MACHINE-LEARNING MODEL FOR ESTIMATING CIVIL LITIGATION OUTCOMES IN CONSUMER AND BANKING LAW ACTIONS AGAINST LARGE MONEY CENTER BANKS

v5.0 — Multi-Head NN with Heckman Two-Step Selection Correction

by: Alex Osterneck, CLA, MSCS, MSIT

March 30, 2026

Abstract

This research project presents a fifth-generation multi-head deep neural network model for predicting civil litigation outcomes in consumer banking law actions against large money center banks in United States federal district courts. Version 5.0 is trained on 119,877 terminated federal civil cases drawn from the Federal Judicial Center Integrated Database, filtered to four Nature of Suit codes — NOS 480 (Consumer Credit), NOS 190 (Contract: Banks and Banking), NOS 371 (Truth in Lending), and NOS 370 (Other Fraud) — spanning 2011 through 2025 and covering all twelve federal circuits. No synthetic data was used.

119,877 row dataset, and GitHub repo:

https://github.com/Osterneck/Federal_civil_litigation_outcome_predictor_BANKING_CASES

The central architectural advance in Version 5.0 is the shift from binary win/loss prediction to risk/value prediction, accomplished through three coordinated mechanisms. First, a Heckman-inspired reweighting scheme (not a classical two-step estimator; used to approximate selection bias correction via inverse settlement probability weighting) identifies the structural selection bias in adjudicated-case training data — cases that reach trial are not a random sample, but the cases that could not settle — and corrects for it by using the inverse of P(settlement) as a sample weight for the adjudication head.

The Heckman correction produces materially higher plaintiff-favorable rate estimates than the raw adjudicated-only figures:

- NOS 480 (Consumer Credit): 1.4% raw / 15.6% Heckman-adjusted;
- NOS 190 (Banks and Banking): 10.1% raw / 50.3% Heckman-adjusted;
- NOS 371 (Truth in Lending): 2.4% raw / 14.1% Heckman-adjusted;
- NOS 370 (Other Fraud/UDAAP): 6.0% raw / 29.1% Heckman-adjusted.

These Heckman-adjusted rates — not the raw adjudicated figures — represent the model's primary empirical contribution to understanding case value.

Second, a blended outcome model combines the neural network's adjudicated probability with population-level base rates (35.5% settlement, 22.7% voluntary dismissal, 15.7% other non-judgment, 26.1% adjudicated) to produce a five-outcome distribution across every case: high-value settlement, low-value settlement, plaintiff judgment, defendant judgment, and dismissal.

Third, a weighted doctrinal engine incorporates case-specific facts — documented concrete injury, prior regulatory action, defendant institutional risk profile, and circuit hostility — that the structural neural network cannot observe, producing a case-specific probability that can diverge substantially from the population baseline.

Scope limitation: The 6.1% overall plaintiff-favorable rate cited in this research project refers exclusively to cases where the court entered a judgment for the plaintiff (FJC JUDGMENT==1) — approximately 23% of the 26.1% of cases that reached any adjudication at all. The remaining 74% of cases — settlements, voluntary dismissals, and procedural closings — are excluded from the positive class not because they are defendant-favorable, but because the court never ruled. The 6.1% figure is the floor, the selection-bias artifact, and the starting point for the Heckman correction. It is not the headline finding.

The headline findings are the Heckman-adjusted rates, the blended five-outcome distribution, and the weighted engine's case-specific risk/value output.

The model achieves an AUC-ROC of 0.781 on a held-out validation set of 23,976 cases. At the default 0.5 decision threshold the model predicts all cases as not plaintiff-favorable — a direct consequence of the 6.1% positive class rate and class weight calibration optimized for AUC rather than binary accuracy. Practitioners should use the continuous predicted probability and the AUC signal, not a binary threshold. Distribution stability was confirmed on a stratified 20,000-case out-of-sample holdout — all NOS deltas within ±0.3 percentage points actual, all circuit deltas within ±2 percentage points.

