Special Teams Workload Volatility and Next Week Injury Risk in the NFL

Independent Student Research

Siddharth Korukonda (Stony Brook University)

January 6, 2026

Abstract

Injury risk in the NFL is commonly discussed through the lens of contact intensity, practice scheduling, and total snap counts, but special teams creates a distinct workload channel, it concentrates high speed collisions among a subset of players whose roles can fluctuate sharply across weeks. This report studies whether week to week instability in special teams workload, including short window volatility and spike style shock events, is systematically associated with injury outcomes in the following week. Using a 2012 to 2024 team week panel constructed from public NFL play by play and roster data, we model next week injuries for offense and defense with fixed effects count models and fixed effects logistic models, and we translate coefficients into football interpretable quantities. The core empirical pattern is consistent across descriptive views and model based translations, non score linked special teams shock weeks are associated with higher next week injuries on both sides of the ball. Volatility in recent special teams workload is more relevant for offense than for defense, while slower moving cumulative exposure behaves as a smaller contributor in the aggregate. We conclude by connecting these results to the Collective Bargaining Agreement (CBA), which heavily regulates practice structure and contact but provides limited trigger based safeguards for game driven special teams workload spikes.

1 Introduction

Player health has become one of the defining constraints of modern NFL decision making, it shapes roster construction, competitive balance, and the economic life cycle of careers. The last decade of sports science has also made a clear conceptual point, injury risk is not only about how much load a player carries in total, it is also about how that load changes. When workload ramps faster than the body can adapt, the probability of soft tissue injury, joint stress, and compensatory movement increases, even if the absolute volume looks normal on paper. This creates a natural tension in football, where coaches optimize for weekly matchup advantages, while players experience the season as a sequence of short recovery windows.

Special teams sits at the intersection of these forces. It is a phase of the game with unusually high speed, unusual angles, and an outsized rate of open field contact relative to many offensive and defensive situations. It also has a different participation logic than offense and defense. Many special teams snaps are taken by backups, depth players, and hybrid role athletes, meaning the set of exposed bodies can change sharply as injuries occur, as tactical plans adjust, or as game scripts shift. In other words, special teams is not just additional load, it is irregular load, and irregular load is precisely what many injury frameworks highlight as risky.

The intuition can be grounded in a standard acute versus chronic framing. If an athlete has adapted to a stable chronic workload, then a sudden acute spike (an extra set of repeated sprints, additional collisions, or more change of direction demands) can exceed tolerance, especially when paired with limited recovery and travel. In the NFL, this logic is typically applied to practice intensity and total snap counts, but special teams can create spikes that are not well approximated by offensive and defensive snap totals alone. A team can have a normal offensive and defensive workload profile but still experience a chaotic special teams week, for example due to a weather driven punt heavy game, a returner who consistently brings kicks out, repeated penalties that force re kicks, or overtime sequences that add extra special teams plays.

Historically, special teams has been recognized as both strategically valuable and physically costly. Kickoff and punt plays create long run ups, compressed lanes, and collisions that often occur at higher approach speeds than many routine offensive and defensive snaps. The league has introduced multiple safety oriented adjustments over time, but many of those changes are framed around specific play types or isolated mechanisms rather than the broader question of week to week load management. Even in a modern environment with fewer returns, the burden of coverage, protection, and return roles still concentrates on a limited group of players, and that burden can expand quickly when a team is short handed, when the opponent forces field position exchanges, or when game conditions demand repeated special teams sequences.

From an analytics perspective, this setting is a textbook reminder that totals are not the whole story. Two weeks can have the same average special teams volume, but one can be smooth and predictable while the other is spiky and disruptive. If physiological adaptation is gradual and recovery is constrained, then the spiky sequence should be more hazardous, even when the season total workload looks similar. This motivates our focus on shocks and volatility as primary features, with cumulative exposure treated as an additional context signal.

This report asks a targeted question that is practical for both analysts and decision makers. Are team weeks with unusually spiky or volatile special teams workload associated with measurably higher injury outcomes in the following week, even after controlling for stable team differences and league wide time variation. The answer matters for two reasons. First, if special teams spikes have a stable relationship with next week injuries, then they are a candidate for monitoring and trigger based recovery protocols, similar to how some clubs already manage pitch counts or sprint loads in other sports. Second, because the NFL labor framework strongly regulates practice structure, any meaningful game driven workload channel that is not captured by practice rules is a natural policy gap to examine.

2 Theory and measurement

Why shocks and volatility are meaningful workload signals

Two related concepts motivate our feature design.

Shocks capture acute spikes. In workload science, a key risk driver is an acute increase relative to what has been typical for the athlete or the system. At the team week level, we approximate this by defining a special teams shock as a week in which a non score linked special teams workload measure is unusually high relative to a rolling historical baseline. Conceptually, if L_{it} is the team i special teams workload in week t, and μ_{it} and σ_{it} are rolling mean and standard deviation over a lookback window, then a standardized spike score can be written as

$$Z_{it} = \frac{L_{it} - \mu_{it}}{\sigma_{it}},$$

and a shock indicator is

$$S_{it} = \mathbb{1}\{Z_{it} \geq \tau\},$$

where τ is a threshold chosen to represent a meaningful spike.

Volatility captures instability. Even if a team does not cross an explicit spike threshold, frequent swings in workload can be disruptive. Volatility is measured as the rolling standard deviation of the same workload signal,

$$V_{it} = \sqrt{\frac{1}{K-1}\sum_{k=1}^{K}\left(L_{i,t-k} - \bar{L}_{it}\right)^2},$$

where K is the lookback horizon and $\bar{L}_{it}$ is the rolling mean. Volatility is not intended to represent more workload in expectation, it represents unpredictability, which can matter because it interacts with preparation, role assignments, and recovery planning.

Cumulative exposure as a slower moving contributor

A complementary hypothesis is that repeated spike exposure accumulates. We therefore compute a cumulative shock count,

$$C_{it} = \sum_{s<t} S_{is},$$

or the same object within a restricted season window. If acute spikes matter most, then S_{it} should dominate. If chronic accumulation matters more, then C_{it} should be the more informative signal. In practice, both can contribute, and the key is to estimate their relative importance.

3 Data

3.1 Unit of analysis and time span

The analysis is conducted at the team week level over 2012 to 2024. The team week unit is a practical compromise, it is granular enough to capture schedule and workload variation, while remaining stable to roster churn and injury reporting noise. Play by play information is used to construct special teams workload features and contextual controls, and injury outcomes are aggregated to the team week.

