**HEDONIX RESEARCH**

WORKING PAPER · No. 2026-01

# A Hedonic Pricing Model for Graded Pokémon Trading Cards

*Evidence from the Scarlet & Violet Era (2023–2026)*

*The methodology described in this paper underlies the H6 Hedonic Engine, the pricing core of Hedonix.*

**Philipp Baro**

Hedonix Research, Frankfurt am Main

baro.philipp@gmail.com · hedonix.tech

*First version: April 2026 · This version: April 2026*

## Abstract

We apply hedonic pricing methodology to the secondary market for graded Pokémon trading cards, focusing on Special Illustration Rare and Illustration Rare cards from fourteen Scarlet & Violet expansion sets released between 2023 and 2026. Using transaction-level data on 360 cards with verifiable PSA-graded sales activity, we estimate log-linear hedonic regressions in which the dependent variable is the natural logarithm of average eBay PSA 10 sale prices, and explanatory variables include raw market price, rarity status, set fixed effects, character popularity tiers, an interaction term capturing differential elasticities across rarity classes, an empirically identified top-tier artist indicator, a hype-divergence proxy derived from price-tracker market signals, and a within-set sales-velocity z-score. The fully specified model achieves an $R^2$ of 0.879 (adjusted $R^2$ of 0.870). Raw card price exhibits a robust elasticity of approximately 0.48, with significantly steeper price transmission for Special Illustration Rares than for Illustration Rares. Top character tiers, top artist identification, and the hype-divergence proxy all contribute statistically and economically significant explanatory power. An appendix documents an experimental extension using CLIP image embeddings, which yields only marginal improvement in fit and motivates a discussion of the limits of computer-vision features in this market segment. The methodology described in this paper underlies the H6 Hedonic Engine, the pricing core of the Hedonix platform.

**Keywords:** *hedonic pricing, collectibles, alternative assets, trading cards, cross-sectional asset pricing, hype proxies, image embeddings, Pokémon TCG*

**JEL Classification:** *G11, G12, G14, Z11*

# 1. Introduction

The market for collectible trading cards has grown to multi-billion-dollar scale over the past decade, with Pokémon trading cards (TCG) constituting the largest single franchise segment. Despite this scale, the application of rigorous quantitative pricing methods to the modern graded-card secondary market has remained limited, with most analytical work conducted by hobbyist communities rather than financial economists. This paper contributes to closing that gap by applying a hedonic pricing framework to a curated cross-section of modern Pokémon TCG cards, demonstrating that observable card characteristics, when combined with empirically constructed demand and quality proxies, explain a substantial share of cross-sectional price variation.

Three contributions distinguish this work. First, we extend prior preliminary analyses by approximately tripling the sample size (n=360 across fourteen Scarlet & Violet sets) and introducing three empirically constructed explanatory variables: a top-tier artist indicator derived from a residual-based identification procedure, a market-divergence proxy capturing short-term hype, and a within-set sales-velocity z-score capturing relative liquidity. Second, we report the results of an experimental extension using CLIP image embeddings (Radford et al., 2021), providing a transparent stress-test of whether modern vision-language models contribute additional explanatory power beyond structured features in this domain. Third, we discuss the implications of the cross-sectional residual structure for downstream applications, including the construction of mispricing-based investment signals.

The methodology developed here forms the core of the H6 Hedonic Engine, which powers the Hedonix platform (hedonix.tech). Hedonix provides institutional-grade fair-value estimates for graded trading cards to a community of investors, collectors, and market professionals. The decision to publish the underlying methodology in working-paper form reflects a commitment to transparency and to academic engagement with the broader collectibles-pricing literature, while specific implementation details, parameter values, and live mispricing rankings remain proprietary to the platform.

The paper proceeds as follows. Section 2 reviews related literature. Section 3 describes the data and sample-construction procedure. Section 4 introduces the empirical specification. Section 5 reports estimation results, model diagnostics, and a residual-based cross-sectional mispricing analysis presented in anonymized form. Section 6 discusses limitations and extensions. Section 7 concludes. Appendix A documents the CLIP-embedding extension.

# 2. Literature Review

The methodological foundation of this study is the hedonic pricing framework introduced by Rosen (1974), in which the price of a heterogeneous good is decomposed into the implicit prices of its observable attributes. Rosen's formulation has been applied extensively in markets for differentiated goods, with particular relevance to collectibles, where individual items vary along multiple observable dimensions but no two items are perfect substitutes.

In the literature on collectible-asset pricing, Mandel (2009) examines art as an investment, documenting that aesthetic and conspicuous-consumption considerations introduce systematic deviations from purely return-driven pricing. Renneboog and Spaenjers (2013) provide a comprehensive hedonic analysis of the global art market, identifying robust effects from artist identity, medium, size, and provenance, while noting that idiosyncratic residuals remain economically large even in well-specified models. Dimson and

Spaenjers (2014) survey the broader landscape of "emotional assets," noting modest financial returns combined with attractive diversification properties. Closer to the present subject, the literature on sports cards has documented systematic pricing relationships with rookie status, player performance, and condition, though most studies precede the modern third-party-grading era.

