Discrete Time-to-Event Regression Analysis Under Left-Truncation with Applications to Consumer Finance

Jackson P. Lautier*† Vladimir Pozdnyakov‡ Jun Yan‡

June 15, 2026

Abstract

Asset-backed securities (ABS) play a vital role in financing American consumer automobile debt. Recently, economic analysis into ABS has benefited from the public release of data, which provides a new, rich source of loan level consumer auto loan information. Because this loan lifetime data is discrete-time and subject to random left-truncation, however, it is nontrivial to analyze. This has attracted recent study, but there is still no suitable approach to model this ABS loan lifetime data that incorporates regression coefficients for the lifetime of interest. We thus generalize the conditional, bivariate distribution to link to covariates, while keeping the left-truncation distribution unspecified. We solve the high-dimensional, constrained likelihood-based parameter estimation problem numerically, using a block coordinate descent design. Under suitable regularity conditions, we provide the complete large sample, multivariate normal distribution of the estimators. This allows for large sample inference into variable and model selection. We prove all results, verify with simulation, and generalize for right-censoring. We then apply our methods in an economic study of borrower prepayment behavior for 19,772 consumer auto loans from the 2017-3 Ally Auto Receivables Trust ABS bond. We find that borrowers with pick-up trucks prepay more slowly, all else equal, among other consumer finance insights.

Keywords: Asset-level disclosures, consumer behavior, length-biased sampling, household finance, Reg AB II, survival analysis

1 Introduction

The American consumer automobile debt marketplace is an economic behemoth, with total outstanding American consumer automobile debt exceeding $1,400 billion (Federal Reserve,

---

*Department of Mathematical Sciences, Bentley University, Waltham, Massachusetts, USA

†Corresponding to Jackson P. Lautier, Morison Hall, 175 Forest Street, Waltham, MA 02452, USA; email: jlautier@bentley.edu.

‡Department of Statistics, University of Connecticut, Storrs, Connecticut, USA.

2023). Despite its size, there is a dearth of economic analysis into the behavior of consumers with auto loans, especially in comparison to residential mortgages (Lautier et al., 2024). This has recently begun to change, with improvements in data availability (Securities and Exchange Commission, 2014), statistical methods (Lautier et al., 2024, 2025, 2026), and a growing interest among economic researchers (e.g., Katcher et al., 2024; Lautier et al., 2024, 2025, 2026). A major hurdle still outstanding, however, is to perform an economic study of consumer auto loan lifetimes using this newly available consumer auto debt data that considers heterogeneity among the borrowers, such as differences in credit, vehicle type, and loan contract details. In this paper, therefore, we present the first known regression-based economic analysis of the Securities and Exchange Commission (2014) loan level data.

Statistically, we desire to model consumer auto loan lifetimes using a survival model capable of handling regressors. This is a robust literature base that includes the famous Cox (1972) proportional hazards model and has long since been canonized in classical textbooks (e.g., Klein and Moeschberger, 2003). Despite this, the Securities and Exchange Commission (2014) data presents two key challenges that, together, require new methods. First, it is monthly loan data, and this requires an assumption of a discrete-time survival model (Lautier et al., 2026). This reduces the literature base substantially because ties must now be permitted, as many loans will terminate in the same month. Discrete-time, by itself, still produces excellent reference textbooks (e.g., Tutz and Schmid, 2016). Therefore, it is the combination of the second challenge that requires our forthcoming effort. Specifically, the Securities and Exchange Commission (2014) data consists of observational discrete lifetime data contained within asset-backed securities (ABS). Critically, this means that left-truncation is also present (Lautier et al., 2025, 2026), and it is this combination of discrete lifetime data subject to random left-truncation that has received limited study (Lautier et al., 2026).

When regressors are assumed present, which is the aim of this paper, the set of related literature narrows even further. Kalbfleisch and Lawless (1991) presents a discrete-time, semi-parametric regression model for right-truncated data, which can conceivably be applied to left-truncated data by reversing the timeline (see also the closely related work of Gross and Huber-Carol (1992)). We prefer to avoid consideration of a baseline hazard, however, because Lautier et al. (2024) demonstrate that the difference between estimated hazard rates by risk group for consumer auto loans are not constant with time (i.e, the proposed *credit risk convergence* is inferred from a convergence of hazard rates by risk group as loan age increases). We also find the proportional hazard assumption to be questionable within our application to loan prepayments in Section 4. This need to avoid the proportional hazards model limits the applicability of other related studies (e.g., Alioum and Commenges, 1996; Pan and Chappell, 2002; Qin and Shen, 2010; Teodorescu et al., 2010; Su and Wang, 2012; Rennert and Xie, 2018; Wu et al., 2018; Chen and Yi, 2021). An additional challenge is that the nature of auto ABS data does not allow for the typical uniformity assumption of the left-truncation random variable, known as *length-biased sampling* (e.g., Wang, 1996; Qin and Shen, 2010) nor rank regression (e.g., Lai and Ying, 1991). Given this outstanding gap in suitable methods to properly analyze the Securities and Exchange Commission (2014) economic ABS data with regressors, therefore, a new formulation is needed.

Our approach will utilize the known, finite term of consumer auto loans to generalize the conditional bivariate density function of the left-truncation and lifetime random variables studied in Lautier et al. (2025, 2026) to consider covariates. This formulation embeds left-truncation and discrete-time at the onset of the problem, whereas classical approaches typically first focus on discrete-time and right-censoring (e.g., Tutz and Schmid, 2016). This implies we are sampling jointly from a conditional bivariate distribution of left-truncated and lifetime random variables, which is better aligned to the Securities and Exchange Commission (2014) data than the classical approach of generating N pairs and keeping only those (i.e., random $n \leq N$) that fit the left-truncation requirements (e.g., Woodroofe, 1985). In other words, the latter sampling scheme would erroneously imply ABS investors also have access to loans that terminated prior to being included in the investment security.

The closest relative to what we propose is Lautier et al. (2025), whereby the lifetime of interest follows a traditional parametric lifetime and the left-truncation random variable follows an unspecified, discrete distribution. In this paper, we will allow this lifetime parameter to be linked to a set of covariates, akin to a *generalized linear model* (GLM) (e.g., Ravishanker and Dey, 2002, §11.4, pg. 422), while maintaining the desired flexibility in the left-truncation random variable. The problem of parameter estimation in this setting is nontrivial, however, because the presence of left-truncation creates a complex, constrained optimization problem over a large dimension in the likelihood. Even in the simpler setting of Lautier et al. (2025), which does not admit heterogeneity via regressors, direct numerical approaches can fail to yield a viable estimate over the dimension needed to model the auto ABS lifetime data. Utilizing this model will require overcoming these challenges, and it is our proposed method to do so that forms the basis of this paper.

Our parameter estimation method begins by proving that a stationary point of the constrained optimization problem created by the likelihood is equivalent to a stationary point of the unconstrained version of the same optimization problem. This is a generalization of a case specific result first appearing in Lautier et al. (2026) and is a valuable simplification of the problem. Specifically, with this established, we may then pursue the unconstrained numerical technique of locating extrema via the Newton-Raphson method, suitably adjusted to handle the large dimension of the parameter set and specifics of the likelihood. For example, we will utilize *coordinate* or *block coordinate* descent to split the higher dimension problem into two smaller dimension problems (Wright, 2015). This creates a novel iterative algorithm to locate zeros of the gradient of the loglikelihood, from which parameter estimates follow.

Attractively, this approach maintains desirable statistical properties under mild regularity conditions, from which large sample inference for variable and model selection follow. We present all calculation details, including closed-form formulas for the update steps of each block iterative optimization and the resulting asymptotic covariance matrix to assess estimation uncertainty. Complete proofs may be found in the Supplemental Material, and all results are also verified numerically via simulation studies in Section 3. Throughout, we leave the forms of the lifetime density and link functions unspecified, which allows for a wide generalization of our results. We then generalize further to accommodate the type of right-censoring commonly found in ABS data (e.g., Lautier et al., 2023), which greatly expands the practical utility of these results. For full methodological details, see Section 2.

We then utilize our methods for an economic study into borrower prepayment behavior for consumer auto loans contained within the AART (2017) ABS bond. This is the first known study of prepayment behavior using the Securities and Exchange Commission (2014) ABS data with regressors. It provides new economic insight into this nascent field, especially in comparison to mortgages (e.g., Deng et al., 2000; Calhoun and Deng, 2002; Ambrose and Sanders, 2003; Jones and Sirmans, 2019). We find expected drivers of prepayment behavior, such as borrower credit score, loan terms (e.g., interest rate, co-obligor) and subvention (e.g., cash or interest rate rebates). We also find the interesting result that drivers of pick-up trucks tend to prepay more slowly, all else equal. These results provide updated comparison points to early, related economic literature (e.g., Ohta and Griliches, 1986; Ohta, 1987; Thompson and Noordewier, 1992). For much more discussion, please see Section 4.

The Supplemental Material provides complete proofs and many additional details. For complete data and replication code, please see the public `git` repository found here: https://github.com/jackson-lautier/left-truncation-regression.

2 Methods

We first consider the likelihood-based estimation problem. Next, we provide results for large sample inference. Finally, we generalize all results to handle right-censoring.

2.1 Likelihood Estimation

This is a three-part subsection that first formally states the optimization problem, then proposes a numerical solution, and closes by providing the details of its implementation.

2.1.1 Optimization Problem

Let the lifetime of interest, denoted X, be a discrete random variable with support, $\mathcal{U} = \{\Delta+1, \Delta+2, \ldots, \omega\}$, where $\Delta \in \mathbb{N}$ and $\omega \in \mathbb{N}^+$ are nonrandom (i.e., known at the onset of the problem). Denote the discrete, left-truncation random variable as Y, with support $\mathcal{V} = \{\Delta+1, \Delta+2, \ldots, \Delta+m\}$, where $m \in \mathbb{N}^+$ is nonrandom. We observe X if and only if $X \geq Y$, and so the support of the joint, conditional distribution of $(X, Y \mid X \geq Y)$ is a trapezoid, $\mathcal{A} = \{(u, v) \in \mathcal{U} \times \mathcal{V} : u \geq v\}$ (Lautier et al., 2026). We further assume that, conditioning on the observed covariates to be introduced next, X and Y are independent (a reasonable assumption for ABS data (Lautier et al., 2023, §4.3)).

It is desirable to allow X to depend on a set of observed explanatory variables. Denote the observed sample as $\mathcal{S}_n = \{(X_i, Y_i \mid \boldsymbol{z}_i)\}_{1 \leq i \leq n}$, where $\boldsymbol{z}_i = (1, Z_{1i}, Z_{2i}, \ldots, Z_{ki})^\top$ is a vector of observed explanatory variables for each left-truncated sample pair, (X_i, Y_i). Denote the probability mass function (pmf) of X as $f(u; p_i)$ for $u \in \mathcal{U}$, where the parameter $p_i \in \mathcal{P}$, where $\mathcal{P}$ is a convex set (e.g., $\mathbb{R}^+$ or $\mathbb{R}$), depends on $\boldsymbol{z}_i$ in the following manner. Define a *link function*, μ_i, that provides a functional relationship between p_i and a linear function of the explanatory variables, $\eta_i(\boldsymbol{z}_i, \boldsymbol{\beta}) \equiv \eta_i = \boldsymbol{z}_i^\top \boldsymbol{\beta}$, where $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \ldots, \beta_k)^\top \in \mathbb{R}^{k+1}$ is a set of unknown coefficient parameters. This is the familiar *generalized linear model* (GLM) form (e.g., Ravishanker and Dey, 2002, §11.4, pg. 422),

$$f(X_i; p_i = \mu_i(\eta_i(\boldsymbol{z}_i, \boldsymbol{\beta}))) \equiv f(X_i; \boldsymbol{z}_i, \boldsymbol{\beta}) \equiv f_i \circ \mu_i \circ \eta_i. \tag{1}$$

Remark. *The different notation in* (1) *oscillates between over-explained and overly casual. We make this stylistic choice for the purposes of narrative clarity, as each form will ease exposition at various points in the coming discourse. Furthermore, it will at times become necessary to clarify the input of f (i.e., either the observed X_i or the more general u), and we will do so by writing f_i for $f(X_i; \cdot)$ and f_u for $f(u; \cdot)$.*

The pmf of Y is denoted by g, and it is a parametric distribution such that $\Pr(Y = v) =$

$g(v) = g_v$, where each probability point mass, $g_v \in \boldsymbol{g} = (g_{\Delta+1}, \ldots, g_{\Delta+m})^\top$, is a parameter to be estimated for $v \in \mathcal{V}$. This, along with the assumption $X \perp Y$, allows us to write the conditional bivariate pmf that generates the sample, $\mathcal{S}_n$, denoted h_*, as

$$h_*(u, v; \boldsymbol{z}_i, \boldsymbol{\beta}, \boldsymbol{g}) = \frac{f(u; \boldsymbol{z}_i, \boldsymbol{\beta}) g_v}{\alpha(\boldsymbol{z}_i, \boldsymbol{\beta}, \boldsymbol{g})}, \quad (u, v) \in \mathcal{A}, \tag{2}$$

where $\alpha(\boldsymbol{z}_i, \boldsymbol{\beta}, \boldsymbol{g}) \equiv \alpha_i \equiv \Pr(X \geq Y; \boldsymbol{z}_i, \boldsymbol{\beta}, \boldsymbol{g})$,

$$\alpha_i = \sum_{u=\Delta+1}^{\omega} f(u; \boldsymbol{z}_i, \boldsymbol{\beta}) \Bigg(\sum_{v=\Delta+1}^{\min(u,\Delta+m)} g_v \Bigg) = \sum_{v=\Delta+1}^{\Delta+m} g_v \Bigg(\sum_{u=v}^{\omega} f(u; \boldsymbol{z}_i, \boldsymbol{\beta}) \Bigg).$$

Thus, if we define the parameter space, $\boldsymbol{\Theta} = (\boldsymbol{g}, \boldsymbol{\beta})^\top \in \mathbb{R}^{k+m+1}$, then, along with (2), the likelihood equation is

$$\mathcal{L}(\boldsymbol{\Theta} \mid \mathcal{S}_n) = \prod_{i=1}^{n} h_*(X_i, Y_i; \boldsymbol{z}_i, \boldsymbol{\beta}, \boldsymbol{g}). \tag{3}$$

It is our objective to find the maximum likelihood estimate (MLE) of $\boldsymbol{\Theta}$ within (3). Because we require $\sum_v g_v = 1$, however, maximizing (3) over $\boldsymbol{\Theta}$ subject to this linear constraint becomes a complex, multidimensional constrained optimization problem,

$$\hat{\boldsymbol{\Theta}}_{\text{MLE}} = \arg\max_{\boldsymbol{\Theta}} \Bigg\{ \mathcal{L}(\boldsymbol{\Theta} \mid \mathcal{S}_n) : \sum_{v=\Delta+1}^{\Delta+m} g_v = 1 \Bigg\}, \tag{4}$$

where uniqueness and existence are assumed to avoid technicalities. To find $\hat{\boldsymbol{\Theta}}_{\text{MLE}}$, it is helpful to show that the unconstrained and linearly constrained stationary points must coincide.

Lemma 1. *Denote* $\boldsymbol{u} = (u_1, \ldots, u_k)^\top$ *and* $\boldsymbol{v} = (v_1, \ldots, v_m)^\top$ *and let* $\varphi(\boldsymbol{u}, \boldsymbol{v}) : \mathbb{R}^{k+m} \mapsto \mathbb{R}$ *be a smooth function with domain (i)* $u_i \geq 0$ *or* $u_i \leq 0$ *for all* $1 \leq i \leq k$, *(ii)* $v_j \geq 0$ *or* $v_j \leq 0$ *for all* $1 \leq j \leq m$, *(iii)* $\sum_i u_i \neq 0$, *and (iv)* $\sum_j v_j \neq 0$. *Further assume that for any nonzero scalars,* π_1, π_2, $\varphi(\pi_1 \boldsymbol{u}, \pi_2 \boldsymbol{v}) = \varphi(\boldsymbol{u}, \boldsymbol{v})$ *and define the following Lagrangian-type function,* Λ, *with constraints,* $\sum_i u_i = \sum_j v_j = 1$ *(e.g., Ravishanker and Dey, 2002, §2.9, pg. 69),*

$$\Lambda(\boldsymbol{u}, \boldsymbol{v}, \lambda_1, \lambda_2) = \varphi(\boldsymbol{u}, \boldsymbol{v}) - \lambda_1\left(1 - \sum_{i=1}^{k} u_i\right) - \lambda_2\left(1 - \sum_{j=1}^{m} v_j\right),$$

with $\lambda_1 \in \mathbb{R}$ *and* $\lambda_2 \in \mathbb{R}$*. If* $\boldsymbol{u}^*$*,* $\boldsymbol{v}^*$*,* λ_1^**, and* λ_2^* *is a stationary point of* Λ*, then* $\lambda_1^* = \lambda_2^* = 0$*.*

Proof. See the Supplemental Material, Section A.1. ☐

Lemma 1 is a generalization of an observation that first appeared in Lautier et al. (2026). Its practical value stems from simplifying the search implied by (4) because h_* defined in (2) is a smooth function that meets the domain and scaling property requirements of Lemma 1. (In the interest of precision, Lemma 1 is a more general statement than what is needed here because it assumes two constraints versus the needed one. For an expanded discussion, see the Supplemental Material, Section B.) That is, again assuming uniqueness and existence,

$$\hat{\boldsymbol{\Theta}}_{\text{MLE}} = \arg\max_{\boldsymbol{\Theta}}\left\{\mathcal{L}(\boldsymbol{\Theta} \mid \mathcal{S}_n) : \sum_{v=\Delta+1}^{\Delta+m} g_v = 1\right\} = \arg\max_{\boldsymbol{\Theta}}\mathcal{L}(\boldsymbol{\Theta} \mid \mathcal{S}_n). \tag{5}$$

This equivalence holds because the likelihood is homogeneous of degree zero in the truncation probabilities $\boldsymbol{g}$ when $\boldsymbol{\beta}$ is held fixed. This is a convenient simplification because it smooths the searching region for $\boldsymbol{\Theta}$, reduces computation demands, and helps numerical methods more reliably recover global extrema (Bertsekas, 1999; Wright, 2015). We thus estimate $\boldsymbol{\Theta}$ using a novel, iterating unconstrained numerical optimization method on the log-likelihood,

$$\ell(\boldsymbol{\Theta} \mid \mathcal{S}_n) \equiv \ell \equiv \log \mathcal{L}(\boldsymbol{\Theta} \mid \mathcal{S}_n) = \sum_{i=1}^{n} \log h_*(X_i, Y_i, \boldsymbol{z}_i, \boldsymbol{\beta}, \boldsymbol{g}). \tag{6}$$

2.1.2 Numerical Solution

The (now unconstrained) objective function in (6) has enough similarity to a GLM that it motivates pursuing typical numeric techniques to estimate the coefficient parameters, such as the *Newton-Raphson* (NR) method (e.g., Hardin and Hilbe, 2007). There is an important added complexity created by the left-truncation random variable, however, in the presence of the $\boldsymbol{g}$ parameters. In addition to adding a large dimension to $\boldsymbol{\Theta}$, the $\boldsymbol{g}$ parameters are also embedded within the $\alpha(\cdot)$ term that is in the denominator of h_* and differs for each observation, $1 \leq i \leq n$. That is, if we consider again the likelihood equation in (3),

$$\mathcal{L}(\boldsymbol{\Theta} \mid \mathcal{S}_n) = \prod_{i=1}^{n} \frac{f(X_i; \boldsymbol{z}_i, \boldsymbol{\beta}) g(Y_i)}{\alpha(\boldsymbol{z}_i, \boldsymbol{\beta}, \boldsymbol{g})} = \prod_{v=\Delta+1}^{\Delta+m} g_v^{\hat{\gamma}_n(v)} \prod_{i=1}^{n} \frac{f(X_i; \boldsymbol{z}_i, \boldsymbol{\beta})}{\alpha(\boldsymbol{z}_i, \boldsymbol{\beta}, \boldsymbol{g})},$$

where $\hat{\gamma}_n(v) = \sum_i \mathbf{1}(Y_i = v)$, then we obtain the detailed objective function,

$$\ell(\boldsymbol{\Theta} \mid \mathcal{S}_n) = \sum_{v=\Delta+1}^{\Delta+m} \hat{\gamma}_n(v) \log g_v + \sum_{i=1}^{n} \log f(X_i; \boldsymbol{z}_i, \boldsymbol{\beta}) - \sum_{i=1}^{n} \log \alpha(\boldsymbol{z}_i, \boldsymbol{\beta}, \boldsymbol{g}). \tag{7}$$

An approach to estimate $\boldsymbol{\Theta}$ from (7) is not immediate. Closed-form solutions are difficult to obtain because $\log \alpha(\cdot)$ is a logarithm of a summation that consists of all parameters. Numerically, we must take care to handle the difficulties created by the $\boldsymbol{g}$ parameters. For example, even in the setting where f depends only on a single parameter, p, for all $1 \leq i \leq n$, the potentially large dimension of $\boldsymbol{\Theta}$ stemming from $\boldsymbol{g}$ can cause direct numeric optimization approaches on (7) to fail (Lautier et al., 2025). Beyond this feasibility question, there may also be settings where the left-truncation distribution is known or it is not a priority of the economic investigation (i.e., a nuisance parameter). Additionally, we will demonstrate in Section 2.1.3 that taking a partial derivative of (7) with respect to $\boldsymbol{g}$ does not provide a viable candidate function for the update step of the NR method.

