Working Paper (2024) 1: 1–24
doi: 123.45.6789
Advance Access publication: 1 October 2024

Pricing D6 RINs Under Imperfect Fungability*

Tucker Strow[1] and Todd Griffith[2]
[1] *Utah State MFE Grad Student;* [2] *Utah State Finance Faculty*

Abstract. This paper examines the pricing dynamics of D6 Renewable Identification Numbers (RINs), highlighting the complexities that arise due to imperfect fungibility. While existing literature often focuses on the price spread between ethanol and gasoline as a primary determinant of RIN prices, this study introduces a more nuanced model that accounts for additional factors such as RIN generation year, Quality Assurance Program (QAP) verification status, volatility in ethanol and gasoline prices, and seasonal trading patterns. The results suggest that past RIN prices, energy price volatility, verification status, RIN age, and transaction month all play significant roles in determining current RIN values, offering important insights for traders and policymakers alike. This research challenges traditional pricing models and provides a framework for better understanding the dynamics of the D6 RIN market.

JEL Classification: NA, NaN, NULL

1. Introduction

Renewable Identification Numbers (RINs) represent the currency of a multi-billion-dollar market that, despite its size and significance, has received surprisingly little attention from academics and financial professionals. Born out of the compliance requirements set by the Renewable Fuel Standard (RFS), the RIN market offers a wealth of untapped data, revealing pricing inefficiencies, risk management strategies, and the profound impact of regulatory frameworks on market behavior.

*The authors would like to thank the Center for Growth and Opportunity at Utah State for support.

For scholars and researchers in finance — especially those focused on environmental economics, market design, and trading strategies — studying the RIN market presents a unique opportunity to explore how regulation-driven markets function, evolve, and influence broader economic systems. With its intricate interplay between policy and finance, the RIN market is an ideal subject for deeper academic exploration.

As an example of the insufficiency of current literature, RINs — specifically D6 RINs — are often viewed as having intrinsic prices attached to the spread between ethanol and gasoline. While this is a valid assumption given that RINs are intended to measure blending compliance, it ignores the primary exigency of RINs, which is to facilitate Business-to-Business (B2B) cooperation in achieving environmental compliance goals. The B2B aspect of these commodities is highly relevant as it introduces conflation between cyclical business objectives and RIN trading patterns. Furthermore, pricing models that focus nearly exclusively on ethanol and gasoline prices neglect that RINs are not perfectly fungible due to differences in generation years and verification status. Taking this into consideration, it would be more appropriate to estimate D6 RIN price as a function of ethanol and gasoline prices, QAP-Verification status, RIN generation year, and when RINs are traded.

2. Literature Review

The existing literature offers limited exploration of RIN pricing as a function of business cycles and the intrinsic attributes of RINs themselves. Instead, the focus tends to be on pricing in relation to gasoline and ethanol spreads (Pouliot and Babcock 2016), (Whistance and Thompson 2014), (Thompson et al. 2010), on the intrinsic and option values of RINs (Afkhami and Ghoddusi 2022), on the relationship between different RIN categories (Irwin et al. 2020), or as a tax on domestic fossil fuel consumption (Lade and Lawell 2015), (Pouliot and Babcock 2016). While many of these works acknowledge the possibility of speculation in the RIN market, they stop short of investigating the arbitrage opportunities that might arise or the differences in separated, tradable credits. Instead, most studies use their models to forecast the compliance costs associated with the Renewable Fuel Standard (RFS).

Previous studies have examined how variations in mandates, ethanol supply, and gasoline prices influence RIN pricing, offering valuable models for understanding price fluctuations and compliance costs (Pouliot and Babcock 2016), (Whistance and Thompson 2014), (Thompson et al. 2010). These studies underscore the dual role of RINs as both a compliance mechanism and an economic incentive, driving behavior among market participants. This understanding is critical for analyzing RIN price formation and policy impacts, and lays a great foundation for understanding intrinsic RIN pricing. However, in the context of RIN trading, these models often overlook important factors, such as seasonal differences in RIN production and trading, speculation and hoarding of RINs, or the imperfect fungibility of different RINs. Addressing these gaps is essential for a more comprehensive understanding of market dynamics and pricing.

Building on this foundational work, more recent contributions, such as Afkhami and Ghoddusi (2021), have begun to explore RIN dynamics in greater detail. Their theoretical model investigates how blenders strategically respond to ethanol price volatility by adjusting their RIN generation and usage. Specifically, they argue that when ethanol prices are low, blenders overproduce and accumulate RINs, which they can later use when prices rise. Their model also highlights the option-like value of RINs, demonstrating how price volatility in ethanol and corn increases RIN prices because they offer certainty in uncertain markets. These findings suggest that, in studying RIN price dynamics and trading behavior, it is crucial to account for price uncertainty and the inherent option value of RINs as a risk management tool. As a whole, this work underscores the importance of considering feedstock price volatility when determining a RIN's value, particularly in relation to D6 RIN pricing.

An example of a further RIN market complication is the influence of other renewable fuel categories, such as biodiesel, on D6 RIN prices. Parts of the literature explore the relationship between biodiesel RINs (D4 RINs) and ethanol RINs, noting that the price of D4 RINs effectively sets a ceiling on D6 RIN prices (Irwin et al. 2020). This relationship exists because obligated parties under the RFS have some flexibility in meeting their RVOs through the blending of different types of renewable fuels, in accordance with the RIN nesting structure (Figure 1). If D4 RINs become cheaper than D6 RINs, obligated parties may choose to comply with their RVOs through biodiesel instead of ethanol, thus capping the demand for D6 RINs. In this context, the volatility observed in D6 RIN markets can stem from external factors influencing the supply and demand of biodiesel, such as soybean oil prices or changes in the biodiesel tax credit.