3.2 Outcomes

We study two types of next week injury outcomes.

Count outcomes. Let $Y^{\text{def}}_{i,t+1}$ be the number of defensive injuries recorded for team i in week $t+1$, and let $Y^{\text{off}}_{i,t+1}$ be the analogous offensive injury count. Count models are appropriate because the data are nonnegative integers, and the conditional mean can vary with covariates.

Any injury outcomes. Let $A^{\text{def}}_{i,t+1} = \mathbb{1}\{Y^{\text{def}}_{i,t+1} \geq 1\}$ and similarly for offense. Logistic models translate naturally into risk probabilities that practitioners can communicate.

3.3 Key exposure features

We focus on a non score linked special teams workload signal because it better isolates special teams volume that is not mechanically driven by offensive scoring. We then derive three exposures from it.

- S_{it}, a shock indicator for unusually high non score linked special teams workload.
- V_{it}, a rolling volatility measure (instability) for the same workload signal.
- C_{it}, a cumulative prior shock measure.

The remainder of the feature set includes controls for schedule and season timing, and fixed effects absorb persistent team differences and week specific league wide conditions.

4 Statistical methodology and safeguards

4.1 Count models with fixed effects

For injury counts we estimate fixed effects generalized linear models. The canonical specification is

$$\mathbb{E}\left[Y_{i,t+1} \mid X_{it}\right] = \exp\left(\alpha_i + \gamma_t + \beta_1 S_{it} + \beta_2 V_{it} + \beta_3 C_{it} + \delta^\top W_{it}\right),$$

where α_i are team fixed effects, γ_t are season week fixed effects, and W_{it} are additional controls. The exponential mean ensures positivity. Coefficients are interpreted through incidence rate ratios (IRRs), for example $\exp(\beta_1)$ is the multiplicative change in expected injuries associated with a shock week.

Count data in football are often overdispersed, meaning variance exceeds the mean. We therefore compare Poisson and Negative Binomial (NB) variants. In practice, Poisson fixed effects models combined with team clustered uncertainty are frequently used because they remain consistent for the conditional mean even when the variance is misspecified, and NB variants provide a complementary check that effect directions are not an artifact of dispersion assumptions.

4.2 Logistic models for any injury risk

For the probability of at least one injury next week we estimate

$$\Pr(A_{i,t+1} = 1 \mid X_{it}) = \Lambda\left(\alpha_i + \gamma_t + \beta_1 S_{it} + \beta_2 V_{it} + \beta_3 C_{it} + \delta^\top W_{it}\right),$$

where $\Lambda(u) = \frac{1}{1+e^{-u}}$ is the logistic link. Logistic coefficients are not directly interpretable in probability space, so we translate them into average marginal effects (AMEs), computed as the average change in predicted probability when toggling S_{it} or shifting V_{it} from the 25th to 75th percentile.

4.3 Predictive integrity and stability checks

Because sports data are noisy and strongly time structured, we emphasize safeguards that reduce the risk of overfitting and spurious narratives.

- **Time aware validation.** Models are evaluated in a time respectful manner (training on earlier weeks and testing on later weeks) so that apparent performance is not inflated by leakage.
- **Cluster robust uncertainty.** Standard errors are clustered at the team level to account for within team dependence across weeks.

- **Robustness variants.** Key effects are checked across closely related specifications (alternative exposure definitions and alternative link functions) to confirm that conclusions are not driven by a single modeling choice.
- **Placebo logic.** We examine whether exposures predict outcomes they should not plausibly affect at the same strength, which helps distinguish stable signals from coincidental correlations.

These checks are not a substitute for causal identification, but they strengthen the credibility of the associations by demonstrating stability.

5 Exploratory patterns and results

5.1 Volatility and shock prevalence

Figure 1 shows the distribution of the special teams volatility measure. The distribution is right skewed, meaning most team weeks have modest instability, while a smaller set of weeks exhibit pronounced swings.

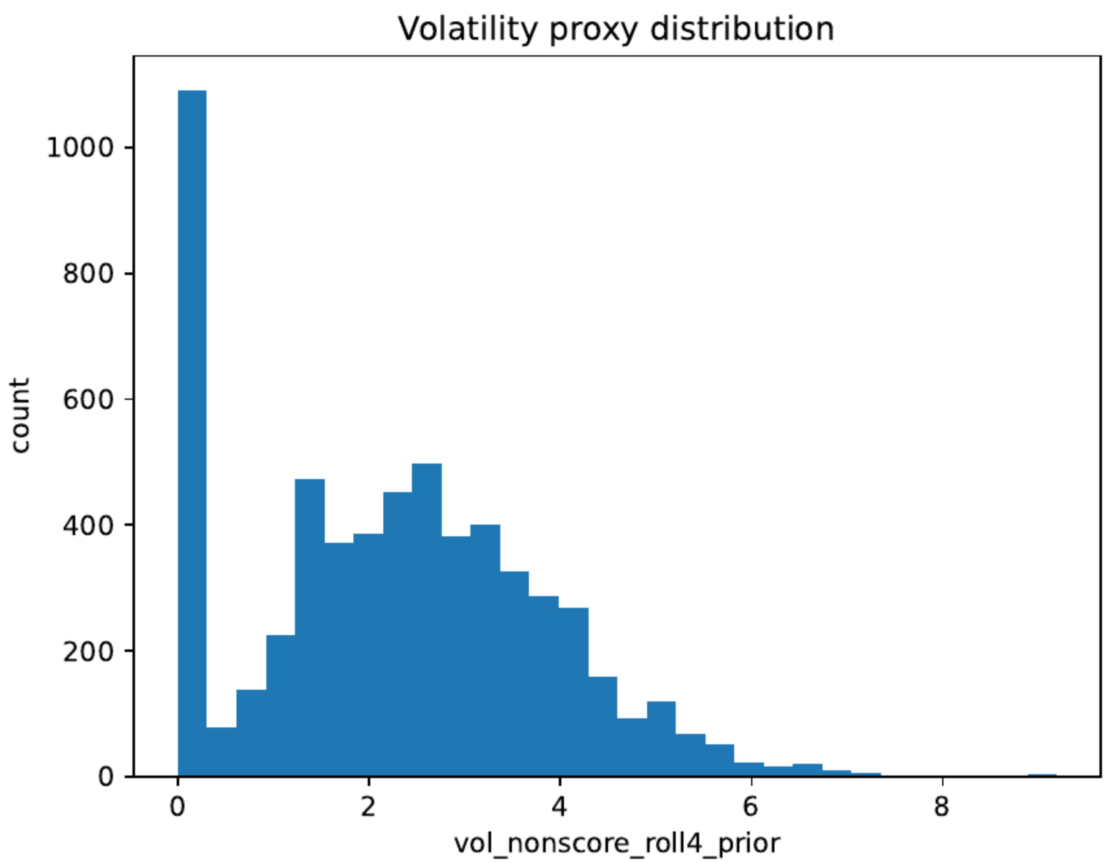

Figure 1: Distribution of special teams volatility across all team weeks.