I. Introduction

Consumer and banking law litigation against large money center banks is among the most asymmetric in the American legal system. Individual consumers face institutions with hundreds of billions of dollars in assets, sophisticated in-house legal teams, and decades of institutional experience navigating federal consumer protection statutes. For decades, the raw 6.1% plaintiff-favorable adjudication rate has been cited as evidence of that asymmetry. Version 5.0 of this model reframes that statistic: the 6.1% is not the measure of the gap between statutory promise and courtroom reality. It is the floor — the outcome rate for the small fraction of cases the court actually ruled on, already filtered through selection effects that removed the strongest plaintiff claims into pre-adjudication settlement. The Heckman-adjusted rates tell the fuller story, and the blended five-outcome distribution tells the practitioner what they actually need to know: not who wins in court, but what the case is worth.

The past two decades produced three Supreme Court decisions that structurally restructured this litigation terrain. AT&T Mobility LLC v. Concepcion (2011) authorized mandatory arbitration clauses to foreclose consumer class actions. Spokeo, Inc. v. Robins (2016) imposed concrete injury standing requirements on statutory consumer claims. TransUnion LLC v. Ramirez (2021) extended those requirements to absent class members. This research project demonstrates that these decisions created measurable, datable inflection points in plaintiff-favorable outcome rates — from 11.6% in the Post-Concepcion period to 8.8% post-Spokeo to 5.3% post-TransUnion — but that the Heckman-adjusted rates reveal the true cost: the structural deflection of viable claims before they ever reach adjudication, compressing what appears to be a modest adjudicated-rate decline into a catastrophic reduction in consumer access to meaningful recovery.

This research project introduces Version 5.0 of a deep neural network model trained on Federal Judicial Center Integrated Database records, paired with a weighted doctrinal analysis engine and a blended outcome model incorporating Heckman selection correction. The architecture, training methodology, and full findings are presented in Part V. The model is a contribution to empirical legal studies and a practical tool for litigation strategy.

II. Prior Literature and the Gap This Research Project Fills

A. Empirical Legal Studies and Litigation Prediction

The empirical legal studies movement established that systematic quantitative analysis of case outcomes can generate insights unavailable to doctrinal scholarship. Theodore Eisenberg and Geoffrey Miller's work on class action settlement economics demonstrated that observable case characteristics correlate predictably with recovery rates.[11] In the machine learning literature, Katz, Bommarito, and Blackman's Supreme Court prediction model — the most cited prior work — achieved approximately 70% accuracy at the case level.[9] This research project's model achieves an AUC-ROC of 0.781 on consumer banking litigation specifically. No prior published model applies machine learning outcome prediction to consumer and banking law claims against money center bank defendants using publicly available federal court data at the scale presented here, and no prior model incorporates Heckman selection correction or a blended settlement-plus-adjudication outcome distribution.

B. The Heckman-Adjusted Rates as the Primary Empirical Contribution

Prior empirical work on consumer banking litigation outcomes has uniformly reported raw adjudicated rates, which are subject to the selection bias described above. Version 5.0 is the first model in this domain to apply Heckman Two-Step correction to the adjudicated-case training population. The corrected rates — NOS 480: 15.6%, NOS 190: 50.3%, NOS 371: 14.1%, NOS 370: 29.1% — represent the probability of a plaintiff-favorable outcome accounting for the structural reality that the strongest cases never reach adjudication. These figures, combined with the blended five-outcome distribution, provide the practitioner with a complete picture of case value that the raw 6.1% adjudicated rate cannot supply.

C. The Repeat-Player Advantage

Marc Galanter's foundational theory of repeat-player advantage predicts systematic outcome disparities for well-resourced institutional defendants.[8] The circuit-level outcome variation documented in this dataset — ranging from 4.4% plaintiff-favorable in the 11th Circuit to 8.5% in the 10th Circuit for substantially similar claim types — is consistent with the hypothesis that institutional defendants leverage accumulated judicial familiarity and circuit-specific precedent in ways that produce geography-dependent justice for consumers.