Within machine-learning-augmented hedonic modeling, several recent contributions employ image-derived features to augment structured pricing variables. Radford et al. (2021) introduce CLIP, a contrastive vision-language model whose pre-trained embeddings have been widely repurposed as transferable image representations across domains, including in nascent applications to art-market valuation. Our appendix follows this methodological tradition, while documenting that the marginal contribution of CLIP features in our specific application is more limited than has been reported in larger-scale art-market settings — a finding consistent with the reduced visual heterogeneity within a tightly defined card-printing era.

# 3. Data

## 3.1 Sources and sample

Card metadata are sourced from the public pokemontcg.io API, which provides authoritative card identifiers, set affiliations, card numbers, rarity classifications, artist attribution, and ancillary attributes. Transaction prices and grading-tier sales data are obtained through a commercial price-tracking service that aggregates TCGplayer raw-card market prices in U.S. dollars and eBay sold-listing data for graded cards across multiple grading services. Data extraction was conducted in April 2026.

The sample is restricted to Special Illustration Rares (SIR) and Illustration Rares (IR) from fourteen Scarlet & Violet expansion sets, spanning the full release window of the era. After applying liquidity filters (a minimum of ten verified PSA 10 sales over the available rolling window), removing two cards identified as data anomalies in the underlying eBay aggregation, and dropping cards with missing key fields, the final analytic sample contains 360 cards. Table 1 reports the cross-set distribution.

*Table 1. Final sample size by expansion set.*

| Set | N |
|---|---|
| Paradox Rift | 46 |
| Paldea Evolved | 45 |
| Destined Rivals | 34 |
| Temporal Forces | 32 |
| Surging Sparks | 32 |
| Twilight Masquerade | 29 |
| Prismatic Evolutions | 29 |
| 151 | 23 |
| Stellar Crown | 19 |
| Obsidian Flames | 18 |

| | |
|---|---|
| Journey Together | 17 |
| Shrouded Fable | 15 |
| Paldean Fates | 11 |
| Scarlet & Violet | 10 |
| Total | 360 |

### 3.2 Variables

The dependent variable is the natural logarithm of the average eBay sold-listing price for PSA 10–graded copies of each card. Logarithmic transformation is motivated by the heavy right-skewness of the raw price distribution and by economic theory predicting multiplicative rather than additive attribute effects.

The explanatory variables fall into three groups.

**Group 1 — Structural variables:**

- log_raw_c: natural logarithm of the TCGplayer raw-card market price, mean-centered. Centering reduces multicollinearity between the main effect and its interaction with the rarity dummy.
- is_sir: binary indicator equal to 1 if the card is a Special Illustration Rare, 0 if an Illustration Rare.
- log_raw_c_x_sir: interaction term permitting the price-transmission elasticity to differ across rarity tiers.
- Set fixed effects: indicator variables for each expansion, with the oldest available set serving as the omitted baseline.
- char_top, char_mid: character-popularity tier indicators. The top tier comprises cards featuring highly traded chase characters such as Charizard, Pikachu, the Eeveelution family, and other persistent franchise icons. The mid tier captures characters with strong but more selective demand. Tier assignments are based on community-recognized character salience.

**Group 2 — Empirically constructed demand and quality proxies:**

- artist_top: a binary indicator equal to 1 for cards illustrated by an empirically identified subset of artists. The set is determined by an iterative procedure: after fitting a baseline model with structural variables only, we identify artists with at least three cards in the sample whose median residual exceeds a pre-specified threshold in log-price space. The construction methodology is fully specified; the specific composition of the resulting artist set, together with the threshold value, is held proprietary as part of the H6 Hedonic Engine.
- smart_divergence: the absolute relative deviation between two market-price estimates produced by the price-tracking service — a robust filtered estimate and a simple sales-average estimate. Higher values indicate higher dispersion between recent and longer-term transaction patterns, which we interpret as a proxy for short-term hype and pricing instability.
- velocity_z: a within-set z-score of monthly sales velocity, capturing relative liquidity within an expansion while controlling for cross-set differences in release timing and overall market activity.

**Group 3 — Auxiliary controls:**

- Pokemon-type indicators for the most common primary types in the sample. These are included as controls but are not the focus of inference.

## 4. Methodology

We estimate two nested ordinary least squares (OLS) specifications with heteroscedasticity-consistent (HC3) standard errors.

*Model A (Structural):* $log(P\_i) = \alpha + \beta\_1\ log\_raw\_c\_i + \beta\_2\ is\_sir\_i + \beta\_3\ (log\_raw\_c \times is\_sir)\_i + \beta\_4\ char\_top\_i + \beta\_5\ char\_mid\_i + \Sigma\_s\ \delta\_s\ Set\_s + \varepsilon\_i$

*Model B (Extended):* $Model\ A + \gamma\_1\ artist\_top\_i + \gamma\_2\ smart\_divergence\_i + \gamma\_3\ velocity\_z\_i + \Sigma\_t\ \theta\_t\ Type\_t + \varepsilon\_i$

where P_i is the average eBay PSA 10 sale price of card i. Statistical significance is reported at the conventional 1%, 5%, and 10% levels. Variance inflation factors are computed for the full Model B specification to assess multicollinearity, and standard residual diagnostics are conducted. Model B forms the production specification underlying the H6 Hedonic Engine.