This shape matters because it implies the most volatile weeks are not typical weeks. When volatility is high, it often corresponds to weeks where special teams is repeatedly pulled into the spotlight by game context. This is precisely the environment where acute load and role disruption mechanisms become plausible, a unit may be asked to run repeated high speed coverage reps with limited recovery, while the coaching staff has less ability to substitute without weakening other phases.

From a modeling perspective, a right tailed exposure is useful because it creates contrast. If every week looked similar, statistical identification would be difficult. Here, volatility provides a natural rank ordering of stable versus unstable special teams environments.

Figure 2 shows the season level prevalence of shock weeks. Rather than being an anomaly, shocks appear regularly, which supports the idea that a trigger based monitoring system could be practically relevant.

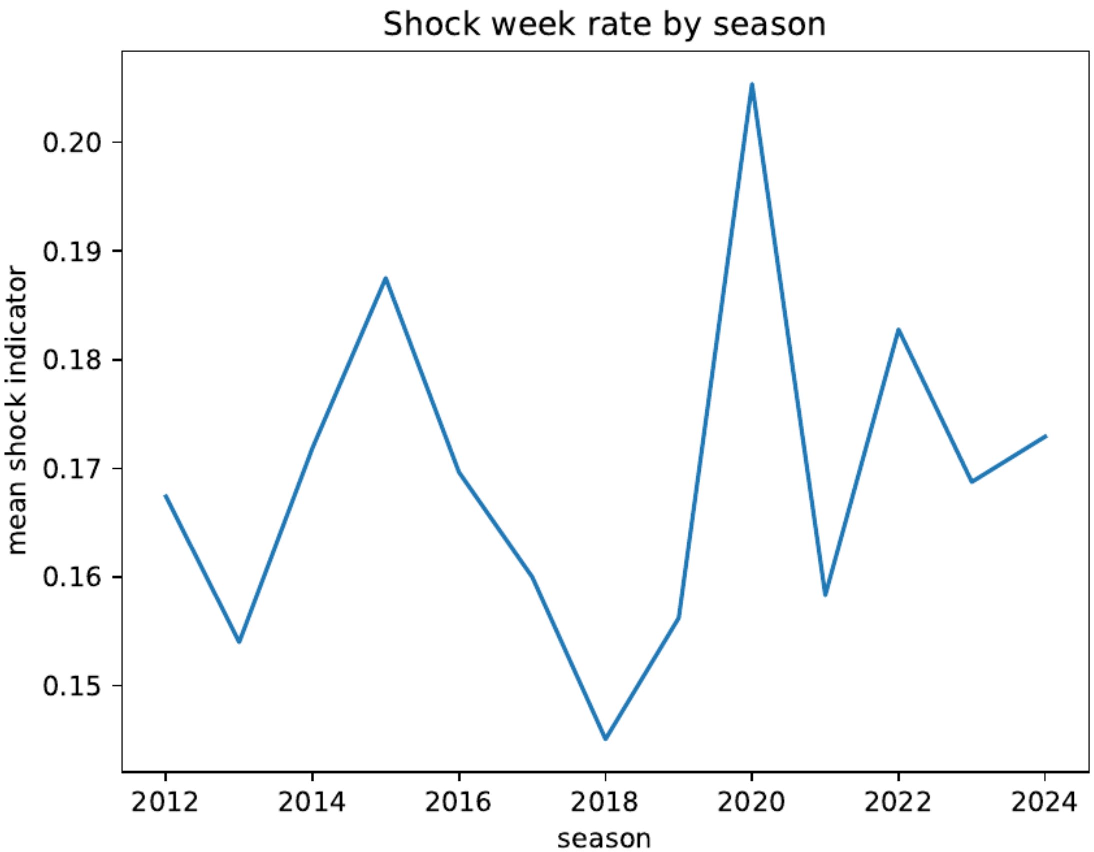

Figure 2: Share of team weeks classified as non score linked special teams shock weeks, by season.

The variation across seasons is also informative. Football is not stationary. Special teams strategy evolves, return behavior shifts with rule emphasis and risk tradeoffs, and teams differ in how aggressively they contest field position. A season with more aggressive returns or more punt heavy game scripts will mechanically raise special teams volume and will also create more opportunities for spikes.

This is one reason we include season week fixed effects in the models. League wide conditions, including season specific changes in tempo, officiating emphasis, and broader strategic shifts, can influence both workload and injuries. Fixed effects absorb these shared shocks, so the estimated relationships are driven by within season, within team deviations that are more plausibly attributable to special teams variation rather than macro trends.

5.2 Volatility deciles and next week injuries

Figure 3 plots offensive injuries next week by volatility decile.