Additionally, some studies analyze RIN prices as a form of tax on gasoline production that simultaneously subsidizes ethanol consumption (Pouliot and Babcock 2016). The higher the RIN price, the greater the cost for refiners to meet their Renewable Volume Obligations (RVOs) by purchasing RINs, thus incentivizing them to blend more ethanol instead. This relationship underscores the exigency of RINs as both a compliance mechanism and an economic lever to promote biofuel blending. However, they argue, the complexity of this system may lead to unintended consequences, such as higher gasoline prices passed on to consumers and supply chain disruptions within the biofuel market.

As this literature review highlights, there has been a multitude of papers exploring D6 RIN pricing mechanisms. However, the dramatic assumption made in all of them is that D6 RINs are perfectly fungible — that is, each D6 RIN is equally as valuable as any other D6 RIN sold in that compliance year (except for the option value explored by Afkhami and Ghoddusi (2021)) . However, this is not the case in practice. While past studies have provided valuable insights into compliance costs and RIN pricing based on fuel spreads, they often fail to explore external factors that may cause deviations from intrinsic values. Our paper will address this gap by focusing on business cycle fluctuations and differences within individual RINs. By doing so, it will contribute to a more nuanced understanding of how RIN prices are influenced by both intrinsic attributes and broader market conditions, such as QAP verification status, RIN age, and the month of the RIN transaction.

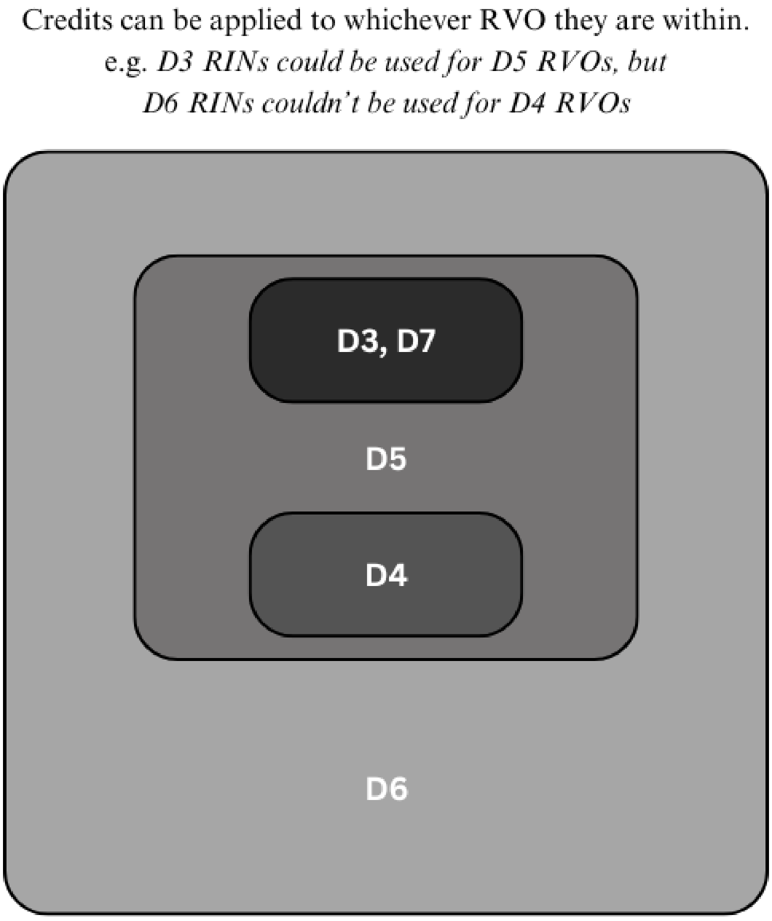

Fig. 1. RIN Nesting Structure

3. Model

3.1 Determining Intrinsic Value

After an extensive review of the literature, it becomes clear that modeling D6 RIN prices as a function of ethanol-gasoline spreads is a widely accepted and practical approach for determining their intrinsic value. In this framework, the price of a RIN reflects the point at which refiners are indifferent between blending more ethanol or buying RINs to fulfill their regulatory obligations. It is worth noting that in our models we chose to use log RIN prices due to the fact that prices are non-negative and greater than zero, distorting their distribution.

The simplest model uses the volumetric price difference between ethanol and gasoline (RBOB), as defined by:

$$lnP_{RIN} = \beta_0 + \beta_1 \times Spread_{Vol} + \epsilon \tag{1}$$

where:

$$Spread_{Vol} = P_{Ethanol} - P_{RBOB}$$

However, because ethanol contains approximately 66.7% of the energy content of gasoline, an energy-equivalent spread may better capture the true substitution relationship between ethanol and gasoline. Therefore an adjusted model would be as follows:

$$lnP_{RIN} = \beta_0 + \beta_1 \times Spread_{En} + \epsilon \tag{2}$$

where:

$$Spread_{En} = P_{Ethanol} - P_{RBOB} \times 0.667$$

An alternative approach, sometimes used in the literature (Pouliot and Babcock 2016), is to use a more aggressive energy content adjustment, treating ethanol as equivelent to 110% of gasoline's price. This decision is arrived to due to the equilibrium price at which it's efficient to achieve 87 octane gas (the anti-knock minimum for most low-elevation states) by blending ethanol. Beyond this point, it becomes uneconomical to blend ethanol and companies instead use alternative methods to achieve higher octanes (absent a binding mandate). This adjustment gives the following equation:

$$lnP_{RIN} = \beta_0 + \beta_1 \times Spread_{EnAdj} + \epsilon \tag{3}$$

where:

$$Spread_{EnAdj} = P_{Ethanol} - P_{RBOB} \times 1.1$$

We initially felt it would be appropriate to employ a composite model that incorporates both volumetric and energy-adjusted spreads. This in theory would allow the model to capture both the straightforward price difference and the energy-equivalent substitution cost. The equation for this model would be:

$$lnP_{RIN} = \beta_0 + \beta_1 \times Spread_{vol} + \beta_2 \times Spread_{En} + \epsilon \qquad (4)$$

