(u = v+\tau)\}\\
&= \frac{1}{\alpha}\sum_{v=\Delta+1}^{m+\Delta} g_v\left[\sum_{u=v:u=v+\tau} S(u+1;p) + \sum_{u=v:u\leq v+\tau} f(u;p))\right]\\
&= \frac{1}{\alpha}\sum_{v=\Delta+1}^{m+\Delta} g_v\left[S(v+\tau+1;p) + \sum_{u=v}^{v+\tau} f(u;p))\right]\\
&= \frac{1}{\alpha}\sum_{v=\Delta+1}^{m+\Delta} g_v\left(\sum_{u=v}^{\xi} f(u;p)\right)\\
&= 1,
\end{aligned}$$

is a valid probability density. Hence,

$$\mathbf{E}[\psi_\tau(Y_i, Z_i, D_i, p)] = \mathbf{E}[\eta_1(Y_i, Z_i, D_i, p)] - \mathbf{E}[\eta_2(Y_i, Z_i, D_i, p)], \tag{S.10}$$

where

$$\eta_1(Y_i, Z_i, D_i, p) = \sum_{v_*=\Delta+1}^{\Delta+m} \left(\frac{\mathbf{1}(Y_i = v_*)}{\sum_{u=v_*}^{\xi} f(u;p)} \right) \left(\sum_{u=v_*}^{\xi} \frac{\partial}{\partial p} f(u;p) \right),$$

and

$$\eta_2(Y_i, Z_i, D_i, p) = \frac{D_i}{f(Z_i;p)} \frac{\partial}{\partial p} f(Z_i;p) + \frac{1 - D_i}{S(Z_i+1;p)} \frac{\partial}{\partial p} S(Z_i+1;p).$$

We consider each expectation of (S.10) in turn for any i, $1 \leq i \leq n$. Observe,

$$\begin{aligned}
&\mathbf{E}[\eta_1(Y_i, Z_i, D_i, p)] \\
&= \sum_{v=\Delta+1}^{m+\Delta} \sum_{u=v}^{\xi} \sum_{d=0}^{1} \{(d)h_*(u,v)\mathbf{1}(u \leq v+\tau) + (1-d)\bar{h}_*(u,v)\mathbf{1}(u = v+\tau)\}\eta_1(v,u,d,p) \\
&= \sum_{v=\Delta+1}^{m+\Delta} \Bigg\{ \sum_{u=v:u=v+\tau}^{\xi} \frac{S(u+1;p)g_v}{\alpha} \left[\frac{\sum_{u=v}^{\xi} f'(u;p)}{\sum_{u=v}^{\xi} f(u;p)} \right] \\
&\qquad + \sum_{u=v:u\leq v+\tau}^{\xi} \frac{f(u;p)g_v}{\alpha} \left[\frac{\sum_{u=v}^{\xi} f'(u;p)}{\sum_{u=v}^{\xi} f(u;p)} \right] \Bigg\} \\
&= \sum_{v=\Delta+1}^{m+\Delta} \frac{g_v}{\alpha} \left[\frac{\sum_{u=v}^{\xi} f'(u;p)}{\sum_{u=v}^{\xi} f(u;p)} \right] \left\{ S(v+\tau+1) + \sum_{u=v}^{v+\tau} f(u;p) \right\} \\
&= \sum_{v=\Delta+1}^{m+\Delta} \frac{g_v}{\alpha} \left(\sum_{u=v}^{\xi} \frac{\partial}{\partial p} f(u;p) \right).
\end{aligned}$$

Similarly,

$$\begin{aligned}
&\mathbf{E}[\eta_2(Y_i, Z_i, D_i, p)] \\
&= \sum_{v=\Delta+1}^{m+\Delta} \sum_{u=v}^{\xi} \sum_{d=0}^{1} \{(d)h_*(u,v)\mathbf{1}(u \leq v+\tau) + (1-d)\bar{h}_*(u,v)\mathbf{1}(u = v+\tau)\}\eta_2(v,u,d,p) \\
&= \sum_{v=\Delta+1}^{m+\Delta} \Bigg\{ \sum_{u=v:u=v+\tau}^{\xi} \frac{S(u+1;p)g_v}{\alpha} \frac{S'(u+1;p)}{S(u+1;p)} + \sum_{u=v:u\leq v+\tau}^{\xi} \frac{f(u;p)g_v}{\alpha} \frac{f'(u;p)}{f(u;p)} \Bigg\} \\
&= \sum_{v=\Delta+1}^{m+\Delta} \frac{g_v}{\alpha} \left\{ S'(v+\tau+1) + \sum_{u=v}^{v+\tau} f'(u;p) \right\} \\
&= \sum_{v=\Delta+1}^{m+\Delta} \frac{g_v}{\alpha} \left(\sum_{u=v}^{\xi} \frac{\partial}{\partial p} f(u;p) \right).
\end{aligned}$$

Therefore, $\mathbf{E}[\eta_1(Y_i, Z_i, D_i, p)] = \mathbf{E}[\eta_2(Y_i, Z_i, D_i, p)]$ and $\mathbf{E}[\psi_\tau(Y_i, Z_i, D_i, p)] = 0$ for all $1 \leq i \leq n$. Further,

$$\Psi_{\tau,n}(p) = \frac{1}{n}\sum_{i=1}^{n} \psi_\tau(Y_i, Z_i, D_i, p),$$

and so $\Psi_{\tau,n}(p) \xrightarrow{\mathbf{P}} \psi_\tau(Y_i, Z_i, D_i, p)$ by the Law of Large Numbers (Lehmann and Casella, 1998, Theorem 8.2, pg. 54-55). That $\Psi_{\tau,n}(\hat{p}_n) = 0$ is immediate by the conditions of (20). The remainder follows the standard Taylor series analysis (e.g., van der Vaart, 1998, §5.3, pg. 51-52), with $\partial/\partial p(\psi_\tau)$ following by the quotient rule (Rudin, 1976, Theorem 5.3, pg. 104). □

Remark. *The conditions (i) through (iv) in Theorem B.2 may be relaxed. See van der Vaart (1998, Theorems 5.21 and 5.23 pg. 51-53) for details. Further, these results may be extended to higher dimensions of parameters, such as those assumed in Corollary 3.4.1. See the discussion van der Vaart (1998, Equation (5.20) pg. 51-52) and Section G for details.*

B.11 Statement & Proof of Corollary B.2.1

In practical settings, the true parameter, $p \in \mathcal{P}$, will not be known. Hence, we state the equivalent of Corollary B.1.1 in the additional incomplete data setting of right-censoring in Corollary B.2.1 for completeness. The conditions of Corollary B.2.1 are typically satisfied under the standard regularity conditions (e.g., van der Vaart, 1998, §5.3, pg. 51; Mukhopadhyay, 2000, §12.2, pg. 539).

Corollary B.2.1. *Assume the conditions of Theorem B.2 and define* $U_\tau = \mathbf{E}[\partial\psi_\tau(Y_i, Z_i, D_i, p_0)/\partial p]$, $U_{\tau,n} = \partial\Psi_{\tau,n}(\hat{p}_{\tau,n})/\partial p$, $V_\tau = \mathrm{Var}[\psi_\tau(Y_i, Z_i, D_i, p_0)]$, *and* $V_{\tau,n} = \sum_i \psi_\tau(Y_i, Z_i, D_i, \hat{p}_{\tau,n})^2/n$. *If* $U_{\tau,n} \xrightarrow{\mathbf{P}} U_\tau$ *and* $V_{\tau,n} \xrightarrow{\mathbf{P}} V_\tau$, *then*

$$[V_{\tau,n}/U_{\tau,n}^2]^{-1/2}\sqrt{n}(\hat{p}_{\tau,n} - p_0) \xrightarrow{\mathcal{L}} N(0,1). \tag{S.11}$$

Additionally, if the second Bartlett identity (Ferguson, 1996, pg. 120) is also satisfied, then

$U_\tau = V_\tau$ *with* U_τ *symmetric, and so* $[V_{\tau,n}]^{1/2}\sqrt{n}(\hat{p}_{\tau,n} - p_0) \xrightarrow{\mathcal{L}} N(0,1)$.