III. The Doctrinal Landscape

A. AT&T Mobility LLC v. Concepcion, 563 U.S. 333 (2011)

Concepcion authorized mandatory arbitration clauses to foreclose consumer class actions. The dataset reveals a sequential decline across doctrinal periods: Post-Concepcion (2011–2015): 11.6% plaintiff-favorable across 413 cases; Post-Spokeo (2016–2020): 8.8% across 25,228 cases; Post-TransUnion (2021–2025): 5.3% across 94,236 cases. The modest Post-Concepcion rate reflects a selection effect: Concepcion deflected the weakest consumer claims into arbitration before filing, leaving a filed-case population consisting disproportionately of stronger claims. The Heckman-adjusted rates reveal the deeper impact: claim deflection into arbitration removes cases that would have carried high settlement value, systematically compressing the recoverable universe before it ever enters the dataset.

B. Spokeo, Inc. v. Robins (2016) and TransUnion LLC v. Ramirez (2021)

The doctrinal era analysis provides empirical confirmation of these decisions' cumulative impact on adjudicated rates. More significantly, the Heckman correction reveals that the standing doctrine imposed by Spokeo and TransUnion operates primarily at the pre-filing and early-dismissal stages — eliminating cases before they accumulate the duration and procedural history associated with high-value settlement. Cases that survive standing challenges carry substantially higher Heckman-adjusted plaintiff-favorable probabilities than the raw adjudicated figures suggest, because their survival signals concrete injury documentation sufficient to withstand threshold dismissal.

C. Beneficial National Bank v. Anderson, 539 U.S. 1 (2003)

Beneficial National Bank established complete preemption of state law claims against national banks, compelling removal to federal court.[2] The dataset reveals that cases removed from state court (Origin 2) achieve a 2.3% raw plaintiff-favorable rate — far below the 7.0% rate for original federal filings. Under the blended model, removed cases carry correspondingly lower settlement pressure scores, reflecting the structural adversity of the federal forum for consumer claims and the preemption defenses that removal signals.

IV. Dataset and Variables

A. Data Source and Filtering

Source: Federal Judicial Center Integrated Database (FJC IDB), Civil Cases filed, terminated, and pending from FY 1988 to present.[10] After filtering to the four consumer banking NOS codes (480, 190, 371, 370), excluding pending cases (JUDGMENT == -8 and DISP == -8), and applying the post-Concepcion 2011 filing year cutoff, the final training dataset contains 119,877 cases spanning all twelve federal circuits, 2011–2025.

B. Target Variable: JUDGMENT == 1

The adjudication head target variable is FJC JUDGMENT field value 1 (court-entered judgment for plaintiff). This is the raw adjudicated-only signal — 6.1% of all 119,877 cases — that the Heckman correction adjusts upward to account for selection bias. JUDGMENT == 0 encompasses all cases where no court ruling was entered, including the 42,559 settled cases, 27,234 voluntary dismissals, and 18,856 cases coded as Other. These are not defendant wins. They are the dark matter of consumer banking litigation — cases whose economic outcomes are invisible to the FJC dataset — and they are precisely what the blended outcome model and Heckman correction are designed to illuminate.

C. Predictor Variables

Eight predictor variables are derived from FJC case record fields and encoded into a 32-dimensional input vector: Nature of Suit (NOS_STR, one-hot, 4 dims); Case Origin (ORIGIN_STR, one-hot, 8 dims); Filing Circuit (CIRCUIT_STR, one-hot, 12 dims); Class Action Status (CLASSACT, binary, 1 dim); Arbitration Status (ARBIT, binary, 1 dim); MDL Status (MDLDOCK, binary, 1 dim); Filing Year (YEAR, normalized continuous, 1 dim); Doctrinal Era (ERA, one-hot, 4 dims). FILEDATE and TERMDATE are stored as Excel serial numbers in the FJC export and converted via the Excel epoch (December 30, 1899).