Motivated by these challenges, we propose an iterative optimization procedure that is a form of *coordinate descent* or *block* coordinate descent (Wright, 2015). We will treat the $\boldsymbol{g}$ portion of $\boldsymbol{\Theta}$ as one block and the $\boldsymbol{\beta}$ portion as the second block. Specifically, we will hold $\boldsymbol{g}$ fixed and maximize ℓ with respect to $\boldsymbol{\beta}$. Once these estimates of $\boldsymbol{\beta}$ are found, they are then held fixed and ℓ is maximized with respect to $\boldsymbol{g}$. This process repeats until the estimates of each block that maximize ℓ stabilize. Under mild regularity conditions, this blockwise optimization approach can be expected to reliably return extrema of ℓ (e.g., Bertsekas, 1999; Tseng, 2001) (a further practical benefit of performing an unconstrained search). A natural analogy in statistical applications is the EM algorithm, which has a long history of close study (e.g., Wu, 1983).

A final outstanding challenge shared by many numerical optimization procedures is to select initial values. We overcome this issue by using the estimation methods of Lautier et al. (2025). That is, for a given $f(\cdot;p)$, let $\hat{p}_{\texttt{INIT}}$ be any $p \in \mathcal{P}$ such that

$$\sum_{v=\Delta+1}^{\Delta+m}\left(\frac{\hat{h}_{\bullet v}}{\sum_{u=v}^{\omega} f(u;p)}\right)\left(\sum_{u=v}^{\omega}\frac{\partial}{\partial p}f(u;p)\right)=\sum_{v=\Delta+1}^{\Delta+m}\sum_{u=v}^{\omega}\frac{\hat{h}_{uv}}{f(u;p)}\frac{\partial}{\partial p}f(u;p), \tag{8}$$

where $\hat{h}_{\bullet v} \coloneqq \sum_u \hat{h}_{uv} > 0$ and $\hat{h}_{uv} = \sum_i \mathbf{1}_{(X_i,Y_i)=(u,v)}/n$. Furthermore, define

$$\hat{g}_{v,\texttt{INIT}} = \frac{\hat{h}_{\bullet v}}{S(v;\hat{p}_{\texttt{INIT}})}\left[\sum_{k=\Delta+1}^{\Delta+m}\frac{\hat{h}_{\bullet k}}{S(k;\hat{p}_{\texttt{INIT}})}\right]^{-1}, \quad v \in \mathcal{V}, \tag{9}$$

where $S(\cdot)$ denotes the survival function,

$$S(x;p) \coloneqq \Pr(X \geq x; p) = \sum_{u=x}^{\omega} f(u;p). \tag{10}$$

These estimators follow from Lautier et al. (2025, Theorem 3.1) and may be generalized for vector-valued parametric forms of f, such as the two-parameter discrete Weibull of Nakagawa and Osaki (1975). See Lautier et al. (2025) for details, including approaches and replication code to solve (8). We are now ready to state the formal estimation procedure.

Algorithm 1. *The following approach may be used to find* $\arg\max_{\mathbf{\Theta}} \ell(\mathbf{\Theta} \mid \mathcal{S}_n)$ *in* (7).

1. *For a given* $f(\cdot;p)$*, use* (8) *to find* $\hat{p}_{INIT}$ *and* (9) *to find* $\hat{\boldsymbol{g}}_{INIT}$.

2. *For a given link function,* μ*, assign* $\hat{\beta}_{0,INIT} = \mu^{-1}(\hat{p}_{INIT})$ *and* $\hat{\boldsymbol{\beta}}_{INIT} = (\hat{\beta}_{0,INIT}, \mathbf{0}_k)^\top$.

3. *Fix* $\hat{\boldsymbol{g}}_{INIT}$ *and use the NR method per* (11) *to maximize* $\ell(\boldsymbol{\beta} \mid \hat{\boldsymbol{g}}_{INIT}, \mathcal{S}_n)$ *starting from* $\hat{\boldsymbol{\beta}}_{INIT}$ *(for details, see §2.1.3). Call this block extrema* $\hat{\boldsymbol{\beta}}^t$.

4. *Fix* $\hat{\boldsymbol{\beta}}^t$ *and use the NR method per* (13) *to maximize* $\ell(\boldsymbol{g} \mid \hat{\boldsymbol{\beta}}^t, \mathcal{S}_n)$ *starting from* $\hat{\boldsymbol{g}}_{INIT}$ *(for details, see §2.1.3). Call this block extrema* $\hat{\boldsymbol{g}}^t$*.*

5. *Return to Step 3 with* $(\hat{\boldsymbol{\beta}}^t, \hat{\boldsymbol{g}}^t)^\top$ *replacing* $(\hat{\boldsymbol{\beta}}_{INIT}, \hat{\boldsymbol{g}}_{INIT})^\top$*. Repeat Steps 3 and 4 to find* $\hat{\boldsymbol{\Theta}}^{t+1} = (\hat{\boldsymbol{\beta}}^{t+1}, \hat{\boldsymbol{g}}^{t+1})^\top$ *and evaluate* $\left\| \nabla \ell(\hat{\boldsymbol{\Theta}}^{t+1}) \right\| < \epsilon$*, where* ϵ *is a small, predetermined tolerance, such as* $\epsilon = 10^{-7}$*.*

6. *If Step 5 is true, exit the procedure and set* $\hat{\boldsymbol{\Theta}}_{MLE} = (\hat{\boldsymbol{\beta}}^{t+1}, \hat{\boldsymbol{g}}^{t+1})^\top$*. If Step 5 is false, repeat Step 5 starting with* $(\hat{\boldsymbol{\beta}}^{t+1}, \hat{\boldsymbol{g}}^{t+1})^\top$*.*

Under suitable regularity conditions, the NR iteration method is known to be asymptotically efficient (Shao, 2003, Theorem 4.19, pg. 295). In the following sections, we demonstrate that the above approach is effective at finding $\hat{\boldsymbol{\Theta}}_{\text{MLE}}$ and, subsequently, generates reliable large sample inference on the main parameters of interest, $\boldsymbol{\beta}$, as well as the $\boldsymbol{g}$ parameters.

2.1.3 Implementation Details

We first present the update equation needed to use the NR method for the parameter $\boldsymbol{\beta}$ in Step 3 of Algorithm 1. Let $\boldsymbol{\beta}^{(t)}$ be the current estimate of $\boldsymbol{\beta}$ for (7). For the NR method, the update step to find $\boldsymbol{\beta}^{(t+1)}$ is

$$\boldsymbol{\beta}^{(t+1)} = \boldsymbol{\beta}^{(t)} - (\mathbf{Z}^\top \mathbf{A}^{(t)} \mathbf{Z})^{-1} \mathbf{Z}^\top \mathbf{W}^{(t)} \mathbf{J}_n, \tag{11}$$

where $\mathbf{Z}^\top = (\boldsymbol{z}_1, \boldsymbol{z}_2, \ldots, \boldsymbol{z}_n)$ is the $(k+1) \times n$ design matrix, $\mathbf{A} = \text{diag}(A_i)$ is an $n \times n$ diagonal matrix with diagonal elements,

$$A_i = \frac{s(X_i, \eta_i)}{f(X_i; \boldsymbol{z}_i, \boldsymbol{\beta})} - \left(\frac{q(X_i, \eta_i)}{f(X_i; \boldsymbol{z}_i, \boldsymbol{\beta})} \right)^2 - \frac{1}{\alpha_i} \sum_{v=\Delta+1}^{\Delta+m} g_v \left(\sum_{u=v}^{\omega} s(u, \eta_i) \right) + \left(\frac{r(\boldsymbol{g}, \eta_i)}{\alpha_i} \right)^2,$$

where $\alpha_i \equiv \alpha(\boldsymbol{z_i}, \boldsymbol{\beta}, \boldsymbol{g})$,

$$s(X_i, \eta_i) = \left(\frac{\partial^2 f_i}{\partial \mu_i^2} \right) \left(\frac{\partial \mu_i}{\partial \eta_i} \right)^2 + \left(\frac{\partial f_i}{\partial \mu_i} \right) \left(\frac{\partial^2 \mu_i}{\partial \eta_i^2} \right),$$

$$q(u,\eta_i) = \left(\frac{\partial f_u}{\partial \mu_i}\right)\left(\frac{\partial \mu_i}{\partial \eta_i}\right),$$

and

$$r(\boldsymbol{g},\eta_i) = \sum_{v=\Delta+1}^{\Delta+m} g_v \left[\sum_{u=v}^{\omega} q(u,\eta_i)\right],$$

for $1 \le i \le n$, $\mathbf{W} = \text{diag}(W_i)$ is an $n \times n$ diagonal matrix with diagonal elements,

$$W_i = \frac{q(X_i,\eta_i)}{f(X_i;\boldsymbol{z}_i,\boldsymbol{\beta})} - \frac{1}{\alpha_i}\left(\sum_{v=\Delta+1}^{\Delta+m} g_v \left[\sum_{u=v}^{\omega} q(u,\eta_i)\right]\right),$$

for $1 \le i \le n$, and $\mathbf{J}_n$ is an $n \times 1$ unit vector. For details in the derivation of (11), see the Supplemental Material, Section A.2.

The details behind Step 4 of Algorithm 1 are not as direct, however. Observe, by (6),

$$\begin{aligned}\frac{\partial}{\partial g_v}\ell(\boldsymbol{\Theta} \mid \mathcal{S}_n) &= \frac{\hat{\gamma}_n(v)}{g_v} - \sum_{i=1}^{n}\frac{1}{\alpha_i}\left(\sum_{u=v}^{\omega} f(u;\boldsymbol{z}_i,\boldsymbol{\beta})\right) \\ &= \sum_{i=1}^{n}\left\{\frac{\mathbf{1}(Y_i = v)}{g_v} - \frac{1}{\alpha_i}\left(\sum_{u=v}^{\omega} f(u;\boldsymbol{z}_i,\boldsymbol{\beta})\right)\right\}. \qquad (12)\end{aligned}$$

This is problematic because the above optimization function takes the approximate form,

$$\zeta(g_v) \approx \frac{k_1}{g_v} - \frac{k_2}{k_3 g_v + k_4} = 0,$$

where $k_i \ge 0$, $1 \le i \le 4$, are some non-negative constants. The asymptotic behavior of $\zeta(g_v)$ is not a good candidate for the NR method, especially as we also require $0 < g_v < 1$ for all $v \in \mathcal{V}$. In other words, it can be shown that, if left unmodified, the NR method will fail to converge for (12) (for details, see the Supplemental Material, Section C).

We therefore proceed as follows. Because it is not difficult to show

$$\frac{\partial \ell}{\partial g_v} = \frac{\hat{\gamma}_n(v)}{g_v} - \sum_{i=1}^{n}\frac{1}{\alpha_i}\left(\sum_{u=v}^{\omega} f(u;\boldsymbol{z}_i,\boldsymbol{\beta})\right) = 0 \iff \hat{\gamma}_n(v) - g_v\sum_{i=1}^{n}\frac{1}{\alpha_i}\left(\sum_{u=v}^{\omega} f(u;\boldsymbol{z}_i,\boldsymbol{\beta})\right) = 0,$$

it is equivalent to consider the modified optimization function

$$\mathcal{F}(g_v) = \hat{\gamma}_n(v) - g_v \sum_{i=1}^{n} \frac{1}{\alpha_i}\left(\sum_{u=v}^{\omega} f(u;\boldsymbol{z}_i,\boldsymbol{\beta})\right) = \sum_{i=1}^{n}\left(\mathbf{1}(Y_i = v) - \frac{g_v S_{vi}}{\alpha_i}\right),$$

where S_{vi} is a generalized form of (10), $S_{vi} \equiv S(v;\boldsymbol{z}_i,\boldsymbol{\beta}) = \sum_{u=v}^{\omega} f(u;\boldsymbol{z}_i,\boldsymbol{\beta})$. In other words, $\mathcal{F}(\boldsymbol{g}) = \mathbf{0}$ if and only if $\partial\ell/\partial\boldsymbol{g} = \mathbf{0}$, and so we may apply the NR method on $\mathcal{F}$ to find the zeros of $\partial\ell/\partial\boldsymbol{g}$. As a final comment before the formal statement of the sibling result to (11), recall that $\sum_v g_v = 1$ implies there are only $m-1$ remaining free parameters to be estimated (in addition to $\boldsymbol{\beta}$). Hence, for consistency in the remainder of the manuscript and without loss of generality, we assume $g_{\Delta+m} = 1 - \sum_{v=\Delta+1}^{\Delta+m-1} g_v$. Note that (5) ensures this holds, and we can still perform Step 4 of Algorithm 1 as an unconstrained search.

Thus, Step 4 of Algorithm 1 may proceed as follows. Let $\boldsymbol{g}^{(t)}$ be the current estimate of $\boldsymbol{g}$ for (7). For the NR method, the update step to find $(\boldsymbol{g}^*)^{(t+1)}$ is

$$(\boldsymbol{g}^*)^{(t+1)} = (\boldsymbol{g}^*)^{(t)} - ((\mathbf{B}^*)^{(t)})^{-1}((\mathbf{D}^*)^{(t)})^\top \mathbf{J}_n, \tag{13}$$

where $\boldsymbol{g}^* = (g_{\Delta+1},\ldots,g_{\Delta+m-1})^\top$, $\mathbf{B}^*$ is an $(m-1)\times(m-1)$ matrix,

$$\mathbf{B}^* = \sum_{i=1}^{n}\begin{bmatrix} B_i^*(\Delta+1,\Delta+1) & B_i^*(\Delta+1,\Delta+2) & \ldots & B_i^*(\Delta+1,\Delta+m-1) \\ B_i^*(\Delta+2,\Delta+1) & B_i^*(\Delta+2,\Delta+2) & \ldots & \\ \vdots & \vdots & \ddots & \\ B_i^*(\Delta+m-1,\Delta+1) & \ldots & & B_i^*(\Delta+m-1,\Delta+m-1) \end{bmatrix},$$

where, for $\Delta+1 \le v, v^* \le \Delta+m-1$,

$$B_i^*(v,v^*) = \frac{1}{(\alpha_i^*)^2}\begin{cases} g_v S_{vi}(S_{vi} - S_{(\Delta+m)i}) - S_{vi}\alpha_i^*, & v = v^* \\ g_v S_{vi}(S_{v^*i} - S_{(\Delta+m)i}), & v \neq v^* \end{cases},$$

where

$$\alpha_i^* = \sum_{v=\Delta+1}^{\Delta+m-1} g_v S_{vi} + \left(1 - \sum_{j=\Delta+1}^{\Delta+m-1} g_v\right) S_{(\Delta+m)i},$$

and $\mathbf{D}^*$ is an $n \times (m-1)$ matrix,

$$\mathbf{D}^* = \begin{bmatrix} d_1^*(\Delta+1) & d_1^*(\Delta+2) & \dots & d_1^*(\Delta+m-1) \\ d_2^*(\Delta+1) & d_2^*(\Delta+2) & \dots & d_2^*(\Delta+m-1) \\ \vdots & \vdots & & \vdots \\ d_n^*(\Delta+1) & d_n^*(\Delta+2) & \dots & d_n^*(\Delta+m-1) \end{bmatrix},$$

where, for $\Delta + 1 \leq v \leq \Delta + m - 1$, $d_i^*(v) = \mathbf{1}(Y_i = v) - g_v S_{vi}/\alpha_i^*$. To find $\boldsymbol{g}^{(t+1)}$, upon convergence of (13), set $g_{\Delta+m}^{(t+1)} = 1 - \sum_{v=\Delta+1}^{\Delta+m-1} g_v^{(t+1)}$ (again, by (5) and not a search constraint). For details on the derivation of (13), see the Supplemental Material, Section A.3.

2.2 Large Sample Inference

This is a two-part subsection that first considers the problem of variable selection by stating the asymptotic behavior of the parameter estimates. Next, we propose a method for model selection that compares the fit of a model with regressors against a null, intercept only model.

2.2.1 Variable Selection

The methods of Section 2.1 locate zeros of the gradient of the loglikelihood equation, (7). This is the standard approach of maximum likelihood estimation, which is a special case under the broad class of asymptotically consistent M-estimators (e.g., van der Vaart, 1998, §5.3, pg. 51). Therefore, under suitable regularity conditions (e.g., van der Vaart, 1998, §5.3, pg. 51; Mukhopadhyay, 2000, §12.2, pg. 539), $\hat{\boldsymbol{\Theta}}_{\text{MLE}}$ is known to be consistent to $\boldsymbol{\Theta}$, with a distribution that is asymptotically multivariate normal. In the present section, we fully specify this asymptotic distribution and provide expressions that may be used to estimate its covariance matrix in practical data settings. Critically, this allows for inference on the economic variables included in the design matrix, $\mathbf{Z}$.

It is not enough to complete the calculation of the off-diagonals of the Hessian matrix of the log-likelihood, (7), however. (For completeness, the negative of the Hessian of the log-likelihood evaluated at $\hat{\boldsymbol{\Theta}}_{\text{MLE}}$ is the *observed Fisher information* (Pawitan, 2013), which may be used to estimate the asymptotic variance of $\hat{\boldsymbol{\Theta}}_{\text{MLE}}$ (e.g., Shao, 2003, pg. 372).) The complication is again the left-truncation parameters, $\boldsymbol{g}$, in that the linear relationship, $\sum_v g_v = 1$, creates a linear dependence within $\boldsymbol{\Theta}$. Hence, a direct attempt to invert the Hessian of the log-likelihood will fail, as it will not have full rank. Large sample inference, therefore, consists of the vector of free parameters, $\boldsymbol{\Theta}^* = (g_{\Delta+1}, \ldots, g_{\Delta+m-1}, \beta_0, \beta_1, \ldots, \beta_k)^\top$. This updates (7), denoted $\ell(\boldsymbol{\Theta}^* \mid \mathcal{S}_n) \equiv \ell^*$, to

$$\begin{aligned}\ell^* = & \sum_{v=\Delta+1}^{\Delta+m-1} \hat{\gamma}_n(v) \log g_v + \hat{\gamma}_n(\Delta+m) \log \left(1 - \sum_{k=\Delta+1}^{\Delta+m-1} g_k\right) \\ & + \sum_{i=1}^{n} \log f(X_i; \boldsymbol{z}_i, \boldsymbol{\beta}) - \sum_{i=1}^{n} \log \alpha^*(\boldsymbol{z}_i, \boldsymbol{\beta}, \boldsymbol{g}).\end{aligned}$$

Theorem 2.1. *Let* $\hat{\boldsymbol{\Theta}}^*_{\text{MLE}} = \hat{\boldsymbol{\Theta}}_{\text{MLE}} \setminus \hat{g}_{\Delta+m}$ *and assume the suitable regularity conditions (e.g., van der Vaart, 1998, §5.3, pg. 51; Mukhopadhyay, 2000, §12.2, pg. 539) hold. Then,*

$$\boldsymbol{\Sigma}_n^{1/2}(\hat{\boldsymbol{\Theta}}^*_{\text{MLE}})\sqrt{n}\big(\hat{\boldsymbol{\Theta}}^*_{\text{MLE}} - \boldsymbol{\Theta}^*\big) \longrightarrow_d \mathcal{N}(\mathbf{0}, \mathbf{I}_{m+k}), \tag{14}$$

where the closed-form components of $\boldsymbol{\Sigma}_n(\hat{\boldsymbol{\Theta}}^*_{\text{MLE}})$ *are defined in Appendix A.*

Proof. See the Supplemental Material, Section A.4. ☐

The practical value of Theorem 2.1 is that we obtain

$$\widehat{\text{Var}}(\hat{\boldsymbol{\Theta}}^*_{\text{MLE}}) = \big(\boldsymbol{\Sigma}_n(\hat{\boldsymbol{\Theta}}^*_{\text{MLE}})\big)^{-1}.$$

This yields immediate hypothesis tests and confidence intervals for the economic variable coefficients, $\boldsymbol{\beta}$. These are presented without proof for completeness.