## 5. Results

### 5.1 Regression results

Table 2 reports the estimation results. Model A achieves an $R^2$ of 0.845. The introduction of the three empirically constructed proxies in Model B raises $R^2$ to 0.879 (adjusted $R^2$ of 0.870). Several findings deserve emphasis.

First, the centered log raw price exhibits a robust elasticity of approximately 0.479, statistically significant at the 1% level. The interaction with the SIR indicator is positive and significant at the 1% level (coefficient = 0.148), implying that the elasticity for Special Illustration Rares is approximately 0.63, materially steeper than the 0.48 observed for Illustration Rares.

Second, character-tier indicators are both economically and statistically meaningful. The top tier coefficient (0.415) implies a premium of approximately +51% relative to the omitted "other" category, with a p-value below 0.001. The mid tier coefficient implies a premium of approximately +24%.

Third, the empirically constructed proxies all contribute significantly. The top-artist indicator carries a coefficient of 0.421 (implied premium ≈ +52%, $p < 0.001$), the smart-market divergence proxy 0.625 ($p < 0.001$), and the within-set velocity z-score 0.073 ($p < 0.01$). The auxiliary type indicators are not statistically distinguishable from zero at conventional levels and are retained as controls.

Fourth, the maximum variance inflation factor is 4.43, well below standard concern thresholds, indicating that multicollinearity is not driving the reported coefficients.

*Table 2. Hedonic regression coefficient estimates (Model B). Standard errors (HC3) in parentheses. *** p<0.01, ** p<0.05, * p<0.10. Set fixed effects and type controls included but not reported individually.*

| Variable | Coefficient | SE |
|---|---|---|
| log_raw_c | 0.4787*** | (0.0389) |
| is_sir | -0.2844*** | (0.0542) |
| log_raw_c × is_sir | 0.1480*** | (0.0468) |

| | | |
|---|---|---|
| char_top | 0.4145*** | (0.1062) |
| char_mid | 0.2178** | (0.1001) |
| artist_top | 0.4208*** | (0.1031) |
| smart_divergence | 0.6249*** | (0.1853) |
| velocity_z | 0.0730*** | (0.0269) |
| Constant | 5.0667*** | (0.1876) |
| | | |
| N | 360 | |
| $R^2$ | 0.879 | |
| Adjusted $R^2$ | 0.870 | |
| Set fixed effects | Yes | |
| Type controls | Yes | |
| Max VIF | 4.43 | |

### 5.2 Diagnostics

Figure 1 reports four standard diagnostic visualizations: a scatterplot of predicted versus actual log PSA 10 prices, a residual-versus-fitted plot, a normal QQ plot, and a histogram of residuals. The predicted-versus-actual plot shows tight alignment with the 45-degree reference line, consistent with the high $R^2$. The residual-versus-fitted plot shows residuals approximately centered around zero with no obvious heteroscedastic pattern, although a small upper tail remains visible — driven by a class of cards whose realized prices systematically exceed model predictions and which are discussed in Section 5.3 and Section 6.