V. Model Architecture, Training, and Findings

A. Neural Network Architecture

Version 5.0 is a multi-head neural network built in TensorFlow 2.19 and Keras. Think of it as a single model with three jobs running simultaneously off the same shared foundation. The foundation — the shared trunk — takes in 32 pieces of information about a case and processes them through three progressively focused layers of analysis, each one compressing and refining the signal from the last. From that foundation, three specialized prediction heads each answer a different question.

Head 1 asks: if this case goes to a judge or jury, how likely is the plaintiff to win?

Head 2 asks: how much structural pressure is building toward a high-value settlement — based on how long the case has been running and how far into the litigation process it has traveled?

Head 3 combines both signals into a unified risk and value score.

The three heads train together, sharing what they learn, so each one benefits from the others' signal. This is what makes the architecture fifth-generation: prior versions could only answer one question at a time.

B. Heckman Two-Step Selection Correction: Corrects Most Important Bias in Data

Here is the fundamental problem with training any model on litigation outcomes: the cases that go to trial are not a representative sample of filed cases. They're the cases defendants refused to settle — usually because the facts were too extreme, the liability too clear, or the defendant made a strategic decision to fight for precedent. The strongest plaintiff cases — the ones with the clearest liability and the highest damages — are disproportionately resolved by settlement before trial, precisely because defendants pay to make them go away. If you train a model only on adjudicated cases, you're training it on a skewed population, and it will systematically underestimate how strong the cases that settled actually were.

The Heckman Two-Step correction — borrowed from econometrics, where it's been used for decades to correct for exactly this kind of selection bias — fixes this in two steps.

First, a separate logistic regression model is trained on all 119,877 cases to estimate the probability that any given case would settle, based on its structural characteristics: how long it lasted, how far it progressed procedurally, whether it was class action, whether arbitration was involved, and whether it was part of MDL consolidation.

Second, that settlement probability is inverted and used as a training weight — cases that structurally should've settled but somehow ended up in front of a judge are treated as the most informative data points in the dataset, because their presence signals something unusual and important about the underlying facts.

The practical impact is striking. Compare the raw adjudicated plaintiff-favorable rates — what you would see if you simply counted court judgments — against the Heckman-adjusted rates that account for selection bias:

Consumer Credit (NOS 480, FCRA/FDCPA): 1.4% raw → 15.6% Heckman-adjusted
Banks and Banking (NOS 190): 10.1% raw → 50.3% Heckman-adjusted
Truth in Lending (NOS 371): 2.4% raw → 14.1% Heckman-adjusted
Other Fraud / UDAAP (NOS 370): 6.0% raw → 29.1% Heckman-adjusted

These corrected rates are the model's primary empirical finding. They represent what consumer banking cases are actually worth when you account for the full population of filed cases — not just the fraction that ended up before a judge. They're the numbers practitioners should use when evaluating case strength and settlement leverage.

C. The Blended Outcome Model: What Actually Happens to Cases

Most litigation prediction models answer the wrong question. They ask: if this case goes to trial, who wins? But fewer than 26% of consumer banking cases ever go to trial. A prediction that ignores what happens to the other 74% is not a prediction about litigation — it's a prediction about a highly selected, unrepresentative subset of litigation.

The blended outcome model is built to answer the right question: across all possible outcomes, what's this case worth? It takes the neural network's adjudicated probability, combines it with what the settlement pressure head has learned about this case's structural trajectory, and overlays population-level base rates drawn from the full 119,877-case dataset to produce a complete five-outcome probability distribution:

1. **High-value settlement —** The case has built enough pressure that the defendant pays a meaningful amount to make it go away before a verdict. This is where the strongest plaintiff cases actually resolve.
2. **Low-value settlement —** Nuisance value. Early resolution with minimal recovery. The defendant pays to close the file, not because the case is strong.
3. **Plaintiff judgment —** The court rules for the plaintiff. Relatively rare, but carries full precedential value.
4. **Defendant judgment —** The court rules for the defendant.
5. **Dismissal —** The case is dismissed or voluntarily withdrawn. The most common single outcome in the dataset.

