

Ko, Seok Bin

## Article

# Real option valuation using Weibull distribution: A new framework for depreciation risk management

Journal of Derivatives and Quantitative Studies: Seonmul yeon'gu (JDQS)

## Provided in Cooperation with:

Korea Derivatives Association (KDA)

*Suggested Citation:* Ko, Seok Bin (2025) : Real option valuation using Weibull distribution: A new framework for depreciation risk management, Journal of Derivatives and Quantitative Studies: Seonmul yeon'gu (JDQS), ISSN 2713-6647, Emerald, Leeds, Vol. 33, Iss. 2, pp. 110-130, <https://doi.org/10.1108/JDQS-10-2024-0041>

This Version is available at:

<https://hdl.handle.net/10419/337292>

### Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

### Terms of use:

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



<https://creativecommons.org/licenses/by/4.0/>

# Real option valuation using Weibull distribution: a new framework for depreciation risk management

Seok Bin Ko

*College of Business and Commerce, Cheongju University, Cheongju, South Korea*

110

Received 23 October 2024  
Revised 30 December 2024  
15 January 2025  
Accepted 16 January 2025

## Abstract

This study aims to develop an accurate option pricing model for car leases by introducing a put option valuation framework based on the Weibull distribution. Traditional models typically assume asset values follow a lognormal distribution, failing to capture the left-skewed nature and bounded dynamics. To address this limitation, this study compares the performance of the Weibull distribution with that of the lognormal model using residual value data from two popular car models in South Korea, evaluating each model's ability to reflect unique depreciation patterns. The findings demonstrate that the Weibull distribution provides a superior fit to the data, leading to more precise option pricing. This enhanced accuracy is crucial for auto finance companies navigating uncertainties in used car prices, particularly as the mobility services and car leasing markets continue to expand. Moreover, the practical implications of this research extend beyond the auto finance industry; insights from this study can inform sectors dealing with skewed or bounded assets, such as insurance products and financial derivatives, thereby enabling improved risk assessment and decision-making processes. This research introduces a novel approach to modeling put option values using the Weibull distribution, filling a significant gap in the existing literature on car lease option pricing. However, as this is the first study to model put option values using the Weibull distribution, further research is necessary. Specifically, investigating volatility patterns over the lifecycle of used cars could significantly enhance the value of this framework.

**Keywords** Car leases, Real option valuation, Put option formula, Weibull distribution, Left-skewed distribution  
**Paper type** Research paper

## 1. Introduction

As car leases become one of the main businesses in the auto finance industry, accurately pricing the embedded options in car leases has become more critical than ever. However, industry practitioners rely on empirical statistical measures and non-parametric methods for product development, risk management, and funding communication due to the limitations of existing derivative pricing models. While the lognormal distribution effectively captures the potential for significant upward movements in most financial asset prices, it is ill-suited for used car values, which predominantly exhibit left-skewed distribution patterns.

Accurate option valuation is pivotal in the auto finance industry for three key reasons: it provides a cost basis for lease products, where consumers may buy or return the vehicle; it underpins risk management, informing loss reserves or provisions; and it supports funding strategies, enabling investors to evaluate embedded risks and returns on investment. However, the Black-Scholes model largely fails to meet these practical needs, as it assumes the underlying asset follows a lognormal distribution. This misalignment is compounded by

© Seok Bin Ko. Published in *Journal of Derivatives and Quantitative Studies*: *선물연구*. Published by Emerald Publishing Limited. This article is published under the Creative Commons Attribution (CC BY 4.0) licence. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial and non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this licence maybe seen at <http://creativecommons.org/licences/by/4.0/legalcode>

**Financial interests:** The author has no competing interests to declare that are relevant to the content of this article.

**Data availability:** The author confirms that the dataset used and generated during the experiment is available from the author on reasonable requests.



volatility assumptions, which incorrectly presume volatility increases with the square root of time. Such assumptions overestimate option prices in car leases, where residual values exhibit irregular and idiosyncratic volatility patterns.

These theoretical limitations have rendered the Black-Scholes model ineffective for valuing embedded options. Consequently, auto finance companies have relied on empirical statistical measures. Such empirical approaches often fail to meet the rigorous standards investors require. For instance, one U.S.-based auto finance company uses the fixed charge coverage ratio — an internal risk measure agreed upon with major institutional investors — to assess its ability to cover operation costs. Yet, this measure encompasses broader business performance rather than isolating lease-specific risks. This complicates product-specific funding efforts and risk management since investors request greater spread to mitigate the immeasurable uncertainty.

The absence of robust option valuation frameworks also prevents auto finance companies from effectively utilizing derivatives for product optimization. Overestimated option values make lease products less competitive than installment products or loans. Addressing these limitations is essential to align auto finance practices with the unique characteristics of their underlying assets, thereby enabling more precise risk management, investor communication, and competitive product development.

This study aims to bridge that gap by providing a rigorous theoretical framework tailored to the unique characteristics of car residual values. The absence of market-traded option prices does not diminish the importance of accurate valuation; instead, it underscores the critical need for a robust methodology that auto finance companies can rely on. Our approach demonstrates how the Weibull-based model better captures the observed characteristics of residual car values and offers a practical alternative to the lognormal distribution.

Used car prices typically decline over time and exhibit limited potential for extreme price increases. Furthermore, they tend to follow a left-skewed pattern, the opposite of the right-skewed nature of lognormal distributions. Such asymmetry is common not only in residual car values but also in several insurance products. Left-skewed characteristics of such insurance products increase a decision-maker's insurance demand under uncertainty (Hong, 2022). Left-skewed patterns are also observed in stock markets. Recent studies (Kim *et al.*, 2023a) have shown that stocks with high left-tail risk exhibit lower future returns, consistent with the skewness-related anomalies. Considering the Weibull distribution, with its flexibility to model asymmetric risks and heavier left tails, is more appropriate for capturing the dynamics of the assets subject to similar depreciation patterns, this study introduces a Weibull distribution-based put option pricing formula and compares the performance between the Weibull and lognormal distributions.

The remainder of this study is organized as follows. Section Two reviews the literature on option pricing models and their applications in car leases. Section Three introduces the Weibull distribution and its parameter estimation methods. Section Four discusses embedded options in car leases. Section Five presents the Weibull-based put option formula and validates it through proofs and testing. Section Six describes the dataset and filtering methods. Section Seven compares Weibull and lognormal distributions in capturing residual car value dynamics. Finally, Section Eight concludes the study with implications for the auto finance industry and suggestions for future research.

## 2. Literature review

Since Fisher Black and Myron Scholes introduced their option pricing formula in 1973 (Black and Scholes, 1973), option valuation techniques have become widely adopted in financial markets and across various industries. In particular, researchers have applied these methods to analyze embedded options in car leases, a topic that has gained significant attention (Dasgupta, 2004; Gamba and Rigon, 2008; Giaccotto *et al.*, 2007; Li and You, 2022; McConnell and Schallheim, 1983; Miller, 1995; Oppenheimer, 2002; Starkey, 1997; Trigeorgis, 1996).

McConnell and Schallheim (1983) were among the first to conceptualize a lease as a compound option. Later studies extended this approach, incorporating indirect utility methods to explore consumer decision-making and estimate option values (Dasgupta, 2004; Starkey, 1997). Giaccotto *et al.* (2007) and Miller (1995) argued that the call option values in leases account for cost differences between loans and leases. Further research has identified other embedded options within leases. Gamba and Rigon (2008) demonstrated that the early termination option behaves like an American put option when market interest rates fluctuate. Trigeorgis (1996) argued that other options, such as renewal options and contingent claims, have influenced the valuation of leases. Li and You (2022) and Oppenheimer (2002) emphasized that put options play a critical role in car lease structures.

Despite extensive research on embedded options, most studies assume that the underlying assets (used car prices) follow a lognormal distribution. However, unlike stock prices, used car values typically decline over time, exhibiting non-linear depreciation. This behavior mirrors certain insurance products, such as warranties, which also experience predictable depreciation. As a result, used car prices require valuation models that accommodate non-linear patterns and heavy-tailed distributions, which traditional lognormal models fail to capture.

To address these limitations, many researchers have adopted non-parametric Bayesian methods to model the complex dynamics of asset uncertainty. Markov Chain Monte Carlo (MCMC) techniques are particularly useful for estimating posterior distributions without relying on strict parametric assumptions (Speagle, 2019). However, the application of MCMC methods presents specific challenges. These techniques are computationally intensive and sensitive to the selection of initial values and priors, which can lead to convergence issues and biased estimates if not carefully managed (Van Ravenzwaaij *et al.*, 2018). Consequently, auto finance industry practitioners have sought simpler valuation models that offer practical utility while maintaining accuracy.