Corollary 2.1.1. *Suppose we wish to test* $H_0 : \beta_j = 0$*, for* $0 \leq j \leq k$*. Then, under* H_0*, the Wald test statistic (e.g., Kutner et al., 2005, pg. 578),*

$$\frac{\hat{\beta}_j}{\hat{\sigma}_{m+j}} \longrightarrow_d N(0, 1),$$

where $\hat{\sigma}_{m+j}$ *denotes the square root of the* $(m + j)$*th diagonal element of* $\widehat{\mathrm{Var}}(\hat{\boldsymbol{\Theta}}^*_{\mathrm{MLE}})$*. Similarly, a* $100(1 - \theta)\%$ *confidence interval for* β_j *is given by*

$$\hat{\beta}_j \pm \Phi^{-1}(1 - \theta/2)\hat{\sigma}_{m+j},$$

where $\Phi^{-1}(1 - \theta/2)$ *is the* $(1 - \theta/2)$*th percentile of the standard normal distribution.*

2.2.2 Model Selection

In addition to variable selection, it is often desirable to formally assess model selection. For (2), we may perform a test of overall model significance against Lautier et al. (2025), which assumes f takes the form $f(\cdot; p)$. In this setting, we may treat $f(\cdot; p)$ as a *null model*, which does not include the economic variables, $\mathbf{Z}$. In other words, $f(\cdot; p)$ is a model with an intercept only in (2), and it is thus a subset model or restricted parameter space (i.e., the restrictions are $\beta_j = 0$ for $1 \leq j \leq k$). The immediate analog for the implied null and alternative hypothesis in this discussion within classical linear models is the overall F test for model significance (e.g., Kutner et al., 2005, pg. 226). In our setting, we will test this same null and alternative hypothesis, detailed in Corollary 2.2.1, by utilizing the statistical machinery of the classical likelihood ratio test (LRT).

Theorem 2.2. *Define the following null likelihood,*

$$\mathcal{L}_0(\boldsymbol{\Theta}_0 \mid \mathcal{S}_n) = \prod_{v=\Delta+1}^{\Delta+m} \prod_{u=v}^{\omega} \left(\frac{f(u; p) g_v}{\alpha} \right)^{\sum_{i=1}^{n} \mathbf{1}((X_i, Y_i) = (u, v))},$$

where $\boldsymbol{\Theta}_0 = (p, \boldsymbol{g})^\top$ *and*

$$\alpha \equiv \Pr(Y \leq X) = \sum_{u=\Delta+1}^{\omega} f(u;p)\left(\sum_{v=\Delta+1}^{\min(u,\Delta+m)} g_v\right) = \sum_{v=\Delta+1}^{\Delta+m} g_v\left(\sum_{u=v}^{\omega} f(u;p)\right).$$

If $\hat{\boldsymbol{\Theta}}_0 = (\hat{p}, \hat{\boldsymbol{g}})^\top$*, where* $\hat{p}$ *and* $\hat{\boldsymbol{g}}$ *follow from Step 1 of Algorithm 1, then, under suitable regularity conditions (e.g., Lehmann and Romano, 2006, Theorem 12.4.2, pg. 515) with* $\beta_j = 0$ *for* $1 \leq j \leq k$*, the LRT statistic*

$$\Omega_n \equiv 2\log\left(\frac{\mathcal{L}(\hat{\boldsymbol{\Theta}}_{\text{MLE}})}{\mathcal{L}_0(\hat{\boldsymbol{\Theta}}_0)}\right) \longrightarrow_d \chi^2_k,$$

where χ^2_k *denotes a chi-square distribution with* k *degrees of freedom.*

Proof. See the Supplemental Material, Section A.5. ☐

Theorem 2.2 then motivates the following large sample hypothesis test of overall model significance, presented without proof for completeness.

Corollary 2.2.1. *Suppose we wish to test* $H_0 : \beta_j = 0$ *for all* $1 \leq j \leq k$ *against the alternative* H_A *: at least one* $\beta_j \neq 0$, $1 \leq j \leq k$*, at the asymptotic significance level,* $0 < \theta < 1$*. To do so, reject* H_0 *if* $\Omega_n > \chi^2_{1-\theta}$*, where* $\chi^2_{1-\theta}$ *denotes the* $100 \times (1-\theta)$*th percentile of a chi-square distribution with* k *degrees of freedom, and fail to reject* H_0 *otherwise.*

2.3 Right-Censoring

It is desirable to generalize the results of Section 2 to handle right-censoring, a common incomplete data challenge of consumer auto loan lifetime data observed in ABS (e.g., Lautier et al., 2023). Right-censoring manifests in ABS data in two ways. First, at a given month that ABS data is scraped for an active ABS, there will be consumer auto loans known to be actively paying with an as of yet unknown termination time. Second, the loans in an ABS must be *serviced*, which includes handling borrower payments, taking action in the event of nonpayment (e.g., collections calls, repossession, etc.), and ensuring the funds cleanly flow through the priority of payment structure defined in the ABS prospectus. Eventually, as the number of actively paying loans within an ABS dwindles, the administrative cost of servicing the few remaining active loans outweighs the economic benefit to the parent lender. When this point is reached, the ABS issuer will buyback all active loans and issue investors a final cash payment to close out the ABS transaction. For the purposes of econometric analysis, this will create a collection of consumer auto loans subject to right-censoring because the data reporting mechanism will end in conjunction with the ABS. In other words, it is not possible to circumvent exposure to right-censoring simply by waiting.

We may generalize the methods of Section 2 as follows. Let ε, $m + \Delta + 1 \leq \varepsilon \leq m + \omega$ be the nonrandom present time. The right-censoring time is then a constant time added to the start of the trust, $C_i = Y_i + \varepsilon - (m + \Delta + 1) \equiv Y_i + \tau$ (Lautier et al., 2026). When right-censoring is present, the sample takes the form $\mathcal{S}_{\tau,n} = \{(T_i, Y_i, D_i \mid \boldsymbol{z}_i)\}_{1 \leq i \leq n}$, where $T_i = \min(X_i, C_i)$ and $D_i = \mathbf{1}(X_i \leq C_i)$. The likelihood therefore must change accordingly. Specifically, when an observation is right-censored, we have

$$\begin{aligned}\Pr(T_i = u, Y_i = v, D_i = 0 \mid \boldsymbol{z}_i) &= \Pr(X_i > u, v + \tau = u, Y_i = v) \\ &= \Pr(X_i > u, Y_i = v)\mathbf{1}(v + \tau = u).\end{aligned}$$

Hence, for the density, h_* in (2), the probability of a right-censored observation is

$$\bar{h}_*(u, v; \boldsymbol{z}_i, \boldsymbol{\beta}, \boldsymbol{g}) = \frac{S(u+1; \boldsymbol{z}_i, \boldsymbol{\beta}) g_v}{\alpha(\boldsymbol{z}_i, \boldsymbol{\beta}, \boldsymbol{g})}. \tag{15}$$

The generalized likelihood equation for right-censoring is then

$$\begin{aligned}\mathcal{L}_\tau(\boldsymbol{\Theta} \mid \mathcal{S}_{\tau,n}) &= \prod_{\{i:D_i=1\}} h_*(T_i, Y_i; \boldsymbol{z}_i, \boldsymbol{\beta}, \boldsymbol{g}) \prod_{\{i:D_i=0\}} \bar{h}_*(T_i, Y_i; \boldsymbol{z}_i, \boldsymbol{\beta}, \boldsymbol{g}) \\ &= \prod_{\{i:D_i=1\}} \frac{f(T_i; \boldsymbol{z}_i, \boldsymbol{\beta}) g_v}{\alpha(\boldsymbol{z}_i, \boldsymbol{\beta}, \boldsymbol{g})} \prod_{\{i:D_i=0\}} \frac{S(T_i + 1; \boldsymbol{z}_i, \boldsymbol{\beta}) g_v}{\alpha(\boldsymbol{z}_i, \boldsymbol{\beta}, \boldsymbol{g})}\end{aligned}$$

$$= \prod_{v=\Delta+1}^{\Delta+m} g_v^{\hat{\gamma}_n(v)} \prod_{\{i:D_i=1\}} f(T_i; \boldsymbol{z}_i, \boldsymbol{\beta}) \prod_{\{i:D_i=0\}} S(T_i+1; \boldsymbol{z}_i, \boldsymbol{\beta}) \prod_{i=1}^{n} \frac{1}{\alpha(\boldsymbol{z}_i, \boldsymbol{\beta}, \boldsymbol{g})}.$$

Starting from $\mathcal{L}_\tau$, we generalize all methods from Sections 2.1 and 2.2 for right-censoring in the Supplemental Material, Section D. These results are deferred to the Supplemental Material because the main theoretical insight stems from handling left-truncation, and it provides the blueprint to also handle the type of right-censoring that occurs with ABS data.

3 Simulation Studies

We conduct two simulation studies to numerically assess the performance of Theorems 2.1 and 2.2 under a less than typical sample size found in ABS data. For the former, we assign $\omega = 12$, $\Delta = 0$, $m = 8$, $\boldsymbol{\beta} = (\beta_0 = 0.5, \beta_1 = 0.5, \beta_2 = 1, \beta_3 = -1.5, \beta_4 = -0.5)^\top$, and $\boldsymbol{g} = (g_1 = 0.3, g_2 = 0.2, g_3 = 0.13, g_4 = 0.10, g_5 = 0.09, g_6 = 0.07, g_7 = 0.06, g_8 = 0.05)^\top$. With the linear relationship of $\boldsymbol{g}$, this results in 12 free parameters that require estimation. The presence of the $\boldsymbol{g}$ parameters due to random left-truncation can quickly increase the dimension of the unknown parameters such that direct numerical approaches can fail (Lautier et al., 2025). The distribution for $\boldsymbol{g}$ is nonuniform, which is more general than the uniformity assumption required by length-biased sampling (e.g., Wang, 1996; Qin and Shen, 2010). The design matrix was generated by $\mathbf{Z} = (\mathbf{1}_n, \mathcal{N}_1(0, 0.1)_n, \mathcal{N}_2(0, 0.1)_n, \mathcal{N}_3(0, 0.1)_n, \mathcal{N}_4(0, 0.1)_n)$, where $\mathcal{N}_i(0, 1)_n$ represents n independent and identically distributed draws from a standard normal distribution, $1 \leq i \leq 4$. For this simulation study, $n = 1{,}000$, which we choose to be much smaller than the sample sizes found in Section 4. We further assume that $f(X_i) \sim \text{Binomial}(\omega - \Delta - 1, p_i)$, where $p_i = \exp(\eta_i)/(1 + \exp(\eta_i))$ and $\eta_i = \boldsymbol{z}_i^\top \boldsymbol{\beta}$.

The simulation procedure consists of first generating a design matrix, $\mathbf{Z}$, and then a corresponding sample of size 1,000 from (2). Next, we perform the parameter estimation via Algorithm 1 and then calculate the implied variability of the parameter estimates by (14). These calculations are stored, and this process is repeated for 1,000 replicates. The results

Table 1: **Simulation Study Results**. The robustness of Theorems 2.1 and 2.2 for $n = 1{,}000$ with $(\omega = 12, \Delta = 0, m = 8)$ and $(\omega = 8, \Delta = 0, m = 5)$, respectively, for X following a binomial distribution with a logit link function for 1,000 replicates. (Left) We report the absolute value of the empirical bias (|eBias|), the empirical standard error (eSE), an average of the estimated standard error using Theorem 2.1 (Thm 2.1 SE), and a coverage probability (CP) for 95% asymptotic confidence intervals using Corollary 2.1.1. (Right) We report the empirical percentile the test statistic, Ω_n (Emp) against the anticipated percentile of a chi-square distribution with four degrees of freedom (Thm 2.2). For completeness, the observed empirical Type I error of the hypothesis test in Corollary 2.2.1 with significance level $\theta = 0.05$ was 0.047 over these 1,000 replicates.

$\boldsymbol{\Theta}^*$	True	|eBias|*	eSE*	Thm 2.1 SE*	CP	Perc	Emp	Thm 2.2
g_1	0.30	0.60	14.52	14.37	95.4%	0.05	0.762	0.711
g_2	0.20	0.29	12.53	12.48	95.4%	0.10	1.033	1.064
g_3	0.13	0.12	10.46	10.46	95.3%	0.20	1.671	1.649
g_4	0.10	0.05	9.54	9.32	94.8%	0.30	2.137	2.195
g_5	0.09	0.17	8.84	8.97	95.5%	0.40	2.726	2.753
g_6	0.07	0.26	8.14	8.25	94.5%	0.50	3.312	3.357
g_7	0.06	0.49	7.95	8.28	95.4%	0.60	3.941	4.045
β_0	0.50	0.25	17.72	20.58	98.5%	0.70	4.836	4.878
β_1	0.50	19.50	202.28	203.89	94.8%	0.80	6.087	5.989
β_2	1.00	1.72	202.04	204.37	94.9%	0.90	7.673	7.779
β_3	-1.50	8.73	208.92	204.97	94.5%	0.95	8.961	9.488
β_4	-0.50	7.70	202.11	203.93	94.8%	0.99	13.676	13.277

*Results on a 10^{-3} scale (i.e., 7.1 = 0.0071).

may be found in Table 1. The absolute value of the empirical bias is small for all 12 parameters. Furthermore, the empirical standard error is close to the empirical average of the estimated standard error using Theorem 2.1. Together, these results help numerically validate Theorem 2.1. Additionally, we construct 95% confidence intervals using Corollary 2.1.1, and the empirical coverage probabilities are close to the expected 95% for all parameters.

The second simulation study numerically explores Theorem 2.2. In this case, $\omega = 8$, $\Delta = 0$, and $m = 5$. Because H_0 must be true, we assign $\boldsymbol{\beta} = (\beta_0 = 0.5, \beta_1 = 0, \beta_2 = 0, \beta_3 = 0, \beta_4 = 0)^\top$, and $\boldsymbol{g} = (g_1 = 0.3, g_2 = 0.25, g_3 = 0.2, g_4 = 0.15, g_5 = 0.10)^\top$. The design matrix generation, sample size, form of $f(\cdot)$, and linking function are the same as in the first simulation study. With H_0 assumed true, this implies $p = 0.6225$ for $f(\cdot)$. The simulation procedure consists of first generating a design matrix, $\mathbf{Z}$, and then a corresponding sample of size 1,000 from (2). Next, we perform the parameter estimation via Algorithm 1 and then calculate the loglikelihood evaluated at the parameter estimate for this sample. Next, we perform a parameter estimation using the restricted parameter space and approach of Lautier et al. (2025) and then calculate the corresponding loglikelihood from Theorem 2.2. Finally, we calculate Ω_n for this simulated sample and store the results. This process is then repeated for 1,000 replicates. Per Theorem 2.2, we expect a distribution of these 1,000 Ω_n replicates to follow a chi-square distribution with four degrees of freedom. A comparison may be found in Table 1. We can see that the true distribution and empirical distribution of the 1,000 test statistics are close by comparing the empirical and true percentiles. For completeness, we also calculate the empirical Type I error observed over the 1,000 replicates for the hypothesis test described in Corollary 2.2.1 with significance level $\theta = 0.05$ to be 0.047, which is further numerical verification of both Theorem 2.2 and Corollary 2.2.1. For completeness, we conclude this section by commenting that corresponding simulation studies to validate the methods of Section 2.3 may be found in the Supplemental Material, Section D.

4 Application

We consider the Ally Auto Receivables Trust (AART) 2017-3 consumer auto loan ABS bond (AART, 2017). The AART 2017-3 ABS bond was issued in the spring of 2017 and was actively paying for 44 months. It originally contained 67,797 unique consumer auto loans originated between February of 2011 and April of 2017. The average credit score is 725 with a median credit score of 719. This corresponds to a *super-prime* credit risk pool (Lautier et al., 2024). The loan term of each loan ranges from 12 to 78 months, and average outstanding loan balances as of the first ABS payment are \$15,605 with a median of \$14,153. The auto loans within AART 2017-3 are for both new and used vehicles, spanning 42 different auto manufacturers. Additionally, loans originations span 51 states and U.S. territories.

Detailed asset level information, known colloquially as the *loan tape*, provides ABS investors with loan level demographic information as of contract signing (Securities and Exchange Commission, 2016). Specifically, there are four continuous covariates and six indicator covariates. The continuous covariates include the borrower's credit score (`cred.s`), the annual percentage rate (`int.rt`), the payment-to-income (`pti`), and the estimated vehicle value (`v.val`). Within this context, the payment-to-income provides *the scheduled monthly payment amount as a percentage of the total monthly income of the obligor and any other obligor at the origination date* (Securities and Exchange Commission, 2016). It is a common measure of risk when pricing consumer loans (Lautier et al., 2024). Prior to the model fitting, these four continuous time variables are scaled and centered. For the indicator variables, we define whether a co-obligor or co-signer is present on the contract (`co.sgn`), the vehicle is new or used, with new coded as 1, (`nw.usd`), subvention is present with a reduced interest rate (`sbvt.r`) or cash rebate (`sbvt.c`), the vehicle is a pick-up truck (`v.picu`), and the vehicle is a sport-utility-vehicle (`v.suv`). In this setting, subvention is a *form of subsidy received on the loan, such as cash incentives or favorable financing for the buyer* (Securities and Exchange Commission, 2016). An advantage of the methods proposed in Section 2 is that we may explore potential explanatory relationships between these economic variables and borrower behavior. These potential insights are not yet available in related analysis (e.g., Lautier et al., 2024, 2025). Additionally, it is possible to demonstrate that the proportional hazard assumption is questionable for the AART 2017-3 data we consider. Please see the Supplemental Material, Section F for details.

We thus need the methods of Section 2, and we now review four components of this proposed model within the context of the AART 2017-3 data. First, the support of $\mathcal{A}$ is nonrandom and known at the onset of the estimation problem. In other words, the nonrandom support variables Δ, m, ω, and ε are defined using given information from the amortization schedule in combination with the bounds of the observed sample. Practically, this implies that each loan term needs to be model separately. (This is consistent with industry practice, as loans are often priced for borrowers by loan term.) Returning back to AART 2017-3, we observe the expected clustering around standard industry loan terms, such as 36, 48, 60, 72, and 75 months. (Consumer auto loans are recently extending into longer terms (Katcher et al., 2024).) Specifically, these five loan terms account for 2,798, 4,411, 17,741, 30,409, and 8,644 loans, respectively, or 64,003 (94.4%) of the total 67,797 auto loans in AART 2017-3. Therefore, it is not a significant practical constraint to prepare separate models by loan term. (For completeness, a detailed breakdown of AART 2017-3 by loan term and outcome may be found in the Supplemental Material, Section E.)

Second, we need to define an event for a time-to-event analysis. For the super-prime consumer auto loan data with few defaults observed within AART 2017-3, the most natural event is a complete loan repayment. Borrower prepayment behavior is a research topic of interest within economic studies in consumer finance, too, especially given it is understudied (Lautier et al., 2024) in comparison to residential mortgages (e.g., Deng et al., 2000; Calhoun and Deng, 2002; Ambrose and Sanders, 2003; Jones and Sirmans, 2019). The monthly reporting of AART 2017-3 provides detailed payment information (Securities and Exchange Commission, 2016). Therefore, we may compare the outstanding balance at each month with any payments received on a loan level basis. Once the outstanding balance reaches zero, we consider a loan repaid. We then record the loan's age in months, and this age becomes the time-to-event observation. This implies that any partial prepayment will come through in the time-to-event data by shortening the observed lifetime of a loan. This approach is consistent with related literature (e.g., Lautier et al., 2024, 2025).