Proof. The result (S.11) follows from Theorem B.2 and Slutsky's Theorem (Lehmann and Casella, 1998, Theorem 8.10, pg. 58). The final sentence is a classical result of maximum likelihood theory (e.g., van der Vaart, 1998, §5.5). □

B.12 Proof of Corollary 3.4.2

Proof of Corollary 3.4.2. The novelty of this proof in comparison to the proof of Theorem 3.3 is to first derive the equivalent statement of Theorem 3.2 under the additional incomplete data setting of right-censoring. We now do this formally.

Lemma 1. *Assume the conditions of Theorem 3.4. Then* (20) *is satisfied if and only if*

$$\frac{\partial}{\partial p}\frac{\prod_{v=\Delta+1}^{\Delta+m} S(v;p)^{\hat{\gamma}_n(v)}}{\prod_{i=1}^{n} f(Z_i;p)^{D_i/n} S(Z_i+1;p)^{(1-D_i)/n}} = 0. \tag{S.12}$$

Proof of Lemma 1. Observe first

$$\sum_{v=\Delta+1}^{\Delta+m}\left(\frac{\hat{\gamma}_n(v)}{\sum_{u=v}^{\xi} f(u;p)}\right)\left(\sum_{u=v}^{\xi}\frac{\partial}{\partial p} f(u;p)\right) = \frac{\partial}{\partial p}\left(\sum_{v=\Delta+1}^{\Delta+m}\hat{\gamma}_n(v)\ln S(v;p)\right).$$

Similarly,

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n}&\left(\frac{D_i}{f(Z_i;p)}\frac{\partial}{\partial p}f(Z_i;p) + \frac{1-D_i}{S(Z_i+1;p)}\frac{\partial}{\partial p}S(Z_i+1;p)\right)\\ &= \frac{\partial}{\partial p}\left(\frac{1}{n}\sum_{i=1}^{n}\{D_i\ln f(Z_i;p) + (1-D_i)\ln S(Z_i+1;p)\}\right).\end{aligned}$$

Hence, the conditions on p in (20) are equivalent to all $p \in \mathcal{P}$ such that

$$\frac{\partial}{\partial p}\left(\sum_{v=\Delta+1}^{\Delta+m}\hat{\gamma}_n(v)\ln S(v;p) - \right.$$

$$\frac{1}{n}\sum_{i=1}^{n}\{D_i \ln f(Z_i;p) + (1-D_i)\ln S(Z_i+1;p)\}\Bigg) = 0. \tag{S.13}$$

But,

$$\sum_{v=\Delta+1}^{\Delta+m} \hat{\gamma}_n(v)\ln S(v;p) = \ln\left(\prod_{v=\Delta+1}^{\Delta+m} S(v;p)^{\hat{\gamma}_n(v)}\right),$$

and

$$\frac{1}{n}\sum_{i=1}^{n}\{D_i \ln f(Z_i;p) + (1-D_i)\ln S(Z_i+1;p)\}$$
$$= \ln\left(\prod_{i=1}^{n} f(Z_i;p)^{D_i/n} S(Z_i+1;p)^{(1-D_i)/n}\right).$$

Therefore, the conditions on p in (S.13) are equivalent to

$$\frac{\partial}{\partial p}\ln\left(\frac{\prod_{v=\Delta+1}^{\Delta+m} S(v;p)^{\hat{\gamma}_n(v)}}{\prod_{i=1}^{n} f(Z_i;p)^{D_i/n} S(Z_i+1;p)^{(1-D_i)/n}}\right) = 0 \tag{S.14}$$

But $f(\cdot;p), S(\cdot;p) > 0$ for all $p \in \mathcal{P}$, and so (S.14) is true true if and only if (S.12). $\square$

To complete the proof of Corollary 3.4.2, recall (13) and observe

$$\prod_{v=\Delta+1}^{\Delta+m} S(v;p)^{\hat{\gamma}_n(v)} = (1-p)^{\sum_{\mathcal{V}}(v-(\Delta+1))\hat{\gamma}_n(v)},$$

$$\prod_{i=1}^{n} f(Z_i;p)^{D_i/n} = p^{(\sum_i \mathbf{1}(Z_i \neq \xi)D_i)/n}(1-p)^{(\sum_i (Z_i-(\Delta+1))D_i)/n},$$

and

$$\prod_{i=1}^{n} S(Z_i+1;p)^{(1-D_i)/n} = (1-p)^{(\sum_i (Z_i+1-(\Delta+1))(1-D_i))/n}.$$

Thus, we obtain the simplified form of (S.12) in Lemma 1.

$$\frac{\partial}{\partial p}\frac{\prod_{v=\Delta+1}^{\Delta+m} S(v;p)^{\hat{\gamma}_n(v)}}{\prod_{i=1}^{n} f(Z_i;p)^{D_i/n} S(Z_i+1;p)^{(1-D_i)/n}} \equiv \frac{\partial}{\partial p}\frac{(1-p)^{a_\tau}}{p^{b_\tau}}.$$

The remainder of the proof follows the proof of Theorem 3.3. □

C Likelihood with Censoring

In this section, we numerically illustrate how the presence of right-censored data that generates $\bar{h}_*$ and h_* impacts the likelihood, $\mathcal{L}_\tau$ of Section 3.3. Suppose $g(1) = 0.5$, $g(2) = 0.30$, and $g(3) = 0.20$. Hence, $\Delta = 0$ and $m = 3$. Further suppose X follows (13) with $p = 0.6$ and $\omega = 4$. That is, $\Pr(X = 1) = 0.6$, $\Pr(X = 2) = 0.24$, $\Pr(X = 3) = 0.096$, and $\Pr(X = 4) = 0.064$. Finally, set $\varepsilon = 6$, and so right-censoring is present in the data because $\varepsilon < \omega + m$ (Lautier et al., 2023). The complete probability density function for all possible samples of (Y_i, Z_i, D_i) may be found in Table C1.

We can see that not all possible combinations of (Y_i, Z_i, D_i) are observable when $\varepsilon = 6$. For example, $(Y_i = 2, Z_i = 2, D_i = 0)$ is not a possible observation because the censoring time, $Y_i + \varepsilon - (m + \Delta + 1) \equiv Y_i + \tau$, would be $Y_i + \tau = 4 > 2 = Z_i$. Hence, $D_i = \mathbf{1}(X_i \leq C_i)$ cannot be equal to 0. Of the 18 possible combinations of (Y_i, Z_i, D_i), we present all 10 possible observations in Table C1. It may be verified that the sum of the $\bar{h}_*$ and h_* columns in Table C1 taken together is unity. This is a numeric validation that the likelihood under right-censoring, $\mathcal{L}_\tau$, is formed through a valid probability density function. A more formal demonstration may be found in Section B.10, in the lead up to (S.10).