In other industries than auto finance, researchers have explored more flexible distributions to improve option valuation. Savickas (2002) introduced a call option formula based on the Weibull distribution, effectively capturing the negative skewness observed in S&P 500 index options. Building on this work, Mauler and McDonald (2015) developed option pricing formulas using various distributions, including inverse hyperbolic sine (IHS), generalized beta of the second kind,  $g$ -and- $h$ , and generalized gamma. Their findings demonstrated that these flexible distributions improve the accuracy of option pricing beyond what traditional models can achieve.

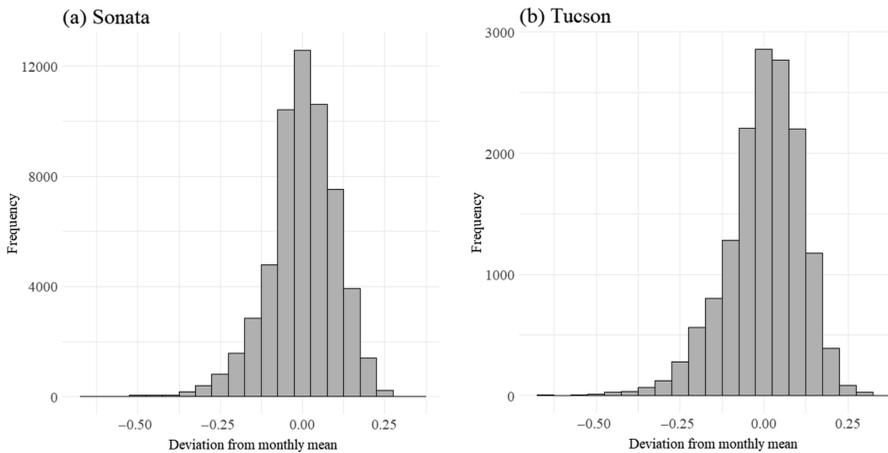
While research on flexible distributions has primarily focused on call options, auto finance and insurance products often require accurate put option valuations to address downside risks. This study proposes to value put options in car leases using the Weibull distribution as a parametric tool, given its flexibility in modeling downside risks. Figure 1 illustrates the left-skewed residual values of used cars, which align well with the Weibull distribution's ability to capture such patterns.

### 3. Weibull distribution

The flexibility of the Weibull distribution helped this study determine more precise option pricing for car leases. The Weibull distribution, introduced by Waloddi Weibull (1951), is a versatile tool widely applied in diverse fields, including reliability engineering, quality control, weather forecasting, electrical and industrial engineering, and more (Gómez *et al.*, 2023). The distribution was also introduced in financial derivatives analysis and applied to various financial product valuation assessments (Chen and Gerlach, 2013; Heo *et al.*, 2012; Li *et al.*, 2024; Lin *et al.*, 2023; Savickas, 2002).

Weibull (1951) introduced the probability density function of the Weibull distribution,

$$f(x; \beta, \delta) = \beta\delta x^{\delta-1} e^{-\beta x^\delta}, x \in \mathbb{R}_{\geq 0}, \beta, \delta \in \mathbb{R}^+ \quad (1)$$



**Note(s):** The residual deviations from the residual mean value of each month were used for the histograms

**Source(s):** Figure by author

**Figure 1.** Histograms of residual value deviations of Hyundai Sonata and Tucson used car prices

If the shape parameter  $\delta$  equals one, the Weibull distribution becomes an exponential distribution with scale parameter  $\beta$ . The distribution pattern can be determined using the size of  $\delta$ . Let  $y = \beta x^\delta$ .

$$dy = \beta \delta x^{\delta-1} dx$$

$$f(x) dx = e^{-y} dy$$

Thus, the cumulative distribution will be

$$F(t) = \int_0^t f(x) dx = \int_0^{\beta t^\delta} e^{-y} dy = 1 - e^{-\beta t^\delta} \quad (2)$$

If the random variable  $X$  is for a lifetime, the reliability function  $R(x)$  is defined as

$$R(x) = P(X > x) = 1 - F(x) = e^{-\beta x^\delta} \quad (3)$$

Sequentially, the failure rate function is determined,

$$\varphi(x) = \frac{f(x)}{R(x)} = \frac{\beta \delta x^{\delta-1} e^{-\beta x^\delta}}{e^{-\beta x^\delta}} = \beta \delta x^{\delta-1} \quad (4)$$

The Weibull distribution's shape is determined through the failure rate function. If  $\delta < 1$ , the shape shows an exponential decrease. If  $\delta = 1$ , the failure rate will be constant. In the meantime, if  $\delta > 1$ , the failure will generate a geometrically increasing pattern, meaning the Weibull distribution is negatively skewed. Savickas (2002) argued that the distribution would have a negative third central moment, or skewness, whenever the parameter  $\delta > 3.60234$ , regardless of the value of the parameter  $\beta$ . This study found distributions of

car residual values are negatively skewed Weibull distributions. The deltas for Sonata and Tucson were 11.422843 and 7.982199, respectively (see Table 2).

The expected value and its associated variance can be determined using the moment-generating function:

$$\begin{aligned}
 E(X^r) &= \int_0^\infty x^r f(x) dx = \int_0^\infty \left(\frac{y}{\beta}\right)^{\frac{r}{\delta}} e^{-y} dy \\
 &= \left(\frac{1}{\beta}\right)^{r/\delta} \int_0^\infty y^{r/\delta} e^{-y} dy = \left(\frac{1}{\beta}\right)^{\frac{r}{\delta}} \Gamma\left(1 + \frac{r}{\delta}\right)
 \end{aligned} \tag{5}$$

Thus, the expected value and variance are

$$E(X) = E(X^1) = \left(\frac{1}{\beta}\right)^{\frac{1}{\delta}} \Gamma\left(1 + \frac{1}{\delta}\right) \tag{6}$$

$$\begin{aligned}
 \text{Var}(X) &= E(X^2) - E(X)^2 \\
 &= \left(\frac{1}{\beta}\right)^{\frac{2}{\delta}} \left[ \Gamma\left(1 + \frac{2}{\delta}\right) - \left\{ \Gamma\left(1 + \frac{1}{\delta}\right) \right\}^2 \right]
 \end{aligned} \tag{7}$$

### 3.1 Parameter estimation

The parameters were estimated using the maximum likelihood estimation method. If data  $x_1, x_2, \dots, x_n$  ( $\forall x_i \geq 0, x_i \in \mathbb{R}$ ) are given, the likelihood function will be:

$$L(\beta, \delta | x_1, x_2, \dots, x_n) = \prod_{i=1}^n f(x; \beta, \delta) = \prod_{i=1}^n \beta \delta x_i^{\delta-1} e^{-\beta x_i^\delta} \tag{8}$$

The associated loglikelihood is:

$$\begin{aligned}
 \ln L(\beta, \delta | x_1, x_2, \dots, x_n) &= \sum_{i=1}^n \ln(\beta \delta x_i^{\delta-1} e^{-\beta x_i^\delta}) \\
 &= n \times \ln(\beta) + n \times \ln(\delta) + (\delta - 1) \sum_{i=1}^n \ln(x_i) - \beta \sum_{i=1}^n x_i^\delta
 \end{aligned} \tag{9}$$

To find parameters  $\beta$  and  $\delta$  maximizing the loglikelihood value, numerical optimization algorithms, such as the Newton-Raphson method, gradient descent, or Broyden-Fletcher-Goldfarb-Shanno (BFGS), are used. This study used the BFGS, one of the most widely used quasi-Newton methods due to its various advantages, such as the Hessian matrix approximation, fast convergence, robustness regardless of function types and initial estimates, and unconstrained optimization (Li, 2020; Xie *et al.*, 2020). More specifically, the BFGS, using the second-order derivative of an object function, performs excellently in self-correcting and approximating through the Hessian matrix approximation even where the second derivative cannot be calculated (Li, 2020; Xie *et al.*, 2020).