Third, brief comments are necessary to clarify that the proposed methods of Section 2 are not designed to handle two time-to-event random variables simultaneously (i.e., a *competing risks* framework). For some consumer auto loan data, this can be problematic because a substantial portion of loans may default (e.g., the *deep subprime* loans considered in Lautier et al. (2024) default at an ∼50% rate). For AART 2017-3, however, only ∼6.1% of loans default in aggregate, with a general expected trend of more defaults as loan terms increase (see the Supplemental Material, Section E). Therefore, the methods of Section 2 are suitable to analyze the AART 2017-3 data with minimal concerns of selection bias. This is a prepayment study for a low-default, super-prime segment, not a competing-risks portfolio model; the latter is outside the scope of the present paper and remains a natural direction for future work. We discuss future generalizations for competing risks further in Section 5.

Fourth, it is left to make a selection for the lifetime density function, $f(\cdot; p_i)$. Because consumer auto loan lifetimes are integer-valued, positive, and have a finite support, a binomial distribution is a logical choice. It also has a history of investigation within the context of ABS consumer auto lifetime data (Lautier et al., 2025). Beyond this, we may also argue in favor of the binomial model from the perspective of the consumer finance application. For example, AART 2017-3 consists of mainly high-credit quality borrowers, who are generally more immune to broader economic shocks and possibly more willing to make higher monthly payments to avoid the perceived hassle of refinancing (Lautier et al., 2024). Thus, prepayments may be more connected to the amortization schedule, on average. Concisely, the typical bell shape for a binomial distribution is reasonable for prepayments, whereas defaults tend to occur more sporadically throughout a loan's lifetime. Additionally, the loan tape described earlier represents borrower information as of the moment a contract is signed. Hence, the parameter p_i is effectively assessing the ability of traditional underwriting components to capture future loan behavior. For ABS investors, this perspective is valuable in helping to compare lenders and even different ABS bonds against each other. In other words, investors will not have future performance when pricing a new ABS bond, and it is helpful to calibrate loan lifetime models as of this initial time. For future work, it may be desirable to allow for time-varying covariates, and we recommend this as an area of future study.

We are now prepared to analyze the AART 2017-3 data. From the discussion of the previous four paragraphs, we will prepare models for 36-month and 60-month consumer auto loans. This will allow us to demonstrate both the model for only left-truncated data (Sections 2.1 and 2.2) via 36-month loans and the generalized model for right-censoring (Section 2.3) via 60-month loans. At the same time, it represents a thorough applied study into borrower prepayment behavior for auto loans using the methodological approaches we propose. After data cleaning, we prepare two models for 2,756 36-month and 17,016 60-month consumer auto loans (see the Supplemental Material, Section E for data notes).

The results are in Table 2. For 36-month loans, we identify $\Delta = 3$, $m = 34$, and $\omega = 41$. For 60-month loans, we identify $\Delta = 2$, $m = 60$, $\omega = 68$, and $\varepsilon = 106$. By Corollary 2.2.1 (and its counterpart in the Supplemental Material, Section D), we reject H_0 at any reasonable significance level for both models. This implies that using regressors has improved the model fit in comparison to Lautier et al. (2025). We may proceed to make inference on prepayment speeds. For the standard logit link function, $p_i = \exp(\eta_i)/(1 + \exp(\eta_i))$, $\eta_i = \boldsymbol{z}_i^\top \boldsymbol{\beta}$, and $\boldsymbol{\beta} = (\beta_0, \beta_1, \ldots, \beta_{11})^\top$. Thus, we can assess the explanatory power of each economic variable by assessing its impact on the log-odds ratio,

$$\log\left(\frac{p_i}{1-p_i}\right) = \boldsymbol{z}_i^\top \boldsymbol{\beta}. \tag{16}$$

That is, each β_j, $1 \leq j \leq 11$, is the additive increase in (16) from a one unit increase in the covariate z_{ij}, if all other covariates, z_{ij*}, $j^* \neq j$, are held fixed. Hence, a positive and significant coefficient can be associated with a slower prepayment speed and vice versa.

With this backdrop, we can provide an economic interpretation of the fitted model in Table 2 with Corollary 2.1.1 (and its counterpart in the Supplemental Material, Section D). We will move sequentially down the ten explanatory variables. First, we see that borrower credit score has a significant, negative effect on loan repayment time for both models. This suggests that a higher borrower credit score will result in a shorter time to loan repayment, all else equal. One interpretation of this result is that a higher credit score provides a borrower with easier access to credit. Thus, there is less friction in the credit approval process to buy another car or refinance, and this may explain why prepayment speeds increase with borrower credit scores. The contract interest rate is also a negative and significant effect in both models, as expected. This aligns with rational borrower behavior in that a higher interest rate increases the cost of borrowing. Hence, all else equal, borrowers paying a higher

Table 2: **AART 2017-3 Prepayment Behavior Regression Results**. The fitted model of Section 2 for the time-to-loan-repayment response for 36-month loans ($n = 2{,}756$, $\Delta = 3$, $m = 34$, $\omega = 41$) and 60-month loans ($n = 17{,}016$, $\Delta = 2$, $m = 60$, $\omega = 68$, $\varepsilon = 106$) from the AART 2017-3 ABS bond. Results assume a binomial lifetime with a logit-link function.

	36-Month Loans					60-Month Loans				
Param.	Est.	S.E.	Wald	*p*-val.	Sig.	Est.	S.E.	Wald	*p*-val.	Sig.
β_0	1.1056	0.0242	45.65	<0.0001	***	0.6022	0.0064	94.07	<0.0001	***
`cred.s`	-0.4547	0.0118	-38.55	<0.0001	***	-0.2361	0.0028	-84.24	<0.0001	***
`int.rt`	-0.2125	0.0118	-17.97	<0.0001	***	-0.2306	0.0033	-69.74	<0.0001	***
`pti`	0.0881	0.0103	8.52	<0.0001	***	0.0835	0.0025	33.93	<0.0001	***
`v.val`	0.0008	0.0151	0.05	0.9588		-0.1053	0.0038	-28.01	<0.0001	***
`co.sgn`	0.2104	0.0178	11.82	<0.0001	***	0.1434	0.0045	31.63	<0.0001	***
`nw.usd`	0.0527	0.0316	1.67	0.0949	.	0.0065	0.0085	0.77	0.4432	
`sbvt.r`	1.1142	0.0507	21.96	<0.0001	***	0.7890	0.0120	65.86	<0.0001	***
`sbvt.c`	0.0293	0.0256	1.15	0.2520		0.0163	0.0069	2.36	0.0183	*
`v.picu`	0.1275	0.0225	5.67	<0.0001	***	0.1837	0.0061	30.27	<0.0001	***
`v.suv`	-0.0195	0.0222	-0.88	0.3778		0.0119	0.0057	2.08	0.0374	*

Significance levels: $\alpha = 0.001$ (***), $\alpha = 0.01$ (**), $\alpha = 0.05$ (*), and $\alpha = 0.1$ (.).

borrowing cost will attempt to pay down their loans sooner.

The payment-to-income (PTI) of a borrower has a positive and significant effect in both models. This implies that a higher PTI will decrease an auto loan's prepayment rate. A higher PTI is typically associated with a higher perceived credit risk, as a larger share of a borrower's income is required to meet the monthly payment. Phrased differently, a borrower with a higher PTI may have less economic freedom to replace an existing auto loan with an alternative option, which would work to lower the prepayment rate. Vehicle value at contract origination is not significant for 36-month loans, and it has a negative and significant effect for 60-month loans. This implies that borrowers purchasing a more expensive vehicle, all else equal, will pay down their loans faster. It is reasonable to interpret vehicle purchase price as a proxy for a borrower's affluence, which implies greater flexibility to obtain financing. Similarly, a higher initial vehicle value may be associated with a higher future trade-in value, which can facilitate the purchase of a new vehicle (and hence a prepayment).

Next, we see that the presence of a co-obligor or co-signer has a positive and significant effect on the time-until-loan-repayment for both models. Consumer loan contracts that include a co-borrower often indicate that the vehicle's driver could obtain improved financing terms with a second party bearing legal responsibility to repay the loan. This is likely associated with a borrower in a weaker financial position that may not have access to the resources to more quickly pay down an auto loan. A new vehicle purchase has a slightly positive and significant effect for 36-month loans, and it is not significant for 60-month loans. All else equal, borrowers that finance a new vehicle prepay more slowly. Over a 36-month period, which is consistent with a typical auto lease contract (Lautier et al., 2023), a borrower may feel that the car retains enough modern features to not need replacing within a three-year window. In comparing the two models, information about a car being new or used has explanatory power on prepayment speeds for 36-month loans while vehicle value does not and vice versa for 60-month loans. Hence, over a longer period of time, it is the monetary initial value of the auto that carries a higher importance than whether it is new or used, which is consistent with a rational economic actor.

For both models, the largest magnitude, significant effect on borrower prepayment speed is the presence of subvention on the contractual interest rate. This is a validating result for the modeling process we propose because it indicates that borrowers paying a below market borrowing cost rationally prepay more slowly, all else equal. This aligns with the expectation that borrowers with a higher credit score are generally more financially sophisticated (e.g., Lautier et al., 2024), which comes through via a demonstrated slower prepayment when it is economically efficient to do so. For cash subvention, the effect is slightly positive in both models, and it is only significant for 60-month loans. The presence of a cash rebate indicates a loan contract that is signed at a below market cost, from which it is rational to prepay more slowly. Subvention via the interest rate has many times the impact of cash-based subvention, however. These results may be of interest to compare with the ability of subvention programs to improve auto sales (Thompson and Noordewier, 1992).

Perhaps the most interesting result in Table 2 is that consumer auto loan contracts secured by pick-up trucks prepay more slowly than other consumer auto loans, ceteris paribus.

The effect is significant in both models. Possible interpretations for this are that pick-up trucks may be more associated with working vehicles or providing some other benefit to a borrower, like hauling, towing, or traversing off-road terrain. If so, such borrowers may make repayment decisions more dependent on vehicle utility than borrowing terms. Additionally, pick-up trucks are known to retain their value better than other vehicle types (Birch, 2012), and so a slower prepayment rate may reflect a slower than typical depreciation rate on the underlying, secured vehicle. Contrast this to the sport-utility vehicle (SUV) indicator, which is slightly negative and not significant for 36-month loans and just positive and significant for 60-month loans. This may reflect improving fuel efficiency of SUVs, which lessens the impact of gasoline prices on consumer preferences (e.g., Ohta and Griliches, 1986; Ohta, 1987). With the increasing popularity of SUVs among consumers, it is not surprising to find no significant effect in comparison to non-SUVs. That is, the SUV is a modern replacement for the traditional consumer sedan in American families, whereas a pick-up truck still retains a distinctive vehicle type among consumers (e.g., Costa, 2014).

We also provide point estimates plus 95% asymptotic confidence intervals using the methods of Section 2 for the parameters of the left-truncation random variable, $\boldsymbol{g}$, corresponding to Table 2 in Figure 1. From a visual inspection, it is apparent that the distribution of Y is non-uniform for both models. Additionally, we may perform the hypothesis test of Lautier et al. (2026, Corollary 3.5.1) to formally test if Y follows a discrete uniform distribution for 36-month loans. We find a p-value of approximately zero, which is further evidence that Y is non-uniform. This has been documented before for ABS data and is attributed to the seasonality observed in auto sales (Lautier et al., 2026). The non-uniformity of Y violates length-biased sampling-based approaches (e.g., Wang, 1996; Qin and Shen, 2010) and reinforces why the flexibility of the model in (2) is necessary for auto ABS loan data.

We conclude this section with an outline for how the regression model we propose may be used in industry settings. We suggest first setting up a database that is fed by the (Securities and Exchange Commission, 2014) monthly reporting for all publicly issued consumer auto

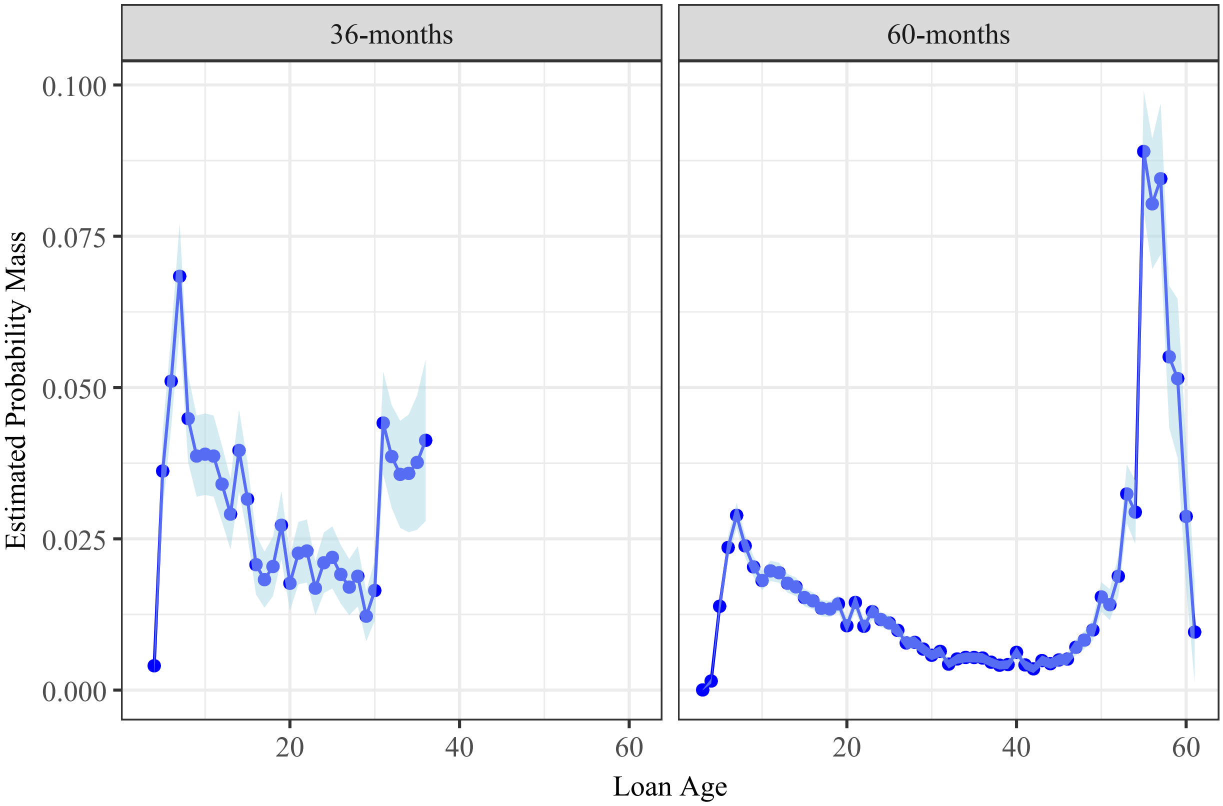

Figure 1: **AART 2017-3 Left-Truncation Distribution**. Point estimates and asymptotic confidence intervals for $\boldsymbol{g}$ follow via Section 2 and correspond to the coefficient estimates in Table 2. These non-uniform distributions are not suitable for length-biased sampling-based approaches.

ABS. Next, this data can feed into a regression model similar to Table 2. We suggest fitting a separate model for the various different issuers, such as AART, CarMax Auto Owner Trust, Drive Auto Receivables Trust, and Santander Drive Auto Receivables Trust (for a complete list, see D'Onofrio et al. (2020)). These issuer specific models may then be compared against each other to search for potential issuer-level differences in the explanatory power of the loan tape variables in Table 2 on underlying loan lifetimes. These insights may then inform internal pricing, risk management, or credit rating models at an issuer level, which can then be compared to market prices and rating agencies. Similarly, a regression model can be fit within one issuer, such as AART, but for different vintages, such as AART 2017-3 versus AART 2019-4 to compare performance over time. Additionally, this ABS-specific regression model may be updated on a monthly basis to compare observed performance versus expectation at issuance. This can help identify material deviations sooner, which can inform more timely buy and sell decisions. In all cases, the regression model we propose can be the stochastic foundation of a cash flow model built using the ABS loan tape, whereas previous iterations were unable to use this loan level information (e.g., Lautier et al., 2023).

5 Discussion

We propose a new lifetime model capable of handling the required combination of discrete-time, left-truncation, and inherent data heterogeneity needed to analyze the recent release of public consumer auto loan data contained within ABS (Securities and Exchange Commission, 2014). Specifically, we generalize Lautier et al. (2025) to allow for regressors, which provides a new level of depth to the growing body of research in this area (e.g., Katcher et al., 2024; Lautier et al., 2024, 2025, 2026). In sum, there are two components to this paper.

First, we review the proposed methods of Section 2. We derive the likelihood function and numerically optimize it to find the coefficient estimates. This is a nontrivial problem because it is a constrained optimization of high-dimension. The procedure we propose is summarized in Algorithm 1, which provides general, closed-form update steps. We next provide a fully-stated, large sample covariance matrix to quantify estimator uncertainty. This allows for inference into variable and model selection, as summarized in Theorems 2.1 and 2.2 and Corollaries 2.1.1 and 2.2.1, respectively. Throughout, we leave the left-truncation random variable completely unspecified. Finally, we generalize all methods for the type of right-censoring found in ABS data, greatly expanding the practical utility of these methods. All methods are verified numerically in Section 3 with proofs in the Supplemental Material.

Second, we utilize these methods for a first-of-its-kind economic study into prepayment behavior for consumer auto loans using the Securities and Exchange Commission (2014) data from the AART (2017) consumer ABS bond. We fit two regression models spanning 19,772 consumer auto loans with original loan terms of 36 and 60 months. Pleasingly, we find evidence that borrowers prepay rationally, such as prepayment speeds increasing for higher contract interest rates or prepayment speeds decreasing for subvented contract interest rates, all else equal. We also find the interesting result that consumers with pick-up trucks tend to prepay more slowly than consumers without pick-up trucks, ceteris paribus, among other results summarized in Section 4, which concludes with suggestions for industry professionals.

We conclude with some suggested areas of further study. The most immediate is to generalize the methods of Section 2 to handle time-varying covariates, such as current unemployment or interest rates. While the Securities and Exchange Commission (2014) regressors are a point-in-time variable recorded at loan origination, it would likely be of value to overlay current payment data with broader economic indicators. Looking ahead, it would also be desirable to study the complex problem of *competing risks* (i.e., a generalization of Lautier et al. (2024) to handle covariates). A competing risks framework is valuable because of the profitability difference between defaulted and prepaid loans. There are various interpretations and competing risk frameworks, however (Klein and Moeschberger, 2003), and so a careful study will be necessary to properly handle two simultaneous, random event outcomes. Furthermore, we also suggest to expand the prepayment study of Section 4 into a focused, applied economic study. This would help close the gap between the robust economic literature into mortgage prepayments (e.g., Deng et al., 2000; Calhoun and Deng, 2002; Ambrose and Sanders, 2003; Jones and Sirmans, 2019) and minimal literature into auto loan prepayments (Lautier et al., 2024). Finally, we postulate that our survival methods may find suitable applications in other fields, such as medicine, insurance, or engineering.