However, when we checked variance inflation factors on our end model, Energy Spread had a VIF of 20.12, and Volume Spread had a VIF of 16.20, both of which are outside the acceptable range. To adjust for this, we decided to only use the Volume Spread. We did this for a few different reasons. The first of which is that using an energy-equivalent equation overlooks the average consumer's ignorance on the energy difference between ethanol and gasoline, and thus fails to account for their unwillingness to pay a large premium for E0 gasoline over E10, changing the blending economics. Furthermore, this volumetric difference avoids the over-weighted valuation of ethanol in model 3. Even though ethanol is important for raising octane levels, it hardly justifies a valuation that places it at a multiple of gasoline, especially once one considers the lower energy and MPG it provides to the end consumer. Thus, we've decided to move forward with just the Volume Spread and omit the two adjusted spreads.

3.2 Accounting for Volatility

As mentioned in the literature review, Afkhami and Ghoddusi (2021) emphasized that RINs gain additional value during periods of higher volatility, due to the option-like characteristics they provide. In this view, RINs offer refiners flexibility in deciding when to blend ethanol or purchase RINs, making them more valuable when ethanol and gasoline prices fluctuate. This optionality becomes more pronounced during times of heightened volatility, as refiners might hold off on blending and purchase RINs in anticipation of more favorable market conditions. To account for this, we expanded the base model to include measures of volatility in ethanol and gasoline prices.

It is crucial to account for both gasoline and ethanol volatility because these two commodities behave differently in response to market forces, and this divergence directly impacts RIN pricing. Ethanol prices are largely driven by agricultural factors—such as corn yields, weather conditions, and biofuel policy—making them susceptible to supply-side shocks that are less relevant to gasoline, which is influenced more by global oil markets and geopolitical events.

The revised model incorporating volatility is as follows:

$$lnP_{RIN} = \beta_0 + \beta_1 \times Spread_{Vol} + \beta_2 \times \sigma_{ETH} + \beta_3 \times \sigma_{RBOB} + \epsilon \quad (5)$$

By incorporating these volatility measures, we account for the option value that makes RINs more attractive during periods of uncertainty, when refiners might delay their compliance strategies. This provides a more comprehensive understanding of how RIN prices are influenced by market conditions beyond just the ethanol-gasoline spread.

3.3 Additional Factors

Consistent with our thesis, the previous models capture both the intrinsic value of D6 RINs and their option-like characteristics, primarily driven by ethanol-gasoline spreads and the volatility of these commodities. However, this approach treats RINs as perfectly fungible, which is an oversimplification. In practice, RINs vary in value based on several key factors beyond price spreads, such as their QAP (Quality Assurance Program) verification status, the month of the transaction, the age of the RIN, and price momentum.

First, QAP verification status indicates whether a RIN has been third-party verified, adding credibility and reducing compliance risk for the buyer. Verified RINs would theoretically command a higher price due to the lower risk of rejection or penalty during compliance audits (the RFS operates under a buyer-beware system, refusing to provide any credit for falsified RINs). Additionally, the month in which a RIN is traded can affect its price, as seasonal factors — such as blending requirements, fuel demand, and business or regulatory deadlines — can cause supply and demand fluctuations. Furthermore, the age of a RIN plays a critical role, as older RINs are closer to expiration and may be less valuable compared to newer ones. Lastly, including the most recent RIN price allows the model to capture momentum and trends in RIN pricing, acknowledging that past prices can influence current valuations.

These additional factors necessitate an expanded model that incorporates both market price dynamics and specific attributes of individual RINs. The composite model is as follows:

$$\begin{aligned} lnP_{RIN,t} =& \beta_0 + \beta_1 \times lnP_{RIN,t-1} + \beta_2 \times Spread_{Vol,t} \\ &+ \beta_3 \times \sigma_{ETH,t} + \beta_4 \times \sigma_{RBOB,t} + \beta_5 \times IsVerfied \\ &+ \beta_6 \times IsCurrent + \beta_7 \times TransactionMonth + \epsilon \end{aligned} \tag{6}$$

Where IsVerified is binary variable indicating whether a RIN is QAP verified, IsCurrent is a binary variable indicating whether the RIN is from the current compliance year, and TransactionMonth is a factorized categorical variable representing the month in which the RIN transaction occurred.

This composite model provides a more nuanced view of RIN pricing by accounting for not only the economic drivers like price spreads and volatility, but also the specific characteristics of RINs that affect their market value. By incorporating these factors, we better capture the complexities of RIN trading dynamics.

4. Data Description

The data used for model testing is available for download from https://www.tuckerstrow.com, along with the R code used to run regressions. This is to encourage replicability and transparency. There were a total of 5,885 observations in our dataset.

4.1 Data Sources (See Appendix for links)

FRED: Monthly ethanol price data
Market Insider by Business Insider: Daily RBOB prices
NASDAQ: Daily CORN ETF prices, CBOT Ethanol Futures
EPA: RIN Price and Transaction Data

4.2 Data Preprocessing and Aggregation

The majority of the data pre-processing was performed using Microsoft Excel. Data from the aforementioned sources were imported into individual sheets within a single Excel workbook. Where applicable, close prices were used and all units were converted into dollars. For RIN data, only D6 Separated RINs were analyzed. From this dataset, weekly and monthly standard deviations were calculated for RBOB and CBOT EH, representing gasoline and ethanol volatility, respectively. CBOT EH was selected for ethanol volatility due to the limited availability of daily ethanol price data. Where possible, monthly prices from FRED were incorporated; however, CBOT EH was utilized for volatility calculations. CORN was explored as an alternative, but rejected due to a lack of significant correlation to actual ethanol prices. To align with the EPA's weekly reporting on RIN transactions, weekly averages of RBOB were computed. For the purpose of spread calculations, monthly ethanol prices were employed as a simplification in cases where weekly data was unavailable. The spreads were calculated as RBOB minus Ethanol multiplied by factors of 0.667, 1, and 1.1. Binary variables, IsCurrent and VerStatus were then generated based on the RIN Year, Transfer Year, and QAP Verification Status. The month of the transfer date was extracted and encoded as a categorical variable. Previous RIN price was calculated as a simple average of the previous week's prices. Finally, the data was consolidated into a single sheet, with the timeframe adjusted to ensure all variables contained data within the analyzed period. See Table 3 in the appendix for a description of variables.