D Implementation Reference

Recall the PL geometric distribution with parameter, $0 < p < 1$, defined in Theorem 3.3,

$$f_G(u; p) = \begin{cases} p(1-p)^{u-(\Delta+1)} & \Delta + 1 \leq u \leq \omega - 1, \\ (1-p)^{u-(\Delta+1)} & u = \omega. \end{cases}$$

Y_i	Z_i	D_i	$g(Y_i)$	$f(Z_i)$	$S(Z_i+1)$	$\bar{h}_*(Z_i, Y_i)$	$h_*(Z_i, Y_i)$
1	3	0	0.50	0.096	0.064	0.0491	—
2	4	0	0.30	0.064	0.000	0.0000	—
1	1	1	0.50	0.600	0.064	—	0.4601
1	2	1	0.50	0.240	0.064	—	0.1840
1	3	1	0.50	0.096	0.064	—	0.0736
2	2	1	0.30	0.240	0.000	—	0.1104
2	3	1	0.30	0.096	0.000	—	0.0442
2	4	1	0.30	0.064	0.000	—	0.0294
3	3	1	0.20	0.096	0.000	—	0.0294
3	4	1	0.20	0.064	0.000	—	0.0196

Table C1: **Complete Density Right-Censoring**. The complete density function for all possible sampling triples (Y_i, Z_i, D_i) under right-censoring and the density assumptions of Section C with $\varepsilon = 6$. The probability mass function $\bar{h}_*$ is only valid when $Y_i + \tau = Z_i$. The probability mass function h_* is only valid when $Z_i \leq Y_i + \tau$. This implies not all triples of (Y_i, Z_i, D_i) are possible observations. It may be verified that the sum of the $\bar{h}_*$ and h_* columns together is unity.

Then,

$$\frac{\partial}{\partial p} f_G(u;p) = f_G(u;p)\left(\frac{\mathbf{1}(u \neq \omega)}{p} - \frac{u-(\Delta+1)}{1-p}\right),$$

$$\frac{\partial^2}{\partial p^2} f_G(u;p) = f_G(u;p)\left[\frac{u-(\Delta+1)}{1-p}\left(\frac{u-\Delta-2}{1-p} - \frac{2\times\mathbf{1}(u\neq\omega)}{p}\right)\right],$$

$$\frac{\partial}{\partial p} S_G(u;p) = (\Delta+1-u)(1-p)^{u-\Delta-2},$$

and

$$\frac{\partial^2}{\partial p^2} S_G(u;p) = (u-\Delta-2)(u-\Delta-1)(1-p)^{u-\Delta-3}.$$

For a shifted binomial distribution over the support $\{\Delta+1, \ldots, \omega\}$ with probability of success $0 < \theta < 1$, we have the probability density function

$$f(u;\theta) = \binom{\omega-(\Delta+1)}{u-(\Delta+1)}\theta^{u-(\Delta+1)}(1-\theta)^{\omega-u}, \quad u \in \{\Delta+1, \ldots, \omega\}.$$

Thus,

$$\frac{\partial}{\partial \theta} f(u;\theta) = f(u;\theta)\left(\frac{u-(\Delta+1)}{\theta} - \frac{\omega-u}{1-\theta}\right),$$

and

$$\frac{\partial^2}{\partial\theta^2} f(u;\theta)$$
$$= f(u;\theta)\left(\frac{(u-\Delta-1)(u-\Delta-2)}{\theta^2} - 2\frac{u-(\Delta+1)}{\theta}\frac{\omega-u}{1-\theta} + \frac{(\omega-u)(\omega-u-1)}{(1-\theta)^2}\right).$$

Recall the PL discrete Weibull distribution defined in (25):

$$f_W(u;p_1,p_2) = \begin{cases} p_1^{(u-(\Delta+1))^{p_2}} - p_1^{(u-\Delta)^{p_2}}, & \Delta+1 \le u \le \omega-1 \\ p_1^{(u-(\Delta+1))^{p_2}}, & u=\omega. \end{cases}$$

Thus,

$$\frac{\partial}{\partial p_1} f_W(u;p_1,p_2) = \begin{cases} (u-(\Delta+1))^{p_2} {p_1}^{(u-(\Delta+1))^{p_2}-1} - (u-\Delta)^{p_2} {p_1}^{(u-\Delta)^{p_2}-1}, & \Delta+1 \le u \le \omega-1 \\ (u-(\Delta+1))^{p_2} {p_1}^{(u-(\Delta+1))^{p_2}-1}, & u=\omega. \end{cases}$$

$$\frac{\partial^2}{\partial p^2} f_W(u;p_1,p_2) = \begin{cases} ((u-(\Delta+1))^{p_2}-1)(u-(\Delta+1))^{p_2} p_1^{(u-(\Delta+1))^{p_2}-2} & \\ \quad - ((u-\Delta)^{p_2}-1)(u-\Delta)^{p_2} p_1^{(u-\Delta)^{p_2}-2}, & \Delta+1 \le u \le \omega-1 \\ ((u-(\Delta+1))^{p_2}-1)(u-(\Delta+1))^{p_2} p_1^{(u-(\Delta+1))^{p_2}-2}, & u=\omega. \end{cases}$$

$$\frac{\partial^2}{\partial p_2 \partial p_1} f_W(u;p_1,p_2) = \begin{cases} -p_1^{(u-\Delta)^{p_2}-1} \ln(p_1) \ln(u-\Delta)(u-\Delta)^{2p_2} \\ \quad - p_1^{(u-\Delta)^{p_2}-1} \ln(u-\Delta)(u-\Delta)^{p_2}, & u=\Delta+1 \\ -p_1^{(u-\Delta)^{p_2}-1} \ln(p_1) \ln(u-\Delta)(u-\Delta)^{2p_2} \\ \quad - p_1^{(u-\Delta)^{p_2}-1} \ln(u-\Delta)(u-\Delta)^{p_2} \\ \quad + p_1^{(u-(\Delta+1))^{p_2}-1} \ln(p_1) \ln(u-(\Delta+1))(u-(\Delta+1))^{2p_2} \\ \quad + p_1^{(u-(\Delta+1))^{p_2}-1} \ln(u-(\Delta+1))(u-(\Delta+1))^{p_2}, & \Delta+2 \le u \le \omega-1 \\ p_1^{(u-(\Delta+1))^{p_2}-1} \ln(p_1) \ln(u-(\Delta+1))(u-(\Delta+1))^{2p_2} \\ \quad + p_1^{(u-(\Delta+1))^{p_2}-1} \ln(u-(\Delta+1))(u-(\Delta+1))^{p_2}, & u=\omega. \end{cases}$$

$$\frac{\partial}{\partial p_2} f_W(u;p_1,p_2) = \begin{cases} 0, & u=\Delta+1 \\ p_1^{(u-(\Delta+1))^{p_2}} \ln(p_1) \ln(u-(\Delta+1))(u-(\Delta+1))^{p_2} \\ \quad - p_1^{(u-\Delta)^{p_2}} \ln(p_1) \ln(u-\Delta)(u-\Delta)^{p_2}, & \Delta+1 \le u \le \omega-1 \\ p_1^{(u-(\Delta+1))^{p_2}} \ln(p_1) \ln(u-(\Delta+1))(u-(\Delta+1))^{p_2}, & u=\omega. \end{cases}$$

$$\frac{\partial^2}{\partial p_2^2} f_W(u;p_1,p_2) = \begin{cases} 0, & u=\Delta+1 \\ -p_1^{(u-\Delta)^{p_2}} \ln^2(p_1) \ln^2(u-\Delta)(u-\Delta)^{2p_2} \\ \quad - p_1^{(u-\Delta)^{p_2}} \ln(p_1) \ln^2(u-\Delta)(u-\Delta)^{p_2} \\ \quad + p_1^{(u-(\Delta+1))^{p_2}} \ln^2(p_1) \ln^2(u-(\Delta+1))(u-(\Delta+1))^{2p_2} \\ \quad + p_1^{(u-(\Delta+1))^{p_2}} \ln(p_1) \ln^2(u-(\Delta+1))(u-(\Delta+1))^{p_2}, & \Delta+2 \le u \le \omega-1 \\ p_1^{(u-(\Delta+1))^{p_2}} \ln^2(p_1) \ln^2(u-(\Delta+1))(u-(\Delta+1))^{2p_2} \\ \quad + p_1^{(u-(\Delta+1))^{p_2}} \ln(p_1) \ln^2(u-(\Delta+1))(u-(\Delta+1))^{p_2}, & u=\omega. \end{cases}$$