A Theorem 2.1 Statement Details

For the covariance matrix, $\mathbf{\Sigma}_n(\hat{\mathbf{\Theta}}^*_{\text{MLE}})$, defined in Theorem 2.1, we have

$$-\mathbf{\Sigma}_n(\hat{\mathbf{\Theta}}^*_{\text{MLE}}) = \begin{bmatrix} \mathbf{\Sigma}_{11} & \mathbf{\Sigma}_{12} \\ \mathbf{\Sigma}_{12}^{\top} & \mathbf{\Sigma}_{22} \end{bmatrix},$$

where

$$\mathbf{\Sigma}_{11} = \sum_{i=1}^{n} \begin{bmatrix} \tilde{B}_i^*(\Delta+1, \Delta+1) & \tilde{B}_i^*(\Delta+1, \Delta+2) & \dots & \tilde{B}_i^*(\Delta+1, \Delta+m-1) \\ \tilde{B}_i^*(\Delta+2, \Delta+1) & \tilde{B}_i^*(\Delta+2, \Delta+2) & \dots & \\ \vdots & \vdots & \ddots & \\ \tilde{B}_i^*(\Delta+m-1, \Delta+1) & \dots & & \tilde{B}_i^*(\Delta+m-1, \Delta+m-1) \end{bmatrix},$$

for $\Delta + 1 \le v, v^* \le \Delta + m - 1$, with

$$\tilde{B}_i^*(v, v^*) = \begin{cases} -\dfrac{\mathbf{1}(Y_i = \Delta+m)}{(1 - \sum_{k=\Delta+1}^{\Delta+m-1} g_v)^2} + \dfrac{(S_{vi} - S_{(\Delta+m)i})(S_{v^*i} - S_{(\Delta+m)i})}{(\alpha_i^*)^2}, & v \neq v^* \\ -\dfrac{\mathbf{1}(Y_i = v)}{g_v^2} - \dfrac{\mathbf{1}(Y_i = \Delta+m)}{(1 - \sum_{k=\Delta+1}^{\Delta+m-1} g_v)^2} + \dfrac{(S_{vi} - S_{(\Delta+m)i})^2}{(\alpha_i^*)^2}, & v = v^*, \end{cases}$$

$$\mathbf{\Sigma}_{12} = \begin{bmatrix} \mathbf{J}_n^\top \tilde{\mathbf{D}}_{\Delta+1}^* \mathbf{Z} & \mathbf{J}_n^\top \tilde{\mathbf{D}}_{\Delta+2}^* \mathbf{Z} & \dots & \mathbf{J}_n^\top \tilde{\mathbf{D}}_{\Delta+m-1}^* \mathbf{Z} \end{bmatrix}^\top,$$

for $\Delta + 1 \le v, v^* \le \Delta + m - 1$, where $\tilde{\mathbf{D}}_v^* = \operatorname{diag}(\tilde{d}_i^*(v))$ is an $n \times n$ diagonal matrix with

$$\tilde{d}_i^*(v) = \frac{r^*(\boldsymbol{g}^*, \eta_i)(S_{vi} - S_{(\Delta+m)i})}{(\alpha_i^*)^2} - \frac{1}{\alpha_i^*}\left(\sum_{u=v}^{\omega} q(u, \eta_i) - \sum_{u=\Delta+m}^{\omega} q(u, \eta_i) \right),$$

$$r^*(\boldsymbol{g}^*, \eta_i) = \sum_{v=\Delta+1}^{\Delta+m-1} g_v \left(\sum_{u=v}^{\omega} q(u, \eta_i) \right) + \left(1 - \sum_{k=\Delta+1}^{\Delta+m-1} g_k \right) \sum_{u=\Delta+m}^{\omega} q(u, \eta_i),$$

and $\mathbf{\Sigma}_{22} = \mathbf{Z}^\top (\mathbf{A}^*) \mathbf{Z}$, where $\mathbf{A}^* = \operatorname{diag}(A_i^*)$ is an $n \times n$ diagonal matrix with

$$\begin{aligned} A_i^* = {} & \frac{s(X_i, \eta_i)}{f(X_i; \boldsymbol{z}_i, \boldsymbol{\beta})} - \left(\frac{q(X_i, \eta_i)}{f(X_i; \boldsymbol{z}_i, \boldsymbol{\beta})} \right)^2 + \frac{r^*(\boldsymbol{g}_*, \eta_i)^2}{(\alpha_i^*)^2} \\ & - \frac{1}{\alpha_i^*} \left(\sum_{v=\Delta+1}^{\Delta+m-1} g_v \left(\sum_{u=v}^{\omega} s(u, \eta_i) \right) + \left(1 - \sum_{k=\Delta+1}^{\Delta+m-1} g_k \right) \sum_{u=\Delta+m}^{\omega} s(u, \eta_i) \right). \end{aligned}$$

References

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

A. Alioum and D. Commenges (1996). "A proportional hazards model for arbitrarily censored and truncated data." *Biometrics* **52**, 512–524.

B. W. Ambrose and A. B. Sanders (2003). "Commercial mortgage-backed securities: Prepayment and default." *The Journal of Real Estate Finance and Economics* **26**, 179–196.

D. Bertsekas (1999). *Nonlinear Programming*. Athena Scientific.

R. Birch (2012). "Pick-ups hold up, despite high price of keeping them fueled up." *Credit Union Journal* **16**, 19.

C. A. Calhoun and Y. Deng (2002). "A dynamic analysis of fixed- and adjustable-rate mortgage terminations." *The Journal of Real Estate Finance and Economics* **24**, 9–33.

L.-P. Chen and G. Y. Yi (2021). "Semiparametric methods for left-truncated and right-censored survival data with covariate measurement error." *Annals of the Institute of Statistical Mathematics* **73**, 481–517.

J. Costa (2014). *The American Pick-up Truck – How deep is your love?*, pp. 109–120. Berlin, Heidelberg: Springer Berlin Heidelberg.

D. R. Cox (1972). "Regression models and life-tables." *Journal of the Royal Statistical Society: Series B (Methodological)* **34**, 187–202.

Y. Deng, J. M. Quigley and R. Van Order (2000). "Mortgage terminations, heterogeneity and the exercise of mortgage options." *Econometrica* **68**, 275–307.

C. D'Onofrio, C. O'Connell and M. Reph (2020). "US ABS figures on automotive securitization tapes (FAST) report." Data and Analytics, Morningstar DBRS.

Federal Reserve (2023). "Statistical release, consumer credit (G.19)." Accessed: 2023-07-14.

S. T. Gross and C. Huber-Carol (1992). "Regression models for truncated survival data." *Scandinavian Journal of Statistics* **19**, 193–213.

J. Hardin and J. Hilbe (2007). *Generalized Linear Models and Extensions, Second Edition*. A Stata Press publication. Taylor & Francis.

T. Jones and G. S. Sirmans (2019). "Understanding subprime mortgage default." *Journal of Real Estate Literature* **27**, 27–52.

J. D. Kalbfleisch and J. F. Lawless (1991). "Regression models for right truncated data with applications to AIDS incubation times and reporting lags." *Statistica Sinica* **1**, 19–32.

B. Katcher, G. Li, A. Mezza and S. Ramos (2024). "One month longer, one month later? Prepayments in the auto loan market." Finance and Economics Discussion Series (FEDS) 2024.056, Federal Reserve Board, Washington, D.C.

J. P. Klein and M. L. Moeschberger (2003). *Survival Analysis: Techniques for Censored and Truncated Data, Second Edition*. Springer Science + Business Media, Inc.

M. H. Kutner, C. J. Nachtsheim, J. Neter and W. Li (2005). *Applied Linear Statistical Models*. McGraw-Hill Irwin.

T. L. Lai and Z. Ying (1991). "Rank regression methods for left-truncated and right-censored data." *The Annals of Statistics* **19**, 531 – 556.

J. P. Lautier, V. Pozdnyakov and J. Yan (2023). "Pricing time-to-event contingent cash flows: A discrete-time survival analysis approach." *Insurance: Mathematics and Economics* **110**, 53–71.

J. P. Lautier, V. Pozdnyakov and J. Yan (2024). "On the convergence of credit risk in current consumer automobile loans." *Journal of the Royal Statistical Society Series A: Statistics in Society* **189**, 196–221.

J. P. Lautier, V. Pozdnyakov and J. Yan (2025). "Estimating the time-to-event distribution for loan-level data within an asset-backed security." *Annals of Applied Statistics* **19**, 2831–2851.

J. P. Lautier, V. Pozdnyakov and J. Yan (2026). "Estimating a discrete distribution subject to random left-truncation with an application to structured finance." *Econometrics and Statistics* **37**, 174–198.

E. Lehmann and J. Romano (2006). *Testing Statistical Hypotheses*. Springer Texts in Statistics. Springer New York.

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

T. Nakagawa and S. Osaki (1975). "The discrete Weibull distribution." *IEEE Transactions on Reliability* **R-24**, 300–301.

M. Ohta (1987). "Gasoline cost and hedonic price indexes of U.S. used Cars for 1970–1983 a note." *Journal of Business & Economic Statistics* **5**, 521–528.

M. Ohta and Z. Griliches (1986). "Automobile prices and quality: Did the gasoline price increases change consumer tastes in the U.S.?" *Journal of Business & Economic Statistics* **4**, 187–198.

W. Pan and R. Chappell (2002). "Estimation in the Cox proportional hazards model with left-truncated and interval-censored data." *Biometrics* **58**, 64–70.

Y. Pawitan (2013). *In All Likelihood: Statistical Modelling and Inference Using Likelihood*. OUP Oxford.

J. Qin and Y. Shen (2010). "Statistical methods for analyzing right-censored length-biased data under Cox model." *Biometrics* **66**, 382–392.

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

L. Rennert and S. X. Xie (2018). “Cox regression model with doubly truncated data.” *Biometrics* **74**, 725–733.

Securities and Exchange Commission (2014). “17 CFR Parts 229, 230, 232, 239, 240, 243, and 249 Asset-Backed Securities Disclosure and Registration.”

Securities and Exchange Commission (2016). “17 CFR S 229.1125 (Item 1125) Schedule AL - Asset-Level Information.”

J. Shao (2003). *Mathematical Statistics*. Springer New York, NY.

Y.-R. Su and J.-L. Wang (2012). “Modeling left-truncated and right-censored survival data with longitudinal covariates.” *The Annals of Statistics* **40**, 1465–1488.

B. Teodorescu, I. Van Keilegom and R. Cao (2010). “Generalized time-dependent conditional linear models under left truncation and right censoring.” *Annals of the Institute of Statistical Mathematics* **62**, 465–485.

P. A. Thompson and T. Noordewier (1992). “Estimating the effects of consumer incentive programs on domestic automobile sales.” *Journal of Business & Economic Statistics* **10**, 409–417.

P. Tseng (2001). “Convergence of a block coordinate descent method for nondifferentiable minimization.” *Journal of Optimization Theory and Applications* **109**, 475–494.

G. Tutz and M. Schmid (2016). *Modeling Discrete Time-to-Event Data*. Springer.

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

M.-C. Wang (1996). “Hazards regression analysis for length-biased data.” *Biometrika* **83**, 343–354.

M. Woodroofe (1985). “Estimating a distribution function with truncated data.” *The Annals of Statistics* **13**, 163–177.

S. J. Wright (2015). “Coordinate descent algorithms.” *Mathematical Programming* **151**, 3–34.

C. F. J. Wu (1983). “On the convergence properties of the EM algorithm.” *Annals of Statistics* **11**, 95–103.

F. Wu, S. Kim, J. Qin, R. Saran and Y. Li (2018). “A pairwise likelihood augmented Cox estimator for left-truncated data.” *Biometrics* **74**, 100–108.

Left-Truncation Regression: Supplemental Material

This is an online companion supplement to the manuscript, *Discrete time-to-event regression analysis under left-truncation with applications to consumer finance.* Please attribute any citations to the original manuscript. This companion includes proofs and derivations of all major results, comments on the applicability of Lemma 1 when covariates are assumed present, additional explanation for the Newton Raphson optimization approach employed, complete details on generalizing Section 2 for right-censoring, additional details on the application in Section 4, and further explanation on why the consumer ABS data likely violates the proportional hazard assumption. For reference, all data and replication code are publicly available at https://github.com/jackson-lautier/left-truncation-regression.

A Proofs

Please see Section 2 for complete statements.

A.1 Proof of Lemma 1

Proof. We will proceed using a proof by contradiction. Without loss of generality, denote the partial derivative with respect to the ith component of φ, $1 \leq i \leq k$,

$$\varphi_i^{(u)}(\boldsymbol{u}, \boldsymbol{v}) = \frac{\partial}{\partial u_i}\varphi(\boldsymbol{u}, \boldsymbol{v}),$$

and assume $\lambda_1^* \neq 0$. Then, for any λ_1 in an interval that contains λ_1^* but not 0, the assumed scaling property of φ in Lemma 1 yields

$$\begin{aligned}\Lambda(\boldsymbol{u}, \boldsymbol{v}, \lambda_1, \lambda_2) &= \varphi(\boldsymbol{u}, \boldsymbol{v}) - \lambda_1\left(1 - \sum_{i=1}^{k} u_i\right) - \lambda_2\left(1 - \sum_{j=1}^{m} v_j\right) \\ &= \varphi(\lambda_1\boldsymbol{u}, \boldsymbol{v}) - \lambda_1\left(1 - \sum_{i=1}^{k} u_i\right) - \lambda_2\left(1 - \sum_{j=1}^{m} v_j\right). \qquad \text{(S.1)}\end{aligned}$$

Therefore, from (S.1) and with the multivariate chain rule (e.g., Rudin, 1976, Theorem 9.15, pg. 214), we have, at the stationary point $\boldsymbol{u}^*$, $\boldsymbol{v}^*$, λ_1^*, and λ_2^*,

$$0 = \frac{\partial}{\partial \lambda_1}\Lambda(\boldsymbol{u}^*, \boldsymbol{v}^*, \lambda_1^*, \lambda_2^*) = \sum_{i=1}^{k} u_i^* \varphi_i^{(u)}(\lambda_1^* \boldsymbol{u}^*, \boldsymbol{v}^*) - 1 + \sum_{i=1}^{k} u_i^*, \tag{S.2}$$

and, for $i = 1, \ldots, k$,

$$0 = \frac{\partial}{\partial u_i}\Lambda(\boldsymbol{u}^*, \boldsymbol{v}^*, \lambda_1^*, \lambda_2^*) = \lambda_1^* \varphi_i^{(u)}(\lambda_1^* \boldsymbol{u}^*, \boldsymbol{v}^*) + \lambda_1^*.$$

Hence,

$$\sum_{i=1}^{k} u_i^* \biggl(\frac{\partial}{\partial u_i}\Lambda(\boldsymbol{u}^*, \boldsymbol{v}^*, \lambda_1^*, \lambda_2^*) \biggr) = \lambda_1^* \Biggl(\sum_{i=1}^{k} u_i^* \bigl(\varphi_i^{(u)}(\lambda_1^* \boldsymbol{u}^*, \boldsymbol{v}^*) + 1 \bigr) \Biggr) = 0.$$

But, from (S.2),

$$\sum_{i=1}^{k} u_i^* \bigl(\varphi_i^{(u)}(\lambda_1^* \boldsymbol{u}^*, \boldsymbol{v}^*) + 1 \bigr) = 1,$$

which implies $\lambda_1^* = 0$, a contradiction. □

A.2 Derivation of Equation (11)

Proof. It is our objective to differentiate $\ell \equiv \ell(\boldsymbol{\Theta} \mid \mathcal{S}_n)$ with respect to $\boldsymbol{\beta}$, per the NR method (Hardin and Hilbe, 2007), to demonstrate

$$\begin{aligned} \boldsymbol{\beta}^{(t+1)} &= \boldsymbol{\beta}^{(t)} - \left(\frac{\partial^2 \ell}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^\top} \right)^{-1} \left(\frac{\partial \ell}{\partial \boldsymbol{\beta}} \right) \\ &= \boldsymbol{\beta}^{(t)} - (\mathbf{Z}^\top \mathbf{A}^{(t)} \mathbf{Z})^{-1} \mathbf{Z}^\top \mathbf{W}^{(t)} \mathbf{J}_n. \end{aligned}$$

Per the notation introduced in (1), we have

$$\frac{\partial f(X_i; \boldsymbol{z}_i, \boldsymbol{\beta})}{\partial \beta_j} = \left(\frac{\partial f_i}{\partial \mu_i} \right) \left(\frac{\partial \mu_i}{\partial \eta_i} \right) \left(\frac{\partial \eta_i}{\partial \beta_j} \right),$$

and

$$\frac{\partial \alpha(\boldsymbol{z_i}, \boldsymbol{\beta}, \boldsymbol{g})}{\partial \beta_j} = \sum_{v=\Delta+1}^{\Delta+m} g_v \Bigg[\sum_{u=v}^{\omega} \left(\frac{\partial f_u}{\partial \mu_i} \right) \left(\frac{\partial \mu_i}{\partial \eta_i} \right) \left(\frac{\partial \eta_i}{\partial \beta_j} \right) \Bigg].$$

Thus, with $\alpha_i \equiv \alpha(\boldsymbol{z_i}, \boldsymbol{\beta}, \boldsymbol{g})$, we obtain via Rudin (1976, Theorem 5.5, pg. 105)

$$\frac{\partial \ell}{\partial \beta_j} = \sum_{i=1}^{n} \frac{q(X_i, \eta_i)}{f(X_i; \boldsymbol{z}_i, \boldsymbol{\beta})} \left(\frac{\partial \eta_i}{\partial \beta_j} \right) - \sum_{i=1}^{n} \frac{1}{\alpha_i} \Bigg\{ \sum_{v=\Delta+1}^{\Delta+m} g_v \Bigg[\sum_{u=v}^{\omega} q(u, \eta_i) \left(\frac{\partial \eta_i}{\partial \beta_j} \right) \Bigg] \Bigg\}.$$

Because,

$$\frac{\partial \eta_i}{\partial \beta_j} = z_{ij},$$

we may write

$$\frac{\partial \ell}{\partial \beta_j} = \sum_{i=1}^{n} \Bigg\{ \frac{q(X_i, \eta_i)}{f(X_i; \boldsymbol{z}_i, \boldsymbol{\beta})} - \frac{1}{\alpha_i} \Bigg(\sum_{v=\Delta+1}^{\Delta+m} g_v \Bigg[\sum_{u=v}^{\omega} q(u, \eta_i) \Bigg] \Bigg) \Bigg\} z_{ij}.$$

Therefore,

$$\frac{\partial \ell}{\partial \boldsymbol{\beta}} = \mathbf{Z}^{\top} \mathbf{W} \mathbf{J}_n. \tag{S.3}$$

Next, we require

$$\begin{aligned} \frac{\partial \ell}{\partial \beta_j \partial \beta_{j^*}} &= \frac{\partial}{\partial \beta_{j^*}} \sum_{i=1}^{n} W_i z_{ij} \\ &= \sum_{i=1}^{n} \Bigg\{ \frac{\partial}{\partial \beta_{j^*}} \frac{q(X_i, \eta_i)}{f(X_i; \boldsymbol{z}_i, \boldsymbol{\beta})} - \frac{\partial}{\partial \beta_{j^*}} \frac{1}{\alpha_i} \Bigg(\sum_{v=\Delta+1}^{\Delta+m} g_v \Bigg[\sum_{u=v}^{\omega} q(u, \eta_i) \Bigg] \Bigg) \Bigg\} z_{ij}. \end{aligned}$$

We work from left to right. By Rudin (1976, Theorem 5.3(c), pg. 104),

$$\frac{\partial}{\partial \beta_{j^*}} \frac{q(X_i, \eta_i)}{f(X_i; \boldsymbol{z}_i, \boldsymbol{\beta})} = \frac{1}{f_i^2} \Bigg[\left(\frac{\partial q_i}{\partial \beta_{j^*}} \right) f_i - q_i \left(\frac{\partial f_i}{\partial \beta_{j^*}} \right) \Bigg].$$

Continuing, with Rudin (1976, Theorem 5.3(b), pg. 104),

$$\frac{\partial q_i}{\partial \beta_{j^*}} = \frac{\partial}{\partial \beta_{j^*}} \left(\frac{\partial f_i}{\partial \mu_i} \right) \left(\frac{\partial \mu_i}{\partial \eta_i} \right) = \Bigg[\left(\frac{\partial^2 f_i}{\partial \mu_i^2} \right) \left(\frac{\partial \mu_i}{\partial \eta_i} \right)^2 + \left(\frac{\partial f_i}{\partial \mu_i} \right) \left(\frac{\partial^2 \mu_i}{\partial \eta_i^2} \right) \Bigg] z_{ij^*}.$$

But,

$$\frac{\partial f_i}{\partial \beta_{j^*}} = \left(\frac{\partial f_i}{\partial \mu_i}\right)\left(\frac{\partial \mu_i}{\partial \eta_i}\right) z_{ij^*},$$

and so,

$$\frac{\partial}{\partial \beta_{j^*}} \frac{q(X_i, \eta_i)}{f(X_i; \boldsymbol{z}_i, \boldsymbol{\beta})} = \left[\frac{s(X_i, \eta_i)}{f(X_i; \boldsymbol{z}_i, \boldsymbol{\beta})} - \left(\frac{q(X_i, \eta_i)}{f(X_i; \boldsymbol{z}_i, \boldsymbol{\beta})}\right)^2\right] z_{ij^*}. \tag{S.4}$$

Proceeding, define $r_i \equiv r(\boldsymbol{g}, \eta_i)$ and use Rudin (1976, Theorem 5.3(c), pg. 104) to write

$$\frac{\partial}{\partial \beta_{j^*}} \frac{1}{\alpha_i}\left(\sum_{v=\Delta+1}^{\Delta+m} g_v\left[\sum_{u=v}^{\omega} q(u, \eta_i)\right]\right) = \frac{1}{\alpha_i^2}\left[\left(\frac{\partial r_i}{\partial \beta_{j^*}}\right)\alpha_i - r_i\left(\frac{\partial \alpha_i}{\partial \beta_{j^*}}\right)\right].$$