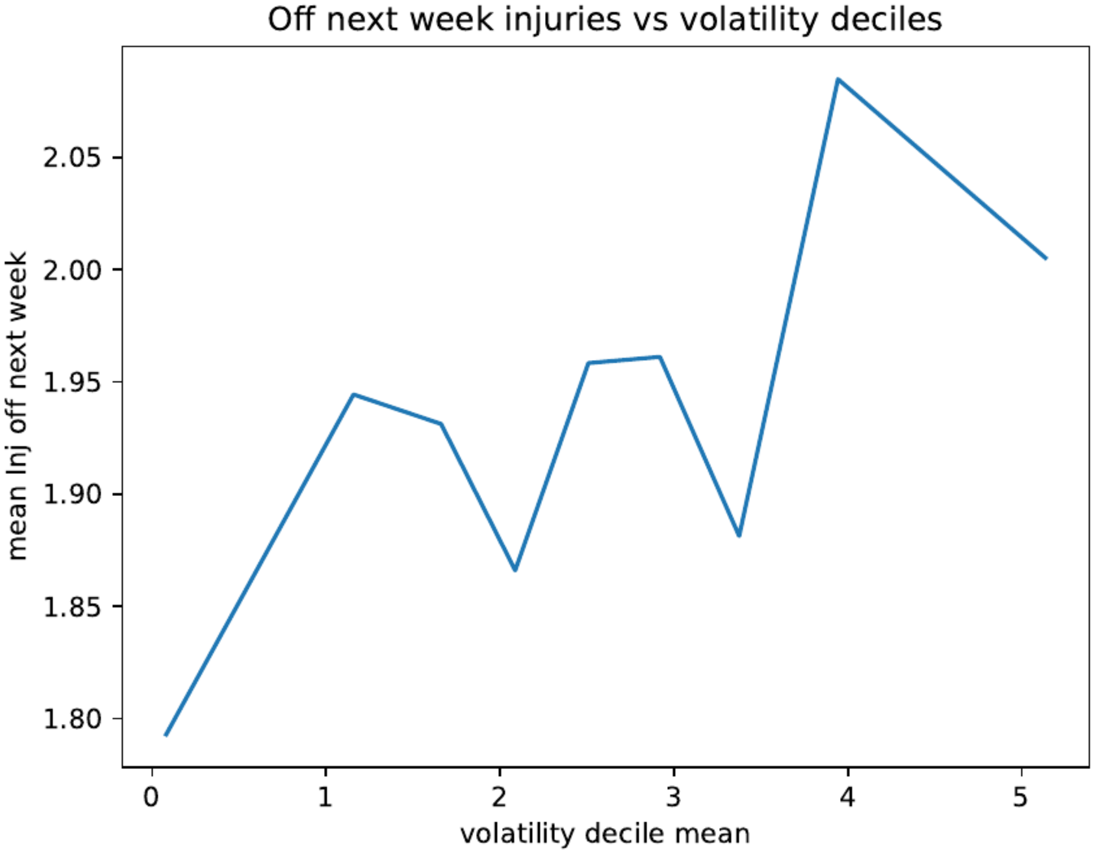

Figure 3: Offensive next week injuries by volatility decile. Higher deciles correspond to more volatile recent special teams workload.

The upward pattern is consistent with the hypothesis that instability in special teams workload can translate into next week offensive injury counts. The key is not that special teams directly injures offensive starters on every snap, it is that special teams volatility can be a marker for broader role disruption. When the special teams unit is unstable, clubs may lean more on offensive skill position backups, tight ends, or running backs for coverage and return roles, and those players are precisely the ones whose injuries are recorded as offensive injuries.

A second mechanism is recovery budgeting. Coaching staffs typically structure weekly practice reps and recovery sessions with an implicit expectation of what Sunday will look like. A week that unexpectedly requires far more coverage reps, more returns, or longer sequences of special teams play can create an acute fatigue hangover that shows up in the following week, even if the next game's offensive snap count is normal.

Figure 4 shows the analogous relationship for defensive injuries.

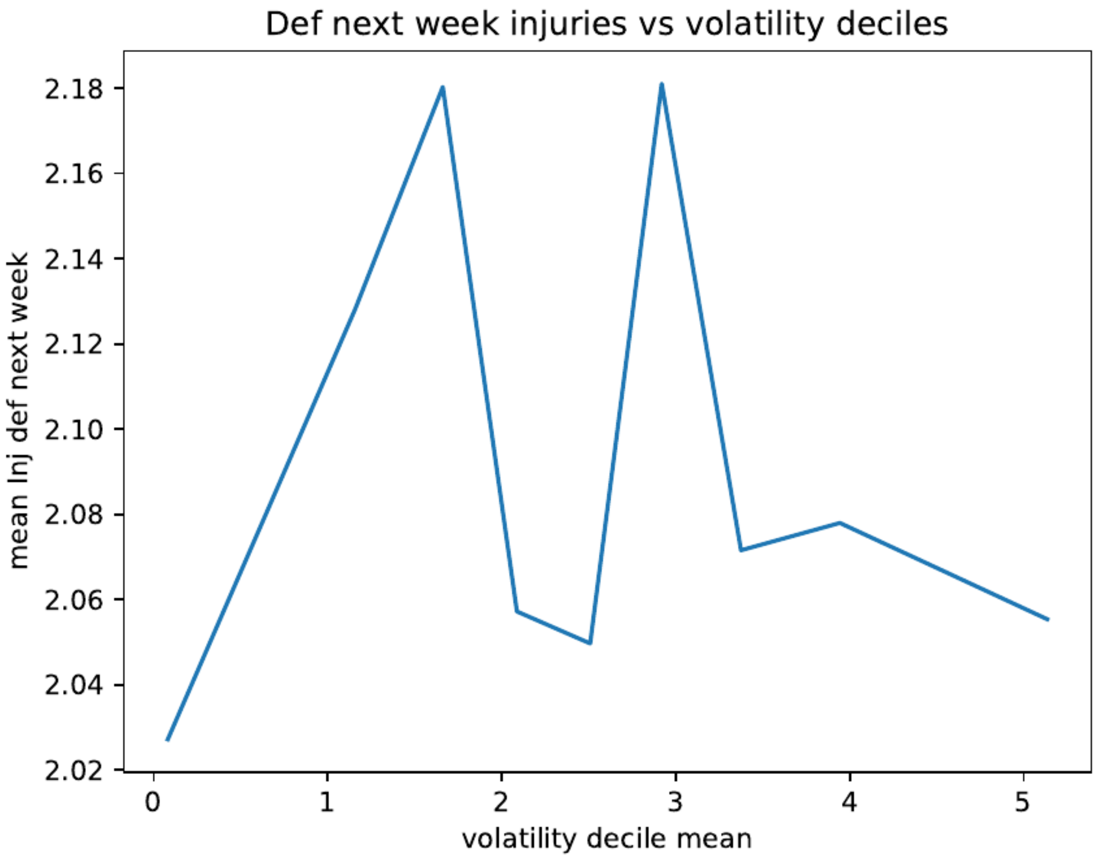

Figure 4: Defensive next week injuries by volatility decile. The relationship is flatter relative to offense.

The defensive relationship is comparatively flat, which foreshadows the model based result that volatility is a smaller contributor for defense. This does not mean defense is protected from special teams. It suggests the volatility channel may not be the primary way special teams exposure translates into next week defensive injury counts at the aggregated team week level.

A plausible football interpretation is selection and specialization. Defensive depth players are often core special teamers, and their baseline exposure may already be high and consistent. If the same set of defenders is repeatedly used on special teams, then volatility across weeks might not change their exposure as much as it changes exposure for offensive role players who are rotated based on game plan and availability.

5.3 Cumulative shock exposure

Figure 5 plots next week injuries by bins of cumulative prior shock exposure.

