STATE TAX SERVICE
UNDER THE MINISTRY OF ECONOMY
OF THE REPUBLIC OF AZERBAIJAN

# Entity Risk Score (ERS) Methodology. A PoC for Transforming Tax Administration through Data-Driven Risk Scoring.

**Roman Bandurin**
**Head of Data Analytics Department**

## What is Entity Risk Score (ERS)

The Entity Risk Score (ERS) is a composite, dynamic, data-driven indicator assigned to every registered taxpayer. It serves as the central decision-support layer of a risk-based operating model, enabling proportional compliance management across every stage of the taxpayer lifecycle.

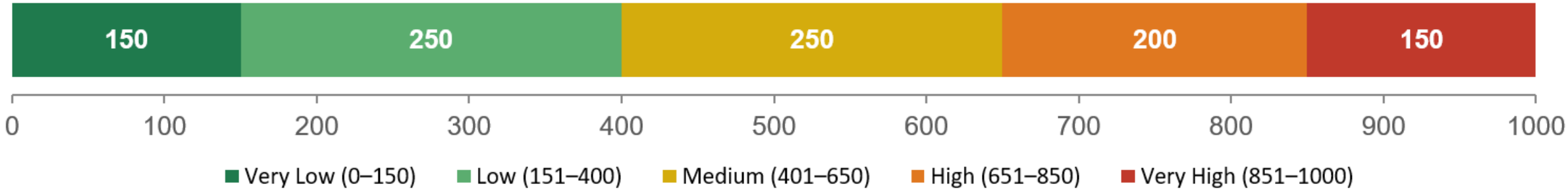
## Design Principles

- **Proportionality:** Administrative response severity directly matches the level of indicated risk, ensuring balanced intervention.
- **Dynamism:** The score evolves continuously as new data becomes available - compliant behavior earns score improvement, while performance deterioration triggers escalation.
- **Auditability:** Every system calculation, manual override, and threshold decision is strictly logged and fully reviewable to ensure complete transparency.
- **Non-discrimination:** The model relies exclusively on objective, compliance-relevant data. Sensitive or non-tax personal indicators are strictly restricted, utilized solely for nominee-detection analytics, and authorized through appropriate consent frameworks.

# Score Range and Segmentation



**The ERS operates on a 0–1000 integer scale mapped to five tiers aligned with the OECD compliance pyramid.**




| Tier | Range | Compliance Pyramid | Taxpayer Attitude | Response Strategy (Treatment) |
|---|---|---|---|---|
| **Very Low** | 0–150 | Make it easy | Willing to comply | **Green Channel:** Whitelisting & Fast-track refunds |
| **Low** | 151–400 | Make it easy | Tries to comply, but fails | **Self-Service:** Standard support & Auto-monitoring |
| **Medium** | 401–650 | Help to comply | Complies if monitored | **Nudge & Educate:** Proactive alerts & Guidance |
| **High** | 651–850 | Detect | Reluctant to comply | **Closer Scrutiny:** Active review & Manual checks |
| **Very High** | 851–1000 | Enforce | Deliberately non-compliant | **Enforcement:** Direct intervention & Restrictions |

# Composite Score Formula & Formula Components



**ERS is computed as a weighted sum of four risk pillars, modified by a structural decay multiplier. The formula integrates four conceptually distinct dimensions of compliance risk into a single, auditable integer score.**

- $ERS(t) = wS(t) \cdot S(t_0) \cdot \mathbf{d(t)} + wT(t) \cdot T(t) + wC(t) \cdot C(t) + wN(t) \cdot N(t)$*
- $wS(t) = 1 - wT(t) - wC(t) - wN(t)$