*Figure 1. Diagnostic plots for Model B.*

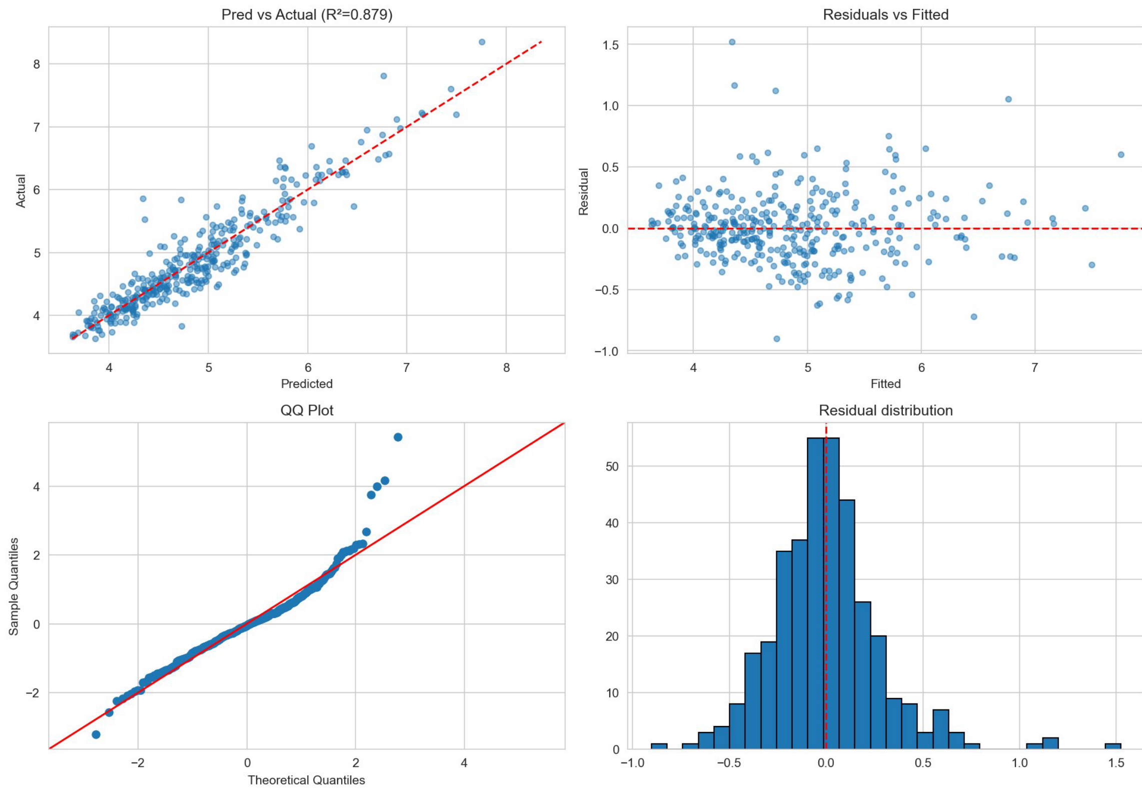
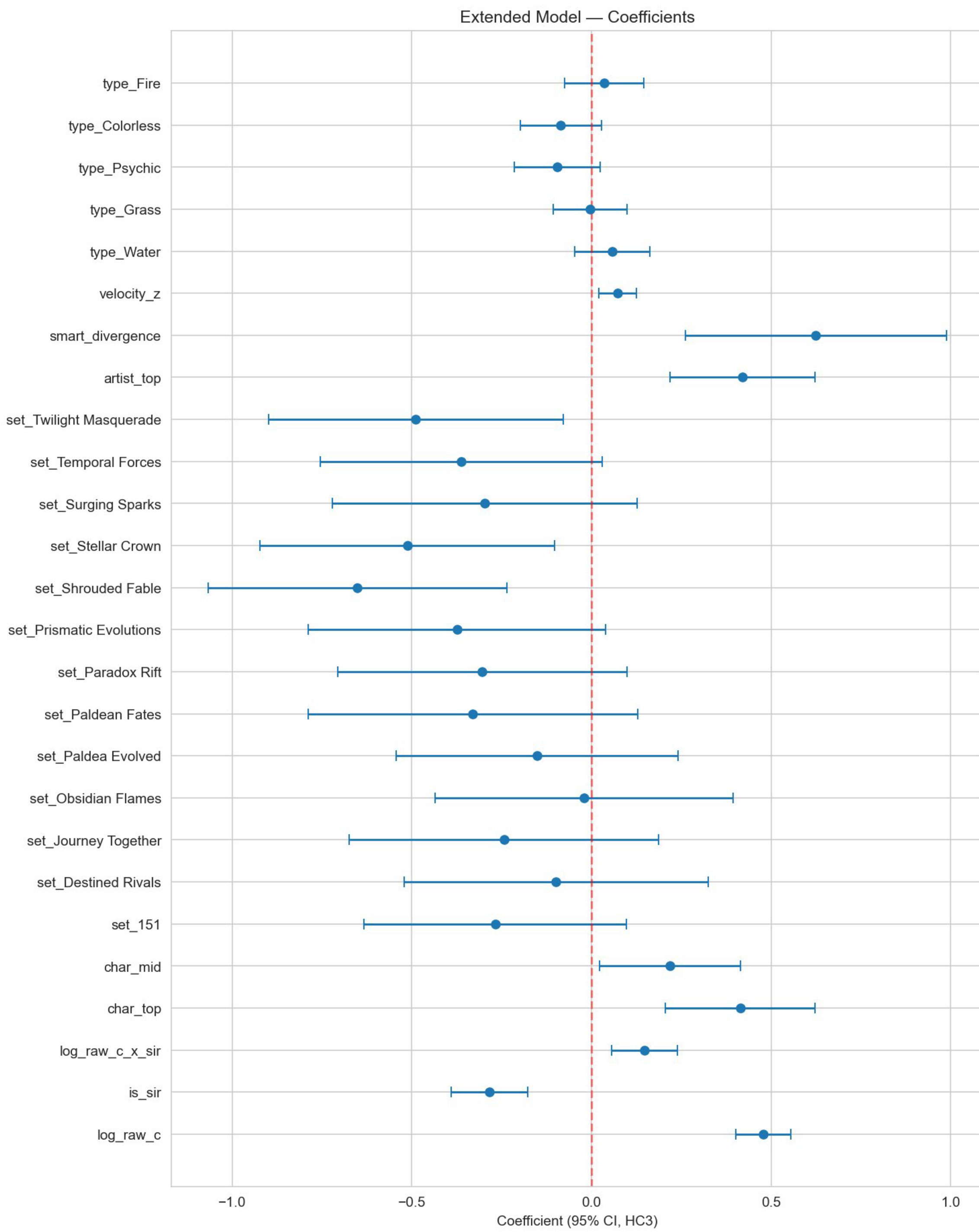