Now,

$$\begin{aligned}\frac{\partial}{\partial \beta_{j^*}} r(\boldsymbol{g}, \eta_i) &= \sum_{v=\Delta+1}^{\Delta+m} g_v\left(\sum_{u=v}^{\omega} \frac{\partial}{\partial \beta_{j^*}} q(u, \eta_i)\right) \\ &= \left[\sum_{v=\Delta+1}^{\Delta+m} g_v\left(\sum_{u=v}^{\omega} s(u, \eta_i)\right)\right] z_{ij^*},\end{aligned}$$

and

$$\frac{\partial}{\partial \beta_{j^*}} \alpha_i = \left[\sum_{v=\Delta+1}^{\Delta+m} g_v\left(\sum_{u=v}^{\omega} q(u, \eta_i)\right)\right] z_{ij^*} = r(\boldsymbol{g}, \eta_i) z_{ij^*}.$$

Hence,

$$\frac{\partial}{\partial \beta_{j^*}} \frac{1}{\alpha_i}\left(\sum_{v=\Delta+1}^{\Delta+m} g_v\left[\sum_{u=v}^{\omega} q(u, \eta_i)\right]\right) = \left\{\frac{1}{\alpha_i}\sum_{v=\Delta+1}^{\Delta+m} g_v\left(\sum_{u=v}^{\omega} s(u, \eta_i)\right) - \left(\frac{r(\boldsymbol{g}, \eta_i)}{\alpha_i}\right)^2\right\} z_{ij^*}. \tag{S.5}$$

Therefore, combine (S.4) and (S.5) to obtain

$$\frac{\partial}{\partial \beta_{j^*}} W_i = A_i z_{ij^*} \implies \frac{\partial \ell}{\partial \beta_j \partial \beta_{j^*}} = \sum_{i=1}^{n} A_i z_{ij^*} z_{ij}.$$

That is,

$$\frac{\partial^2 \ell}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^\top} = \mathbf{Z}^\top \mathbf{A} \mathbf{Z},$$

which, along with (S.3), completes the proof. □

A.3 Derivation of Equation (13)

Proof. It is our objective to differentiate $\mathcal{F}(\boldsymbol{g}^*)$ with respect to $\boldsymbol{g}^*$ per the NR method (Hardin and Hilbe, 2007), to demonstrate

$$\begin{aligned}(\boldsymbol{g}^*)^{(t+1)} &= (\boldsymbol{g}^*)^{(t)} - \left(\frac{\partial \mathcal{F}}{\partial \boldsymbol{g}^*}\right)^{-1} \mathcal{F}(\boldsymbol{g}^*) \\ &= (\boldsymbol{g}^*)^{(t)} - ((\mathbf{B}^*)^{(t)})^{-1}((\mathbf{D}^*)^{(t)})^\top \mathbf{J}_n.\end{aligned}$$

It follows through standard matrix operations that $\mathcal{F}(\boldsymbol{g}^*) = (\mathbf{D}^*)^\top \mathbf{J}_n$. Hence, it is left to show

$$\frac{\partial \mathcal{F}}{\partial \boldsymbol{g}^*} = \mathbf{B}^*.$$

Observe, for $\Delta + 1 \leq v, v^* \leq \Delta + m - 1$,

$$\frac{\partial \mathcal{F}(g_v)}{\partial g_{v^*}} = \sum_{i=1}^{n} \frac{\partial}{\partial g_{v^*}} d_i^*(v) = -\sum_{i=1}^{n} \frac{\partial}{\partial g_{v^*}} \frac{g_v S_{vi}}{\alpha_i^*}.$$

Therefore, by Rudin (1976, Theorem 5.3(c), pg. 104) and with

$$\frac{\partial \alpha_i^*}{\partial g_v^*} = S_{v^* i} - S_{(\Delta+m)i}, \tag{S.6}$$

we have

$$-\frac{\partial}{\partial g_{v^*}} \frac{g_v S_{vi}}{\alpha_i^*} = B_i^*(v, v^*),$$

which completes the proof. □

A.4 Proof of Theorem 2.1

Proof. Per Shao (e.g., 2003, pg. 372), under the assumption that the suitable regularity conditions (e.g., van der Vaart, 1998, §5.3, pg. 51; Mukhopadhyay, 2000, §12.2, pg. 539) hold, (14) follows if we can demonstrate

$$-\mathbf{\Sigma}_n(\mathbf{\Theta}^*) = \begin{bmatrix} \mathbf{\Sigma}_{11} & \mathbf{\Sigma}_{12} \\ \mathbf{\Sigma}_{12}^\top & \mathbf{\Sigma}_{22} \end{bmatrix} = \begin{bmatrix} \dfrac{\partial^2 \ell^*}{\partial \boldsymbol{g}^* \partial (\boldsymbol{g}^*)^\top} & \dfrac{\partial \ell^*}{\partial \boldsymbol{g}^* \partial \boldsymbol{\beta}} \\ \left(\dfrac{\partial \ell^*}{\partial \boldsymbol{g}^* \partial \boldsymbol{\beta}}\right)^\top & \dfrac{\partial^2 \ell^*}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^\top} \end{bmatrix}.$$

We proceed to do so by working through each of the four blocks of $-\mathbf{\Sigma}_n$. Observe first,

$$\begin{aligned} \frac{\partial \alpha^*}{\partial \beta_j} &= \frac{\partial}{\partial \beta_j}\left\{ \sum_{v=\Delta+1}^{\Delta+m-1} g_v S_{vi} + \left(1 - \sum_{j=\Delta+1}^{\Delta+m-1} g_v\right) S_{(\Delta+m)i} \right\} \\ &= \left(\sum_{v=\Delta+1}^{\Delta+m-1} g_v \left(\sum_{u=v}^{\omega} q(u,\eta_i)\right) + \left(1 - \sum_{k=\Delta+1}^{\Delta+m-1} g_k\right) \sum_{u=\Delta+m}^{\omega} q(u,\eta_i)\right) z_{ij}, \end{aligned}$$

by (S.6) and

$$\frac{\partial}{\partial \beta_j} S_{vi} = \sum_{u=v}^{\omega} q(u,\eta_i) z_{ij}. \tag{S.7}$$

Thus,

$$\frac{\partial \ell^*}{\partial g_v} = \sum_{i=1}^{n} \left(\frac{\mathbf{1}(Y_i = v)}{g_v} - \frac{\mathbf{1}(Y_i = \Delta + m)}{1 - \sum_{k=\Delta+1}^{\Delta+m-1} g_v} - \frac{S_{vi} - S_{(\Delta+m)i}}{\alpha_i^*} \right),$$

and

$$\frac{\partial \ell^*}{\partial \boldsymbol{\beta}} = \mathbf{Z}^\top \mathbf{W}^* \mathbf{J}_n,$$

where $\mathbf{W}^* = \text{diag}(W_i^*)$ is an $n \times n$ dimension matrix with diagonal elements

$$\begin{aligned} W_i^* &= \frac{q(X_i,\eta_i)}{f(X_i;\boldsymbol{z}_i,\boldsymbol{\beta})} - \frac{1}{\alpha_i^*}\left(\sum_{v=\Delta+1}^{\Delta+m-1} g_v \left(\sum_{u=v}^{\omega} q(u,\eta_i)\right) + \left(1 - \sum_{k=\Delta+1}^{\Delta+m-1} g_k\right) \sum_{u=\Delta+m}^{\omega} q(u,\eta_i)\right) \\ &= \frac{q(X_i,\eta_i)}{f(X_i;\boldsymbol{z}_i,\boldsymbol{\beta})} - \frac{r^*(\boldsymbol{g}^*,\eta_i)}{\alpha_i^*}. \end{aligned}$$

Hence,

$$\frac{\partial^2 \ell^*}{\partial g_v \partial g_{v^*}} = \frac{\partial}{\partial g_{v^*}} \sum_{i=1}^{n} \left(\frac{\mathbf{1}(Y_i = v)}{g_v} - \frac{\mathbf{1}(Y_i = \Delta + m)}{1 - \sum_{k=\Delta+1}^{\Delta+m-1} g_v} - \frac{S_{vi} - S_{(\Delta+m)i}}{\alpha_i^*} \right) = \sum_{i=1}^{n} \tilde{B}_i^*,$$

and so,

$$\frac{\partial^2 \ell^*}{\partial \boldsymbol{g}^* \partial (\boldsymbol{g}^*)^\top} = \mathbf{\Sigma}_{11}.$$

Next, for ease of exposition, we proceed to the lower right quadrant of $-\mathbf{\Sigma}_n$. That is, we require

$$\frac{\partial^2 \ell^*}{\partial \beta_j \partial \beta_{j^*}} = \frac{\partial}{\partial \beta_{j^*}} \sum_{i=1}^{n} W_i^* z_{ij} = \sum_{i=1}^{n} \left(\frac{\partial}{\partial \beta_{j^*}} W_i^* \right) z_{ij}.$$

By (S.4), we obtain

$$\frac{\partial}{\partial \beta_{j^*}} W_i^* = \left[\frac{s(X_i, \eta_i)}{f(X_i; \boldsymbol{z}_i, \boldsymbol{\beta})} - \left(\frac{q(X_i, \eta_i)}{f(X_i; \boldsymbol{z}_i, \boldsymbol{\beta})} \right)^2 \right] z_{ij^*} - \frac{\partial}{\partial \beta_{j^*}} \left(\frac{r^*(\boldsymbol{g}_*, \eta_i)}{\alpha_i^*} \right). \tag{S.8}$$

Now,

$$\frac{\partial}{\partial \beta_{j^*}} r^*(\boldsymbol{g}_*, \eta_i) = \left(\sum_{v=\Delta+1}^{\Delta+m-1} g_v \left(\sum_{u=v}^{\omega} s(u, \eta_i) \right) + \left(1 - \sum_{k=\Delta+1}^{\Delta+m-1} g_k \right) \sum_{u=\Delta+m}^{\omega} s(u, \eta_i) \right) z_{ij^*},$$

and

$$\frac{\partial}{\partial \beta_{j^*}} \frac{1}{\alpha_i^*} = -\frac{r^*(\boldsymbol{g}_*, \eta_i)}{(\alpha_i^*)^2} z_{ij^*}. \tag{S.9}$$

Therefore, by Rudin (1976, Theorem 5.3(b), pg. 104),

$$\begin{aligned} \frac{\partial}{\partial \beta_{j^*}} \left(\frac{r^*(\boldsymbol{g}_*, \eta_i)}{\alpha_i^*} \right) = & - \frac{r^*(\boldsymbol{g}_*, \eta_i)^2}{(\alpha_i^*)^2} z_{ij^*} \\ & + \frac{1}{\alpha_i^*} \left(\sum_{v=\Delta+1}^{\Delta+m-1} g_v \left(\sum_{u=v}^{\omega} s(u, \eta_i) \right) + \left(1 - \sum_{k=\Delta+1}^{\Delta+m-1} g_k \right) \sum_{u=\Delta+m}^{\omega} s(u, \eta_i) \right) z_{ij^*}. \end{aligned}$$

Hence, returning to (S.8), we obtain

$$\frac{\partial^2 \ell^*}{\partial \beta_j \partial \beta_{j^*}} = \sum_{i=1}^{n} \left(\frac{\partial}{\partial \beta_{j^*}} W_i^* \right) z_{ij} = \sum_{i=1}^{n} A_i^* z_{ij^*} z_{ij}.$$

That is,

$$\frac{\partial^2 \ell^*}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^\top} = \mathbf{Z}^\top (\mathbf{A}^*) \mathbf{Z} = \mathbf{\Sigma}_{22}.$$

For the off-diagonals of $\mathbf{\Sigma}_n$, it is left to calculate

$$\begin{aligned}\frac{\partial \ell^*}{\partial g_v \partial \beta_j} &= \frac{\partial}{\partial \beta_j}\sum_{i=1}^{n}\left(\frac{\mathbf{1}(Y_i = v)}{g_v} - \frac{\mathbf{1}(Y_i = \Delta + m)}{1 - \sum_{k=\Delta+1}^{\Delta+m-1} g_v} - \frac{S_{vi} - S_{(\Delta+m)i}}{\alpha_i^*}\right)\\ &= -\sum_{i=1}^{n}\frac{\partial}{\partial \beta_j}\left(\frac{S_{vi} - S_{(\Delta+m)i}}{\alpha_i^*}\right).\end{aligned}$$

By (S.7), (S.9), and Rudin (1976, Theorem 5.3(b), pg. 104), we obtain

$$\frac{\partial \ell^*}{\partial \boldsymbol{g}^* \partial \boldsymbol{\beta}} = \begin{bmatrix}\mathbf{J}_n^\top \tilde{\mathbf{D}}_{\Delta+1}^* \mathbf{Z} & \mathbf{J}_n^\top \tilde{\mathbf{D}}_{\Delta+2}^* \mathbf{Z} & \dots & \mathbf{J}_n^\top \tilde{\mathbf{D}}_{\Delta+m-1}^* \mathbf{Z}\end{bmatrix}^\top = \mathbf{\Sigma}_{12}.$$

The proof is then complete by the symmetry of the second derivatives (e.g., Rudin, 1976, Theorem 9.41, pg. 235-6). □

A.5 Proof of Theorem 2.2

Proof. With the regularity conditions of Lehmann and Romano (2006, Theorem 12.4.2, pg. 515) assumed satisfied, it is only necessary to demonstrate $\mathbf{\Theta}_0$ is a composite parameter space of $\mathbf{\Theta}$. This is done most directly through using Lehmann and Romano (2006, (12.84), pg. 515). That is,

$$\mathbf{\Theta}_0 = \{\mathbf{\Theta} = (\boldsymbol{g}, \beta_0, \beta_1, \dots, \beta_k)^\top : \mathbb{A}\mathbf{\Theta} = \mathbf{0}_k\},$$

where $\mathbb{A}$ is the dimension $k \times (k+1+m)$ matrix,

$$\mathbb{A} = \begin{bmatrix}\mathbf{0}_{k\times(m+1)} & \mathbf{I}_k\end{bmatrix}.$$

For completeness, we also remark that the estimator $\hat{p}$ from Lautier et al. (2025) is equivalent in efficiency (under the terminology of Lehmann and Romano (2006, Theorem 12.4.2, pg. 515)) to $\hat{\beta} = \mu^{-1}(\hat{p})$ by the invariance property of the maximum likelihood estimate (e.g., Mukhopadhyay, 2000, Theorem 7.2.1, pg. 250). □

B Lemma 1 Single Constraint

Lemma 1, as stated, is a more general case than required in Section 2. Specifically, the optimization presented in (4) has a single restriction on $\boldsymbol{g}$, whereas Lemma 1 assumes two restrictions in line with Lautier et al. (2026). For completeness and clarity, therefore, we present Lemma B.1, which has a clearer connection to the methodological developments in the main manuscript.

Lemma B.1. *Denote* $\boldsymbol{u} = (u_1, \ldots, u_k)^\top$ *and let* $\varphi(\boldsymbol{u}) : \mathbb{R}^k \mapsto \mathbb{R}$ *be a smooth function with domain (i)* $u_i \geq 0$ *or* $u_i \leq 0$ *for all* $1 \leq i \leq k$ *and (ii)* $\sum_i u_i \neq 0$*. Further assume that for any nonzero scalar,* π_1*,* $\varphi(\pi_1 \boldsymbol{u}) = \varphi(\boldsymbol{u})$ *and define the following Lagrangian-type function,* Λ*, with constraints,* $\sum_i u_i = 1$ *(e.g., Ravishanker and Dey, 2002, §2.9, pg. 69),*

$$\Lambda(\boldsymbol{u}, \lambda_1) = \varphi(\boldsymbol{u}) - \lambda_1 \left(1 - \sum_{i=1}^{k} u_i\right),$$

with $\lambda_1 \in \mathbb{R}$*. If* $\boldsymbol{u}^*$ *and* λ_1^* *is a stationary point of* Λ*, then* $\lambda_1^* = 0$*.*

Proof. Lemma B.1 is a special case of Lemma 1 where $\varphi(\boldsymbol{u}, \boldsymbol{g}) = \varphi(\boldsymbol{u})$, for all $\boldsymbol{g}$, and $\lambda_2 = 0$. □

Because $f(\cdot;\cdot,\boldsymbol{\beta})$ takes the form of (1), h_* will remain a smooth function that meets the domain and scaling property requirements of Lemma B.1 for all $\boldsymbol{\beta}$.

C Newton Raphson Optimization Comments

We demonstrate the Newton Raphson method will fail to converge for the optimization function (12). Define, for $v \in \mathcal{V}$,

$$d_i(v) = \frac{\mathbf{1}(Y_i = v)}{g_v} - \frac{1}{\alpha_i}\left(\sum_{u=v}^{\omega} f(u; \boldsymbol{z}_i, \boldsymbol{\beta})\right),$$

and the $n \times m$ dimension matrix,

$$\mathbf{D} = \begin{bmatrix} d_1(\Delta+1) & d_1(\Delta+2) & \dots & d_1(\Delta+m) \\ d_2(\Delta+1) & d_2(\Delta+2) & \dots & d_2(\Delta+m) \\ \vdots & \vdots & & \vdots \\ d_n(\Delta+1) & d_n(\Delta+2) & \dots & d_n(\Delta+m) \end{bmatrix}.$$

Then,

$$\frac{\partial \ell}{\partial g_v} \equiv \ell'(\boldsymbol{g}) = \mathbf{D}^\top \mathbf{J}_n.$$

Furthermore,

$$\begin{aligned} \frac{\partial^2 \ell}{\partial g_v \partial g_{v^*}} &= \frac{\partial}{\partial g_{v^*}} \sum_{i=1}^{n} d_i(v) \\ &= \frac{\partial}{\partial g_{v^*}} \sum_{i=1}^{n} \left\{ \frac{\mathbf{1}(Y_i = v)}{g_v} - \frac{1}{\alpha_i} \left(\sum_{u=v}^{\omega} f(u; \boldsymbol{z}_i, \boldsymbol{\beta}) \right) \right\} \\ &= \sum_{i=1}^{n} \left\{ \frac{\partial}{\partial g_{v^*}} \frac{\mathbf{1}(Y_i = v)}{g_v} - \left(\sum_{u=v}^{\omega} f(u; \boldsymbol{z}_i, \boldsymbol{\beta}) \right) \frac{1}{\alpha_i^2} \frac{\partial}{\partial g_{v^*}} \alpha_i \right\} \\ &= \sum_{i=1}^{n} \left\{ \frac{\partial}{\partial g_{v^*}} \frac{\mathbf{1}(Y_i = v)}{g_v} - \left(\sum_{u=v}^{\omega} f(u; \boldsymbol{z}_i, \boldsymbol{\beta}) \right) \frac{1}{\alpha_i^2} \left(\sum_{u=v^*}^{\omega} f(u; \boldsymbol{z}_i, \boldsymbol{\beta}) \right) \right\}. \end{aligned}$$

Thus, if we define

$$B_i(v, v^*) = \begin{cases} -\dfrac{\mathbf{1}(Y_i = v)}{g_v^2} + \dfrac{1}{\alpha_i^2} \left(\displaystyle\sum_{u=v}^{\omega} f(u; \boldsymbol{z}_i, \boldsymbol{\beta}) \right)^2, & v = v^* \\ \left(\displaystyle\sum_{u=v}^{\omega} f(u; \boldsymbol{z}_i, \boldsymbol{\beta}) \right) \dfrac{1}{\alpha_i^2} \left(\displaystyle\sum_{u=v^*}^{\omega} f(u; \boldsymbol{z}_i, \boldsymbol{\beta}) \right), & v \neq v^* \end{cases},$$

and the $n \times m$ dimension matrix,

$$\mathbf{B} = \sum_{i=1}^{n} \begin{bmatrix} B_i(\Delta+1,\Delta+1) & B_i(\Delta+1,\Delta+2) & \dots & B_i(\Delta+1,\Delta+m) \\ B_i(\Delta+2,\Delta+1) & B_i(\Delta+2,\Delta+2) & \dots & \\ \vdots & \vdots & \ddots & \\ B_i(\Delta+m,\Delta+1) & \dots & & B_i(\Delta+m,\Delta+m) \end{bmatrix},$$

then

$$\frac{\partial^2 \ell}{\partial \boldsymbol{g} \partial \boldsymbol{g}^\top} \equiv \ell''(\boldsymbol{g}) = \mathbf{B}.$$