| Term | Formula Component | Interpretation & Drivers |
|---|---|---|
| $S(t_0)$ | Structural sub-score (0–1000), fixed at registration | Evaluates **initial risk on Day 1** using static registration data (e.g., owner details, previous employment, income, and industry sector). |
| d(t) | Decay multiplier (0.30–1.00) | Reduces the weight of the structural sub-score as the **entity progressively accumulates a business history (capped at a maximum 70% reduction).** |
| T(t) | Transactional sub-score (0–1000), updated continuously | **Near real-time risk profile** driven by behavioral anomalies in transaction patterns, e-invoices, and VAT sub-account movements. |
| C(t) | Compliance sub-score (0–1000), updated on filing events | Risk profile derived from **historical tax declaration behavior**, including filing timeliness, payment records, and desk audit outcomes. |
| N(t) | Network sub-score (0–1000), updated from graph database | Risk profile based on **graph database analysis**, detecting proximity to fraud, circular trading patterns, and suspicious ownership concentration. |
| wS, wT, wC, wN | Dynamic weights (sum to 1.0) | Governed by **sigmoid activation functions** to shift the model from structural dominance at registration toward behavioral weighting as data accumulates. |

** Note: This is a conceptually simplified model view. After the Decay multiplier is applied, the released weight is proportionally distributed among the remaining weights according to their target levels.*

# Dynamic Weighting Mechanism. Sigmoid function



**ERS does not apply fixed weights to its four pillars. Instead, weights are determined dynamically by sigmoid (S-shaped) activation functions that respond to accumulating data volume.**

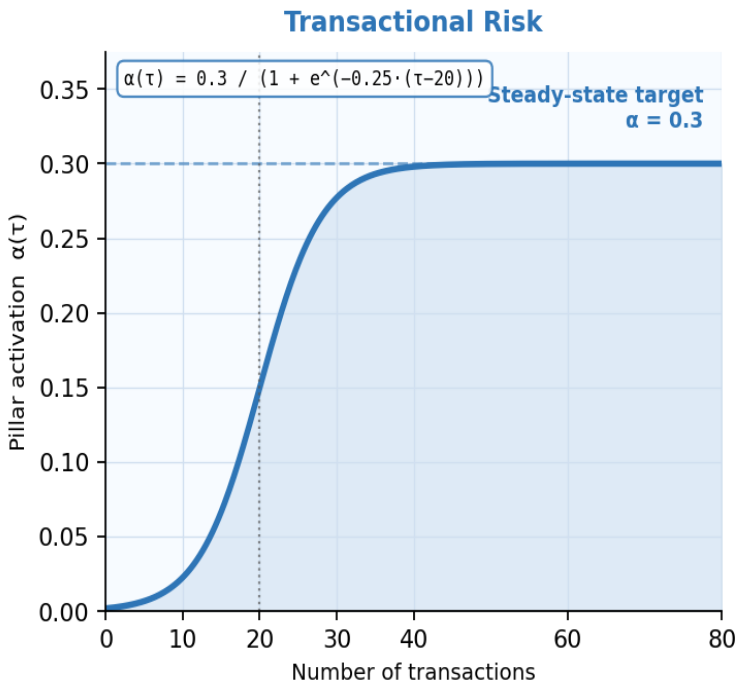
- $\alpha(\tau) = \text{TARGET} / (1 + e^{(-k \cdot (\tau - \tau^*))})$

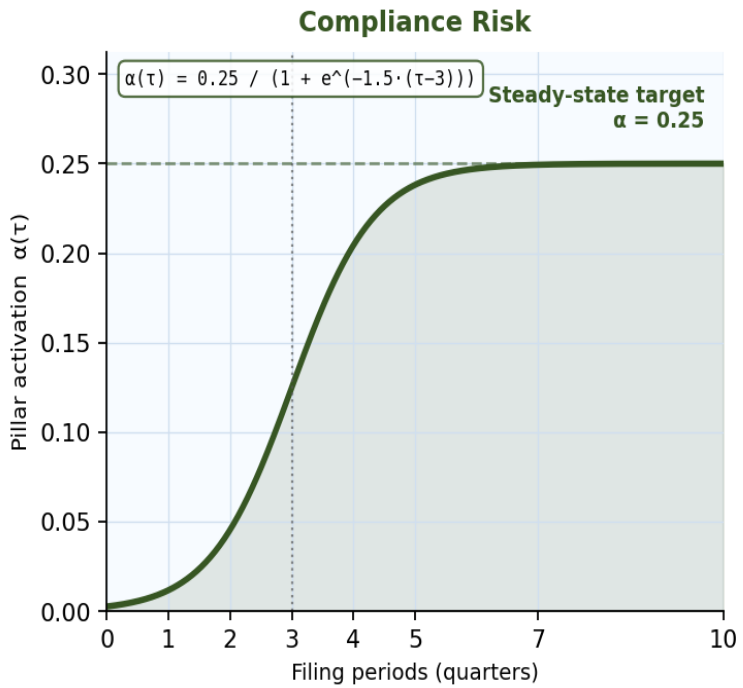
| Parameter | Symbol | Meaning in ERS Context |
|---|---|---|
| **Target weight** | TARGET | The maximum weight this pillar can achieve at full activation. Determined by policy (T=0.30, C=0.25, N=0.20). |
| **Data trigger** | $\tau$ | The quantity of data accumulated: number of transactions (T), number of filing quarters (C), or binary graph availability (N). |
| **Midpoint** | $\tau^*$ | The data volume at which activation reaches exactly half its target weight. Chosen to reflect the point at which the data volume is deemed statistically meaningful. |
| **Steepness** | k | Controls how sharply the transition happens. High k = abrupt transition; low k = gradual ramp-up. Calibrated to prevent gaming. |