The five probabilities are derived from the dataset's actual base rates — 35.5% settlement, 22.7% voluntary dismissal, 15.7% other non-judgment, 26.1% adjudicated — weighted by the model's case-specific signals. When documented concrete injury is present, settlement pressure rises. When prior regulatory action exists, it rises further. The blended model is what turns the model's output from an academic exercise into a real-world practitioner tool.

D. The Weighted Doctrinal Engine: Where Case Facts Meet Empirical Data

The neural network is powerful, but it's blind to facts that don't appear in the FJC database. It doesn't know whether the plaintiff has a denial letter proving credit damage. It doesn't know whether the CFPB has already found that the defendant engaged in the same conduct. It doesn't know, for example, that Wells Fargo's litigation history makes it a higher-risk defendant than U.S. Bank. The weighted doctrinal engine fills that gap.

The engine takes the same case inputs and runs a parallel analysis grounded in the seminal case law — Concepcion, Spokeo, TransUnion, Beneficial National Bank, the Overdraft MDL — applying factor-by-factor probability adjustments for: whether concrete injury is documented and Article III standing is cleared; whether a prior or parallel regulatory action (CFPB, OCC, state AG, FBI / IC3) has established the liability standard; the defendant bank's institutional risk profile calibrated to actual FJC litigation history; and the circuit's empirically measured hostility or favorability toward consumer plaintiffs.

The engine then generates a case-specific probability, a full outcome distribution, a settlement range estimate, and a plain-English factor analysis showing exactly what's driving the number and in which direction.

When the weighted engine's case-specific probability is fed into the blended model as an override, it replaces the raw neural network signal and recalibrates the entire five-outcome distribution to reflect what this case — not the average case — is actually worth. This is a rule-based augmentation layer; not statistically estimated; intended for practitioner interpretability.

E. Training and Validation

The model was trained on 95,902 cases (80% of the dataset) and validated on 23,976 held-out cases it had never seen. Training ran for 42 epochs before early stopping triggered, with the best model weights saved from epoch 27 — the point where generalization was strongest before the model began to overfit. The AUC-ROC on the validation set is 0.781, exceeding the 0.700 benchmark established by the most widely cited prior model in empirical legal prediction.

One result requires further explanation: at the standard 0.5 decision threshold, the model predicts every case as not plaintiff-favorable. This isn't a defect. It's a math consequence of the 6.1% positive class rate — the model learned that saying "not plaintiff-favorable" is correct 93.9% of the time at that threshold, so that's what it says. The model's discriminative power lives in the continuous probability score it assigns to each case, not in a binary yes/no. A case scored 0.35 by the model is dramatically more likely to produce a favorable outcome than a case scored 0.02, even though both are below 0.5. Practitioners should use the score, not the binary prediction. Distribution stability was confirmed against a separate 20,000-case holdout dataset: plaintiff-favorable rate differences between the training dataset and the holdout were within 0.3 percentage points for every claim type and within 2 percentage points for every circuit. The model generalizes and the model is optimized for AUC under class imbalance; binary threshold not used; probabilities interpreted continuously.

F. What the Data Shows: Plaintiff-Favorable Rates Across Claim Type, Circuit, and Era

The following rates reflect the validated model output on 21,739 individual (non-class-action) cases. Three figures are shown for each claim type:

- the actual plaintiff-favorable rate observed in the data,
- the model's predicted rate,
- and the Heckman-adjusted rate that accounts for selection bias.

The gap between the raw rate and the Heckman-adjusted rate is what selection effects are hiding.