Therefore, the update step of the Newton Raphson method is

$$\begin{aligned} \boldsymbol{g}^{(t+1)} &= \boldsymbol{g}^{(t)} - (\ell''(\boldsymbol{g}^{(t)}))^{-1} \ell'(\boldsymbol{g}^{(t)}) \\ &= \boldsymbol{g}^{(t)} - (\mathbf{B}^{(t)})^{-1} \mathbf{D}^{(t)} \mathbf{J}_n. \end{aligned} \tag{S.10}$$

But, if we set $\Delta = 0$ without loss of generality and let $v \in \mathcal{V}$,

$$\begin{aligned} \sum_{k=1}^{m} B_i(v,k) g_k &= B_i(v,1)g_1 + B_i(v,2)g_2 + \cdots + B_i(v,m)g_m \\ &= \frac{S_i(v)S_i(1)g_1}{\alpha_i^2} + \cdots + \left(-\frac{\mathbf{1}(Y_i = v)}{g_v^2} + \frac{S_i(v)^2}{\alpha_i^2} \right) g_v + \cdots + \frac{S_i(v)S_i(m)g_m}{\alpha_i^2} \\ &= -\frac{\mathbf{1}(Y_i = v)}{g_v} + \frac{S_i(v)S_i(1)g_1}{\alpha_i^2} + \cdots + \frac{S_i(v)S_i(v)g_v}{\alpha_i^2} + \cdots + \frac{S_i(v)S_i(m)g_m}{\alpha_i^2} \\ &= -\frac{\mathbf{1}(Y_i = v)}{g_v} + \frac{S_i(v)}{\alpha_i} \left(\frac{S_i(1)g_1 + S_i(2)g_2 + \cdots + S_i(m)g_m}{\alpha_i} \right) \\ &= -\frac{\mathbf{1}(Y_i = v)}{g_v} + \frac{S_i(v)}{\alpha_i} \frac{\alpha_i}{\alpha_i} \\ &= -d_i(v), \end{aligned}$$

where

$$S_i(v) = \sum_{u=v}^{\omega} f(u; \boldsymbol{z}_i, \boldsymbol{\beta}).$$

This implies

$$\mathbf{D}^\top \mathbf{J}_n = -\mathbf{B}\boldsymbol{g},$$

and, therefore, convergence of (S.10) is not possible because

$$\boldsymbol{g}^{(t+1)} = \boldsymbol{g}^{(t)} + (\mathbf{B}^{(t)})^{-1}\mathbf{B}^{(t)}\boldsymbol{g}^{(t)} = 2\boldsymbol{g}^{(t)}.$$

D Right-Censoring Additional Details

The purpose of the present section is to provide all relevant details to generalize the results of Sections 2.1 and 2.2 to also handle right-censoring. We will employ the subscript τ to indicate that right-censoring is assumed to be present in the data. Depending on the censoring rate, the observational support of X becomes $\{\Delta+1,\dots,\xi\}$, where $\xi = \min(\omega, \varepsilon - 1)$ (Lautier et al., 2023). To proceed to estimation, observe that when $D_i = 0$, $\bar{h}_*$ defined in (15) is a smooth function that meets the domain and scaling property requirements of Lemma 1 (i.e., Lemma B.1). As in Section 2.1, therefore, it is sufficient to locate stationary points of the unconstrained optimization problem. This allows us to proceed with the objective function,

$$\begin{aligned}\log \mathcal{L}_\tau(\boldsymbol{\Theta} \mid \mathcal{S}_{\tau,n}) \equiv \ell_\tau(\boldsymbol{\Theta} \mid \mathcal{S}_{\tau,n}) \equiv \ell_\tau & \\ &= \sum_{v=\Delta+1}^{\Delta+m} \hat{\gamma}_n(v)\log g_v + \sum_{\{i:D_i=1\}} \log f(T_i;\boldsymbol{z}_i,\boldsymbol{\beta}) \\ &+ \sum_{\{i:D_i=0\}} \log S(T_i+1;\boldsymbol{z}_i,\boldsymbol{\beta}) - \sum_{i=1}^{n}\log \alpha(\boldsymbol{z}_i,\boldsymbol{\beta},\boldsymbol{g}). \qquad \text{(S.11)}\end{aligned}$$

To find $\arg\max_{\boldsymbol{\Theta}} \ell_\tau$, we may again use Algorithm 1 but with modified implementation details to reflect the differences between (S.11) and (7). Specifically, we need to modify the initial values, (8) and (9), and the two update steps, (11) and (13).

Per Lautier et al. (2025, Theorem 3.1), for a given $f(\cdot;p)$, let $\hat{p}_{\tau,\texttt{INIT}}$ be any $p \in \mathcal{P}$ such that

$$\sum_{v=\Delta+1}^{\Delta+m}\left(\frac{\hat{\gamma}_n(v)}{\sum_{u=v}^{\xi} f(u;p)}\right)\left(\sum_{u=v}^{\xi}\frac{\partial}{\partial p}f(u;p)\right)$$
$$=\sum_{i=1}^{n}\left(\frac{D_i}{f(T_i;p)}\frac{\partial}{\partial p}f(T_i;p)+\frac{1-D_i}{S(T_i+1;p)}\frac{\partial}{\partial p}S(T_i+1;p)\right), \qquad \text{(S.12)}$$

and define

$$\hat{g}_{\tau,v,\texttt{INIT}}=\frac{\hat{\gamma}_n(v)}{S(v;\hat{p}_{\tau,\texttt{INIT}})}\left[\sum_{k=\Delta+1}^{\Delta+m}\frac{\hat{\gamma}_n(k)}{S(k;\hat{p}_{\tau,\texttt{INIT}})}\right]^{-1}, \quad v\in\mathcal{V}. \qquad \text{(S.13)}$$

As with (8) and (9), these estimators may be generalized for vector-valued parametric forms of f, such as the two-parameter discrete Weibull of Nakagawa and Osaki (1975). See Lautier et al. (2025) for details, including approaches and replication code to solve (S.12).

We now proceed to modify (11). Per the NR method (Hardin and Hilbe, 2007), we seek

$$\boldsymbol{\beta}_{\tau}^{(t+1)}=\boldsymbol{\beta}_{\tau}^{(t)}-\left(\frac{\partial^2\ell_\tau}{\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}^\top}\right)^{-1}\left(\frac{\partial\ell_\tau}{\partial\boldsymbol{\beta}}\right).$$

To begin, observe

$$\frac{\partial\ell_\tau}{\partial\beta_j}=\sum_{i=1}^{n}\left\{\frac{\mathbf{1}(D_i=1)q(T_i,\eta_i)}{f(T_i;\boldsymbol{z}_i,\boldsymbol{\beta})}\right.$$
$$\left.+\frac{\mathbf{1}(D_i=0)\sum_{u=T_i+1}^{\xi}q(u,\eta_i)}{S(T_i+1;\boldsymbol{z}_i,\boldsymbol{\beta})}-\frac{1}{\alpha_i}\left(\sum_{v=\Delta+1}^{\Delta+m}g_v\left[\sum_{u=v}^{\xi}q(u,\eta_i)\right]\right)\right\}z_{ij}. \qquad \text{(S.14)}$$

Remark. *The indicator convention in* (S.14) *follows for ease of exposition. That is, it is understood in* (S.14) *and all following statements that* $q(T_i,\eta_i)/f(T_i;\boldsymbol{z}_i,\boldsymbol{\beta})$ *appears only when* $D_i=1$ *and* $\sum_{u=T_i+1}^{\xi}q(u,\eta_i)/S(T_i+1;\boldsymbol{z}_i,\boldsymbol{\beta})$ *appears only when* $D_i=0$. *This avoids any complications when* $D_i=1$ *and* $S(\cdot;\boldsymbol{z}_i,\boldsymbol{\beta})=0$.

Therefore,

$$\frac{\partial\ell_\tau}{\partial\boldsymbol{\beta}}=\mathbf{Z}^\top\mathbf{W}_\tau\mathbf{J}_n,$$

where $\mathbf{W}_\tau = \mathrm{diag}(W_{\tau,i})$ is an $n \times n$ diagonal matrix with diagonal elements,

$$W_{\tau,i} = \frac{\mathbf{1}(D_i = 1) q(T_i, \eta_i)}{f(T_i; \boldsymbol{z}_i, \boldsymbol{\beta})} + \frac{\mathbf{1}(D_i = 0) \sum_{u=T_i+1}^{\xi} q(u, \eta_i)}{S(T_i + 1; \boldsymbol{z}_i, \boldsymbol{\beta})} - \frac{1}{\alpha_i}\Bigg(\sum_{v=\Delta+1}^{\Delta+m} g_v \Bigg[\sum_{u=v}^{\xi} q(u, \eta_i) \Bigg] \Bigg),$$

for $1 \leq i \leq n$. Next, we require

$$\frac{\partial \ell_\tau}{\partial \beta_j \partial \beta_{j^*}} = \frac{\partial \ell_\tau}{\partial \beta_{j^*}} \sum_{i=1}^{n} W_{\tau,i} z_{ij}.$$

By Rudin (1976, Theorem 5.3(c), pg. 104), we may write

$$\frac{\partial \ell_\tau}{\partial \beta_{j^*}} \frac{\sum_{u=T_i+1}^{\xi} q(u, \eta_i)}{S(T_i + 1; \boldsymbol{z}_i, \boldsymbol{\beta})} = \Bigg[\frac{\sum_{u=T_i+1}^{\xi} s(u, \eta_i)}{S(T_i + 1; \boldsymbol{z}_i, \boldsymbol{\beta})} - \Bigg(\frac{\sum_{u=T_i+1}^{\xi} q(u, \eta_i)}{S(T_i + 1; \boldsymbol{z}_i, \boldsymbol{\beta})} \Bigg)^2 \Bigg] z_{ij^*},$$

which, along with (S.4) and (S.5), yields

$$\frac{\partial \ell_\tau}{\partial \beta_j \partial \beta_{j^*}} = \mathbf{Z}^\top \mathbf{A}_\tau \mathbf{Z},$$

where $\mathbf{A}_\tau = \mathrm{diag}(A_{\tau,i})$ is an $n \times n$ diagonal matrix with diagonal elements,

$$\begin{aligned} A_{\tau,i} = \Bigg\{ & \mathbf{1}(D_i = 1) \Bigg[\frac{s(T_i, \eta_i)}{f(T_i; \boldsymbol{z}_i, \boldsymbol{\beta})} - \Bigg(\frac{q(T_i, \eta_i)}{f(T_i; \boldsymbol{z}_i, \boldsymbol{\beta})} \Bigg)^2 \Bigg] \\ & + \mathbf{1}(D_i = 0) \Bigg[\frac{\sum_{u=T_i+1}^{\xi} s(u, \eta_i)}{S(T_i + 1; \boldsymbol{z}_i, \boldsymbol{\beta})} - \Bigg(\frac{\sum_{u=T_i+1}^{\xi} q(u, \eta_i)}{S(T_i + 1; \boldsymbol{z}_i, \boldsymbol{\beta})} \Bigg)^2 \Bigg] \\ & - \frac{1}{\alpha_i} \sum_{v=\Delta+1}^{\Delta+m} g_v \Bigg(\sum_{u=v}^{\xi} s(u, \eta_i) \Bigg) + \Bigg(\frac{r(\boldsymbol{g}, \eta_i)}{\alpha_i} \Bigg)^2 \Bigg\}. \end{aligned}$$

Thus, the modified update step to replace (11) becomes

$$\boldsymbol{\beta}_\tau^{(t+1)} = \boldsymbol{\beta}_\tau^{(t)} - (\mathbf{Z}^\top \mathbf{A}_\tau^{(t)} \mathbf{Z})^{-1} \mathbf{Z}^\top \mathbf{W}_\tau^{(t)} \mathbf{J}_n. \tag{S.15}$$

For the modifications to (13), it is sufficient to observe

$$\frac{\partial}{\partial g_v}\ell_\tau(\boldsymbol{\Theta}) = \frac{\partial}{\partial g_v}\ell(\boldsymbol{\Theta}).$$

In other words, no modifications are necessary for (13). This produces a complete modification to Algorithm 1 for administrative right-censoring. We state it now for completeness.

Algorithm D.1. *The following approach may be used to find* $\arg\max_{\boldsymbol{\Theta}} \ell_\tau(\boldsymbol{\Theta} \mid \mathcal{S}_{\tau,n})$ *in* (S.11).

1. *For a given* $f(\cdot;p)$*, use* (S.12) *to find* $\hat{p}_{\tau,INIT}$ *and* (S.13) *to find* $\hat{\boldsymbol{g}}_{\tau,INIT}$*.*
2. *For a given link function,* μ*, assign* $\hat{\beta}_{0,\tau,INIT} = \mu^{-1}(\hat{p}_{\tau,INIT})$ *and* $\hat{\boldsymbol{\beta}}_{\tau,INIT} = (\hat{\beta}_{0,\tau,INIT}, \mathbf{0}_k)^\top$*.*
3. *Fix* $\hat{\boldsymbol{g}}_{\tau,INIT}$ *and use the NR method per* (S.15) *to maximize* $\ell_\tau(\boldsymbol{\beta} \mid \hat{\boldsymbol{g}}_{\tau,INIT}, \mathcal{S}_{\tau,n})$ *starting from* $\hat{\boldsymbol{\beta}}_{\tau,INIT}$*. Call this block extrema* $\hat{\boldsymbol{\beta}}_\tau^t$*.*
4. *Fix* $\hat{\boldsymbol{\beta}}_\tau^t$ *and use the NR method per* (13) *to maximize* $\ell_\tau(\boldsymbol{g} \mid \hat{\boldsymbol{\beta}}_\tau^t, \mathcal{S}_{\tau,n})$ *starting from* $\hat{\boldsymbol{g}}_{\tau,INIT}$*. Call this block extrema* $\hat{\boldsymbol{g}}_\tau^t$*.*
5. *Return to Step 3 with* $(\hat{\boldsymbol{\beta}}_\tau^t, \hat{\boldsymbol{g}}_\tau^t)^\top$ *replacing* $(\hat{\boldsymbol{\beta}}_{\tau,INIT}, \hat{\boldsymbol{g}}_{\tau,INIT})^\top$*. Repeat Steps 3 and 4 to find* $\hat{\boldsymbol{\Theta}}_\tau^{t+1} = (\hat{\boldsymbol{\beta}}_\tau^{t+1}, \hat{\boldsymbol{g}}_\tau^{t+1})^\top$ *and evaluate* $\left\|\nabla\ell_\tau(\hat{\boldsymbol{\Theta}}_\tau^{t+1})\right\| < \epsilon$*, where* ϵ *is a small, predetermined tolerance, such as* $\epsilon = 10^{-7}$*.*
6. *If Step 5 is true, exit the procedure and set* $\hat{\boldsymbol{\Theta}}_{\tau,MLE} = (\hat{\boldsymbol{\beta}}_\tau^{t+1}, \hat{\boldsymbol{g}}_\tau^{t+1})^\top$*. If Step 5 is false, repeat Step 5 starting with* $(\hat{\boldsymbol{\beta}}_\tau^{t+1}, \hat{\boldsymbol{g}}_\tau^{t+1})^\top$*.*

For variable selection, it is necessary to modify Theorem 2.1. This begins by first updating (S.11) for the vector of free parameters. Denote $\ell_\tau(\boldsymbol{\Theta}^* \mid \mathcal{S}_{\tau,n}) \equiv \ell_\tau^*$ and write

$$\ell_\tau^* = \sum_{v=\Delta+1}^{\Delta+m-1} \hat{\gamma}_n(v)\log g_v + \hat{\gamma}_n(\Delta+m)\log\left(1 - \sum_{k=\Delta+1}^{\Delta+m-1} g_k\right)$$

$$+ \sum_{\{i:D_i=1\}} \log f(T_i; \boldsymbol{z}_i, \boldsymbol{\beta}) + \sum_{\{i:D_i=0\}} \log S(T_i+1; \boldsymbol{z}_i, \boldsymbol{\beta}) - \sum_{i=1}^{n} \log \alpha^*(\boldsymbol{z}_i, \boldsymbol{\beta}, \boldsymbol{g}).$$

Because

$$\frac{\partial \ell_\tau^*}{\partial g_v} = \frac{\partial \ell^*}{\partial g_v},$$

we have

$$\begin{bmatrix} \dfrac{\partial^2 \ell_\tau^*}{\partial \boldsymbol{g}^* \partial (\boldsymbol{g}^*)^\top} & \dfrac{\partial \ell_\tau^*}{\partial \boldsymbol{g}^* \partial \boldsymbol{\beta}} \\ \left(\dfrac{\partial \ell_\tau^*}{\partial \boldsymbol{g}^* \partial \boldsymbol{\beta}}\right)^\top & \dfrac{\partial^2 \ell_\tau^*}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^\top} \end{bmatrix} = \begin{bmatrix} \dfrac{\partial^2 \ell^*}{\partial \boldsymbol{g}^* \partial (\boldsymbol{g}^*)^\top} & \dfrac{\partial \ell^*}{\partial \boldsymbol{g}^* \partial \boldsymbol{\beta}} \\ \left(\dfrac{\partial \ell^*}{\partial \boldsymbol{g}^* \partial \boldsymbol{\beta}}\right)^\top & \dfrac{\partial^2 \ell_\tau^*}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^\top} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{12}^\top & \dfrac{\partial^2 \ell_\tau^*}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^\top} \end{bmatrix}.$$

In other words, we need only modify the lower right quadrant of the Hessian matrix. Because

$$\frac{\partial \ell_\tau^*}{\partial \boldsymbol{\beta}} = \mathbf{Z}^\top \mathbf{W}_\tau^* \mathbf{J}_n,$$

where $\mathbf{W}_\tau^* = \text{diag}(W_{\tau,i}^*)$ is an $n \times n$ diagonal matrix with diagonal elements,

$$W_{\tau,i}^* = \mathbf{1}(D_i = 1)\frac{q(T_i, \eta_i)}{f(T_i; \boldsymbol{z}_i, \boldsymbol{\beta})} + \mathbf{1}(D_i = 0)\frac{\sum_{u=T_i+1}^{\xi} q(u, \eta_i)}{S(T_i+1; \boldsymbol{z}_i, \boldsymbol{\beta})} - \frac{r^*(\boldsymbol{g}^*, \eta_i)}{\alpha_i^*},$$

it follows that

$$\boldsymbol{\Sigma}_{\tau,22} \equiv \frac{\partial^2 \ell_\tau^*}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^\top} = \mathbf{Z}^\top \mathbf{A}_\tau^* \mathbf{Z},$$

where $\mathbf{A}_\tau^* = \text{diag}(A_{\tau,i}^*)$ is an $n \times n$ diagonal matrix with diagonal elements,

$$\begin{aligned} A_{\tau,i}^* = {} & \mathbf{1}(D_i = 1)\left[\frac{s(T_i, \eta_i)}{f(T_i; \boldsymbol{z}_i, \boldsymbol{\beta})} - \left(\frac{q(T_i, \eta_i)}{f(T_i; \boldsymbol{z}_i, \boldsymbol{\beta})}\right)^2\right] \\ & + \mathbf{1}(D_i = 0)\left[\frac{\sum_{u=T_i+1}^{\xi} s(u, \eta_i)}{S(T_i+1; \boldsymbol{z}_i, \boldsymbol{\beta})} - \left(\frac{\sum_{u=T_i+1}^{\xi} q(u, \eta_i)}{S(T_i+1; \boldsymbol{z}_i, \boldsymbol{\beta})}\right)^2\right] \\ & + \frac{r^*(\boldsymbol{g}_*, \eta_i)^2}{(\alpha_i^*)^2} - \frac{1}{\alpha_i^*}\left(\sum_{v=\Delta+1}^{\Delta+m-1} g_v \left(\sum_{u=v}^{\xi} s(u, \eta_i)\right) + \left(1 - \sum_{k=\Delta+1}^{\Delta+m-1} g_k\right) \sum_{u=\Delta+m}^{\xi} s(u, \eta_i)\right). \end{aligned}$$

Hence, the modified Theorem 2.1 is as follows. The proof closely follows Section A.4 and is therefore omitted for brevity.