*Figure 2. Coefficient estimates with 95% confidence intervals (HC3 robust SE) for Model B.*




## 5.3 Cross-sectional mispricing analysis

Defining the mispricing residual as the deviation between observed log PSA 10 price and the model-predicted value, we obtain a cross-sectional ranking of cards by relative expensiveness or cheapness conditional on the explanatory variables. The transformation exp(residual) − 1 expresses this deviation as an approximate percentage of the predicted price.

Across the 360 cards in the final sample, the mean mispricing percentage is 4.51% and the median is -1.32%, with a standard deviation of 37.34% and a range from -59.5% to 358.5%. The asymmetric upper tail is the most economically salient feature of the distribution.

Individual card identities are not disclosed in the public version of this paper, as they form the live signal layer of the Hedonix platform. Tables 3 and 4 instead report the distributional summaries and pattern characterizations of the most extreme cards in each direction.

*Table 3. Anonymized profile of the fifteen most negatively mispriced cards (model predicts higher PSA 10 price than observed).*

| ID | Set | Rarity | Char tier | Raw | Actual | Predicted | Mispricing |
|---|---|---|---|---|---|---|---|
| Card U01 | Paldea Evolved | Special Illustr | other | $19.44 | $46 | $113 | -59.5% |
| Card U02 | Twilight Masquerad | Special Illustr | other | $178.21 | $311 | $640 | -51.4% |
| Card U03 | Destined Rivals | Special Illustr | other | $32.20 | $86 | $162 | -46.8% |
| Card U04 | Paldea Evolved | Special Illustr | other | $31.06 | $89 | $165 | -45.8% |
| Card U05 | Stellar Crown | Special Illustr | other | $79.32 | $118 | $212 | -44.3% |
| Card U06 | Paradox Rift | Illustration Ra | other | $33.11 | $109 | $193 | -43.5% |
| Card U07 | Scarlet & Violet | Illustration Ra | other | $22.12 | $112 | $195 | -42.5% |
| Card U08 | Paldea Evolved | Special Illustr | other | $57.45 | $124 | $214 | -41.9% |
| Card U09 | Prismatic Evolutio | Special Illustr | mid | $129.06 | $216 | $371 | -41.8% |
| Card U10 | Scarlet & Violet | Illustration Ra | other | $35.33 | $184 | $297 | -38.0% |
| Card U11 | Temporal Forces | Illustration Ra | other | $30.78 | $85 | $136 | -37.9% |
| Card U12 | Paldea Evolved | Illustration Ra | other | $10.38 | $66 | $106 | -37.7% |
| Card U13 | Prismatic Evolutio | Special Illustr | mid | $31.17 | $83 | $129 | -35.6% |
| Card U14 | Temporal Forces | Illustration Ra | other | $19.08 | $94 | $145 | -35.3% |
| Card U15 | Twilight Masquerad | Special Illustr | other | $34.95 | $77 | $119 | -35.3% |

*Table 4. Anonymized profile of the fifteen most positively mispriced cards (model predicts lower PSA 10 price than observed).*

| ID | Set | Rarity | Char tier | Raw | Actual | Predicted | Mispricing |
|---|---|---|---|---|---|---|---|
| Card O01 | Destined Rivals | Illustration Ra | other | $2.77 | $352 | $77 | +358.5% |
| Card O02 | Scarlet & Violet | Illustration Ra | other | $2.69 | $251 | $78 | +221.8% |
| Card O03 | Shrouded Fable | Illustration Ra | other | $23.40 | $345 | $112 | +206.7% |
| Card O04 | Paldea Evolved | Illustration Ra | other | $378.16 | $2476 | $864 | +186.5% |
| Card O05 | Temporal Forces | Special Illustr | other | $65.70 | $643 | $303 | +112.4% |
| Card O06 | Stellar Crown | Special Illustr | mid | $34.34 | $311 | $162 | +92.1% |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Card O07 | Prismatic Evolutio | Special Illustr | mid | $155.51 | $805 | $419 | +91.9% |
| Card O08 | Paldea Evolved | Illustration Ra | other | $76.08 | $582 | $305 | +90.8% |
| Card O09 | Paradox Rift | Special Illustr | other | $7.87 | $193 | $105 | +84.9% |
| Card O10 | Prismatic Evolutio | Special Illustr | top | $1466.72 | $4254 | $2325 | +83.0% |
| Card O11 | Paldea Evolved | Illustration Ra | other | $17.58 | $262 | $145 | +81.4% |
| Card O12 | 151 | Special Illustr | other | $108.58 | $581 | $320 | +81.3% |
| Card O13 | 151 | Special Illustr | other | $19.93 | $147 | $82 | +79.9% |
| Card O14 | Paldea Evolved | Illustration Ra | other | $9.86 | $164 | $91 | +79.7% |
| Card O15 | Paldea Evolved | Illustration Ra | other | $87.97 | $566 | $323 | +75.4% |