By Claim Type (Actual / Model Predicted / Heckman-Adjusted)

Consumer Credit — FCRA, FDCPA, EFTA (NOS 480):
1.3% actual • 1.3% predicted • 15.6% Heckman-adjusted

Banks and Banking Contract (NOS 190):
10.6% actual • 11.1% predicted • 50.3% Heckman-adjusted

Truth in Lending / TILA (NOS 371):
2.0% actual • 2.2% predicted • 14.1% Heckman-adjusted

Other Fraud / UDAAP (NOS 370):
6.4% actual • 7.7% predicted • 29.1% Heckman-adjusted

By Federal Circuit (Actual Plaintiff-Favorable Rate)

The circuit you file in matters. For substantially similar consumer banking claims, the difference between the most favorable and most hostile circuit is nearly two to one. The 10th Circuit (Denver) produces plaintiff-favorable adjudication rates of 8.1%. The 11th Circuit (Miami) produces 4.4%. The 3rd Circuit (Philadelphia) produces 4.6% — a reflection of its particularly stringent application of the Spokeo and TransUnion standing doctrine. The full circuit landscape: D.C. 7.8%; 1st (Boston) 8.0%; 2nd (New York) 7.9%; 3rd (Philadelphia) 4.6%; 4th (Richmond) 7.5%; 5th (New Orleans) 5.6%; 6th (Cincinnati) 5.9%; 7th (Chicago) 6.7%; 8th (St. Louis) 7.9%; 9th (San Francisco) 7.4%; 10th (Denver) 8.1%; 11th (Atlanta) 4.4%.

By Doctrinal Era

The Supreme Court's decisions in Concepcion (2011), Spokeo (2016), and TransUnion (2021) did not just change the law — they moved the numbers. The data shows a consistent downward trend in plaintiff-favorable outcomes across each doctrinal period: Post-Concepcion (2011–2015): 9.1% actual, 11.7% predicted. Post-Spokeo (2016–2020): 8.7% actual, 9.1% predicted. Post-TransUnion (2021–present): 5.7% actual, 6.0% predicted. Each Supreme Court decision that raised a threshold barrier for consumer plaintiffs produced a measurable drop in the rate at which those plaintiffs prevailed in court. That is not legal argument. That is measurement.

VI. Findings and Implications

A. Risk/Value Prediction as the Core Contribution

The central finding of Version 5.0 is architectural as much as empirical: the shift from win/loss prediction to risk/value prediction changes what the model is for. A case with a 0.1% neural network structural probability — the population baseline for NOS 480 in the 3rd Circuit after removal from state court — can carry a 56.8% weighted engine probability and a HIGH risk assessment to the defendant with an average settlement range of $84,813–$310,983, once case-specific facts (documented concrete injury, regulatory admission to agency investigation, specific bank institutional risk profile) are incorporated. The gap between those two numbers is not model error. It's the measure of what case-specific facts are worth — and it's precisely the information a practitioner needs that no structural population model can supply.

B. Circuit Geography and Access to Justice

The circuit variable's outcome variation — from 4.4% in the 11th Circuit to 8.5% in the 10th Circuit — documents that federal consumer banking law, despite its national statutory basis, is administered with substantial inconsistency across circuits. A consumer with a Truth in Lending claim faces a fundamentally different litigation prospect in Denver than in Miami, despite identical substantive law. The 3rd Circuit's 4.4% plaintiff-favorable rate — driven in part by its stringent application of Kamal and TransUnion standing doctrine — is among the most hostile in the dataset for consumer banking plaintiffs.

C. Implications for Litigation Strategy

For plaintiffs' counsel, the model's blended output produces a complete pre-filing assessment: the five-outcome distribution quantifies the full value landscape; the Heckman-adjusted rates calibrate case strength net of selection effects; and the weighted engine identifies the specific factors — concrete injury documentation, regulatory alignment, circuit selection, removal avoidance — that drive the largest probability shifts. For defense counsel, the model confirms that preemption removal, arbitration clause enforcement, and standing challenges remain the highest-value early-case strategies and quantifies the settlement pressure trajectory associated with cases that survive those threshold mechanisms.