$$\frac{\partial}{\partial p_1} S_W(u;p_1,p_2) = (u-(\Delta+1))^{p_2}\, p_1^{(u-(\Delta+1))^{p_2}-1}$$

$$\frac{\partial^2}{\partial p^2} S_W(u;p_1,p_2) = \left((u-(\Delta+1))^{p_2}-1\right)(u-(\Delta+1))^{p_2}\, p_1^{(u-(\Delta+1))^{p_2}-2}$$

$$\frac{\partial^2}{\partial p_2 \partial p_1} S_W(u;p_1,p_2) = \begin{cases} 0, & u = \Delta+1 \\ p_1^{(u-(\Delta+1))^{p_2}-1} \ln(p_1) \ln(u-(\Delta+1)) (u-(\Delta+1))^{2p_2} & \\ \quad + p_1^{(u-(\Delta+1))^{p_2}-1} \ln(u-(\Delta+1)) (u-(\Delta+1))^{p_2}, & \Delta+1 \le u \le \omega \end{cases}$$

$$\frac{\partial}{\partial p_2} S_W(u;p_1,p_2) = \begin{cases} 0, & u = \Delta+1 \\ p_1^{(u-(\Delta+1))^{p_2}} \ln(p_1) \ln(u-(\Delta+1)) (u-(\Delta+1))^{p_2}, & \Delta+1 \le u \le \omega \end{cases}$$

$$\frac{\partial^2}{\partial p_2^2} S_W(u;p_1,p_2) = \begin{cases} 0, & u = \Delta+1 \\ p_1^{(u-(\Delta+1))^{p_2}} \ln^2(p_1) \ln^2(u-(\Delta+1)) (u-(\Delta+1))^{2p_2} & \\ \quad + p_1^{(u-(\Delta+1))^{p_2}} \ln(p_1) \ln^2(u-(\Delta+1)) (u-(\Delta+1))^{p_2} & \Delta+1 \le u \le \omega \end{cases}$$

E Simulation Procedure Outline

To simulate left-truncated data from the distribution h_* defined in (1), the following procedure may be employed.

1. Select values for Δ, m, and ω and create a pairwise mapping for all possible pairs $(u,v) \in \mathcal{A}$, where $\Delta+1 \le v \le \Delta+m$, $\Delta+1 \le u \le \omega$, and $u \le v$.

2. Select a distribution and parameters for the lifetime distribution, X, $f(\cdot;p)$ and the left-truncation distribution, Y, $\boldsymbol{g}$.

3. Using the choices in the previous step, calculate (1) over all pairs $(u,v) \in \mathcal{A}$. This will require calculating the probability α.

4. Starting with the pair $(\Delta+1, \Delta+1)$ and ending with the pair $(\omega, \Delta+m)$, create a one-to-one lower bound mapping from 0 by cumulative sums to $\sum_{\mathcal{A}\setminus(\omega,\Delta+m)} h_*(u,v)$. Call this lower bound $\lfloor H_*(u,v) \rfloor$ for $(u,v) \in \mathcal{A}$.

5. Starting with the pair $(\Delta+1, \Delta+1)$ and ending with the pair $(\omega, \Delta+m)$, create a

u	v	$\lfloor H_*(u,v) \rfloor$	$\lceil H_*(u,v) \rceil$
1	1	0.0000000	0.1856436
2	1	0.1856436	0.3155941
2	2	0.3155941	0.3935644
3	1	0.3935644	0.4845297
3	2	0.4845297	0.5391089
3	3	0.5391089	0.5754950
4	1	0.5754950	0.7877475
4	2	0.7877475	0.9150990
4	3	0.9150990	1.0000000

Table E1: **Illustrative Simulation Mapping**. The above table illustrates how to simulate left-truncated data from the bivariate distribution, h_* defined in (1) for f following (13) with $p = 0.30$ and $\boldsymbol{g} = (0.5, 0.3, 0.2)^\top$. For example, a random uniform number from the interval $(0, 1)$ of 0.4000497 would result in the simulated pair $(3, 1)$.

one-to-one upper bound mapping from $h_*(\Delta+1, \Delta+1)$ by cumulative sums to 1. Call this upper bound $\lceil H_*(u,v) \rceil$ for $(u,v) \in \mathcal{A}$.

6. Simulate a continuous uniform random number in the interval $(0, 1)$, say ρ. The simulated pair $(u, v) \in \mathcal{A}$ is the pair such that $\lfloor H_*(u,v) \rfloor \leq \rho \leq \lceil H_*(u,v) \rceil$. Repeat as needed for the desired sample size.

F Simulation Study Additional Details

This section closely follows the order of Section 4 and may be read concurrently, as needed. For reference, Sections D and E provide helpful background material on preparing the results of this section.

Section 4 opens with an example in which a direct numerical optimization on (4) fails to recover the true parameter values. For the reported time of 25.07 minutes, the computer was a `Surface Pro 9` with a `12th Gen Intel(R) Core(TM) i5-1235U 2.50 GHz` processor, 8.00GB installed RAM, 10 cores, and 12 logical processors. The resulting parameter estimates are summarized in Table F1.

We now provide numerical verification of Theorems 3.1, B.1, 3.3, 3.4, B.2, and Corol-

	$\boldsymbol{\Theta}$	`contstrOptim`	Theorem 3.1	Theorem 3.3
p	0.05000	0.47904	0.05318	0.05318
g_1	0.00828	0.00029	0.00868	0.00868
g_2	0.04015	0.00158	0.03210	0.03210
g_3	0.08648	0.00669	0.08025	0.08025
g_4	0.10865	0.04471	0.10449	0.10449
g_5	0.08775	0.04168	0.07332	0.07332
g_6	0.04725	0.05606	0.04891	0.04891
g_7	0.01696	0.05335	0.02497	0.02497
g_8	0.00391	0.06518	0.00182	0.00182
g_9	0.00053	0.07223	0.00000	0.00000
g_{10}	0.00003	0.03974	0.00000	0.00000
g_{11}	0.01243	0.02852	0.01607	0.01607
g_{12}	0.06022	0.03099	0.06449	0.06449
g_{13}	0.12971	0.07770	0.13025	0.13025
g_{14}	0.16297	0.08743	0.15144	0.15144
g_{15}	0.13163	0.09085	0.15062	0.15062
g_{16}	0.07088	0.05174	0.06898	0.06898
g_{17}	0.02544	0.09321	0.03568	0.03568
g_{18}	0.00587	0.05682	0.00628	0.00628
g_{19}	0.00079	0.04465	0.00166	0.00166
g_{20}	0.00005	0.05755	0.00000	0.00000

Table F1: **Limits of Numeric Optimization**. A direct numerical optimization on (4) fails to recover the true parameter values, $\boldsymbol{\Theta}$, after 25 minutes, whereas the results of Section 3 do recover close estimates of $\boldsymbol{\Theta}$ in under one second. For details, see Section 4.

lary 3.4.2. We assume $m = 3$, $\Delta = 0$, and $\omega = 4$. This results in a 4×3 trapezoid, $\mathcal{A}$. For the lifetime random variable, we consider two parametric distributions. The first assumes X takes the form of (13) with $p = 0.3$. This results in the pmf of X as $\Pr(X = 1) = 0.3$, $\Pr(X = 2) = 0.21$, $\Pr(X = 3) = 0.147$, and $\Pr(X = 4) = 0.343$. We then assume the pmf of Y is $g_1 \equiv \Pr(Y = 1) = 0.5$, $g_2 \equiv \Pr(Y = 2) = 0.3$, and $g_3 \equiv \Pr(Y = 3) = 0.2$. This results in $\alpha = 0.808$. Observe that Y is non-uniform, which demonstrates our results may be applied outside the domain of length-biased sampling. The second assumes X takes the form of a shifted binomial distribution with the number of successes equal to $\omega - (\Delta + 1) = 3$ and a probability of success equal to $0 < \theta = 0.75 < 1$. This results in the pmf of X as $\Pr(X = 1) = 0.016$, $\Pr(X = 2) = 0.141$, $\Pr(X = 3) = 0.422$, and $\Pr(X = 4) = 0.422$.