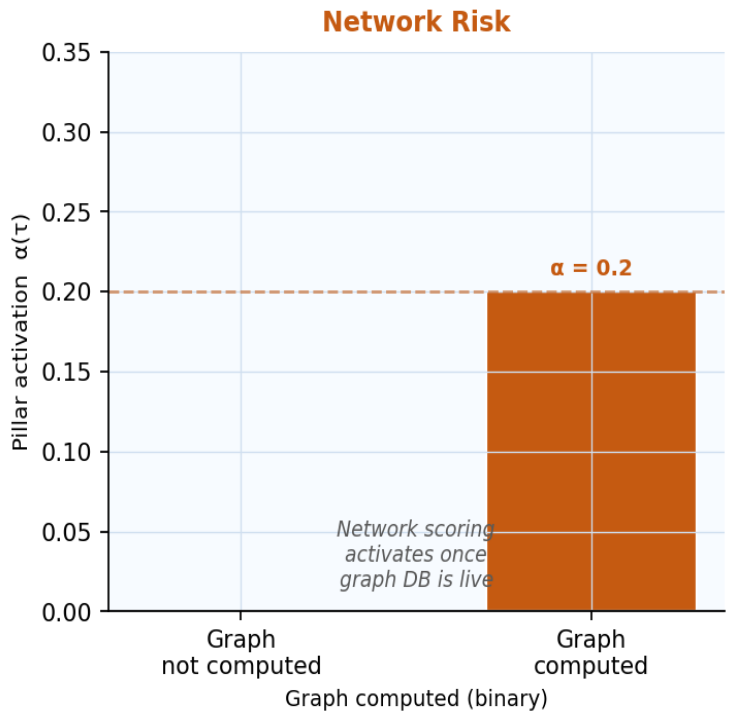
# Dynamic Weighting Mechanism. Reasoning



**We chose the sigmoid function because it avoids the statistical insignificance of linear weighting and the gameable jumps of step functions by smoothly transitioning from zero when data is limited to a stable target weight as volume grows, ensuring the score reflects genuine statistical intuition and remains robust against single-event manipulation.**