#### 4. Options embedded in car leases

Since lessees can purchase leased cars at maturity or early terminate the contracts, many researchers argued that leases have call options that can be exercised at maturity and put options to terminate before the contracts expire (Dasgupta, 2004; Gamba and Rigon, 2008; Giaccotto *et al.*, 2007; Miller, 1995; Oppenheimer, 2002; Starkey, 1997; Trigeorgis, 1996). Li and You (2022) added that lessees also have a put option that can be exercised at maturity. This study contends that the only put option at maturity is significant to both lessees and lessors.

Some may argue that call options are valuable because lessees benefit from lower monthly payments and the option to purchase the car at maturity. However, consumers sensitive to payments can opt for long-term loans, like 60-month loans. Instead, their primary concerns are used car depreciation and uncertainty at maturity. They can return the cars or exercise put options when market conditions are unfavorable or when they prefer not to sell the used cars themselves. Auto lease companies also worry about used car price uncertainty and aim to manage this while offering competitive lease products (Pierce, 2007).

The early termination option is also insignificant because the penalty typically outweighs its benefits. In general, auto finance companies charge lessees for all losses incurred from early terminations. Giaccotto *et al.* (2007) argued that the early termination penalty in the U.S. is significantly larger than the theoretical prices of termination options. Hyundai Capital Korea also imposes an early termination penalty ranging from 19% to 40% of the outstanding balance, substantially greater than theoretical option prices, when lease consumers return vehicles within the contract's first two years. This makes it nearly impossible for lease consumers to profit from early termination.

#### 5. Weibull distribution-based put option formula

Although managing downside risks is crucial for real-world businesses like auto finance and insurance companies, the valuation of put options using flexible distributions has been largely overlooked. A standalone put option formula may not be necessary since put-call parity always enables the calculation of put option prices. However, while this argument holds in idealized markets where call option prices are reliably available, practical and theoretical limitations render put-call parity insufficient in many real-world scenarios. Bastianello *et al.*'s (2024) findings reinforce the argument that the validity of put-call parity can be undermined by non-linear pricing rules, market frictions such as bid-ask spreads, and the absence of additivity in probability measures. These factors often disrupt the theoretical foundation of put-call parity, leading to deviations from its expected outcomes. Their work highlights scenarios where these deviations occur due to market imperfections.

Krause (2019) also examined put-call parity violations in the U.S. equity options markets, identifying non-synchronous trading hours between stock and options markets as a key factor contributing to these anomalies. While aligning trading hours mitigated many of these deviations, his findings suggest that market imperfections, rather than arbitrage opportunities, often drive apparent violations. These insights underline the necessity for alternative pricing methods, reinforcing the importance of a standalone put option framework for specific asset classes like those discussed in this study.

Let the risk-neutral Weibull density function be (Weibull, 1951):

$$\tilde{f}(x; \beta, \delta) = \beta\delta x^{\delta-1} e^{-\beta x^\delta}, x \in \mathbb{R}_{\geq 0}, \beta, \delta \in \mathbb{R}^+ \quad (10)$$

If  $\beta$  and  $\delta$  are scale and shape parameters of the Weibull distribution under the equivalent martingale measure, the current asset price and variance will be the same as those at time  $T$ . If  $S_0$  is the current asset price and  $\Omega_t$  is the information available at time  $t$ , then the risk-neutral mean and variance will be

$$\tilde{E}(S_T : \Omega_0) = \left(\frac{1}{\beta}\right)^{\frac{1}{\delta}} \Gamma\left(1 + \frac{1}{\delta}\right) \tag{11}$$

$$\widetilde{Var}(S_T : \Omega_0) = \left(\frac{1}{\beta}\right)^{\frac{2}{\delta}} \left[ \Gamma\left(1 + \frac{2}{\delta}\right) - \left\{ \Gamma\left(1 + \frac{1}{\delta}\right) \right\}^2 \right] \tag{12}$$

$$\text{where } \Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt, x \in \mathbb{R}^+$$

When the strike price is  $K$ , and the maturity is  $T$ , the put option value  $P$  is determined by calculating the present value of the payoffs at maturity  $T$ , represented by  $\max(K - S_T, 0)$  where  $S_T$  is the asset price at time  $T$ . If there is no price-changing intervention like a dividend, we have

$$P = e^{-rT} \tilde{E}[\max(K - S_T, 0) | \Omega_0] \tag{13}$$

where  $\tilde{E}$  is the expectation under the risk-neutral measure, and  $\Omega_0$  represents the information available at time zero. If dividends exist, the  $S_T$  is replaced with  $S_T^D (= S_T - D_T)$  where  $D_T$  is the present value of the expected dividends paid from time 0 to  $T$ . Using risk-neutral density function  $\tilde{f}(x; \beta, \delta)$ , we can express  $P$  as

$$\begin{aligned} P &= e^{-rT} \int_0^K (K - x) \tilde{f}(x; \beta, \delta) dx \\ &= e^{-rT} \int_0^K (K - x) \beta \delta x^{\delta-1} e^{-\beta x^\delta} dx \\ &= e^{-rT} \left[ K \int_0^K \beta \delta x^{\delta-1} e^{-\beta x^\delta} dx - \int_0^K \beta \delta x^\delta e^{-\beta x^\delta} dx \right] \\ &= e^{-rT} \left[ KF_{Weibull}(K; \beta, \delta) - \int_0^K \beta \delta x^\delta e^{-\beta x^\delta} dx \right] \end{aligned} \tag{14}$$

where  $F_{Weibull}(K; \beta, \delta)$  is the cumulative distribution function of the Weibull distribution evaluated at  $K$ . For the second integral,

$$dt = \beta \delta x^{\delta-1} dx$$

if  $t$  is replaced with

$$\beta x^\delta.$$

Then we have:

$$dx = \frac{dt}{\beta\delta\left(\frac{t}{\beta}\right)^{\frac{\delta-1}{\delta}}} = \frac{dt}{\beta\delta}\left(\frac{\beta}{t}\right)^{\frac{\delta-1}{\delta}} = \frac{dt}{\beta\delta}\beta^{\frac{\delta-1}{\delta}}t^{-\frac{\delta-1}{\delta}} = \frac{dt}{\beta^{1/\delta}\delta}t^{-\frac{\delta-1}{\delta}}$$

Therefore,

$$\int_0^{\beta K^\delta} \beta\delta \frac{t}{\beta} e^{-t} \times \frac{1}{\beta\delta} t^{-(\delta-1)/\delta} dt = \beta^{-\frac{1}{\delta}} \int_0^{\beta K^\delta} t^{1/\delta} e^{-t} dt \quad (15)$$

The integral  $\int_0^\infty t^{1/\delta} e^{-t} dt$  can be split into two parts:  $\int_0^{\beta K^\delta} t^{1/\delta} e^{-t} dt$  and  $\int_{\beta K^\delta}^\infty t^{1/\delta} e^{-t} dt$   
From the gamma function definition,

$$\int_0^\infty t^{1/\delta} e^{-t} dt = \Gamma\left(1 + \frac{1}{\delta}\right)$$

The gamma function can be split into lower and upper incomplete gamma functions:

$$\gamma\left(1 + \frac{1}{\delta}, \beta K^\delta\right) = \int_0^{\beta K^\delta} t^{1/\delta} e^{-t} dt \quad (16)$$

$$\Gamma\left(1 + \frac{1}{\delta}, \beta K^\delta\right) = \int_{\beta K^\delta}^\infty t^{1/\delta} e^{-t} dt \quad (17)$$

Thus,

$$\beta^{-\frac{1}{\delta}} \int_0^{\beta K^\delta} t^{1/\delta} e^{-t} dt = \beta^{-\frac{1}{\delta}} \gamma\left(1 + \frac{1}{\delta}, \beta K^\delta\right) \quad (18)$$

Combining the first and second integrals,

$$P = e^{-rT} KF_{Weibull}(K; \beta, \delta) - \beta^{-\frac{1}{\delta}} \gamma\left(1 + \frac{1}{\delta}, \beta K^\delta\right) \quad (19)$$

where  $\gamma(s, x)$  is the lower incomplete gamma function defined as  $\int_0^x t^{s-1} e^{-t} dt$ . Using the substitution  $\tilde{E}(S_T : \Omega_0) = \left(\frac{1}{\beta}\right)^{\frac{1}{\delta}} \Gamma\left(1 + \frac{1}{\delta}\right)$ , we have:

$$P = e^{-rT} KF_{Weibull}(K; \beta, \delta) - \tilde{E}(S_T : \Omega_0) \gamma\left(1 + \frac{1}{\delta}, \beta K^\delta\right) / \Gamma\left(1 + \frac{1}{\delta}\right) \quad (20)$$

Since  $\tilde{E}(S_T : \Omega_0)$  is the same as  $S_0$  under the equivalent martingale measure, we arrive at

$$P = e^{-rT} KF_{Weibull}(K; \beta, \delta) - S_0 \gamma\left(1 + \frac{1}{\delta}, \beta K^\delta\right) / \Gamma\left(1 + \frac{1}{\delta}\right) \quad (21)$$

Appendix A.2 provides proof of the formula using the put-call parity.