Theorem D.1. *Let* $\hat{\mathbf{\Theta}}^*_{\tau,\text{MLE}} = \hat{\mathbf{\Theta}}_{\tau,\text{MLE}} \setminus \hat{g}_{\tau,\Delta+m}$ *and assume the suitable regularity conditions (e.g., van der Vaart, 1998, §5.3, pg. 51; Mukhopadhyay, 2000, §12.2, pg. 539) hold. Then,*

$$\mathbf{\Sigma}^{1/2}_{\tau,n}(\hat{\mathbf{\Theta}}^*_{\tau,\text{MLE}})\sqrt{n}\big(\hat{\mathbf{\Theta}}^*_{\tau,\text{MLE}} - \mathbf{\Theta}^*\big) \longrightarrow_d \mathcal{N}(\mathbf{0}, \mathbf{I}_{m+k}),$$

where

$$\mathbf{\Sigma}_{\tau,n}(\hat{\mathbf{\Theta}}^*_{\tau,\text{MLE}}) = \begin{bmatrix} \mathbf{\Sigma}_{11} & \mathbf{\Sigma}_{12} \\ \mathbf{\Sigma}_{12}^\top & \mathbf{\Sigma}_{\tau,22} \end{bmatrix}.$$

The modified Corollary 2.1.1 then follows.

Corollary D.1.1. *Suppose we wish to test* $H_0 : \beta_j = 0$*, for* $0 \leq j \leq k$*. Then, under* H_0*, the Wald test statistic (e.g., Kutner et al., 2005, pg. 578),*

$$\frac{\hat{\beta}_{\tau,j}}{\hat{\sigma}_{\tau,m+j}} \longrightarrow_d N(0,1),$$

where $\hat{\sigma}_{\tau,m+j}$ *denotes the square root of the* $(m+j)$*th diagonal element of*

$$\widehat{\text{Var}}(\hat{\mathbf{\Theta}}^*_{\tau,\text{MLE}}) = \big(\mathbf{\Sigma}_{\tau,n}(\hat{\mathbf{\Theta}}^*_{\tau,\text{MLE}})\big)^{-1}.$$

Similarly, a $100(1-\theta)\%$ *confidence interval for* β_j *is given by*

$$\hat{\beta}_{\tau,j} \pm \Phi^{-1}(1-\theta/2)\hat{\sigma}_{\tau,m+j},$$

where $\Phi^{-1}(1-\theta/2)$ *is the* $(1-\theta/2)$*th percentile of the standard normal distribution.*

The final theoretical components from Section 2 left to modify for right-censoring are Theorem 2.2 and Corollary 2.2.1. We do so now by providing the complete statements, presented without proof because of its similarity to Section A.5.

Theorem D.2. *Define the following null likelihood,*

$$\mathcal{L}_{\tau,0}(\mathbf{\Theta}_0 \mid \mathcal{S}_{\tau,n}) = \alpha^{-n} \prod_{v=\Delta+1}^{\Delta+m} g_v^{\hat{\gamma}_n(v)} \prod_{\{i:D_i=1\}} f(T_i; p) \prod_{\{i:D_i=0\}} S(T_i+1; p).$$

If $\hat{\mathbf{\Theta}}_{\tau,0} = (\hat{p}_\tau, \hat{\boldsymbol{g}}_\tau)^\top$*, where* $\hat{p}_\tau$ *and* $\hat{\boldsymbol{g}}_\tau$ *follow from Step 1 of Algorithm D.1, then, under suitable regularity conditions (e.g., Lehmann and Romano, 2006, Theorem 12.4.2, pg. 515) with* $\beta_j = 0$ *for* $1 \leq j \leq k$*, the LRT statistic*

$$\Omega_{\tau,n} \equiv 2 \log \left(\frac{\mathcal{L}(\hat{\mathbf{\Theta}}_{\tau,\mathrm{MLE}})}{\mathcal{L}_0(\hat{\mathbf{\Theta}}_{\tau,0})} \right) \longrightarrow_d \chi^2_k,$$

where χ^2_k *denotes a chi-square distribution with* k *degrees of freedom.*

The modified Corollary 2.2.1 then follows.

Corollary D.2.1. *Suppose we wish to test* $H_0 : \beta_j = 0$ *for all* $1 \leq j \leq k$ *against the alternative* H_A *: at least one* $\beta_j \neq 0$*,* $1 \leq j \leq k$*, at the asymptotic significance level,* $0 < \theta < 1$*. To do so, reject* H_0 *if* $\Omega_{\tau,n} > \chi^2_{1-\theta}$*, where* $\chi^2_{1-\theta}$ *denotes the* $100 \times (1-\theta)$*th percentile of a chi-square distribution with* k *degrees of freedom, and fail to reject* H_0 *otherwise.*

As in Section 3 and for completeness, we may numerically assess the performance of Theorems D.1 and D.2. For consistency, we use the same study design for each suitably modified for right-censoring. For the former, we assign $\varepsilon = 15$. This yields a censoring rate of 28.44%. The results may be found in Table D1, and they confirm a consistent performance to Section 3. For the latter, we assigned $\varepsilon = 10$. This yields a censoring rate of 20.36%. The results may be found in Figure D1, and they confirm a consistent performance to Section 3.

E Application Additional Details

We present a complete count of the 67,797 consumer auto loans contained in the AART (2017) ABS bond by original loan term in Table E1. We can observe the expected clustering

Table D1: **Simulation Study (Thm D.1)**. The robustness of Theorem D.1 for $n = 1{,}000$ with $\omega = 12$, $\Delta = 0$, $m = 8$, $\varepsilon = 15$, for X following a binomial distribution with a logit link function for 1,000 replicates. We report the absolute value of the empirical bias (|eBias|), the empirical standard error (eSE), an average of the estimated standard error using Theorem D.1 (Thm D.1 SE), and a coverage probability (CP) for 95% asymptotic confidence intervals using Corollary D.1.1.

$\boldsymbol{\Theta}^*$	True	|eBias|*	eSE*	Thm D.1 SE*	CP
g_1	0.30	0.10	13.70	14.38	95.7%
g_2	0.20	0.16	11.86	12.47	96.1%
g_3	0.13	0.13	10.35	10.45	95.1%
g_4	0.10	0.09	9.51	9.32	94.2%
g_5	0.09	0.43	8.76	9.00	95.4%
g_6	0.07	0.14	8.63	8.25	94.1%
g_7	0.06	0.18	8.32	8.26	94.0%
β_0	0.50	0.26	20.48	21.68	96.3%
β_1	0.50	7.46	224.56	214.06	93.3%
β_2	1.00	9.93	213.84	214.86	94.6%
β_3	-1.50	0.38	220.00	215.88	94.0%
β_4	-0.50	2.05	216.74	214.11	95.0%

*Results on a 10^{-3} scale (i.e., 7.1 = 0.0071).

around standard industry loan terms, such as 36, 48, 60, and 72 months. In some cases, lenders may defer the first payment a month or two, which may explain the unexpected record keeping (i.e., reported loan terms for these clusters are (37, 38), (49, 50), (61, 62), and (73, 74) months, respectively). For the purposes of the economic analysis in Section 4, therefore, we treat $(37, 38) \equiv 36$ and $(61, 62) \equiv 60$ for consistency with industry standards.

For the analysis in Section 4, we present the following data notes. Of the 2,798 36-month loans, we remove five loans with data irregularities. We also remove one right-censored loan and five defaulted loans. Next, we remove 17 loans with repayment times that extend beyond 41 months, which are likely loan extensions. Finally, we remove two loans with an unreported credit score and twelve loans with atypical income verification records. This leaves 2,756 36-month auto loans to analyze. For the 17,741 60-month loans, we remove 57 loans with data irregularities. We also remove 570 (~3.22%) defaulted loans. Next, we remove 25 loans with repayment times that extend beyond 68 months, which are likely loan extensions. We also remove one loan with a calculated left-truncation time outside of

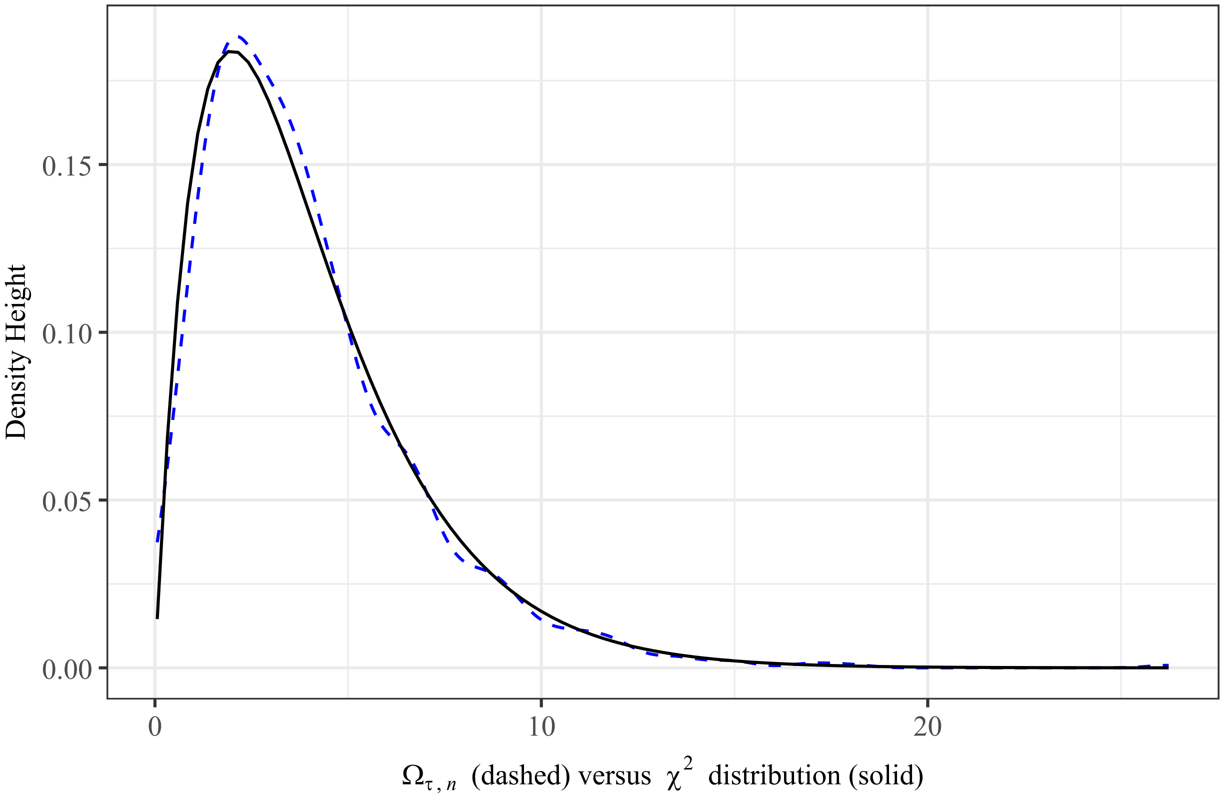

Figure D1: **Simulation Study (Thm D.2)**. The robustness of Theorem D.2 for $n = 1{,}000$ with $\omega = 8$, $\Delta = 0$, $m = 5$, $\varepsilon = 10$ for X following a binomial distribution with a logit link function for 1,000 replicates. We display the empirical density of the test statistic, $\Omega_{\tau,n}$ (dashed line), against the anticipated density of a chi-square distribution with four degrees of freedom (solid line). For completeness, the observed empirical Type I error of the hypothesis test in Corollary D.2.1 with significance level $\theta = 0.05$ was 0.045 over these 1,000 replicates.

the support space (an additional data irregularity). Finally, we remove 19 loans with an unreported credit score and 56 loans with atypical income verification records. This leaves 17,016 60-month auto loans to analyze. For reference, complete data processing details may be found in the publicly available replication code.

Table E1: **AART 2017-3 Loan Term and Outcome Summary**. Select counts of original loan terms for AART (2017) by outcome: right-censored, repaid in full, and defaulted. Loan terms of (37, 38), (49, 50), (61, 62), and (73, 74) are assumed to be industry standard loan terms of 36, 48, 60, and 72 month loans, respectively. Not reported in this table are 346 loans with original loan terms of 12, 13, 14, 15, 19, 20, 21, 23, 28, 29, 30, 31, 32, 33, 35, 36, 39, 40, 41, 42, 45, 47, 48, 51, 52, 53, 54, 57, 58, 59, 60, 63, 66, 69, 70, 71, 72, 75, and 78 months.

Term	Count	Censored	Repaid	Defaulted	Censored (%)	Repaid (%)	Defaulted (%)
25	151	0	151	0	0.00	100.00	0.00
26	115	0	115	0	0.00	100.00	0.00
37	1,569	4	1,562	3	0.25	99.55	0.19
38	1,229	2	1,225	2	0.16	99.67	0.16
43	126	3	121	2	2.38	96.03	1.59
44	84	2	81	1	2.38	96.43	1.19
49	2,531	52	2,448	31	2.05	96.72	1.22
50	1,880	48	1,806	26	2.55	96.06	1.38
55	202	26	166	10	12.87	82.18	4.95
56	111	15	93	3	13.51	83.78	2.70
61	10,287	1,723	8,225	339	16.75	79.96	3.30
62	7,454	1,254	5,969	231	16.82	80.08	3.10
64	864	198	631	35	22.92	73.03	4.05
65	723	202	503	18	27.94	69.57	2.49
67	630	156	438	36	24.76	69.52	5.71
68	442	113	297	32	25.57	67.19	7.24
73	17,757	4,720	11,533	1,504	26.58	64.95	8.47
74	12,652	3,462	8,073	1,117	27.36	63.81	8.83
76	5,204	1,112	3,653	439	21.37	70.20	8.44
77	3,440	781	2,378	281	22.70	69.13	8.17

Note: Percentages may not sum to 100% due to rounding.

F Proportional Hazard Assumption

Per Klein and Moeschberger (2003, (8.1.1), pg. 244), the classical Cox proportional hazard model typically takes the form,

$$h(t \mid \boldsymbol{Z}) = h_0(t)\exp(\boldsymbol{\beta}^\top \boldsymbol{Z}) = h_0(t)\exp\left(\sum_{k=1}^{p} \beta_k Z_k\right), \tag{S.16}$$

where $h(t \mid \boldsymbol{Z})$ is the hazard rate at time t, $\boldsymbol{Z}$ is a vector of covariates (i.e., risk factors), and $h_0(t)$ is an arbitrary baseline hazard function. (We adopt the notation of Klein and Moeschberger (2003) for convenience of reference.) In Section 4, consumer auto loans were

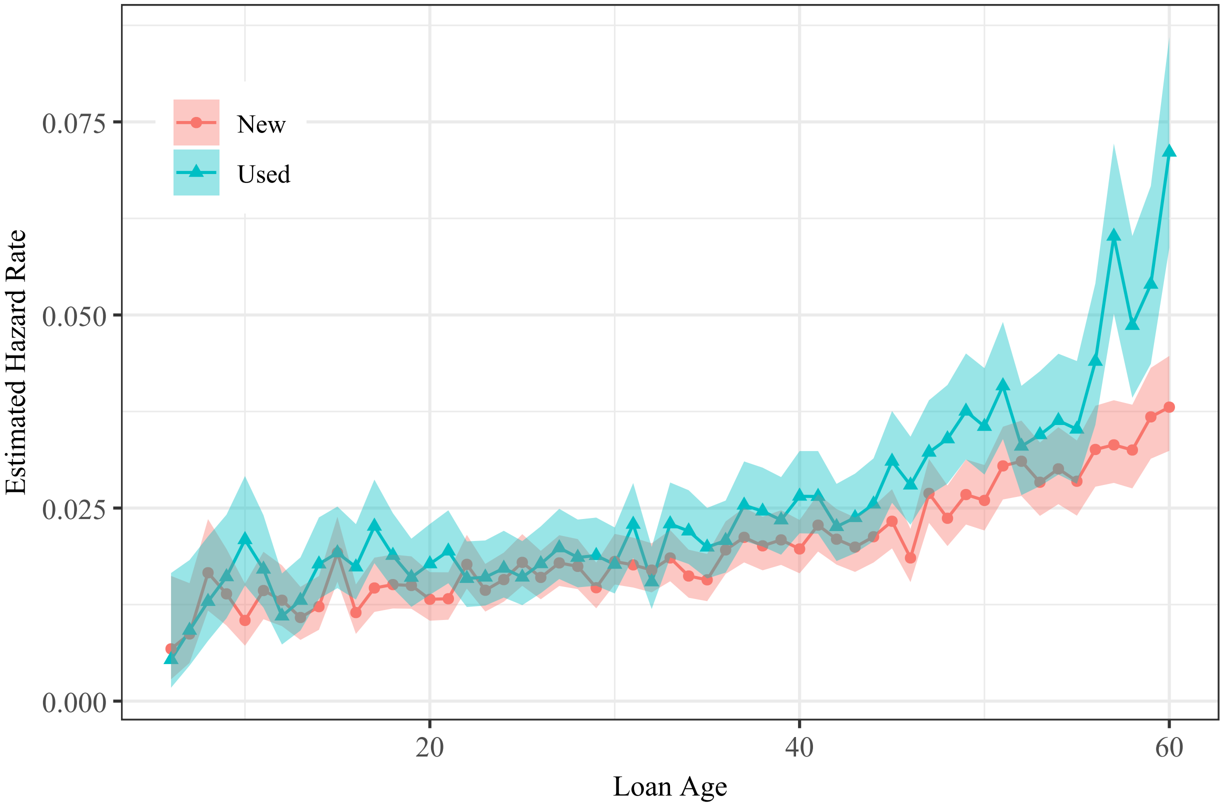

Figure F1: **AART 2017-3 Proportional Hazard Assumption Assessment**. Estimated hazard rates by loan age plus 95% asymptotic confidence intervals via Lautier et al. (2023) for 60-month consumer auto loans by New (10,974 loans) and Used (6,042 loans) vehicle. In it is evident the ratio of hazard rates for each group varies by time, which violates the implied constant ratio of the proportional hazard assumption derived in (S.17).

grouped by various indicator variables. Let us consider the `nw.usd` indicator, which, using the notation of (S.16) is,

$$\boldsymbol{Z} = Z_1 = \begin{cases} 1, & \text{New Vehicle} \\ 0, & \text{Used Vehicle.} \end{cases}$$

This implies that a pairwise ratio of hazards under (S.16) yields,

$$\frac{h(t \mid Z_1 = 1)}{h(t \mid Z_1 = 0)} = \frac{h_0(t)\exp(\beta_1)}{h_0(t)\exp(0)} = \exp(\beta_1) \equiv K, \tag{S.17}$$

where K is some constant free from t. We can assess whether or not this assumption is reasonable with a visual inspection of the estimated hazard for each group using Lautier et al. (2023). This is done in Figure F1 for the 60-month time-to-event data studied in Section 4. It is immediate that the ratio of hazard rates as a function of time will not be constant. In younger loan ages, for example, the two hazard rates cross frequently, and so the ratio will sometimes be above and below one. As the loans get older, we see that the estimated hazard rate for Used vehicles grows at a faster rate than the estimated hazard rate for New vehicles. For loan ages closer to 60 months, the asymptotic confidence intervals no longer overlap. Hence, the ratio of the hazard rates for the Used to New vehicle group will grow as loan age increases, an additional example that the ratio will not be constant as a function of time. This implies that alternative methods will be necessary to recover the loan lifetime distribution, such as those proposed in Section 2.

References

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

J. Hardin and J. Hilbe (2007). *Generalized Linear Models and Extensions, Second Edition*. A Stata Press publication. Taylor & Francis.

J. P. Klein and M. L. Moeschberger (2003). *Survival Analysis: Techniques for Censored and Truncated Data, Second Edition*. Springer Science + Business Media, Inc.

M. H. Kutner, C. J. Nachtsheim, J. Neter and W. Li (2005). *Applied Linear Statistical Models*. McGraw-Hill Irwin.

J. P. Lautier, V. Pozdnyakov and J. Yan (2023). "Pricing time-to-event contingent cash flows: A discrete-time survival analysis approach." *Insurance: Mathematics and Economics* **110**, 53–71.

J. P. Lautier, V. Pozdnyakov and J. Yan (2025). "Estimating the time-to-event distribution for loan-level data within an asset-backed security." *Annals of Applied Statistics* **19**, 2831–2851.

J. P. Lautier, V. Pozdnyakov and J. Yan (2026). "Estimating a discrete distribution subject to random left-truncation with an application to structured finance." *Econometrics and Statistics* **37**, 174–198.

E. Lehmann and J. Romano (2006). *Testing Statistical Hypotheses*. Springer Texts in Statistics. Springer New York.

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

T. Nakagawa and S. Osaki (1975). "The discrete Weibull distribution." *IEEE Transactions on Reliability* **R-24**, 300–301.

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

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

J. Shao (2003). *Mathematical Statistics*. Springer New York, NY.

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