5. Results

5.1 Full Model

In accordance with the theory outlined above, we tested the following model:

$$\begin{aligned} lnP_{RIN,t} =& \beta_0 + \beta_1 \times lnP_{RIN,t-1} + \beta_2 \times Spread_{Vol,t} \\ & + \beta_3 \times \sigma_{ETH,t} + \beta_4 \times \sigma_{RBOB,t} + \beta_5 \times IsVerfied \\ & + \beta_6 \times IsCurrent + \beta_7 \times TransactionMonth + \epsilon \end{aligned} \tag{7}$$

This model had low Variance Inflation Factors, with the max being 2.36 on one of the factorized transaction months.

The results from our model — when run on the described data (with robust standard errors) — are in table 1.

Table 1 Full Regression Results

Variable	**Estimate**	**Std. Error**	**t value**
(Intercept)	−0.423 121	0.034 275	−12.3450
`Previous RIN Price`	1.059 809	0.019 069	55.5780
`Volume Spread`	0.187 779	0.020 643	9.0965
`Ethanol Weekly SD`	0.268 250	0.518 787	0.5171
`RBOB Weekly SD`	0.688 157	0.199 716	3.4457
`is_verified`	0.276 152	0.014 154	19.5107
`is_current`	0.076 560	0.013 797	5.5492
`is_january`	0.068 203	0.036 252	1.8813
`is_february`	0.101 668	0.034 555	2.9422
`is_march`	0.049 204	0.036 146	1.3612
`is_april`	0.064 537	0.036 020	1.7917
`is_may`	0.083 329	0.034 449	2.4189
`is_june`	0.067 784	0.035 005	1.9364
`is_july`	0.054 197	0.033 936	1.5971
`is_august`	0.058 597	0.034 525	1.6972
`is_september`	0.046 307	0.034 342	1.3484
`is_october`	0.045 188	0.034 864	1.2961
`is_november`	0.007 892	0.034 604	0.2281

From these results it is apparent that the previous RIN price, Spread between Ethanol and Gasoline, Volatility of Gasoline, RIN verification status, RIN age, and even RIN transaction month are all highly relevant when estimating the price of a RIN. Interestingly, Ethanol Volatility wasn't a significant estimator of RIN price.

For the Ethanol-Gasoline Spread, the estimate of 0.1878 matches the proposed theory that a larger spread would indicate a greater RIN price (due to the fact that RINs are intended to reflect blending costs). Similarly, RBOB Weekly SD matches the intuited sign, with a more volatile feedstock environment correlating to higher RIN prices, likely due to their option characteristics

The positive estimates on the dummy variables all align with the proposed theory as well. The results suggest an existence of a RIN price premium on verified RINs (likely due to buyer-beware enforcement) as well as a premium on RINs that were generated in the current compliance year — likely because these current RINs can be rolled over to the next year, providing flexibility to the owners in meeting environmental obligations.

Lastly, the seasonal variables suggest that there is a meaningful difference in RIN prices across months. We tend to see higher prices in February and May as compared to December, which could have to do with year-end sell offs of RINs to pad financial statements or the differences in tax strategies between the beginning and ending of years.

5.2 Partial Model

We were interested in seeing the strength of time-related variables that previous RIN price could be masking, so we decided to run a partial regression that omitted the Previous RIN Price variable. The results from our partial model — when run on the described data (with robust standard errors) — are in table 2.

The major revelation from these results is the true significance of seasonal trends on RIN prices. We see a large increase in RIN prices around March, but this is unsurprising given that March 31st is the EPA's compliance deadline for the RFS. What is even more interesting is the persistence of elevated prices up through July, suggesting that there are external or cognitive factors driving higher RIN prices during these months. It would be interesting for future studies to explore the behavior of compliance departments during this time frame. These increased prices could suggest a desire for companies to begin a compli-

Table 2 Partial Regression Results

Variable	Estimate	Std. Error	t value
(Intercept)	−0.462 312	0.045 020	−10.2691
Volume Spread	0.387 387	0.025 346	15.2840
Ethanol Weekly SD	−16.102 166	0.710 135	−22.6748
RBOB Weekly SD	4.097 078	0.238 700	17.1641
is_verified	0.292 601	0.017 674	16.5557
is_current	0.047 526	0.017 504	2.7152
is_january	0.131 938	0.047 629	2.7701
is_february	0.069 105	0.044 704	1.5458
is_march	0.200 828	0.046 642	4.3057
is_april	0.228 774	0.048 464	4.7205
is_may	0.273 091	0.046 196	5.9116
is_june	0.257 550	0.045 170	5.7018
is_july	0.239 665	0.044 333	5.4060
is_august	0.147 881	0.046 201	3.2008
is_september	0.144 943	0.046 968	3.0860
is_october	0.110 887	0.045 294	2.4482
is_november	0.045 009	0.045 514	0.9889

ance period with a surplus of RINs, an adjusted appetite to the raised prices experienced in March, or simply the shortage of RIN supply in the overall market due to the mass retiring that occurs in March.