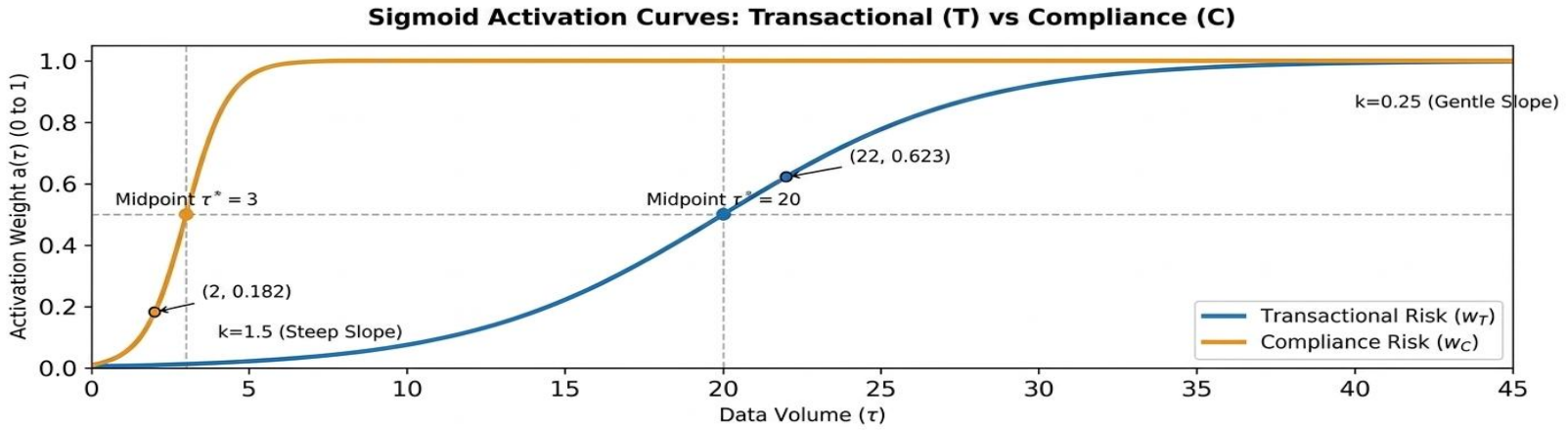



# Pillar-Specific Activation Parameters



**Each pillar uses a distinct data trigger and calibration, reflecting the different rates at which meaningful evidence accumulates across the three behavioral dimensions.**

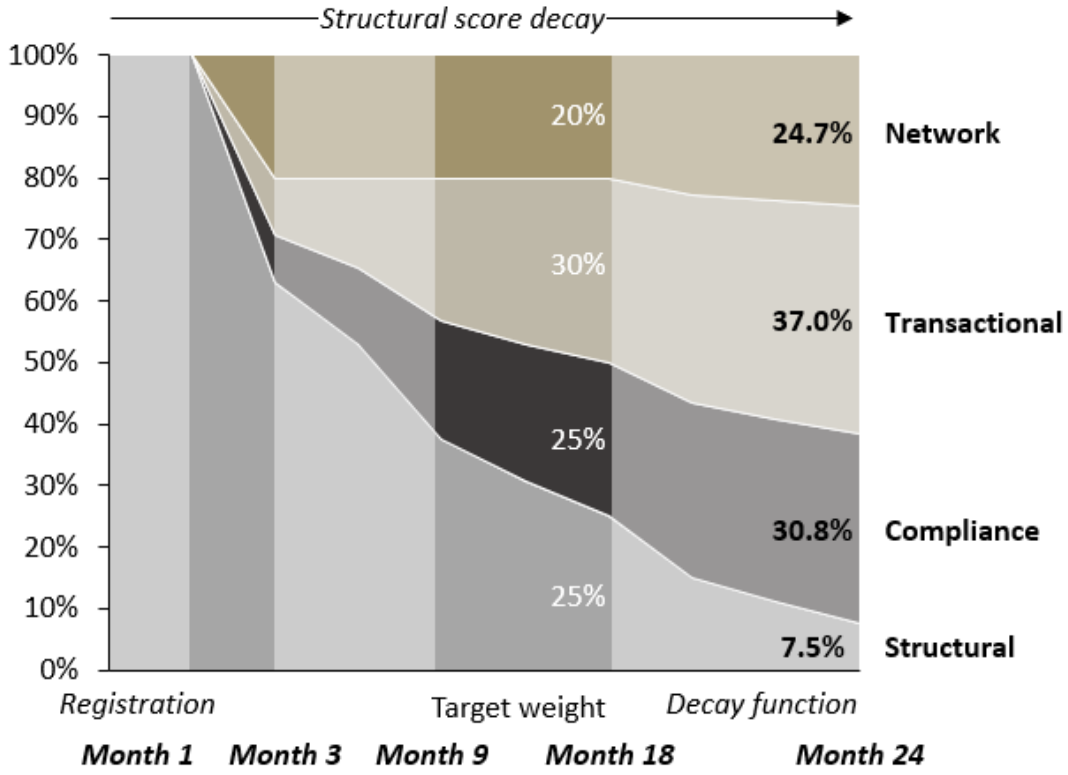
| Pillar | Data Trigger | Midpoint τ* | Steepness k | Target Weight | Rationale |
|---|---|---|---|---|---|
| **Transactional** | # transactions | 20 | 0.25 | 0.30 | **Transaction data** accumulates quickly. **τ*=20 transactions is about 6–8 weeks for an active trader.** Gentle k=0.25 to prevent a single active month from over-weighting. |
| **Compliance** | # filing periods (quarters) | 3 | 1.5 | 0.25 | **Filing periods** are discrete and slow. 3 quarters provides a meaningful compliance track record. High k=1.5 reflects that a full cycle of three filings is qualitatively informative. |
| **Network** | Graph computed (binary) | n/a | n/a | 0.20 | **Network scoring** requires the graph database to be operational. Activation is binary: 0.20 when graph is activated, 0.0 otherwise. |
| **Structural** | Residual (1−wT−wC−wN) | n/a | n/a | ~0.25 at steady state | **Structural weight** is not directly activated, it absorbs whatever weight the behavioral pillars have not yet claimed. Always ≥ minimum to prevent behavioral pillars from entirely suppressing structural assessment. |