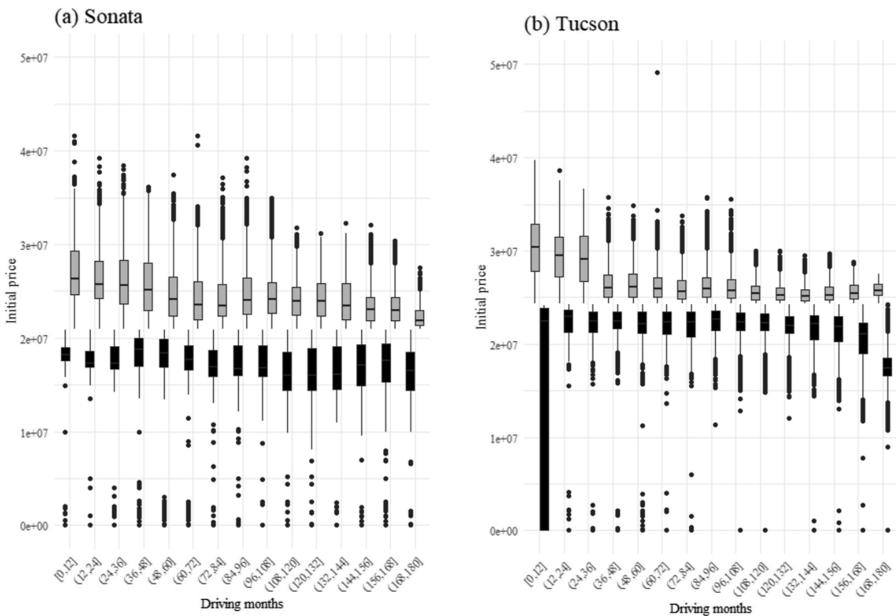
**6. Data**

This study examined the two popular car models traded in South Korea: Hyundai Sonata and Tucson. The data were officially obtained from the South Korean Transportation Safety Authority of the Ministry of Land, Infrastructure and Transport. The data is open to the public as a paid service. The total annual transaction units of the two car models were 241,404 in 2023. This study filtered 169,001 outliers in four steps.

First, cars sold immediately after purchase were eliminated. Selling cars immediately after purchase in South Korea is not rare because of huge discounts for employees and retired employees. Hyundai and Kia’s employees buy their brand cars at a maximum 30% discount, even after retirement, every two years in South Korea. Employees at their affiliate companies also buy cars at a maximum 25% discount when employed. Although Hyundai and Kia prohibit the employees from selling the cars within two years after purchase, controlling it is not always possible.

Second, transactions with original new cars sold to manufacturers’ employees were removed using K-means methods, represented by

$$\operatorname{argmin}_C \sum_{i=1}^n \sum_{x \in C_i} \|x - \mu_i\|^2 \tag{22}$$



**Note(s):** The lower box-plot data were eliminated, assuming that most cars in the lower box plots were cars sold to employees

**Source(s):** Figure by author

**Figure 2.** Box plots by K-means methods

where  $C_i$  is the set of points assigned to cluster  $i$  and  $\mu_i$  is the centroid of cluster  $i$ . Because it is rare to buy cars at more than a five percent discount in South Korea, cars purchased at a significant discount can be said to be cars sold to employees.

The K-means algorithm is an unsupervised learning method that clusters data into K groups by iteratively updating the centroids of each cluster to group the nearest data points together (Steinley and Brusco, 2007). This technique effectively distinguished between cars sold to manufacturer employees and ordinary consumers, although it does not guarantee perfect classification. Figure 2 shows the box plots after classifying the transactions into two groups for Sonata and Tucson.

Silhouette scores were calculated to measure how similar an object is to its own cluster compared to others. The silhouette score, which evaluates the consistency within clusters by considering both cohesion and separation, is represented by Rousseeuw (1987):

$$s(i) = I(C_i > 1) \cdot \frac{b(i) - a(i)}{\max(a(i), b(i))}, \forall i \in C_i \quad (23)$$

where  $a(i)$  is the average distance between object  $i$  and all other points in the same cluster, and  $b(i)$  is the average distance between object  $i$  and all other points in the nearest cluster.  $I(C_i > 1)$  is an indicator function that takes the value one if  $C_i$  is greater than one and zero otherwise. The average silhouette scores for the K-means classifications of Sonata and Tucson were 0.57184 and 0.53849, respectively, indicating that both were moderately positive and that the clustering was reasonably well-structured.

Third, excess mileage and excess aging cars were removed. Considering that most South Korean car lessors set 20,000 km as an annual mileage limit (Kim et al., 2023b), cars with an annual mileage of more than 20,000 km were removed. According to the Korea Transportation Safety Authority, the average passenger car mileage in South Korea was 13,140 km in 2022. This study eliminated those cars because most leasing consumers select leases considering their expected annual mileage. In addition, cars aged more than 15 years were removed because their small transactions made statistical outlier detection methods not converge. Their removal was not a problem in the analysis since the terms of most leasing contracts are less than 60 months in South Korea.

Fourth, statistical outliers were eliminated using the Extended Isolation Forest (EIF). The EIF method begins by building an ensemble of isolation trees through random sampling of data points, selecting a feature, and determining a split value. The isolation score of a point  $x$  is calculated as the average height at which the point is isolated across all trees, represented by the formula:

$$I(x) = \frac{1}{n} \sum_{i=1}^n h(x, T_i) \quad (24)$$

where  $n$  is the number of trees in the ensemble and  $h(x, T_i)$  is the height of the node in the tree  $T_i$  at which point  $x$  is isolated. The isolation score quantifies how quickly a point can be isolated from the rest of the data. Lower isolation scores indicate that the points are outliers since they are isolated more quickly. Next, the anomaly score is calculated based on the average of the points across the isolation trees, expressed as:

$$A(x) = 2^{-\frac{E(h(x))}{c(n)}} \quad (25)$$

where  $A(x)$  represents the anomaly score for point  $x$ ,  $E(h(x))$  is the expected depth of the point  $x$ , and  $c(n)$  is an average path length of all data points. The expected depth of the point  $x$ , or  $E(h(x))$ , decreases to zero, the anomaly score increases to one. If the anomaly score exceeds

the threshold, the point  $x$  is classified as an anomaly. The Isolation Forest technique isolates anomalies instead of profiling normal data points (Liu *et al.*, 2008). It works by recursively splitting the data space with random cuts, and the assumption is that outliers are isolated faster than regular data points because they are fewer and more different (Liu *et al.*, 2008). It estimates the anomaly score based on the number of splits, but there's no closed-form formula for directly computing outlier scores; the score comes from the distribution of paths taken in the isolation trees (Liu *et al.*, 2008).

Before choosing the EIF method, this study compared it with two other outlier detection methods: the Local Outlier Factor (LOF) and the Interquartile Range (IQR). The Local Outlier Factor (LOF) algorithm identifies anomalies by comparing the local density of a point with that of its neighbors (Breunig *et al.*, 2000). The LOF score for a data point  $x$  is defined as (Breunig *et al.*, 2000):

$$LOF_k(x) = \frac{\sum_{o \in N_k(x)} lrd_k(o) / lrd_k(x)}{|N_k(x)|} \quad (26)$$

where the local reachability density of a point  $x$  is defined as:

$$lrd_k(x) = \frac{|N_k(x)|}{\sum_{o \in N_k(x)} reachability\ distance(x, o)} \quad (27)$$

with  $N_k(x)$  being the  $k$ -nearest neighbors of  $x$ , and the reachability distance between points  $x$  and  $o$  defined as:

$$reachability\ distance(x, o) = \max(dist(x, o), dist(x, N_k(o))) \quad (28)$$

The central idea is that points with significantly lower density than their neighbors are considered outliers. The LOF requires calculating each data point's  $k$ -nearest neighbors ( $k$ -NN), which involves measuring distances, often Euclidean, between points (Breunig *et al.*, 2000). Like EIF, no closed-form formula exists; instead, the LOF score is a numerical result derived from these distance comparisons (Breunig *et al.*, 2000).