Because $\boldsymbol{g}$ is unchanged, we obtain $\alpha = 0.964$.

To perform the numeric validation for Theorem 3.1 and Theorem 3.3, we simulate a single sample size of $n = 1{,}000$ under the setting of Section 3.2 (i.e., $\mathcal{S}_n$). We then estimate the parameters in three ways. First, we solve (4) with a direct, constrained numeric optimization using `constroOptim` via R Core Team (2023). Next, we perform a single parameter optimization through Theorem 3.1 using `optimize` via R Core Team (2023) in combination with the closed-form solutions for (5). Finally, we report the parameter estimates with all closed-form solutions using Theorem 3.3. The results are summarized in the top-half of Table F2. All three approaches yield nearly identical solutions, all of which are quite close to the true parameter values. For completeness, we also estimate computational performance statistics for all three parameter estimation approaches with the `microbenchmark` package (Mersmann, 2023). It is immediate that both Theorems 3.1 and 3.3 provide substantial improvements in computational demands. We then repeat this procedure under the sampling assumptions of Section 3.3 (i.e., $\mathcal{S}_{\tau,n}$) with $\varepsilon = 6$ to verify Theorem 3.4 and Corollary 3.4.2. The comparison is again quite close. The full results may be found in Table F2.

We also provide an additional simulation study to supplement the robustness analysis of Table 3 with a larger value of p and non-zero Δ parameter in Table F3.

We next validate Theorem B.1. First, we calculate the true asymptotically normal distribution assuming a sample size of $n = 1{,}000$. Next, we perform 1,000 replicates of simulating a sample size of $n = 1{,}000$ and estimating $\hat{p}_n$ with Theorem 3.3 (for the shifted-binomial distribution, Theorem 3.1). The true asymptotic density and the empirical density of the 1,000 replicates of $\sqrt{n}(\hat{\theta}_n - \theta_0)$ and $\sqrt{n}(\hat{p}_n - p_0)$ may be found in Figure F1. The true density (solid line) and empirical density (dashed line) closely agree. This process is repeated under the sampling assumptions of left-truncation and right-censoring to validate Theorem B.2. Once again, the true density (solid line) and empirical density (dashed line) closely agree. These results may also be found in Figure F1.

The coverage probabilities in Table 3 represent the percentage of 1,000 replicates such

Table F2: **Numeric Validation and Performance Summary**. Numerical verification of Theorems 3.1, 3.3, and 3.4, and Corollary 3.4.2. Trapezoid parameters are $m = 3$, $\Delta = 0$, and $\omega = 4$ (with $\varepsilon = 6$ for right-censoring; $\mathcal{C}_\varepsilon = 21.22\%$ for PL geometric; $\mathcal{C}_\varepsilon = 21.88\%$ for shifted-binomial). Single sample of size $n = 1{,}000$. Performance calculations per the `microbenchmark` package (Mersmann, 2023) evaluated in microseconds (μS). The computer was a `Surface Pro 9` with a `12th Gen Intel(R) Core(TM) i5-1235U 2.50 GHz` processor, 8.00GB installed RAM, 10 cores, and 12 logical processors.

		Left-Trunc. (§3.2)			**Left-Trunc. & Ri.-Cens. (§3.3)**		
		PL Geometric			*PL Geometric*		
Param.	True	`constrOptim`	Thm 3.1	Thm 3.3	`constrOptim`	Thm 3.4	Cor 3.4.2
p	0.30	0.3070043	0.3069670	0.3069670	0.3033170	0.3033175	0.3033175
g_1	0.50	0.4917720	0.4913635	0.4913635	0.5184552	0.5188642	0.5188642
g_2	0.30	0.3039001	0.3036941	0.3036941	0.2818413	0.2820547	0.2820547
g_3	0.20	0.2050589	0.2049424	0.2049424	0.1989716	0.1990811	0.1990811
Speed (μS)		562,996	6,797	861	9,027,715	169,510	12,706
		Shifted-Binomial			*Shifted-Binomial*		
Param.	True	`constrOptim`	Thm 3.1		`constrOptim`	Thm 3.4	
θ	0.75	0.7462976	0.7462779		0.7445825	0.7446186	
g_1	0.50	0.4870210	0.4867887		0.5251005	0.5246253	
g_2	0.30	0.2949194	0.2947756		0.2733774	0.2731187	
g_3	0.20	0.2185810	0.2184357		0.2024117	0.2022560	
Speed (μS)		543,432	8,633		5,100,241	233,909	

that,

$$\hat{p}_n - Z_{0.975}\sqrt{\frac{1/V_n}{n}} \leq p_0 \leq \hat{p}_n + Z_{0.975}\sqrt{\frac{1/V_n}{n}}, \tag{S.15}$$

where $Z_{0.975}$ represents the 97.5th percentile of a standard normal random variable.

G Application Additional Details

To begin, Figure G1 provides a numeric verification of the likelihood ratio test (LRT) proposed in Section 5.1. For details, please the caption of Figure G1.

When the parametric distribution for the lifetime random variable, f, has more than one parameter, the estimating equation theory (e.g., Theorems B.1 and Theorem B.2) will need to be vectorized. We illustrate how to do this in the case f follows two parameters, such as

Table F3: **Robustness Simulation Study, Supplement**. A robustness analysis of estimation methods (i.e., §3.2, §3.3) for (13) with $Y \sim \mathcal{B}$ by sample size (n) and level of right-censoring (ε, $\mathcal{C}_\varepsilon$). Trapezoid parameters are $m = 15$, $\Delta = 5$, and $\omega = 24$. We report the empirical mean (eMean), empirical standard deviation (eSD), theoretical standard deviation (Thm B.1, Thm B.2), and coverage probability (CP), i.e., (S.15), for a 95% asymptotic confidence interval estimated using Corollaries B.1.1 and B.2.1.

		$\varepsilon = 25$ (5.0%)				$\varepsilon = 30$ (0.3%)			
n	p_0	eMean	eSD	Thm B.2	CP	eMean	eSD	Thm B.2	CP
50	0.45	0.4550	0.0489	0.0484	0.944	0.4542	0.0467	0.0473	0.950
100	0.45	0.4522	0.0344	0.0342	0.950	0.4540	0.0346	0.0334	0.944
250	0.45	0.4496	0.0214	0.0217	0.953	0.4523	0.0218	0.0211	0.948
500	0.45	0.4491	0.0153	0.0153	0.950	0.4505	0.0154	0.0149	0.942
		$\varepsilon = 35$ (0.01%)				No Censoring			
n	p_0	eMean	eSD	Thm B.2	CP	eMean	eSD	Thm B.1	CP
50	0.45	0.4554	0.0464	0.0472	0.963	0.4554	0.0454	0.0472	0.961
100	0.45	0.4506	0.0342	0.0334	0.944	0.4527	0.0320	0.0334	0.965
250	0.45	0.4509	0.0214	0.0211	0.946	0.4510	0.0213	0.0211	0.948
500	0.45	0.4507	0.0150	0.0149	0.943	0.4503	0.0150	0.0149	0.948

the PL discrete Weibull of (25). For details, see van der Vaart (1998, Eq. (20), pg. 52).