The negative-residual ranking is dominated by Special Illustration Rares and Illustration Rares featuring less-prominent characters, including Trainer cards and second-tier Pokémon, where the model's structural predictors imply higher prices than the market currently sustains. The positive-residual ranking, in contrast, is dominated by Illustration Rares featuring distinctive artwork, set-mascot Pokémon, and characters that have benefited from concentrated short-term collector enthusiasm. Several cards in the positive-residual ranking exceed the model prediction by more than 200%, with the most extreme outlier exceeding the prediction by approximately 359%.

We caution against interpreting the rankings as direct trading signals. A positive residual indicates that the realized price exceeds the model prediction; the model is partially misspecified along the dimension of aesthetic quality and short-term hype persistence, and a substantial portion of the upper-tail residuals likely reflects real market premia for unmeasured characteristics rather than transient mispricing. The rankings are most usefully interpreted as an attention-prioritizing device for further qualitative research, which is the role they play within the Hedonix platform.

# 6. Discussion and Limitations

## 6.1 Sample composition

The expanded sample of 360 cards spanning fourteen Scarlet & Violet expansion sets provides adequate statistical power for the core specification and allows several effects that were imprecisely estimated in earlier preliminary work — notably character-tier and artist effects — to be identified at conventional significance levels. The sample remains restricted to a single product era and to Western-market transaction venues. Extension to earlier expansion eras and to Japanese-market transaction data is a natural avenue for further work, with the caveat that pre-Scarlet & Violet card frame designs differ structurally and would likely require era-specific fixed effects.

## 6.2 The aesthetic-quality channel

The most economically meaningful limitation of the structural specification is its incomplete measurement of aesthetic quality. While the empirically constructed top-artist indicator and smart-market-divergence proxy capture portions of this channel, the upper-tail residuals continue to be dominated by cards that domain experts identify as exceptionally well-designed or as having become objects of concentrated collector enthusiasm. The hedonic-pricing literature on art markets (Renneboog and

Spaenjers, 2013) explicitly recognizes this as a near-universal challenge: observable attributes capture a meaningful share of price variation, but a structurally important residual component persists in any specification that excludes direct measures of aesthetic appeal. Appendix A documents an experimental extension using CLIP image embeddings, which provides a transparent test of whether modern vision-language representations close part of this gap.

### 6.3 Reprint risk

A second structural limitation is the omission of any direct measure of reprint risk — the probability that a given card will subsequently appear in a future reprint product, thereby expanding effective supply and depressing graded-card prices. Reprint risk is economically important in the modern Pokémon TCG market, where designated reprint products have repeatedly compressed prices for previously scarce cards. A serious analysis of reprint risk requires multi-year time-series data spanning one or more reprint events and is therefore left for subsequent work. Development of a reprint-risk module is on the Hedonix research roadmap.

### 6.4 Sample selection and survivorship

The liquidity filter (PSA 10 sales count $\geq$ 10) excludes a small number of cards with insufficient transaction volume for reliable price estimation. While this filter is methodologically necessary to control measurement error, it introduces a mild survivorship bias toward cards with active collector demand. Robustness checks at lower liquidity thresholds (not reported) confirm that the principal coefficients are qualitatively stable, with confidence intervals widening as expected when noisier observations are retained.

## 7. Conclusion

This paper applies hedonic pricing methodology to the modern Pokémon TCG graded-card market, using a sample of 360 Special Illustration Rare and Illustration Rare cards drawn from fourteen Scarlet & Violet expansion sets. A specification combining structural variables with three empirically constructed demand and quality proxies explains approximately 88% of the cross-sectional variation in log PSA 10 prices, with theoretically expected signs and robust multicollinearity diagnostics. This methodology forms the basis of the H6 Hedonic Engine, the pricing core of the Hedonix platform.

The estimated price-transmission elasticity from raw to graded markets is approximately 0.48 for Illustration Rares and approximately 0.63 for Special Illustration Rares, quantifying the differential responsiveness of higher-rarity cards to underlying market conditions. The empirically constructed top-artist indicator implies a premium of approximately +52% relative to comparable cards, providing initial structural evidence that artist-specific demand is a quantifiable and systematic driver of pricing in this market.

The persistence of a right-tail residual driven by aesthetically distinctive and hype-driven cards remains an open methodological question, partially addressed by the present specification but not eliminated. Appendix A documents that the marginal contribution of CLIP-derived image features in this dataset is modest, suggesting that further progress in this direction will likely require either substantially larger image corpora, domain-fine-tuned vision models, or hybrid approaches combining image features with social-media-derived demand signals. Subsequent Hedonix research papers will document each of these extensions as they are integrated into the production model.