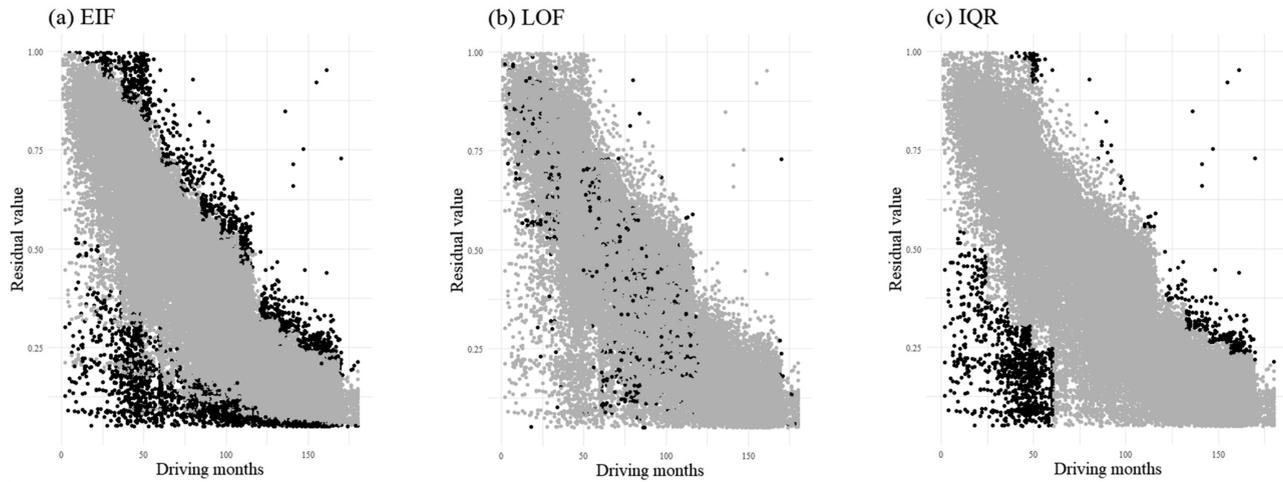
In the meantime, the IQR method is a statistical technique based on quartiles. It identifies outliers by comparing the spread of the middle 50% of the data. The IQR is calculated as:

$$IQR = Q3 - Q1 \quad (29)$$

where  $Q1$  and  $Q3$  are the first and third quartiles, respectively. Outliers are then identified as any data points falling below  $Q1 - 1.5 \times IQR$  or above  $Q3 + 1.5 \times IQR$ . This threshold helps detect points that lie significantly beyond the typical spread of the data.

The Chi-square testing score for the outlier detection methods was 2010.7, with a degree of freedom of 28, calculated by multiplying 14 by 2. The 14 was the number of years or driving month groups minus one, and 2 was the number of outlier detection methods minus one. The  $p$ -value of the Chi-square testing was smaller than  $2.2e-16$ , indicating significant differences in the number of outliers identified by the three methods. This implies that the methods did not perform equally when detecting outliers.

The  $t$ -testing results also showed significant differences among the three methods. The  $t$ -value of EIF was  $-18.601$  with a degree of freedom of 3243.4 and a  $p$ -value lower than  $2.2e-16$ . The  $t$ -value of IQR was  $-31.101$  with a degree of freedom of 1,503 and a  $p$ -value lower than  $2.2e-16$ , too. In contrast, the  $t$ -value of LOF was 8.099 with a degree of freedom of 1883.9 and a  $p$ -value lower than  $9.849e-16$ . The EIF and IQR methods detected lower observations as outliers more than the LOF method.



**Note(s):** Black dots are outliers detected by each method

**Source(s):** Figure by author

**Figure 3.** Comparison of three outlier detection methods

Figure 3 visually compares three outlier detection methods for Hyundai Sonata’s used car transactions. It indicates that EIF detected upper and lower outliers better than the IQR method. The LOF method did not simply detect outliers based on whether a value was too large or too small but primarily focused on density-based detection. Because of this, values appearing in the middle of the data distribution could still be considered outliers if the density of the region they are located in significantly differed from their surrounding points. Considering the Chi-square and *t*-testing results and visual comparison of the three methods, this study selected the EIF as the outlier detection method.

**Table 1.** Examined car models’ transaction volumes and outlier elimination

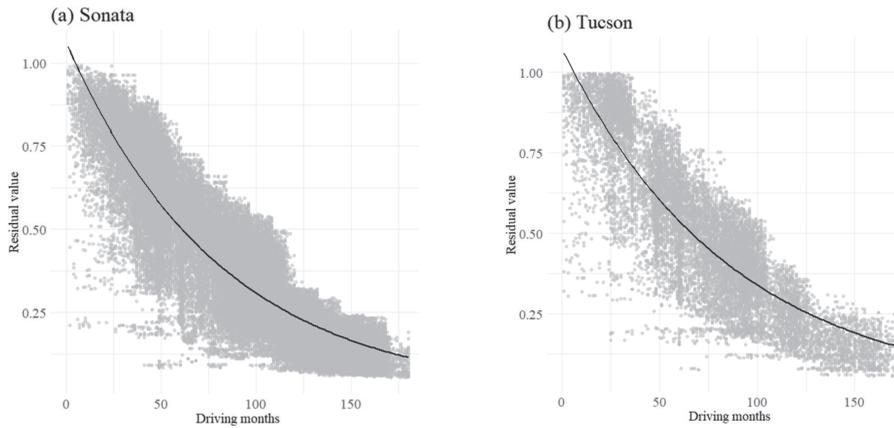
	Sonata	Tucson
① Number of used car transactions in 2023	172,897	68,507
② Immediate sale after purchase	187	246
③ Longer than 180 months (or 15 years)	16,653	4,878
④ Internal sales to employees	93,101	46,756
⑤ Cars driving more than 20,000 km per year	19,296	6,031
⑥ Outliers	3,028	789
Total (① - ② - ③ - ④ - ⑤ - ⑥)	57,472	14,931

**Source(s):** The author

Table 1 shows the details of the data filtering results. A significant number of transactions were outliers. In particular, 93,101 and 46,756 units of Sonata and Tucson transactions were cars sold to employees. The numbers looked reasonable because the two models are among the most popular cars for employees, and the number of Hyundai Motor conglomerates in South Korea was 188,891 as of 2022, according to the Fair Trade Commission of South Korea. Most employees buy cars every two years thanks to the employee discount policy. If retired employees eligible for discounted purchases are included, the number of employee discount candidate buyers will be significantly more considerable. Figure 4 shows the scatter plots of residual values by driving months after eliminating outliers.

**7. Testing results**

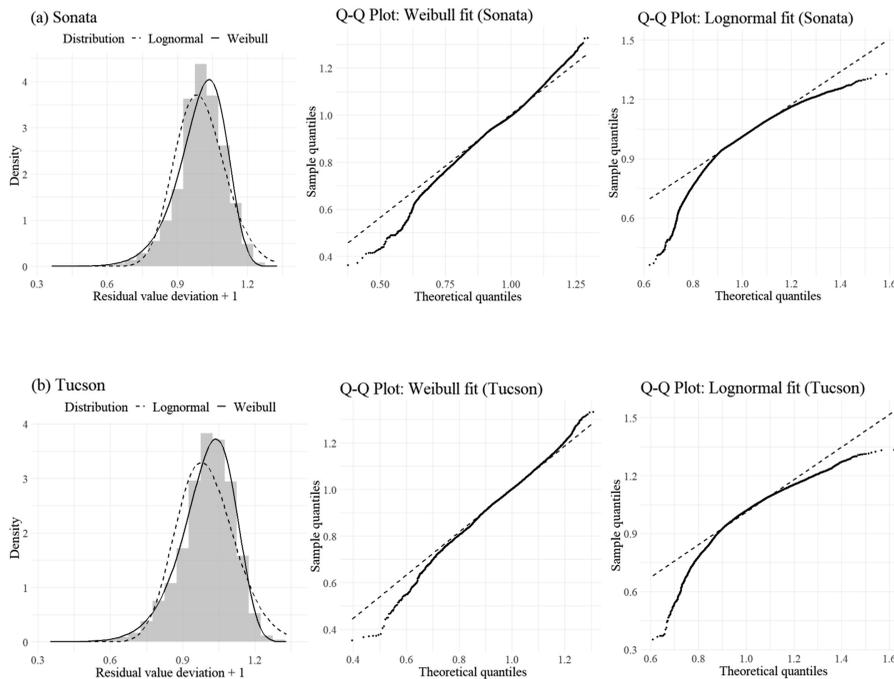
This study tested the model performance in two ways: model fit and prediction accuracy. For this purpose, the Weibull and lognormal parameters were estimated using the maximum likelihood method. The numeric number one was added to all the residual value deviations to make all the values greater than zero, critical to applying lognormal and Weibull distributions. Figure 5 shows the estimated Weibull and lognormal functions with the histograms of the residual value deviations. The details of the Weibull parameter estimation for two car models are in Appendix.