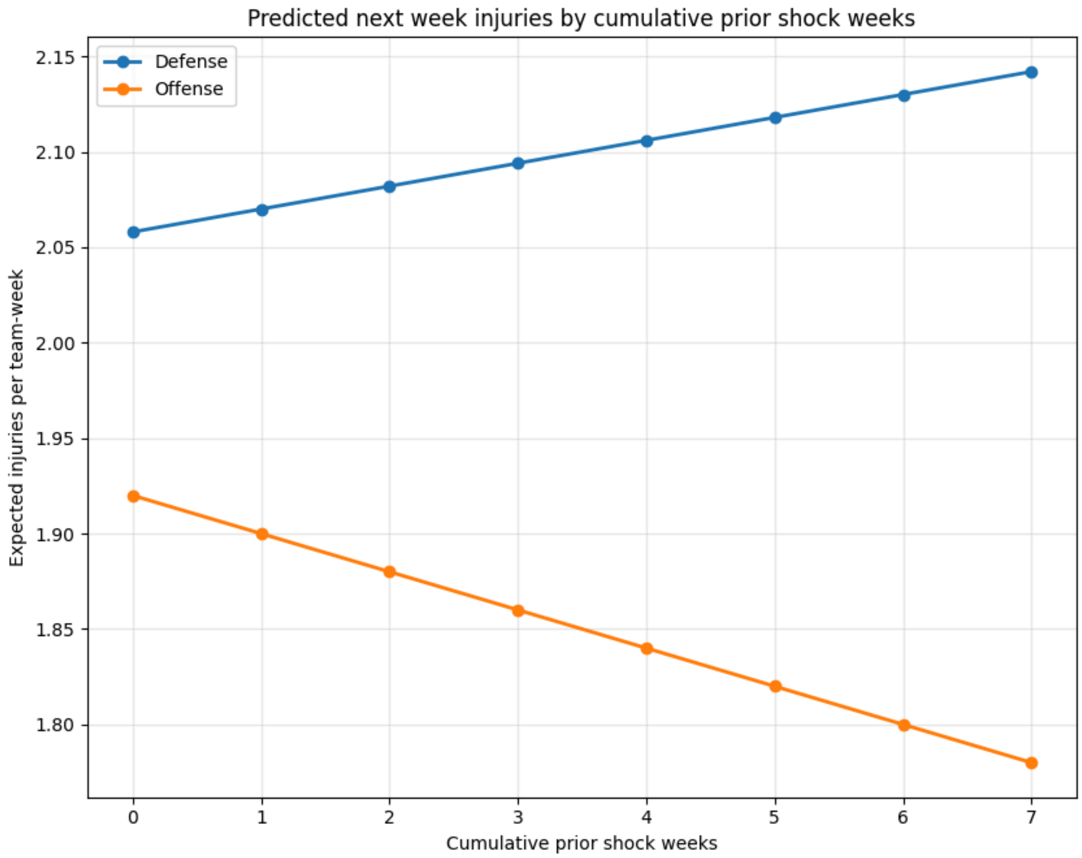

Figure 5: Next week injuries by bins of cumulative prior shock weeks.

The trend is modest, which is consistent with cumulative shocks being a slower moving signal in this aggregated setting. There are two reasons this can happen without contradicting the underlying theory. First, accumulation may be real at the player level but diluted at the team week level, because different players absorb different spikes across the year. Second, chronic accumulation may express itself in performance degradation, availability constraints, and minor injury management rather than in a sharp change in counted next week injuries.

For practical decision making, this pattern suggests a simple prioritization. If a club must pick one monitoring target, acute spikes are the most actionable, because they are immediate and they align with the strongest empirical signal. Cumulative measures can still support longer horizon planning, but they act as a smaller contributor in the weekly injury outcome models.

5.4 Model based incidence rate ratios

Figure 6 summarizes incidence rate ratios for the core special teams terms in the preferred count models.

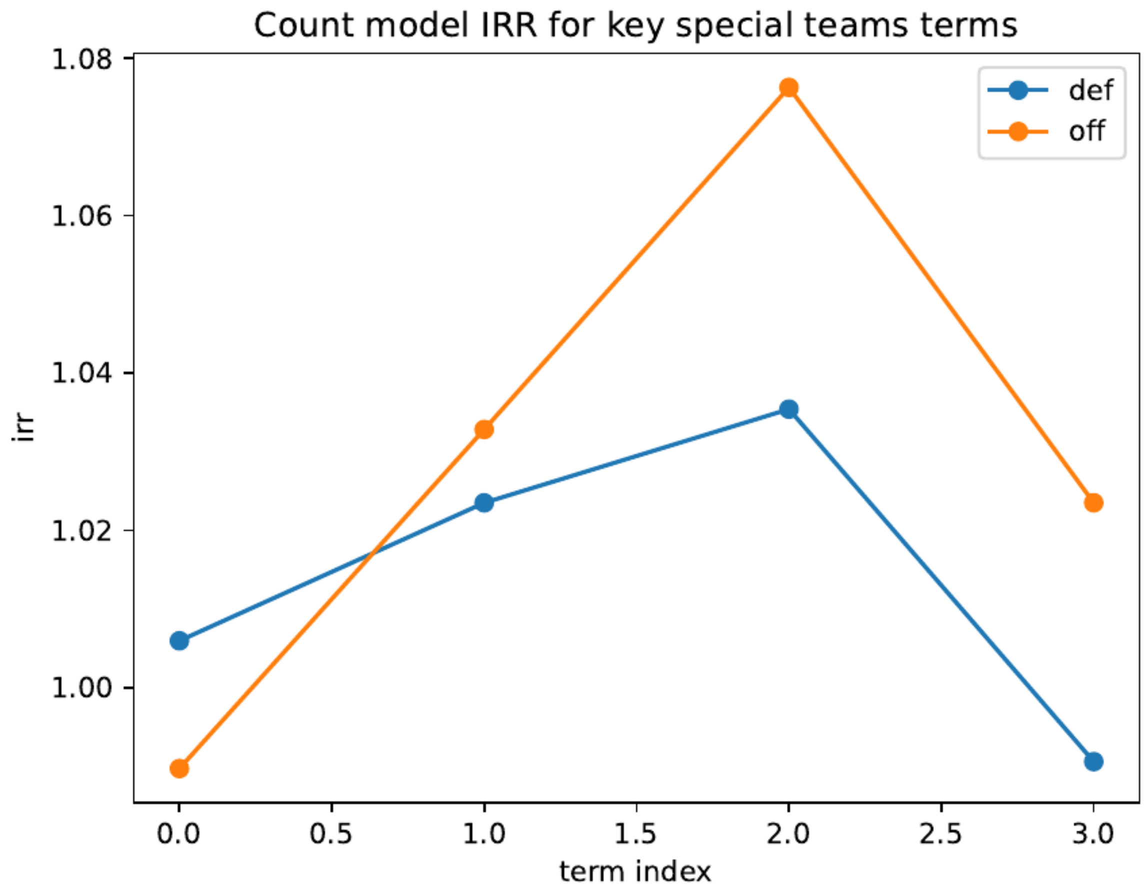

Figure 6: Incidence rate ratios (IRRs) for key special teams terms, with team and season week fixed effects and team clustered uncertainty.