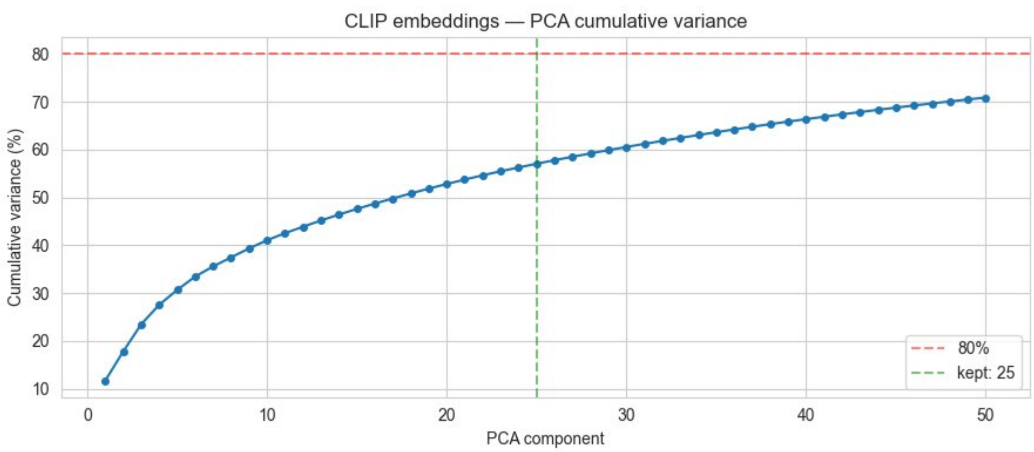
## Appendix A. Experimental Extension Using CLIP Image Embeddings

### A.1 Motivation and methodology

A natural question raised by Section 5.3 is whether modern vision-language models can systematically capture aesthetic-quality variation that the structural specification leaves in residuals. To address this question transparently, we conduct an experimental extension using OpenAI CLIP (ViT-B/32) image embeddings (Radford et al., 2021), a model pre-trained on a large corpus of natural-language-image pairs that has become a standard transferable image representation in applied machine-learning research.

For each card in the sample, we obtain the official high-resolution image from pokemontcg.io, compute the 512-dimensional CLIP image embedding, and L2-normalize the result. We then apply principal-component analysis (PCA) to the matrix of embeddings to reduce dimensionality and mitigate overfitting risk. Twenty-five principal components are retained, capturing approximately 57% of the cumulative embedding variance (Figure A1).

*Figure A1. Cumulative explained variance from principal-component analysis of CLIP image embeddings.*




We then estimate two augmented specifications. Model D adds the five principal components most highly correlated with the residuals from Model B (a forward-selection procedure designed to provide a parsimonious extension). Model C adds all twenty-five retained principal components. Both specifications retain the full structural and empirically constructed feature set from Model B.

### A.2 Results

Table A1 reports the model-fit statistics for the three specifications. $R^2$ rises modestly from 0.876 in Model B to 0.885 in Model D and 0.889 in Model C. Adjusted $R^2$ peaks at 0.876 in Model D and is essentially flat thereafter, indicating that incremental CLIP components beyond the top five contribute no additional signal once parameter penalties are accounted for. Both Akaike and Bayesian information criteria favor Model D over Model C.

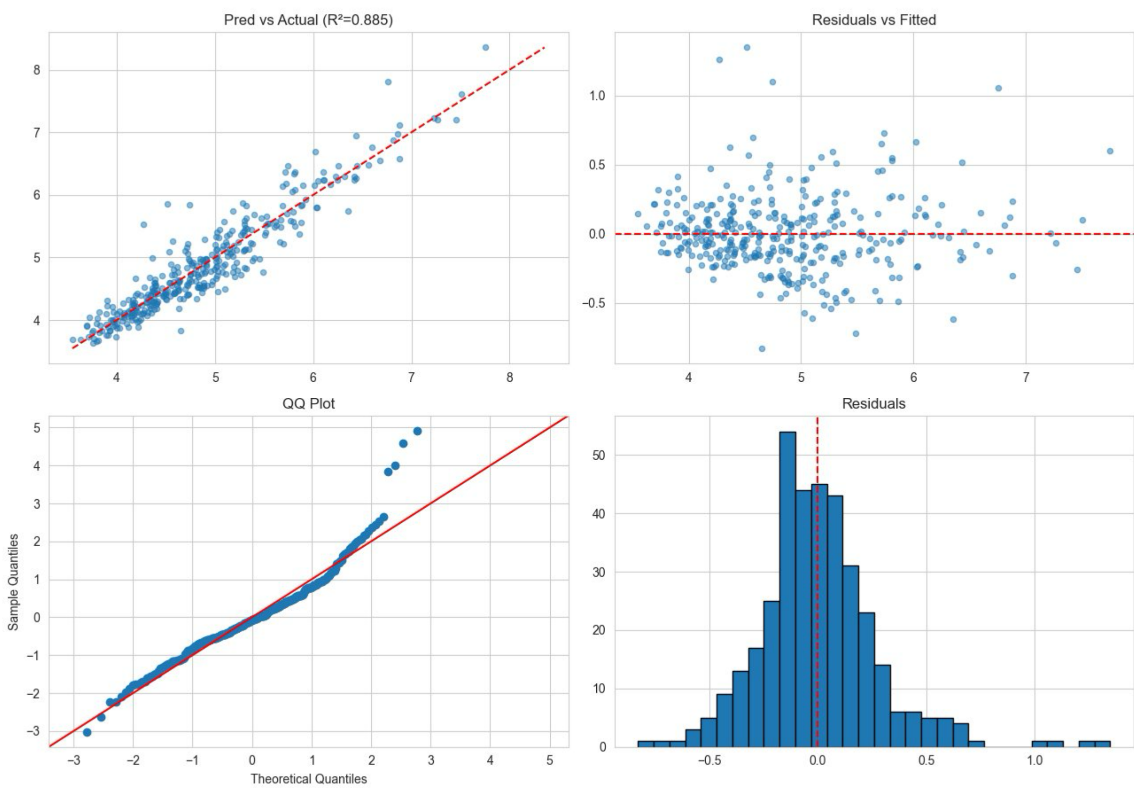
*Table A1. Model-fit comparison across structural and CLIP-augmented specifications.*