When $\boldsymbol{p} = (p_1, p_2)^\top$, there are two estimating equations. For example, following Theorem B.1, for $j = 1, 2$,

$$\Psi_{n,j}(p_j, \mathcal{S}_n) = \sum_{v=\Delta+1}^{\Delta+m} \left(\frac{\hat{h}_{\bullet v}}{\sum_{u=v}^{\omega} f(u; p_j)} \right) \left(\sum_{u=v}^{\omega} \frac{\partial}{\partial p_j} f(u; p_j) \right) - \sum_{v=\Delta+1}^{\Delta+m} \sum_{u=v}^{\omega} \frac{\hat{h}_{uv}}{f(u; p_j)} \frac{\partial}{\partial p_j} f(u; p_j),$$

and

$$\psi_j(X_i, Y_i, p_j) = \sum_{v=\Delta+1}^{\Delta+m} \left(\frac{\sum_{u=v}^{\omega} W_i}{\sum_{u=v}^{\omega} f(u; p_j)} \right) \left(\sum_{u=v}^{\omega} \frac{\partial}{\partial p_j} f(u; p_j) \right) - \sum_{v=\Delta+1}^{\Delta+m} \sum_{u=v}^{\omega} \frac{W_i}{f(u; p_j)} \frac{\partial}{\partial p_j} f(u; p_j).$$

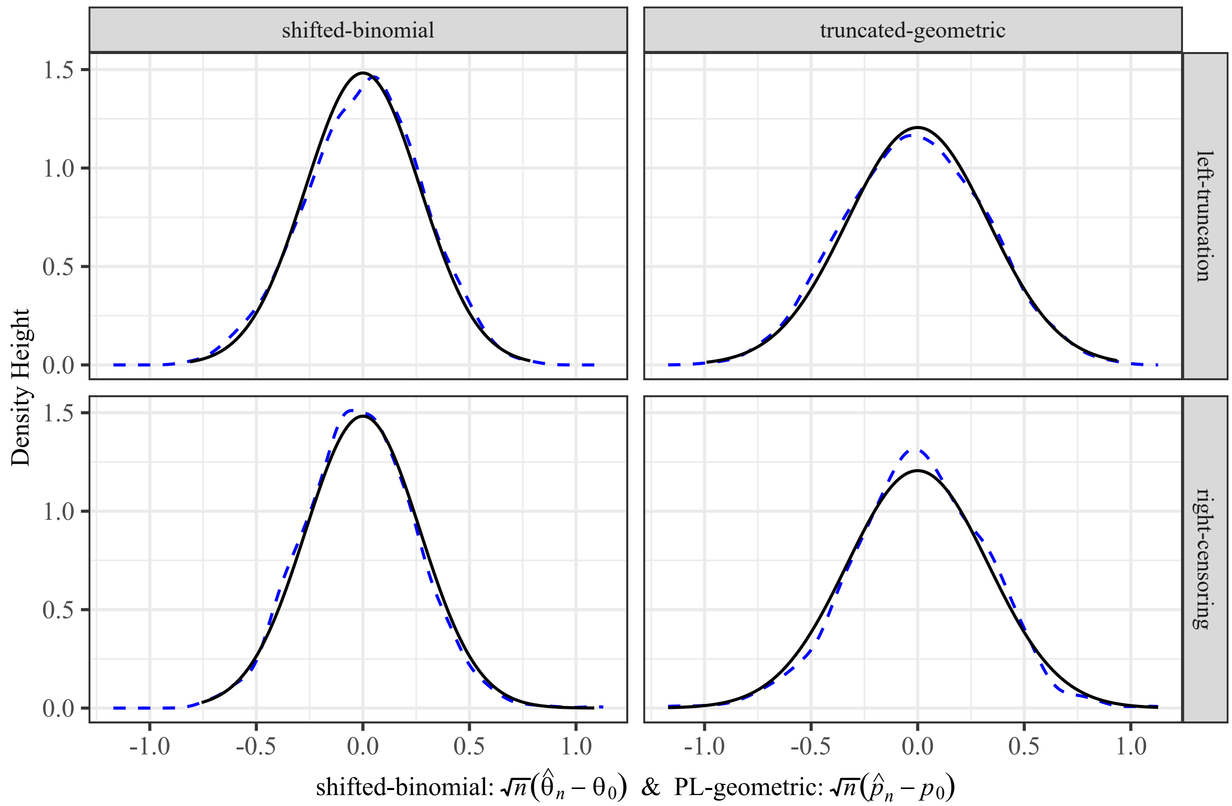

Figure F1: **Asymptotic Normality Verification**. Numeric verification of Theorems B.1 and B.2. Trapezoid parameters are $m = 3$, $\Delta = 0$, and $\omega = 4$ (with $\varepsilon = 6$ for right-censoring; $\mathcal{C}_\varepsilon = 21.22\%$ for PL geometric; $\mathcal{C}_\varepsilon = 21.88\%$ for shifted-binomial). Results for 1,000 replicates each with a sample size of $n = 1{,}000$. For random variable specifications, see Table F2.

Then,

$$\sqrt{n}\begin{bmatrix}\hat{p}_{n,1} - p_{0,1} \\ \hat{p}_{n,2} - p_{0,2}\end{bmatrix} \xrightarrow{\mathcal{L}} \mathbf{N}_2(\mathbf{0}_2, \mathbf{U}_{2\times 2}^{-1}\mathbf{V}_{2\times 2}(\mathbf{U}_{2\times 2}^{-1})^\top),$$

where

$$\mathbf{U}_{2\times 2} = \begin{bmatrix}\mathbf{E}(\partial\psi_1/\partial p_1) & \mathbf{E}(\partial\psi_1/\partial p_2) \\ \mathbf{E}(\partial\psi_2/\partial p_1) & \mathbf{E}(\partial\psi_2/\partial p_2)\end{bmatrix}, \quad \text{and} \quad \mathbf{V}_{2\times 2} = \begin{bmatrix}\mathbf{E}(\psi_1^2) & \mathbf{E}(\psi_1\psi_2) \\ \mathbf{E}(\psi_2\psi_1) & \mathbf{E}(\psi_2^2)\end{bmatrix}.$$