**Note(s):** Sonata's nonlinear regression model is residual value =  $1.06299 \times \exp(-0.01238 \times \text{driving\_months})$ . Tucson's nonlinear regression equation is residual value =  $1.07142 \times \exp(-0.01149 \times \text{driving\_months})$ . The estimated values were used as asset price movements for option valuation

**Source(s):** Figure by author

**Figure 4.** Residual values and fitted curves aging less than or equal to 180 months



**Source(s):** Figure by author

**Figure 5.** Estimated Weibull and lognormal distribution functions fitted to the residual value deviations plus one

**Table 2.** Performance testing results

	Sonata Weibull	Lognormal	Tucson Weibull	Lognormal
New car price*	28,080,000		27,710,000	
Asset price (S0) **	10398579.52		10261561.20	
Strike price (K)	13352040.41		13176105.40	
Annual standard deviation		13.49%		14.59%
Beta	0.6105587		0.597323	
Delta	11.422843		7.982199	
Risk-free interest rate	5%		5%	
Time to maturity (years)***	5		5	
Put option price (Korean Won)	580580.53	1230718.89	722456.98	1329775.42
<i>Prediction accuracy</i>				
Mean square error (MSE)	0.02169159	0.0223690	0.0259642	0.0281327
Mean absolute error (MAE)	0.1150858	0.1180491	0.1258247	0.1325515
<i>Model fit</i>				
AIC	-101533.05	-92581.37	-23685.70	-20459.81
BIC	-101515.13	-92563.45	-23670.48	-20444.59
Log-likelihood	50768.52	46292.68	11844.85	10231.91
<b>Note(s):</b> *The most common price of each model was selected as the new car price. **The asset prices are the present values of the strike prices. ***Five-year leases are the most common leasing terms in South Korea				
<b>Source(s):</b> The author				

Table 2 shows the performance testing results and the put option prices with five-year maturities where the strike price equals the expected values at time  $T$ . The put option price with Weibull distribution for Sonata was 580,580.53 Korean Won, 52.83% lower than the put option price with lognormal distribution, 1,230,718.89 Korean Won. This apparent contradiction arises because Weibull distributions are left-skewed, capturing more extreme losses. The primary reason is the overestimated volatility in the Black-Scholes formula: The Black-Scholes formula assumes that the volatility increases by the square root of time (Hull and Basu, 2016). However, the actual volatility of used car prices did not follow the square root of time relationship.

The Sonata residual values' annual standard deviations for the five years were 0.1581, 0.1333, 0.1242, 0.1174, and 0.1417. Given the relatively stable fluctuations over time, this study applied 0.1349, the arithmetic average of these five years' annual standard deviations, as the annual volatility for the Black-Scholes option valuation. The Tucson residual values also showed a similar volatility pattern as the Sonata's. Their annual standard deviations were 0.1404, 0.1351, 0.1619, 0.1351, and 0.1571. The arithmetic average was 0.1459.

However, the arithmetic average for volatility calculation does not fully align with theoretical assumptions. This demonstrates that the Black-Scholes model is theoretically and practically invalid for valuing options in car lease assets. The observed stability in residual value volatility underscores the mismatch between the Black-Scholes model's volatility assumptions and the actual behavior of used cars. Although applying stochastic volatility could mitigate some errors, it does not fully address the fundamental inaccuracies stemming from the lognormal distribution assumption either. This further emphasizes the need for alternative models tailored to the unique characteristics of depreciating assets like car residual values.

The Weibull distribution provided a better fit than the lognormal distribution for estimating residual value deviations (see Table 2). Model performance was assessed using the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). The AIC and BIC are commonly employed to compare models by balancing goodness of fit with model complexity. A lower AIC or BIC value indicates a better trade-off between model fit and complexity, helping to prevent overfitting.

While AIC imposes a penalty for models with more parameters to prevent overfitting, BIC applies a stricter penalty by accounting for sample size, favoring simpler models as the sample size increases (Ding *et al.*, 2017; Liu and Zhou, 2024). The Weibull distribution is characterized by two parameters: shape and scale. Similarly, the lognormal distribution has two parameters — mean and standard deviation. As a result, the penalty terms for both distributions are equivalent. For the Sonata's residual value distribution, the log-likelihood values for the Weibull and lognormal distributions were 50768.52 and 46292.68, respectively. Consequently, the Weibull distribution yielded lower AIC and BIC values. For Tucson's residual value distribution, the log-likelihood values were 11844.85 for the Weibull distribution and 10231.91 for the lognormal distribution. As with the Sonata, the Weibull distribution resulted in lower AIC and BIC values for the Tucson.

Mean Squared Error (MSE) and Mean Absolute Error (MAE) were employed as evaluation metrics to assess the prediction accuracy, focusing on model performance rather than option valuation accuracy since no market prices exist. MSE measures the average of the squared differences between predicted and actual values, placing a higher penalty on larger errors, while MAE calculates the average of the absolute differences, providing a more straightforward interpretation of the model's performance. By comparing both metrics, it is possible to understand the overall error magnitude and how sensitive the model is to large deviations (Chai and Draxler, 2014; Willmott and Matsuura, 2005). In this study, a Monte Carlo simulation with 1,000 iterations was conducted to ensure robust performance evaluation. The results showed that the Weibull distributions generated more accurate estimations.

## 8. Conclusion

Options valuation is one of the most critical decision-making factors for auto finance companies and their consumers. However, traditional option valuation models have struggled to accurately capture the complex pricing mechanisms embedded in car leases. This limitation has led researchers and industry practitioners to explore non-parametric Bayesian methods and empirical statistical measures to estimate the dynamics of used car value uncertainties instead of relying on the Black-Scholes formula. While Bayesian statistics offer flexibility and powerful tools for uncertainty estimation, they also present challenges, such as high computational demands and sensitivity to prior assumptions. As a result, in some cases, parametric statistical approaches like the Weibull distribution are more suitable for modeling the distinct characteristics of car residual values due to their simplicity and interpretability. This study proposes a put option valuation formula using the Weibull distribution, suggesting that this approach provides a better fit than traditional models based on the lognormal distribution.

Since this study is the first attempt to model put option values under the assumption of a Weibull distribution, the field remains open for further exploration. One critical area for future research is the behavior of volatility. Unlike stock prices, used car values experience heightened volatility in the early stages, eventually stabilizing as the car ages. Understanding such stochastic volatility of residual values could enable the auto finance industry to manage both short- and long-term leases more effectively. Furthermore, while stock prices exhibit unlimited upside potential, making the lognormal distribution a reasonable assumption, used car values are constrained in their movements. They have a maximum value (the price of a new car) and can reach a minimal bottom price (scrap value). Comparing the lognormal distribution with the Weibull distribution could provide more practical insights into the auto leasing market by enhancing the understanding of residual value dynamics.

As many auto manufacturers and finance companies increasingly adopt Mobility as a Service (MaaS) as part of their future strategies (Brown *et al.*, 2022; Weber, 2022), there is a growing imperative within the auto industry to conduct a more comprehensive examination of used car values and associated risk management practices. This shift is not merely a response to evolving consumer preferences but also a recognition of the complexities involved in

valuing used cars in a dynamic market characterized by fluctuating demand and uncertain residual values.

Moreover, the implications of this research extend well beyond the realms of auto finance and MaaS. The insights gained from this study can be applied to various sectors where the underlying assets exhibit skewness or bounded dynamics. For instance, insurance products with asymmetric payout structures, such as life insurance or warranties, often entail risks that can be better modeled using Weibull distribution approaches. Additionally, financial assets that demonstrate non-linear risks, such as certain derivatives or structured products, could greatly benefit from the enhanced modeling capabilities that the Weibull framework provides.