Two results are consistent across the count model family.

First, **shock weeks raise next week injuries for both sides**. The shock IRR exceeds one for both the offense and defense models. Interpreted mechanically, this means that, holding constant the team, the week of the season, and other controls, weeks with an unusually high non score linked special teams workload are followed by more injuries in the subsequent week.

Second, **volatility behaves differently by unit**. Volatility is associated with higher offense injuries but is closer to neutral for defense. This asymmetry matters because it tells us volatility is not simply a proxy for "more football" in general. If volatility were just capturing a generally more physical or chaotic environment, we would expect a similar response on both sides. The data suggest a more targeted exposure pathway, likely tied to who is being used on special teams and how that interacts with offensive depth and role rotation.

The cumulative shock term tends to sit near one. In the language of this report, cumulative exposure functions as a smaller contributor in the weekly injury count model, while still serving as a meaningful context variable when discussing season long workload.

5.5 Football interpretable translations for injury counts

IRRs are interpretable for analysts but can be abstract for football operations. We therefore translate the count model effects into changes in expected injuries per team week.

Figure 7 reports average marginal effects for three changes, switching from no shock to shock, moving from the 25th to 75th percentile of volatility, and adding one additional prior shock week.

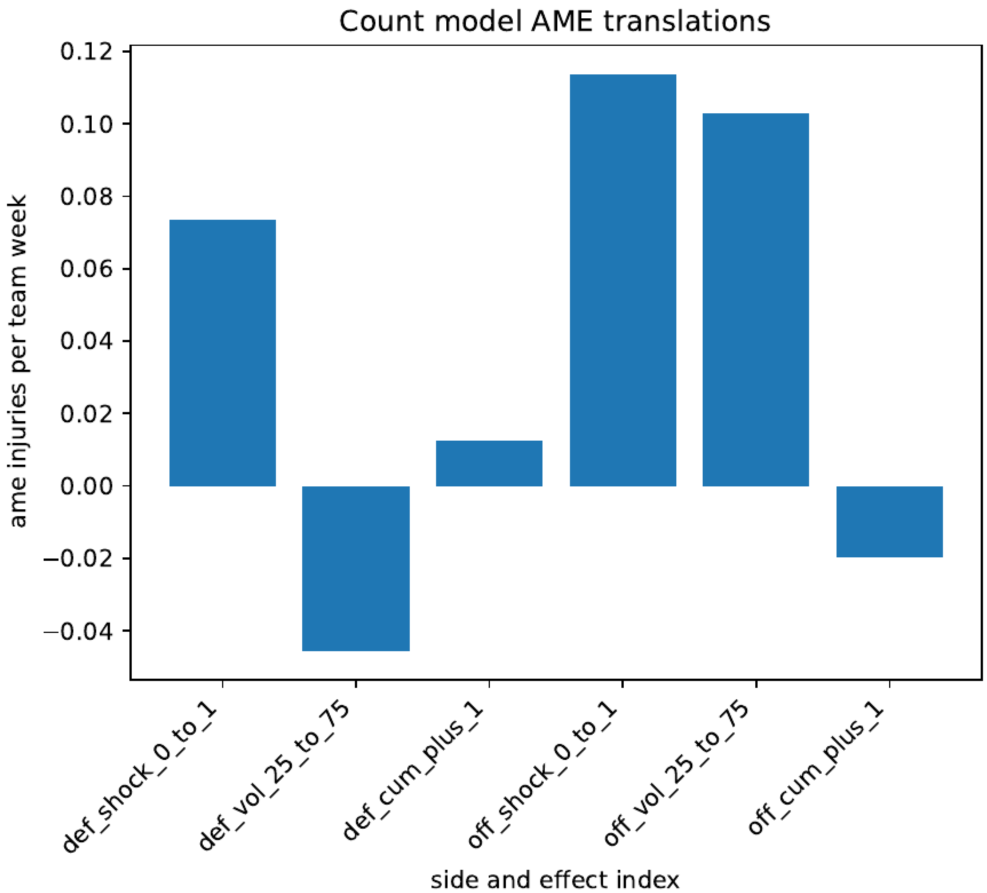

Figure 7: Count model translations as average marginal effects (expected injuries per team week). Baselines are mean predicted injuries holding other features at observed values.

The baseline expected injuries are about 2.07 defensive injuries per team week and about 1.90 offensive injuries per team week under the model's average prediction context. Against that baseline, the translations are concrete.

A shock week is associated with about +**0.074 defensive injuries per team week** (a 3.6% increase) and about +**0.114 offensive injuries per team week** (a 6.0% increase). These are not extreme single week jumps, and that is exactly why translating them is useful. In a sport with many team weeks, small persistent margins are the difference between "noise" and "organizationally relevant".

Volatility from the 25th to the 75th percentile is associated with about +**0.103 offensive injuries per team week** (about 5.5%), while the defense translation is about **-0.046 defensive injuries per team week** (about -2.2%). This is the most direct expression of the unit asymmetry, volatility is a measurable risk signal for offense in this framework, but it is a smaller contributor for defense.

Adding one additional prior shock week is associated with a very small increase for defense (about

+0.012 injuries per team week, about +0.6%) and a small decrease for offense (about -0.020 injuries per team week, about -1.0%). This supports the interpretation that the sharpest weekly signal is acute spikes, while accumulation contributes more slowly, and is partially obscured by player level rotation within the team week aggregation.

5.6 Probability translations for any injury risk

Figure 8 repeats the translation logic for the probability of at least one injury next week.