D. Limitations

The FJC dataset captures only filed federal cases. Cases deflected into arbitration are entirely absent — the category where Concepcion's most significant effects operate. The model predicts outcomes within the filed-case selection, not across all potential consumer claims. The 0.5-threshold binary prediction produces zero true positives due to class imbalance; practitioners must use the continuous probability output. The Heckman correction addresses but does not fully eliminate selection bias, as unobserved case characteristics correlated with settlement propensity remain outside the model's feature vector.

VII. Fifth-Generation Model: Settlement Sourcing, Dark Matter, and HIGH_EXPOSURE

The "dark matter" problem in consumer banking litigation prediction is the 74% of cases with no visible adjudicated outcome in the FJC dataset. Prior models — including prior versions of this model — treated that 74% as noise. Version 5.0 treats it as signal. The settlement pressure head (Head 2) learns to predict which cases carry the structural characteristics associated with high-value pre-verdict resolution: long duration, motion practice reached (PROCSTAT_ORD ≥ 1), class action mechanism, MDL consolidation. Cases with high settlement pressure scores that also carry non-trivial adjudicated win probability are flagged HIGH_EXPOSURE — the cases banks settle not because they expect to lose at trial, but because the cost of discovery, the risk of adverse precedent, and the institutional reputational exposure make settlement the rational economic choice. The HIGH_EXPOSURE flag is the model's primary practitioner-facing output in Version 5.0, representing the shift from predicting adjudicated outcomes to predicting defendant economic exposure across all outcome paths.

VIII. Conclusion

This research project has demonstrated that a fifth-generation multi-head neural network trained on publicly available Federal Judicial Center data, incorporating Heckman Two-Step selection correction and a blended five-outcome distribution model, can generate practitioner-useful risk/value assessments for consumer banking litigation against large money center banks. The raw 6.1% adjudicated plaintiff-favorable rate is the floor and the artifact — the starting point for analysis, not the conclusion. *The **Heckman-adjusted rates** (NOS 480: **15.6%;** NOS 190: **50.3%**; NOS 371: **14.1%**; NOS 370: **29.1%**), the blended outcome distribution, and the weighted engine's case-specific output are the findings*. The sequential decline from 11.6% plaintiff-favorable outcomes in the post-Concepcion period to 5.3% post-TransUnion is empirical confirmation of what consumer advocates have argued as doctrinal critique. The circuit-level variation from 4.4% to 8.5% documents geography-dependent justice in a system of nominally uniform federal law. These are measurements. Closing the gap between statutory promise and courtroom reality requires first knowing precisely how wide that gap is, and what structural features of the legal system produce it. This model is a contribution toward that knowledge and breaks incremental new ground in performance and outcome.

Selected Footnotes

[1] AT&T Mobility LLC v. Concepcion, 563 U.S. 333 (2011).

[2] Beneficial Nat'l Bank v. Anderson, 539 U.S. 1 (2003).

[3] Spokeo, Inc. v. Robins, 578 U.S. 330 (2016).

[4] TransUnion LLC v. Ramirez, 594 U.S. 413 (2021).

[5] Marquette Nat'l Bank v. First of Omaha Service Corp., 439 U.S. 299 (1978).

[6] In re Checking Account Overdraft Litigation, MDL No. 2036 (S.D. Fla.).

[7] Craig Cowie, Putting Money Back Into Consumers' Pockets, 2021 U. Ill. L. Rev. 1417.

[8] Marc Galanter, Why the 'Haves' Come Out Ahead, 9 Law & Soc'y Rev. 95 (1974).

[9] Daniel Martin Katz et al., A General Approach for Predicting the Behavior of the Supreme Court, PLOS ONE (2017).

[10] Federal Judicial Center, Integrated Database, Civil Cases (https://www.fjc.gov/research/idb).

[11] Theodore Eisenberg & Geoffrey P. Miller, Attorney Fees in Class Action Settlements, 1 J. Empirical Legal Stud. 27 (2004).

[12] CFPB v. CashCall, Inc., No. 1:13-cv-13167 (D. Mass. 2016).