## References

- Bastianello, L., Chateaufneuf, A. and Cornet, B. (2024), "Put-call parities, absence of arbitrage opportunities, and nonlinear pricing rules", *Mathematical Finance*, Vol. 34 No. 4, pp. 1242-1262, doi: [10.1111/mafi.12433](https://doi.org/10.1111/mafi.12433).
- Black, F. and Scholes, M. (1973), "The pricing of options and corporate liabilities", *Journal of Political Economy*, Vol. 81 No. 2, pp. 637-654, doi: [10.1142/9789814759588\\_0001](https://doi.org/10.1142/9789814759588_0001).
- Breunig, M.M., Kriegel, H.P., Ng, R.T. and Sander, J. (2000), "LOF: identifying density-based local outliers", *Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data*, ACM, New York, NY, pp. 93-104, doi: [10.1145/335191.335388](https://doi.org/10.1145/335191.335388).
- Brown, C., Hardman, M., Davies, N. and Armitage, R. (2022), "Mobility as a service: defining a transport utopia", *Future Transport*, Vol. 2 No. 1, pp. 300-309, doi: [10.3390/futuretransp2010016](https://doi.org/10.3390/futuretransp2010016).
- Chai, T. and Draxler, R.R. (2014), "Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature", *Geoscientific Model Development*, Vol. 7 No. 3, pp. 1247-1250, doi: [10.5194/gmd-7-1247-2014](https://doi.org/10.5194/gmd-7-1247-2014).
- Chen, Q. and Gerlach, R.H. (2013), "The two-sided Weibull distribution and forecasting financial tail risk", *International Journal of Forecasting*, Vol. 29 No. 4, pp. 527-540, doi: [10.1016/j.ijforecast.2013.01.007](https://doi.org/10.1016/j.ijforecast.2013.01.007).
- Dasgupta, S. (2004), "The determinants and implications of lease and buy strategies for consumer durables", Doctoral dissertation, University of Southern California.
- Ding, J., Tarokh, V. and Yang, Y. (2017), "Bridging AIC and BIC: a new criterion for autoregression", *IEEE Transactions on Information Theory*, Vol. 64 No. 6, pp. 4024-4043, doi: [10.1109/TIT.2017.2717599](https://doi.org/10.1109/TIT.2017.2717599).
- Gamba, A. and Rigon, R. (2008), "The value of embedded real options: evidence from consumer automobile lease contracts—a note", *Finance Research Letters*, Vol. 5 No. 4, pp. 213-220, doi: [10.1016/j.frl.2008.08.003](https://doi.org/10.1016/j.frl.2008.08.003).
- Giacotto, C., Goldberg, G.M. and Hegde, S.P. (2007), "The value of embedded real options: evidence from consumer automobile lease contracts", *The Journal of Finance*, Vol. 62 No. 1, pp. 411-445, doi: [10.1111/j.1540-6261.2007.01211.x](https://doi.org/10.1111/j.1540-6261.2007.01211.x).
- Gómez, Y.M., Gallardo, D.I., Marchant, C., Sánchez, L. and Bourguignon, M. (2023), "An in-depth review of the Weibull model with a focus on various parameterizations", *Mathematics*, Vol. 12 No. 1, p. 56, doi: [10.3390/math12010056](https://doi.org/10.3390/math12010056).
- Heo, S.J., Yeo, S.C. and Kang, T.H. (2012), "Performance analysis of economic VaR estimation using risk neutral probability distributions", *The Korean Journal of Applied Statistics*, Vol. 25 No. 5, pp. 757-773, doi: [10.5351/KJAS.2012.25.5.757](https://doi.org/10.5351/KJAS.2012.25.5.757).
- Hong, J. (2022), "The optimal insurance demand under an ambiguity aversion", *Journal of Derivatives and Quantitative Studies: 선물연구*, Vol. 30 No. 4, pp. 296-308, doi: [10.1108/JDQS-06-2022-0014](https://doi.org/10.1108/JDQS-06-2022-0014).
- Hull, J.C. and Basu, S. (2016), *Options, Futures, and Other Derivatives*, Pearson Education, India.

- Kim, J., Park, Y.J. and Truong, T.T.T. (2023a), "Retail investors and overpricing of left-tail risk: evidence from the Korean stock market", *Journal of Derivatives and Quantitative Studies: 선물연구*, Vol. 31 No. 4, pp. 309-327, doi: [10.1108/JDQS-04-2023-0008](https://doi.org/10.1108/JDQS-04-2023-0008).
- Kim, M., Choi, J., Kim, J., Kim, W., Baek, Y., Bang, G., Son, K., Ryou, Y. and Kim, K. (2023b), "Trustworthy residual vehicle value prediction for auto finance", *AI Magazine*, Vol. 44 No. 4, pp. 394-405, doi: [10.1002/aaai.12136](https://doi.org/10.1002/aaai.12136).
- Krause, T. (2019), "Put-call parity in equity options markets: recent evidence", *Theoretical Economics Letters*, Vol. 9 No. 04, pp. 563-569, doi: [10.4236/tel.2019.94039](https://doi.org/10.4236/tel.2019.94039).
- Li, X. (2020), "A limited memory BFGS subspace algorithm for bound constrained nonsmooth problems", *Journal of Inequalities and Applications*, Vol. 2020 No. 1, p. 135, doi: [10.1186/s13660-020-02398-6](https://doi.org/10.1186/s13660-020-02398-6).
- Li, K. and You, L. (2022), "Flexibility, option value of leasing, and investment", *SSRN*, doi: [10.2139/ssrn.4381925](https://doi.org/10.2139/ssrn.4381925).
- Li, Y., Nolte, I. and Pham, M.C. (2024), "Parametric risk-neutral density estimation via finite lognormal-Weibull mixtures", *Journal of Econometrics*, Vol. 241 No. 2, 105748, doi: [10.1016/j.jeconom.2024.105748](https://doi.org/10.1016/j.jeconom.2024.105748).
- Lin, H., Liu, L. and Zhang, Z. (2023), "Tail risk signal detection through a novel EGB2 option pricing model", *Mathematics*, Vol. 11 No. 14, p. 3194, doi: [10.3390/math11143194](https://doi.org/10.3390/math11143194).
- Liu, S. and Zhou, D.J. (2024), "Using cross-validation methods to select time series models: promises and pitfalls", *British Journal of Mathematical and Statistical Psychology*, Vol. 77 No. 2, pp. 337-355, doi: [10.1111/bmsp.12330](https://doi.org/10.1111/bmsp.12330).
- Liu, F.T., Ting, K.M. and Zhou, Z.H. (2008), "Isolation forest", *2008 Eighth IEEE International Conference on Data Mining*, IEEE, pp. 413-422.
- Mauler, D.J. and McDonald, J.B. (2015), "Option pricing and distribution characteristics", *Computational Economics*, Vol. 45 No. 4, pp. 579-595, doi: [10.1007/s10614-014-9441-z](https://doi.org/10.1007/s10614-014-9441-z).
- McConnell, J.J. and Schallheim, J.S. (1983), "Valuation of asset leasing contracts", *Journal of Financial Economics*, Vol. 12 No. 2, pp. 237-261, doi: [10.1016/0304-405X\(83\)90037-5](https://doi.org/10.1016/0304-405X(83)90037-5).
- Miller, S.E. (1995), "Economics of automobile leasing: the call option value", *Journal of Consumer Affairs*, Vol. 29 No. 1, pp. 199-218, doi: [10.1111/j.1745-6606.1995.tb00045.x](https://doi.org/10.1111/j.1745-6606.1995.tb00045.x).
- Modarres, M., Kaminskiy, M.P. and Krivtsov, V. (2016), *Reliability Engineering and Risk Analysis: A Practical Guide*, CRC Press, Boca Raton, FL.
- Oppenheimer, P. (2002), "Including real options in evaluating terminal cash flows in consumer auto leases", *Financial Services Review*, Vol. 11 No. 2, pp. 135-151.
- Pierce, L. (2007), "Manufacturer lease pricing and adverse selection", available at: *SSRN*, No. 1157086, doi: [10.2139/ssrn.1157086](https://doi.org/10.2139/ssrn.1157086).
- Rousseeuw, P.J. (1987), "Silhouettes: a graphical aid to the interpretation and validation of cluster analysis", *Journal of Computational and Applied Mathematics*, Vol. 20, pp. 53-65, doi: [10.1016/0377-0427\(87\)90125-7](https://doi.org/10.1016/0377-0427(87)90125-7).
- Savickas, R. (2002), "A simple option-pricing formula", *Financial Review*, Vol. 37 No. 2, pp. 207-226, doi: [10.1111/1540-6288.00012](https://doi.org/10.1111/1540-6288.00012).
- Speagle, J.S. (2019), "A conceptual introduction to Markov chain Monte Carlo methods", *arXiv preprint, arXiv:1909.12313*, doi: [10.48550/arXiv.1909.12313](https://doi.org/10.48550/arXiv.1909.12313).
- Starkey, W.S. (1997), "Automobile leasing in the United States: why do consumers lease cars?", [Doctoral dissertation], University of Washington.
- Steinley, D. and Brusco, M.J. (2007), "Initializing K-means batch clustering: a critical evaluation of several techniques", *Journal of Classification*, Vol. 24 No. 1, pp. 99-121, doi: [10.1007/s00357-007-0003-0](https://doi.org/10.1007/s00357-007-0003-0).
- Trigeorgis, L. (1996), "Evaluating leases with complex operating options", *European Journal of Operational Research*, Vol. 91 No. 2, pp. 315-329, doi: [10.1016/0377-2217\(95\)00288-X](https://doi.org/10.1016/0377-2217(95)00288-X).