# ERS changes based on taxpayer behavior



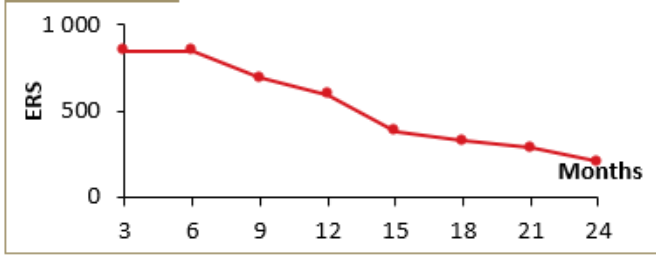
## Evolution of weights of ERS modules




**Impact of the structural data on the ERS is reduced as more information about the taxpayer becomes available**

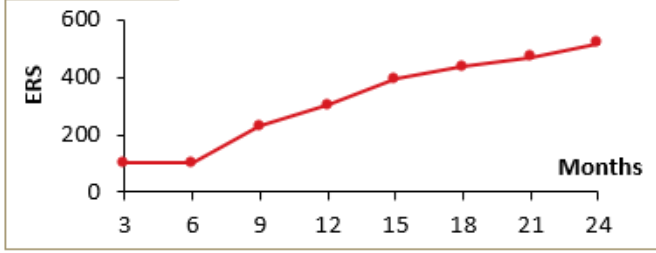
## Scenario simulation based on different examples of taxpayers

**Scenario 1**



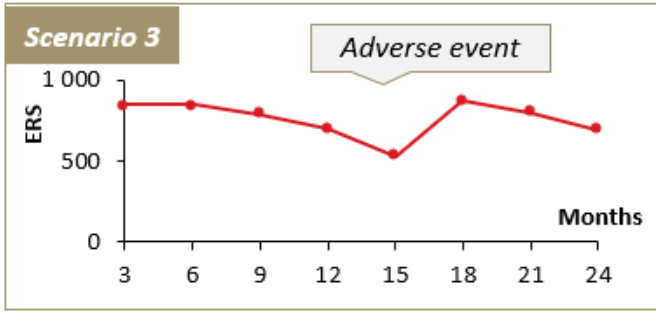

*The taxpayer began operations with a **high structural score**; however, **continued compliance** and the reduction of the structural score's impact, resulted in a **lower ERS.***

**Scenario 2**




*The taxpayer began operations with a **low structural score**; however, **continued incompliance** and the reduction of the structural score's impact, resulted in a **higher ERS.***

**Scenario 3**




*The taxpayer began operations with a **high structural score** and remained **consistently compliant**; however, a **change in management or majority shareholders** triggered a **decay reset, increasing the ERS.***

# Thank you for attention

www.taxes.gov.az