| Specification | $R^2$ | Adjusted $R^2$ | AIC |
|---|---|---|---|

| | | | |
|---|---|---|---|
| (B) Structural | 0.8759 | 0.8682 | 160.5 |
| (D) +Top 5 CLIP PCs | 0.8847 | 0.8757 | 144.1 |
| (C) +All 25 CLIP PCs | 0.8890 | 0.8727 | 170.3 |

Three of the five highest-correlation CLIP principal components are statistically significant at conventional levels in Model D, indicating that some image-derived features do carry independent information beyond the structural variables. However, the incremental $R^2$ gain of approximately one percentage point is modest in absolute terms and does not eliminate the upper-tail residuals identified in the main analysis. The card with the largest residual in Model B — exceeding the model prediction by 359% — remains a substantial outlier in Model D, with its residual reduced only to approximately 285%. Aesthetic and hype-driven premia therefore remain only partially captured.

*Figure A2. Diagnostic plots for the CLIP-augmented Model D.*




## A.3 Interpretation and implications

Three interpretive observations follow from the CLIP-augmentation results. First, the modest marginal contribution of CLIP features in this dataset is consistent with the hypothesis that visual heterogeneity within a tightly defined card-printing era is limited. CLIP embeddings, pre-trained on broadly heterogeneous internet imagery, appear to recover dimensions of variation that are largely already captured by structural variables in our specification. Second, the principal components most strongly correlated with Model B residuals capture only approximately 13% univariate correlation in absolute value, indicating that CLIP image features and aesthetic premia are weakly aligned in this market segment. Third, these results suggest that further progress in modeling aesthetic and hype premia will likely require either domain-specific fine-tuned vision models or alternative data sources capturing demand-side rather than supply-side signal — for example, social-media mention counts, community-curated wishlist data, or competitive-play tournament representation. These directions are on the Hedonix research roadmap but are not part of the production H6 Hedonic Engine at this version.

## References


Dimson, E., and Spaenjers, C. (2014). The investment performance of emotional assets. In V. A. Ginsburgh and D. Throsby (Eds.), Handbook of the Economics of Art and Culture, Vol. 2 (pp. 521–549). Elsevier. https://doi.org/10.1016/B978-0-444-53776-8.00009-9

Mandel, B. R. (2009). Art as an investment and conspicuous consumption good. American Economic Review, 99(4), 1653–1663. https://doi.org/10.1257/aer.99.4.1653

Radford, A., Kim, J. W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., Sastry, G., Askell, A., Mishkin, P., Clark, J., Krueger, G., and Sutskever, I. (2021). Learning transferable visual models from natural language supervision. Proceedings of the 38th International Conference on Machine Learning, PMLR 139, 8748–8763.

Renneboog, L., and Spaenjers, C. (2013). Buying beauty: On prices and returns in the art market. Management Science, 59(1), 36–53. https://doi.org/10.1287/mnsc.1120.1580

Rosen, S. (1974). Hedonic prices and implicit markets: Product differentiation in pure competition. Journal of Political Economy, 82(1), 34–55. https://doi.org/10.1086/260169

## About Hedonix Research

Hedonix Research is the research arm of Hedonix, a platform providing institutional-grade pricing analytics for the collectibles market. Working papers are released periodically to document the methodologies underlying the H6 Hedonic Engine and to engage with the broader academic literature on collectible-asset pricing. The views expressed in this working paper are those of the author and do not constitute investment advice. Past performance of the H6 Hedonic Engine in identifying cross-sectional mispricings does not guarantee future results. Comments are welcome at baro.philipp@gmail.com.