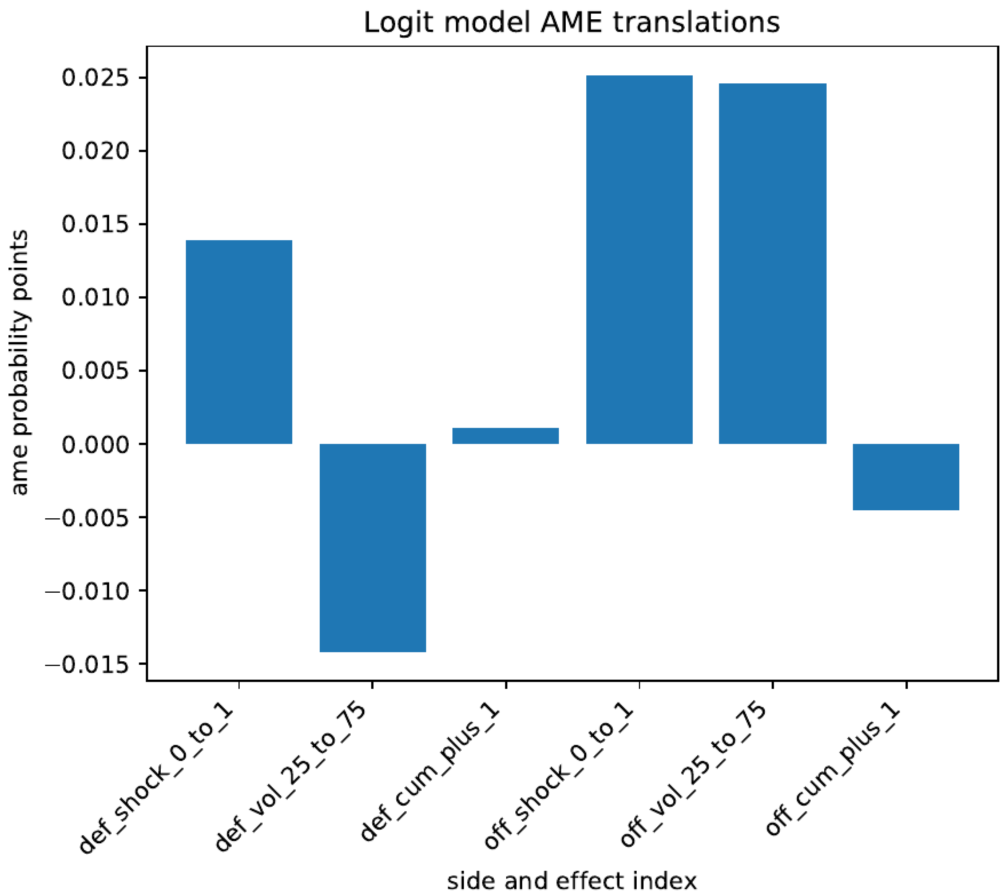

Figure 8: Logit model translations as average marginal effects in probability space.

Baseline probabilities are high in this team week definition, about 0.843 for defensive any injury and about 0.819 for offensive any injury. The relevant question is therefore not "is an injury possible" but "how much does the risk move."

A shock week increases the probability of at least one injury next week by about +**1.39 percentage points on defense** and about +**2.52 percentage points on offense**. Volatility from the 25th to 75th percentile increases the offense probability by about +**2.46 percentage points** and decreases the defense probability by about **-1.42 percentage points**. Cumulative shock effects are close to zero (about +0.11 points for defense and about -0.45 points for offense), reinforcing that accumulation is a smaller contributor for the aggregated probability outcome.

For applied analytics, these numbers are intuitive. They are large enough to matter when scaled across a season, but they are not so large that a single week would obviously "explain" an injury. That is exactly what most credible workload signals look like in elite sport, they shift risk margins rather than deterministically producing outcomes.

5.7 League level scaling and season sustainability framing

To put magnitudes in perspective, we scale the shock week marginal impacts by the observed number of shock team weeks in each season.

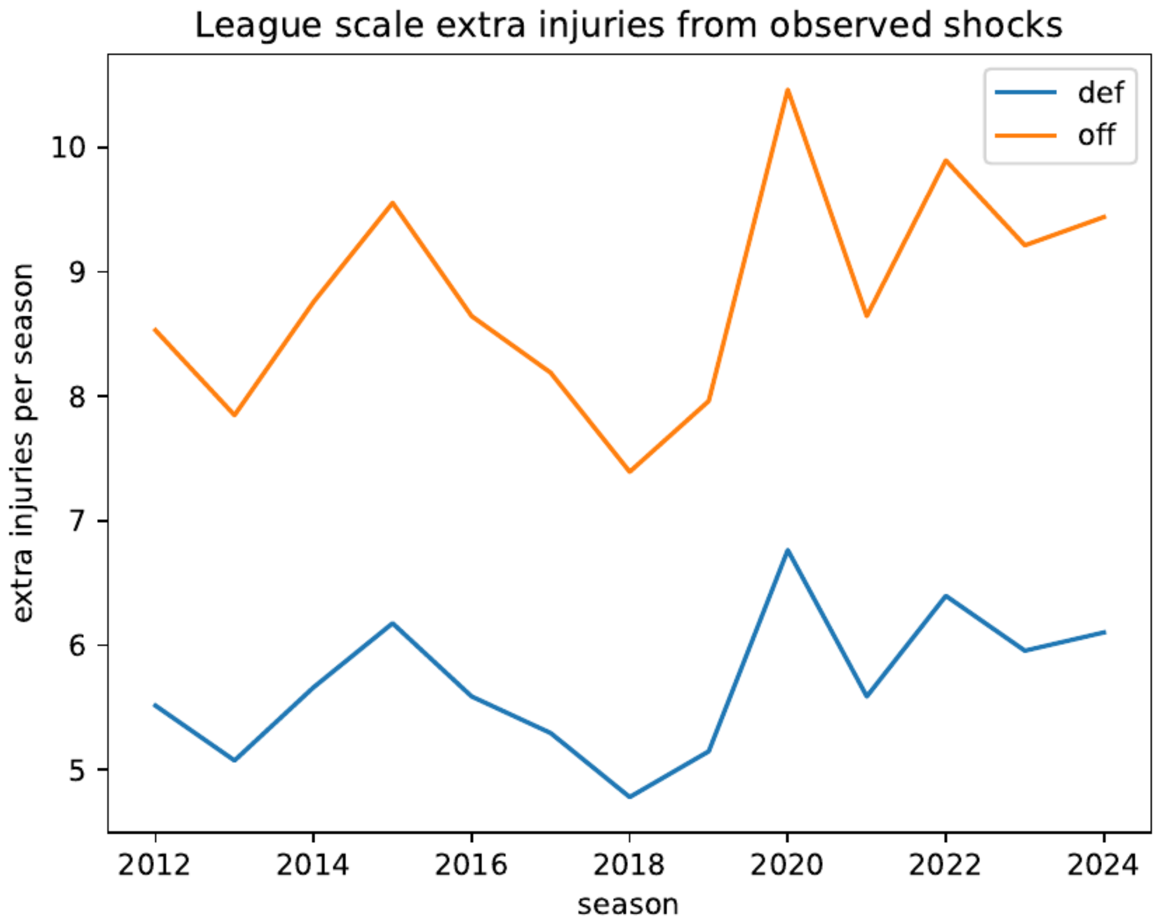

Figure 9: League scaled extra injuries from observed shock weeks, by season (2012 to 2024).