Another incredibly surprising result is the massive, negative coefficient on Ethanol Volatility. If this was the true impact of ethanol volatility it would contradict the prevailing theory in the literature that RINs are used as a hedge in times of feedstock volatility. However, it's much more likely that Ethanol Volatility is accounting for the impact of omitting previous RIN price from our regression.

5.3 Evaluating the Model

Lastly, we used the full and partial models to predict RIN prices. We checked to see whether we needed to constrain estimates based on the explicit constraints the EPA puts on RIN prices; however, none of the predicted values were outside the EPA's mandated range.

In order to facilitate the comparison of prices in dollars we performed the following transformation on the predictions:

$$P^{*}{}_{RIN} = e^{\ln P^{*}{}_{RIN}}$$

We were then free to compare these values to the actual RIN prices of the dataset.

5.3.1 Full Model

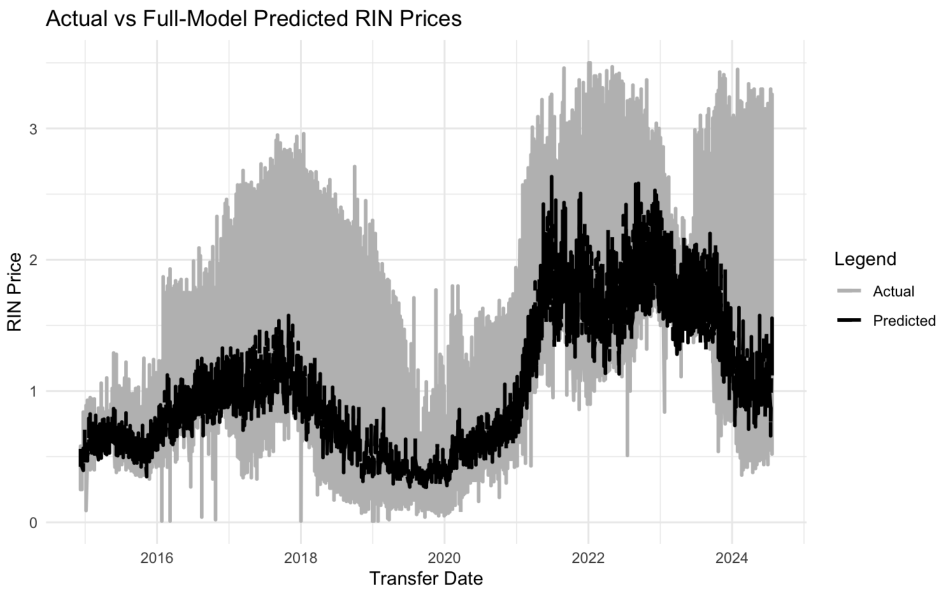

Fig. 2. Actual vs Full-Model Predictions

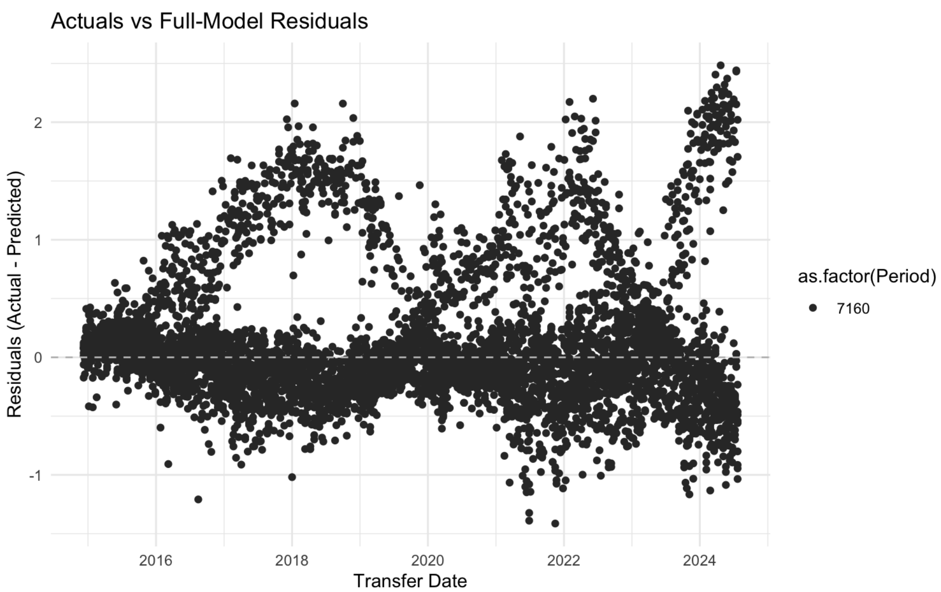

Fig. 3. Actual vs Full-Model Residuals

The full model generated predicted RIN prices in the range of \$0.24 to \$2.63, which indicates that the model captures a substantial portion of the actual price variation. However, the range is somewhat narrower than the actual prices, which spanned from \$0.01 to \$3.50. This discrepancy suggests that while the model performs well for moderate price levels, it underestimates the extremes—particularly the lowest and highest RIN prices. The ability to capture extreme values is crucial for traders and policymakers, as significant price deviations may impact compliance strategies and trading decisions. This underestimation could be addressed by incorporating additional variables that might capture market shocks or extreme volatility, such as policy changes or unexpected supply disruptions.

The mean predicted price (\$1.05) is slightly lower than the actual mean price (\$1.15), while the median predicted price (\$0.918) is almost identical to the actual median (\$0.92). This closeness between the medians suggests that the model accurately predicts central tendencies in RIN prices, which is valuable for market participants who are less concerned with outliers and more focused on typical market conditions. However, the lower mean of the predictions indicates that the model may be systematically underestimating prices, which could result from failing to capture high-price events. This bias towards underprediction could be problematic for traders, as it would imply a less conservative approach to future price forecasts.