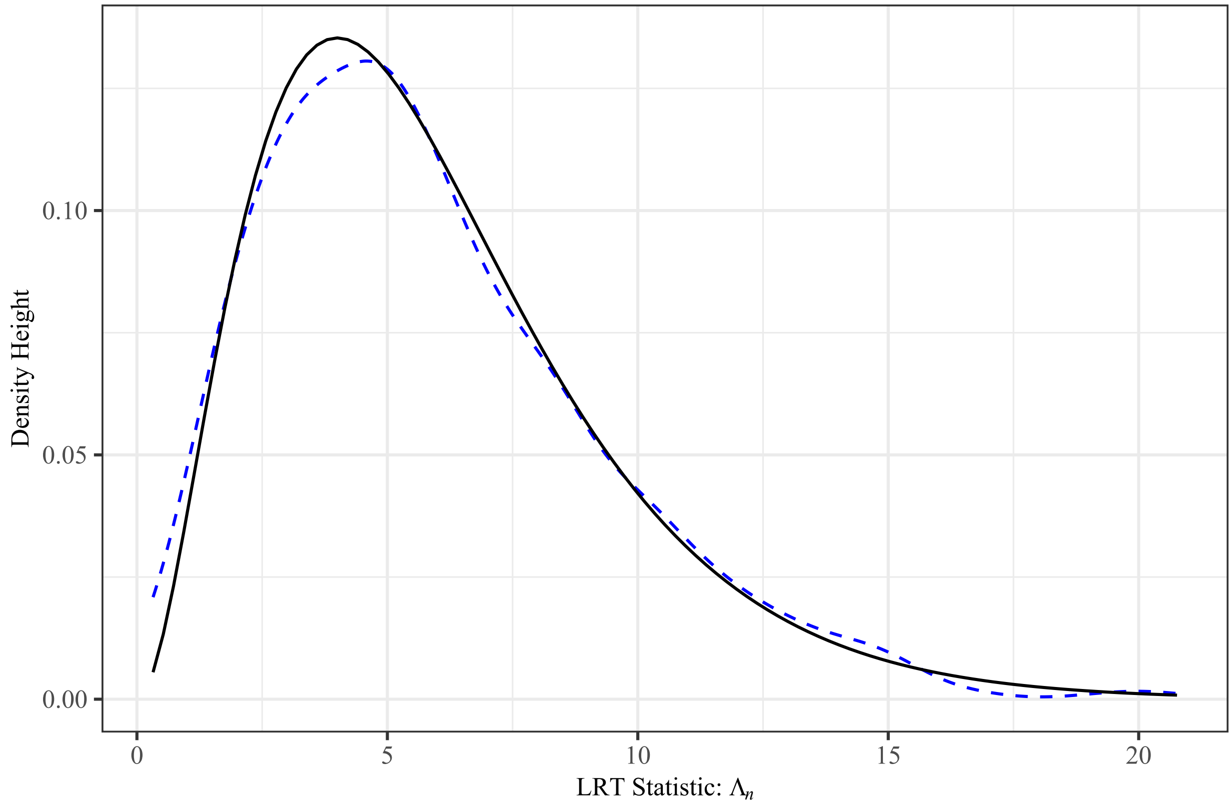

Figure G1: **Likelihood Ratio Test, Simulation Study**. Numeric verification of the proposed LRT in Section 5.1. Trapezoid parameters are $m = 5$, $\Delta = 0$, $\omega = 8$, and $\varepsilon = 10$ ($\mathcal{C}_\varepsilon = 15.25\%$). The true parameters are $p = 0.3$ and $\boldsymbol{g} = (0.35, 0.25, 0.20, 0.15, 0.05)^\top$. The blue, dashed line represents an empirical plot of LRT statistics across 1,000 replicates each with a sample size of $n = 1{,}000$. The black, solid line represents a chi-square distribution with the anticipated 8 degrees of freedom.

The analog to Corollary B.1.1 then follows with

$$\mathbf{U}_n = \begin{bmatrix} \partial\Psi_{n,1}(\hat{\boldsymbol{p}})/\partial p_1 & \partial\Psi_{n,1}(\hat{\boldsymbol{p}})/\partial p_2 \\ \partial\Psi_{n,2}(\hat{\boldsymbol{p}})/\partial p_1 & \partial\Psi_{n,2}(\hat{\boldsymbol{p}})/\partial p_2 \end{bmatrix},$$

and

$$\mathbf{V}_n = \frac{1}{n}\sum_{i=1}^{n} \begin{bmatrix} \psi_1(X_i, Y_i, \hat{p}_1)^2 & \psi_1(X_i, Y_i, \hat{p}_1)\psi_2(X_i, Y_i, \hat{p}_2) \\ \psi_2(X_i, Y_i, \hat{p}_2)\psi_1(X_i, Y_i, \hat{p}_1) & \psi_2(X_i, Y_i, \hat{p}_2)^2 \end{bmatrix}.$$

As in Corollary B.1.1, if the second Bartlett identity (Ferguson, 1996, pg. 120) is also satisfied, then $\mathbf{U} = \mathbf{V}$ with $\mathbf{U}$ symmetric (which is the case here). The extension to Theorem B.2

Table G1: **Robustness Simulation Study, PL Discrete Weibull**. A robustness analysis of estimation methods (i.e., §3.2, §3.3) for (25) with $\boldsymbol{g} = (0.5, 0.3, 0.2)^\top$ by sample size (n). Trapezoid parameters are $m = 3$, $\Delta = 0$, and $\omega = 4$. For right-censoring, we assign $\varepsilon = 6$ ($\mathcal{C}_\varepsilon = 9.92\%$). We report the empirical mean (eMean), empirical standard deviation (eSD), vectorized versions of the theoretical standard deviation (Thm B.1, Thm B.2), and coverage probability (CP), i.e., (S.15), for a 95% asymptotic confidence interval estimated using vectorized versions of Corollaries B.1.1 and B.2.1.

Left-Truncation Only								
	$p_{1,0} = 0.3$				$p_{2,0} = 0.5$			
n	eMean	eSD	Thm B.1	CP	eMean	eSD	Thm B.1	CP
100	0.300	0.052	0.051	0.943	0.505	0.098	0.096	0.935
250	0.300	0.032	0.032	0.947	0.502	0.064	0.061	0.954
1,000	0.301	0.016	0.016	0.945	0.502	0.030	0.030	0.955
Left-Truncation & Right-Censoring								
	$p_{1,0} = 0.3$				$p_{2,0} = 0.5$			
n	eMean	eSD	Thm B.2	CP	eMean	eSD	Thm B.2	CP
100	0.300	0.052	0.051	0.943	0.505	0.098	0.096	0.935
250	0.299	0.032	0.032	0.947	0.501	0.063	0.061	0.954
1,000	0.300	0.016	0.016	0.945	0.500	0.030	0.030	0.955

and Corollary B.2.1 may be found similarly. For completeness, an abbreviated simulation study for the PL discrete Weibull distribution, similar to that of Table 3, may be found in Table G1.

To estimate the point-wise asymptotic confidence intervals for the hazard rates in Figure 2, we may employ the Delta Method (e.g., Lehmann and Casella, 1998, Theorem 8.12, pg. 58). Specifically, observe that from (24)

$$S_W(u; p_1, p_2) = \sum_{k \geq u} f_W(u) = p_1^{(u - (\Delta + 1))^{p_2}}.$$

Thus, the hazard rate, $\lambda_W(u;p_1,p_2)$, for (24) is

$$\lambda_W(u;p_1,p_2) = \frac{f_W(u;p_1,p_2)}{S_W(u;p_1,p_2)} = \begin{cases} 1 - p_1^{(u-\Delta)^{p_2} - (u-(\Delta+1))^{p_2}}, & \Delta+1 \le u \le \omega - 1, \\ 1, & u = \omega. \end{cases}$$

Then, define

$$\mathbf{G}_{1\times 2} = \begin{bmatrix} \frac{\partial \lambda_W}{\partial p_1} & \frac{\partial \lambda_W}{\partial p_2} \end{bmatrix}.$$

Hence, by the Delta Method (e.g., Lehmann and Casella, 1998, Theorem 8.12, pg. 58), in combination with Corollary B.1.1 and the second Bartlett identity (Ferguson, 1996, pg. 120), we obtain

$$\sqrt{n}(\lambda_W(\hat{\boldsymbol{p}_n}) - \lambda_W(\boldsymbol{p}_0)) \xrightarrow{\mathcal{L}} \mathbf{N}_2(\mathbf{0}_2, \mathbf{G}_{1\times 2}\mathbf{V}_{2\times 2}^{-1}\mathbf{G}_{2\times 1}^{\top}).$$

In practical settings, replace $\mathbf{V}$ with $\mathbf{V}_n$ and use $\lambda_W(u;\hat{p}_1,\hat{p}_2)$ within $\mathbf{G}$.