Across 2012 to 2024, the implied contribution from shock weeks sums to roughly 74 additional defensive injuries and roughly 115 additional offensive injuries. Averaged across seasons, that is about 5.7 defensive injuries and about 8.8 offensive injuries per season attributable to the shock exposure channel.

This framing is not a causal claim. It is a policy scale translation of a stable marginal association. If the league can reduce the frequency or severity of shock weeks, or can reduce the post shock vulnerability window through recovery protocols, then the seasonal injury burden can plausibly be shifted in a favorable direction without changing the fundamental structure of football.

6 Putting the findings in league context

The empirical patterns align with several football realities, and grounding the numbers in mechanism is important for credibility.

Special teams shocks are often game script driven. A punt heavy day created by strong defenses or struggling offenses, a weather event that changes field position strategy, or a close game that forces more conservative decisions can all increase non score linked special teams volume. Those are not always predictable on Wednesday.

Volatility is a proxy for disruption, not just intensity. A club with stable special teams usage can train and allocate recovery predictably. A club that alternates between quiet weeks and chaotic weeks is more likely to shift personnel, increase rep counts for specific players, and create mismatches between training load and game load.

Offense sensitivity is plausible through role overlap. Many offensive depth players contribute heavily to special teams, especially running backs, tight ends, and wide receivers on coverage and return units. When special teams becomes volatile, those roles may expand, which naturally maps into offensive injury counts.

Defense still responds to spikes. Even though volatility is flatter for defense, the shock effect is positive, meaning sudden spikes matter regardless of unit. This supports an acute load narrative, a large spike is disruptive even if week to week variability is not the dominant channel.

From a data science perspective, the asymmetry is a feature, not a bug. If every exposure predicted every outcome symmetrically, we would worry the model is capturing a generic confounder like "tough games." Here, the patterns differ by unit in a way that matches plausible exposure pathways.

7 Collective Bargaining Agreement implications and a workload gap

The NFL labor framework contains substantial safeguards related to practice contact, practice time, and prohibited drills. For example, the CBA explicitly limits the number of in season padded practices and defines what counts as a padded practice, including a season wide cap and a definition based on required equipment (Article 24, Section 1(a) and 1(c)). It also constrains daily on field activity time and prohibits a set of high risk padded practice drills (Article 24, Section 1(d) and 1(e)).

These provisions reflect a clear safety philosophy. Reduce unnecessary contact in practices, standardize recovery windows, and eliminate drills that replicate the worst collision dynamics without competitive value.

7.1 Where the gap appears

The results of this report point to a practical gap that is not a violation of practice rules. It is a gap in what the rules are designed to regulate.

Practice rules target what happens from Monday through Saturday. Special teams shocks are game driven, and they occur on Sunday. A club can be fully compliant with practice limits while still exposing a subset of its roster to unusually intense special teams weeks driven by game circumstances and tactical choices. Because shock weeks occur regularly and are associated with a stable next week injury increase for both offense and defense, the gap is not theoretical.

The CBA does contain governance structures that are aligned with modern monitoring. In particular, Article 39 establishes an Accountability and Care Committee, and a joint subcommittee is explicitly tasked with analyzing injury information and performance tracking data to study training methods, practices, and drills that may lead to injuries (Article 39, Section 5(g)). That language supports the idea that workload analytics are part of player safety governance, not merely a competitive tool.

7.2 Policy aligned safeguards motivated by the evidence

We propose four incremental safeguards that align with existing CBA logic and preserve competitive integrity.

1. **Standardized special teams workload reporting.** Require clubs to compute and report a standardized team and player level special teams workload metric each week, including a non score linked component and a short window volatility component. Reporting can go to the joint health and safety governance process in aggregated form.

2. **Shock week trigger protocol.** Define a shock threshold and establish a mandatory review when it is crossed. The protocol can include additional recovery time, targeted medical screening for high exposure special teams players, and explicit limits on adding new special teams roles for those players in the following week.

3. **Roster flexibility after shocks.** Create a narrowly scoped gameday elevation or substitution mechanism for clubs following a shock week, so that rest decisions do not force competitive disadvantages. The purpose is not to expand rosters broadly, it is to make recovery feasible.

4. **Formalize analytics as preventive care.** Clarify through committee guidance that workload monitoring, including game driven special teams exposure, is part of preventive care expectations. This connects directly to the CBA's health and safety governance intent, and it encourages consistent adoption.

These proposals are intentionally modest. They do not change rules of play. They translate a measurable risk signal into a monitoring and response framework, in the same spirit that practice rules translate known risks into structural limits.

8 Conclusion

This report studies how special teams workload shocks, special teams volatility, and cumulative shock exposure relate to next week injury outcomes in a 2012 to 2024 NFL team week panel.

The main result is that non score linked special teams shock weeks are associated with a small but consistent increase in next week injuries for both offense and defense. Volatility is more relevant for offense than for defense, while cumulative shock exposure behaves as a smaller contributor in the weekly outcomes. These patterns are stable across descriptive views, fixed effects count models, and probability translations.

The practical implication is that special teams workload can be treated as a measurable risk factor suitable for monitoring and trigger based safeguards. Because current labor safeguards focus primarily on practice structure, the findings motivate a policy opportunity to incorporate game driven special teams spikes into preventive workload governance.

References

1. National Football League and NFL Players Association. *Collective Bargaining Agreement.* March 15, 2020.

2. `nflverse` contributors. *nflverse: R Packages for Collecting and Analyzing NFL Data.* https://nflverse.nflverse.com.

3. Cameron, A. C., and Trivedi, P. K. *Regression Analysis of Count Data.* Cambridge University Press, 2013.

4. Statsmodels Development Team. *statsmodels: Statistical Modeling and Econometrics in Python.* https://www.statsmodels.org.

5. DuckDB Labs. *DuckDB: An Embeddable Analytical Database.* https://duckdb.org.