The predictions had a Mean Average Error (MAE) of \$0.37 and a Root Mean Squared Error (RMSE) of 0.57. The MAE reflects the average absolute difference between predicted and actual prices, while the RMSE penalizes larger errors more heavily. An RMSE of 0.57 suggests that larger errors are present but not overly dominant. Given that the RMSE is not significantly larger than the MAE, we can infer that the model's prediction errors are fairly consistent and not heavily skewed by a few extreme mispredictions. However, given the financial significance of RIN price deviations, even moderate errors may impact decision-making in this context. Improving the model's ability to predict extreme prices would reduce both the MAE and RMSE, thereby improving its reliability for stakeholders who rely on accurate pricing to hedge against risks in the RIN market.

In all, the model had an R-Squared of 0.4289, meaning it explains 42.89% of the variation in D6 RIN prices. While this indicates that the model captures a substantial portion of the factors influencing RIN prices, there remains a significant amount of unexplained variation. This level of explanatory power is fairly typical for complex financial markets, where price movements are influenced by a wide range of unpredictable factors, such as regulatory shifts, market sentiment, or unforeseen supply chain disruptions. However, to improve the model's predictive accuracy, additional variables could be explored, such as more detailed fuel market indicators or speculative trading activity, which may help to capture the remaining unexplained variation.

To enhance the model's performance, future iterations could incorporate more dynamic market variables. For instance, integrating macroeconomic indicators, such as inflation rates or crude oil price fluctuations, might better capture the broader economic factors that indirectly influence RIN prices. Incorporating policy announcements and market sentiment indicators could also offer deeper insights into price movements, as RIN markets are highly sensitive to regulatory changes. Finally, advanced machine learning techniques, such as random forests or neural networks, could be explored to identify nonlinear relationships between the variables and further improve price predictions.

5.3.2 Partial Model

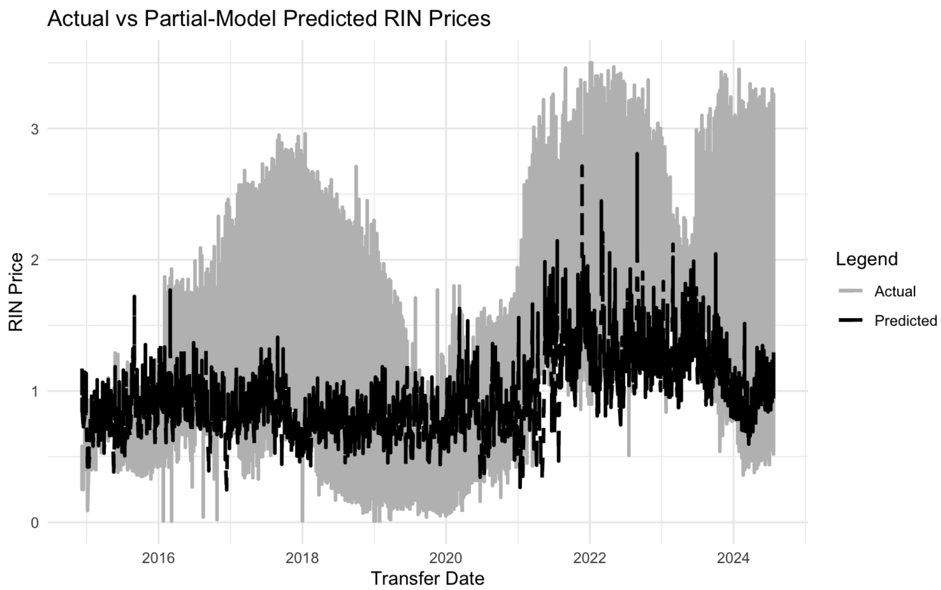

Fig. 4. Actual vs Partial-Model Predictions

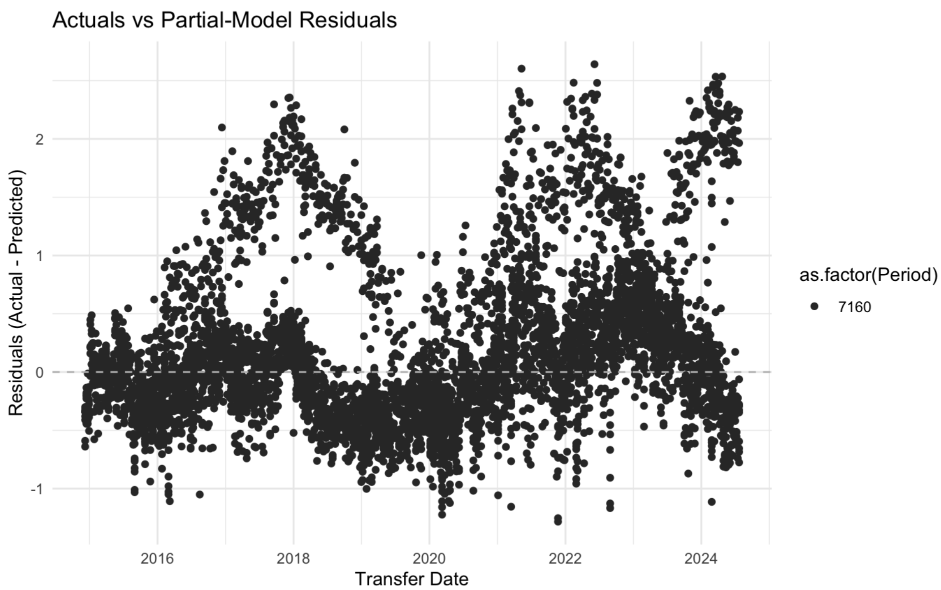

Fig. 5. Actual vs Partial-Model Residuals

The partial model, which omits the previous week's RIN price as a predictor, generated predicted RIN prices in the range of \$0.24 to \$2.81. This wider range of predictions, as compared to the full model, reflects greater variability in the predicted prices. However, the partial model's predicted range still falls short of the actual price range, which spans from \$0.01 to \$3.50. The broader upper limit of the predicted price range (\$2.81) suggests that the model is more responsive to high-price events than the full model, but it still underestimates the highest observed prices in the dataset. Furthermore, a look at the graph confirms that the large price predictions were accurate in the context they were made, but an inability to extend this high-price environment to the next week caused near-immediate reversions back to lower predicted price levels when actual prices remained high, suggesting that the isolated existence of this model is impractical.