We now illustrate the method employed to estimate the parameters of the discrete Weibull lifetime distribution for the subset of 25-month loans from AART (2019). We will utilize the `optim` function from R Core Team (2023). To do so, it is necessary to write an optimization function that has the vector $\boldsymbol{p}$ as an input with a scalar output. It is also helpful to the numeric optimization process to assign search bounds for each component of $\boldsymbol{p}$. There are many methods to write such a function, and we elect to write the somewhat crude

$$\begin{aligned} \mathcal{P}(p_1,p_2) = \Bigg\{ & \sum_{v=\Delta+1}^{\Delta+m} \left(\frac{\hat{h}_{\bullet v}}{\sum_{u=v}^{\omega} f_W(u;p_1,p_2)} \right) \left(\sum_{u=v}^{\omega} \frac{\partial}{\partial p_1} f_W(u;p_1,p_2) \right) \\ & - \sum_{v=\Delta+1}^{\Delta+m} \sum_{u=v}^{\omega} \frac{\hat{h}_{uv}}{f_W(u;p_1,p_2)} \frac{\partial}{\partial p_1} f_W(u;p_1,p_2) \Bigg\}^2 \\ + \Bigg\{ & \sum_{v=\Delta+1}^{\Delta+m} \left(\frac{\hat{h}_{\bullet v}}{\sum_{u=v}^{\omega} f_W(u;p_1,p_2)} \right) \left(\sum_{u=v}^{\omega} \frac{\partial}{\partial p_2} f_W(u;p_1,p_2) \right) \\ & - \sum_{v=\Delta+1}^{\Delta+m} \sum_{u=v}^{\omega} \frac{\hat{h}_{uv}}{f_W(u;p_1,p_2)} \frac{\partial}{\partial p_2} f_W(u;p_1,p_2) \Bigg\}^2. \end{aligned}$$

Hence, the objective is to find p_1 and p_2 such that $\mathcal{P}(p_1,p_2) = 0$ (i.e., it is minimized). One

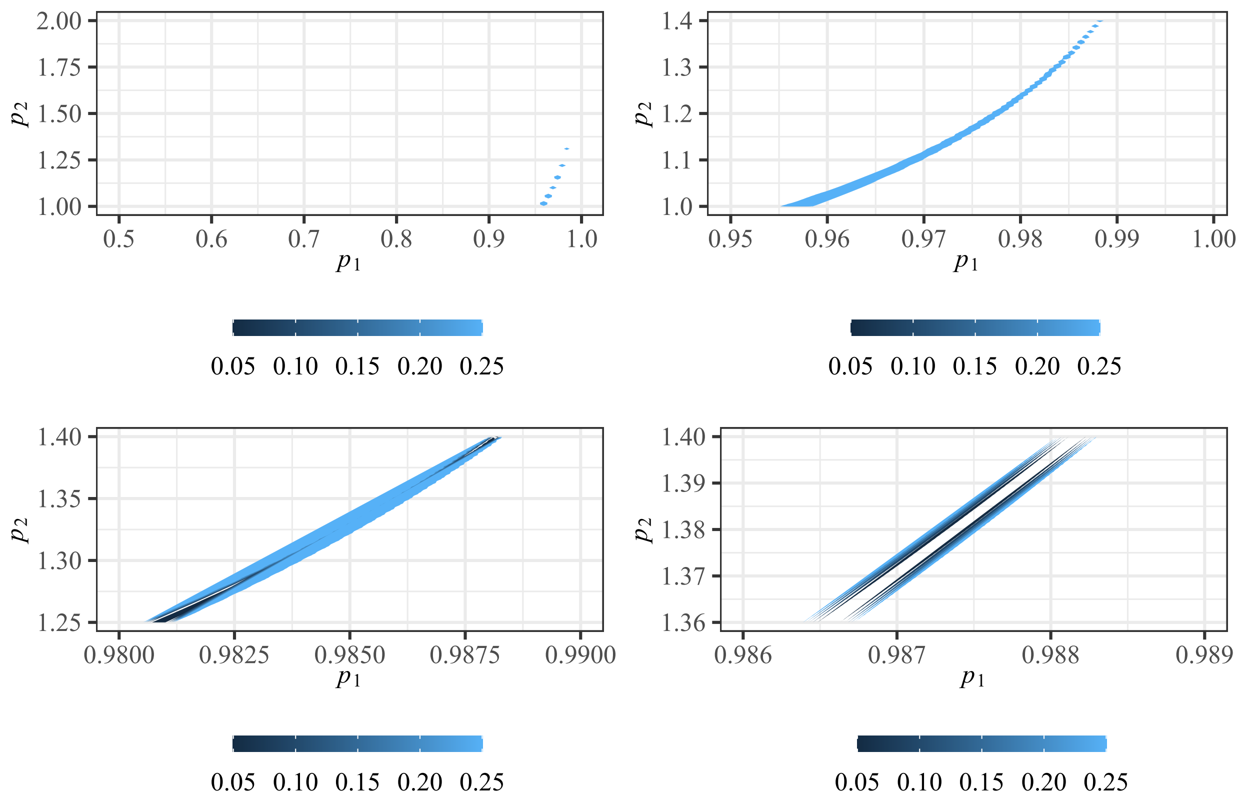

Figure G2: **PL Discrete Weibull, Optimization Search**. Repeated two-dimensional contour plots of the optimization function $\mathcal{P}_*$ for 25-month loans from AART (2019) (moving from top left, top right, bottom left, bottom right). These narrower search regions can then be fed into the `optim` function (R Core Team, 2023) to find $\hat{\boldsymbol{p}}$ for the PL discrete Weibull distribution via Theorem B.1.

option to help establish search bounds for $\boldsymbol{p}$ is to use our desire for an increasing hazard function (i.e., $p_2 > 1$) and slow rate of decay (i.e., both p_1 and p_2 close to 1). For conservatism, we can start by discretizing the intervals $0.5 < p_1 < 1$ and $1 < p_2 < 2$ to each have 100 equally spaced points. Next, we can take a cross product of these two discretized intervals and prepare a contour plot of the adjusted function $\mathcal{P}_*(p_1, p_2) = \min(0.3, \mathcal{P}(p_1, p_2))$. From this visualization, the search bounds can be narrowed again. This process can be repeated until the search bounds are suitably narrow for the `optim` function (R Core Team, 2023) to finish the task. Four such visual searches may be found in Figure G2 for the subset of 25-month loans from AART (2019). We present our proposed parameter estimates for each parametric model evaluated in Table 4 using this approach in Table G2.

Table G2: **Table 4 Parameter Estimates**.

Bond	Loan Term	PL geometric (13)	PL discrete Weibull (25)	
		$\hat{p}$	$\hat{p}_1$	$\hat{p}_2$
AART (2017)	25-months	0.03125	0.98990	1.34001
AART (2019)	25-months	0.04309	0.98774	1.38880
AART (2017)	50-months	0.02535	0.98964	1.23160

References

AART (2017), "Ally Auto Receivables Trust," Prospectus 2017-3, Ally Auto Assets LLC.

— (2019), "Ally Auto Receivables Trust," Prospectus 2019-3, Ally Auto Assets LLC.

Ferguson, T. S. (1996), *A Course in Large Sample Theory*, Chapman & Hall.

Lautier, J. P., Pozdnyakov, V., and Yan, J. (2023), "Pricing Time-to-Event Contingent Cash Flows: A Discrete-Time Survival Analysis Approach," *Insurance: Mathematics and Economics*, 110, 53–71.

Lehmann, E. and Casella, G. (1998), *Theory of Point Estimation, 2nd Edition*, Springer.

Mersmann, O. (2023), *microbenchmark: Accurate Timing Functions*, R package version 1.4.10.

Mukhopadhyay, N. (2000), *Probability and Statistical Inference*, New York, NY: Marcel Dekker.

R Core Team (2023), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria.

Ravishanker, N. and Dey, D. (2002), *A First Course in Linear Model Theory*, Chapman & Hall (CRC).

Rudin, W. (1976), *Principles of Mathematical Analysis*, McGraw-Hill, Inc.

van der Vaart, A. W. (1998), *Asymptotic Statistics*, Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press.