Van Ravenzwaaij, D., Cassey, P. and Brown, S.D. (2018), "A simple introduction to Markov chain Monte Carlo sampling", *Psychonomic Bulletin and Review*, Vol. 25 No. 1, pp. 143-154, doi: [10.3758/s13423-016-1015-8](https://doi.org/10.3758/s13423-016-1015-8).

Weber, J. (2022), "Mobility as a service", in *Moving Times*, Springer, Wiesbaden, pp. 161-228.

Weibull, W. (1951), "A statistical distribution function of wide applicability", *Journal of Applied Mechanics*, Vol. 18 No. 3, pp. 293-297, doi: [10.1115/1.4010337](https://doi.org/10.1115/1.4010337).

Willmott, C.J. and Matsuura, K. (2005), "Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance", *Climate Research*, Vol. 30 No. 1, pp. 79-82, doi: [10.3354/cr030079](https://doi.org/10.3354/cr030079).

Xie, Y., Byrd, R.H. and Nocedal, J. (2020), "Analysis of the BFGS method with errors", *SIAM Journal on Optimization*, Vol. 30 No. 1, pp. 182-209, doi: [10.1137/19M1240794](https://doi.org/10.1137/19M1240794).

**Appendix**

**A.1. Derivation of Weibull parameters**

The Weibull distribution is widely defined as (Modarres *et al.*, 2016):

$$f(x; \theta, \delta) = \frac{\delta}{\theta} \left(\frac{x}{\theta}\right)^{\delta-1} \exp\left[-\left(\frac{x}{\theta}\right)^\delta\right], x \in \mathbb{R}_{\geq 0}, \theta, \delta \in \mathbb{R}^+$$

where  $\theta$  and  $\delta$  are scale and shape parameters, respectively. This formula is a modified version of the one originally presented by Weibull (1951):

$$f(x; \beta, \delta) = \beta \delta x^{\delta-1} e^{-\beta x^\delta}$$

If  $\left(\frac{1}{\theta}\right)^\delta$  is replaced with  $\beta$ , then

$$\begin{aligned} & \frac{\delta}{\theta} \left(\frac{x}{\theta}\right)^{\delta-1} \exp\left[-\left(\frac{x}{\theta}\right)^\delta\right] \\ &= \left(\frac{1}{\theta}\right)^\delta \delta x^{\delta-1} \exp\left[-\left(\frac{1}{\theta}\right)^\delta x^\delta\right] \\ &= \beta \delta x^{\delta-1} e^{-\beta x^\delta} \end{aligned}$$

Table A1 shows the scale ( $\theta$ ) and shape ( $\delta$ ) parameters calculated in R, and  $\beta$  and  $\delta$  used in this study. The variance of the beta ( $\beta$ ) were calculated using the covariance matrix and partial derivative. The partial derivatives are:

$$\begin{aligned} \frac{\partial \beta}{\partial \text{scale}} &= -\text{shape} \times \left(\frac{1}{\text{scale}}\right)^{\text{shape}+1} \\ \frac{\partial \beta}{\partial \text{shape}} &= \left(\frac{1}{\text{scale}}\right)^{\text{shape}} \times \ln\left(\frac{1}{\text{scale}}\right) \end{aligned}$$

The variance of beta is:

$$\left(\frac{\partial\beta}{\partial\text{shape}}\right)^2 \text{var}(\text{shape}) + \left(\frac{\partial\beta}{\partial\text{scale}}\right)^2 \text{var}(\text{scale}) + 2\left(\frac{\partial\beta}{\partial\text{shape}}\right)\left(\frac{\partial\beta}{\partial\text{scale}}\right) \text{cov}(\text{shape}, \text{scale})$$

**Table A1.** Parameter estimation of the Weibull distribution

	Parameters in R		Parameters in this study	
	Scale ( $\theta$ )	Shape ( $\delta$ )	Beta ( $\beta$ )	Delta ( $\delta$ )
<i>Sonata</i>	1.044139	11.422843	0.6105587	11.422843
Standard errors	0.0004021	0.0361197	0.0019063	0.0361197
Z-value	2596.54	316.25	320.29	316.25
<i>Tucson</i>	1.04806	10.55960	0.6091565	10.55960
Standard error	0.0008551	0.0662933	0.0037234	0.0662933
Z-value	1225.62	159.29	163.60	159.29

**Source(s):** The author

### A.2. Proof of put option formula

The put option value derived in this study can be proved using the put-call parity using the call option formula Savickas' (2002) derived. The call option formula is

$$C = S_0 \left[ 1 - F_{2\left(1+\frac{1}{\delta}\right)}(2\beta K^\delta) \right] - Ke^{-rt} e^{-\beta K^\delta}$$

where  $F_{2\left(1+\frac{1}{\delta}\right)}$  is the  $\chi^2$  CDF at point  $2\beta K^\delta$  with the degrees of freedom  $2\left(1 + 1/\delta\right)$ , and no dividend is assumed. The put option formula this study derived is

$$P = e^{-rT} KF_{Weibull}(K; \beta, \delta) - S_0 \gamma \left( 1 + \frac{1}{\delta}, \beta K^\delta \right) / \Gamma \left( 1 + \frac{1}{\delta} \right)$$

Because the put-call parity is  $C - P = S_0 - e^{-rT}K$ ,

$$\begin{aligned} C - P &= S_0 \left[ 1 - F_{2\left(1+\frac{1}{\delta}\right)}(2\beta K^\delta) \right] - Ke^{-rT} e^{-\beta K^\delta} \\ &\quad - e^{-rT} KF_{Weibull}(K; \beta, \delta) + S_0 \gamma \left( 1 + \frac{1}{\delta}, \beta K^\delta \right) / \Gamma \left( 1 + \frac{1}{\delta} \right) \end{aligned}$$

Thus,

$$\begin{aligned} C - P &= S_0 \left[ 1 - F_{2\left(1+\frac{1}{\delta}\right)}(2\beta K^\delta) + \gamma \left( 1 + \frac{1}{\delta}, \beta K^\delta \right) / \Gamma \left( 1 + \frac{1}{\delta} \right) \right] \\ &\quad - e^{-rT} K \left[ F_{Weibull}(K; \beta, \delta) + e^{-\beta K^\delta} \right] \end{aligned}$$

$$\text{Since } F_{2\left(1+\frac{1}{\delta}\right)}(2\beta K^\delta) = \frac{\gamma\left(1+\frac{1}{\delta}, \beta K^\delta\right)}{\Gamma\left(1+\frac{1}{\delta}\right)},$$

$$C - P = S_0 - e^{-rT} K \left[ F_{Weibull}(K; \beta, \delta) + e^{-\beta K^\delta} \right]$$

$$\text{Since } F_{Weibull}(K; \beta, \delta) = 1 - e^{-\beta K^\delta}$$

$$F_{Weibull}(K; \beta, \delta) + e^{-\beta K^\delta} = 1 - e^{-\beta K^\delta} + e^{-\beta K^\delta} = 1$$

In conclusion,

$$C - P = S_0 - e^{-rT} K$$

**Corresponding author**

Seok Bin Ko can be contacted at: [Sbko@cju.ac.kr](mailto:Sbko@cju.ac.kr)