The mean predicted price of \$0.97 is lower than the actual mean of \$1.15, and the median predicted price of \$0.919 is nearly identical to the actual median of \$0.92. The lower mean prediction indicates that the partial model tends to underestimate overall RIN prices, which in part may be due to the exclusion of the previous week's price variable. Without capturing the persistence in RIN price movements over time, the model fails to fully reflect ongoing market trends, leading to a bias towards lower price levels. The accuracy of the median prediction, however, suggests that the model is relatively effective at predicting central price tendencies but less so at analyzing periods with persistent upward price pressures. This limitation in accounting for current market environment may reduce the model's utility for traders or policymakers who need to account for significant price shifts in the RIN market.

The predictions had a Mean Average Error (MAE) of \$0.49 and a Root Mean Squared Error (RMSE) of 0.69, both of which are higher than those of the full model. This indicates that the partial model produces larger average errors and is less accurate overall. The higher RMSE suggests that the model is particularly sensitive to large errors, which may arise due to its inability to capture past price momentum. By omitting the previous week's RIN price, the model loses a key explanatory variable, which helps reduce both small and large errors in the full model. The relatively large difference between the MAE and RMSE also highlights that the partial model is more prone to making large errors, which further limits its predictive reliability, especially during periods of price volatility or rapid market shifts.

The partial model had an R-Squared of 0.1706, meaning it explains only 17.06% of the variation in D6 RIN prices. This lower R-Squared value indicates that the model is unable to account for a substantial portion of the factors driving RIN price movements that the full model captures. The most likely explanation for this drop in explanatory power is the omission of the previous week's RIN price, which serves as a strong predictor of future prices. In the full model, this variable captures much of the inertia and momentum inherent in the RIN market, where price movements are often driven by the continuation of trends from prior periods. Without this variable, the partial model must rely more heavily on other factors, such as energy spreads and market volatility, which alone are insufficient to fully capture the complexity of RIN pricing. This significantly lower explanatory power suggests that while the partial model captures some price movements, it is much less effective at accounting for long-term or sustained trends.

Given the limitations of the partial model, it seems like the full model is a better option for those looking to model appropriate RIN prices. Alternatively, analysts should at least consider using other forms of time-series data to capture price momentum. Time-lagged variables, such as a moving average of past prices, could offer a similar benefit to the previous week's price without making the model overly reliant on a single point of historical data. Additionally, like with the full model, incorporating external market indicators, such as speculative trading volume or shifts in policy, could provide additional context to explain price fluctuations not captured by the current set of variables. As mentioned above, another potential improvement would be to explore machine learning models capable of detecting non-linear patterns between price variables and broader market conditions, which could enhance the model's ability to predict extreme price movements.

6. Conclusion

In conclusion, this paper offers a comprehensive exploration into the complexities of pricing D6 Renewable Identification Numbers (RINs) under the lens of imperfect fungibility. By introducing a model that extends beyond traditional determinants like ethanol and gasoline price spreads, we have integrated factors such as Quality Assurance Program (QAP) verification status, RIN generation year, seasonal trading patterns, and market volatility into the pricing framework. This approach has provided a more nuanced understanding of how RIN prices are shaped in real-world trading environments, offering new insights for both market participants and policymakers.

The findings of this research emphasize that the pricing of D6 RINs is not purely a function of fuel spreads, as commonly assumed in previous literature. Instead, factors like RIN verification, age, and trading month significantly influence their market value. For instance, verified RINs command a premium, likely due to the reduced compliance risks associated with third-party verification, while newer RINs tend to be more valuable given their longer shelf life for compliance purposes. Moreover, the impact of volatility, particularly in gasoline prices, plays a pivotal role in determining RIN values, reaffirming the optional characteristics that make RINs valuable during periods of uncertainty.

This paper also sheds light on the seasonal trends observed in RIN pricing, with certain months experiencing price spikes due to regulatory deadlines, such as the March compliance deadline set by the EPA. These seasonal fluctuations suggest that market participants adjust their behavior based on external regulatory pressures, creating opportunities for strategic trading and compliance planning. The persistence of elevated prices through mid-year further indicates potential speculative or risk-management behaviors among market participants, a topic that warrants further research.

While this study has successfully highlighted the multifaceted nature of D6 RIN pricing, there remain areas for future exploration. For instance, the negative impact of ethanol volatility on RIN prices observed in our partial model contradicts some of the established literature, suggesting the need for deeper investigation into how volatility in feedstock markets influences RIN trading behavior. Additionally, the role of other renewable fuel categories, such as biodiesel RINs, in capping D6 RIN prices offers an interesting avenue for future research, as interactions between different RIN categories could further complicate the pricing dynamics.

Overall, this paper contributes significantly to the understanding of RIN markets by challenging the conventional notion of perfect fungibility and offering a more comprehensive model that accounts for the intrinsic and extrinsic factors driving D6 RIN prices. These insights are critical not only for traders looking to optimize their strategies but also for policymakers aiming to design more efficient and effective renewable fuel standards. The enhanced understanding of RIN pricing dynamics provided by this research will hopefully serve as a foundation for further academic inquiries into regulation-driven markets and their broader economic implications.

References

Afkhami and Ghoddusi 2022; Mohamad Afkhami and Hamed Ghoddusi (2022). *Pricing renewable identification numbers under uncertainty*, Quantitative Finance, 22:4, 725-742, DOI: 10.1080/14697688.2021.1996625 Available at: `https://www.tandfonline.com/action/showCitFormats?doi=10.1080%2F14697688.2021.1996625&mobileUi=0`

Irwin et al. 2020; Irwin, S.H., McCormack, K. and Stock, J.H. (2020), *The Price of Biodiesel RINs and Economic Fundamentals.* Amer. J. Agr. Econ., 102: 734-752. Available at: `https://doi.org/10.1002/ajae.12014`

Korting and Just 2017; Christina Korting, David R. Just (2017), *Demystifying RINs: A partial equilibrium model of U.S. biofuel markets*, Energy Economics, Volume 64, Pages 353-362, ISSN 0140-9883. Available at: `https://doi.org/10.1016/j.eneco.2017.04.004`.

Lade and Lawell 2015; Lade, G. E, and Lawell, C. (2015). *Mandating green: On the Design of Renewable Fuel Policies and Cost Containment Mechanisms*. UC Davis: National Center for Sustainable Transportation. Available at https://escholarship.org/uc/item/5zj382t4

McPhail et al. 2011; Lihong McPhail, Paul Westcott, and Heather Lutman (2011). *The Renewable Identification Number System and U.S. Biofuel Mandates*. Washington, DC. Available at: https://downloads.usda.library.cornell.edu/usda-esmis/files/ft848q60n/79408124f/hh63t019f/BioEnergy-11-08-2011.pdf

Pouliot and Babcock 2016; Pouliot, S. and Babcock, B.A. (2016), *Compliance Path and Impact of Ethanol Mandates on Retail Fuel Market in the Short Run*. American Journal of Agricultural Economics, 98: 744-764. Available at: https://doi.org/10.1093/ajae/aav071

Thompson et al. 2010; Thompson, W., Meyer, S. and Westhoff, P. (2010), *The New Markets for Renewable Identification Numbers*. Applied Economic Perspectives and Policy, 32: 588-603. Available at: https://doi.org/10.1093/aepp/ppq021

Whistance and Thompson 2014; Whistance, J. and Thompson, W. (2014), A Critical Assessment of RIN Price Behavior and the Implications for Corn, Ethanol, and Gasoline Price Relationships. Applied Economic Perspectives and Policy, 36: 623-642. Available at: https://doi.org/10.1093/aepp/ppu012

Appendix

Data Links

Monthly ethanol Prices are available at: https://fred.stlouisfed.org/series/WPU06140341

Daily RBOB Prices are available at: https://markets.businessinsider.com/commodities/rbob-gasoline

CORN prices are available at: `https://www.nasdaq.com/market-activity/etf/corn/historical?page=1&rows_per_page=10&timeline=y10`

CBOT EH prices are available at: `https://www.nasdaq.com/market-activity/commodities/eh/historical`

RIN Data is available at: `https://www.epa.gov/fuels-registration-reporting-and-compliance-help/rin-trades-and-price-information`

Data Description

Table 3 Description of Variables

Variable	Mean	Median	Mode	Min	Max	Range	1st Quartile	3rd Quartile	IQR	Variance	SD	Skewness	Kurtosis	Count
RIN Price	1.1546	0.92	0.75	0.01	3.5	3.49	0.59	1.59	1	0.5728	0.7568	1.0693	0.5611	5885
is_current	0.4744	0	0	0	1	1	0	1	1	0.2494	0.4994	0.1025	-1.9902	5885
is_verified	0.4075	0	0	0	1	1	0	1	1	0.2415	0.4914	0.3767	-1.8587	5885
Volume Spread	0.3172	0.3529	-0.1604	-0.8198	1.5732	2.3930	0.0608	0.5815	0.5207	0.1572	0.3965	-0.0848	0.3726	5885
Energy Spread	0.9634	0.9352	1.4120	0.0948	2.2771	2.1822	0.7389	1.1472	0.4083	0.1240	0.3521	0.5957	0.9525	5885
Energy Spread Adjusted	0.4075	0.5071	1.2899	-0.5704	1.7652	2.3355	0.2479	0.7303	0.4825	0.1400	0.3742	0.0922	0.5092	5885
Ethanol Weekly SD	0.0134	0.0114	0	0	0.0833	0.0833	0	0.0208	0.0208	0.0002	0.0148	1.5216	3.1815	5885
RBOB Weekly SD	0.0421	0.0341	0.0714	0	0.2792	0.2792	0.0220	0.0517	0.0298	0.0010	0.0319	2.6589	11.4144	5885
is_january	0.0996	0	0	0	1	1	0	0	0	0.0897	0.2995	2.6752	5.1587	5885
is_february	0.0970	0	0	0	1	1	0	0	0	0.0876	0.2960	2.7235	5.4196	5885
is_march	0.1118	0	0	0	1	1	0	0	0	0.0993	0.3152	2.4643	4.0741	5885
is_april	0.0822	0	0	0	1	1	0	0	0	0.0755	0.2748	3.0419	7.2559	5885
is_may	0.0826	0	0	0	1	1	0	0	0	0.0758	0.2753	3.0338	7.2062	5885
is_june	0.0800	0	0	0	1	1	0	0	0	0.0736	0.2714	3.0962	7.5892	5885
is_july	0.0799	0	0	0	1	1	0	0	0	0.0735	0.2711	3.1005	7.6156	5885
is_august	0.0712	0	0	0	1	1	0	0	0	0.0661	0.2572	3.3358	9.1308	5885
is_september	0.0681	0	0	0	1	1	0	0	0	0.0635	0.2520	3.4285	9.7582	5885
is_october	0.0775	0	0	0	1	1	0	0	0	0.0715	0.2674	3.1615	7.9975	5885
is_november	0.0698	0	0	0	1	1	0	0	0	0.0650	0.2549	3.3763	9.4028	5885