Building a Custom Prepayment Model

Asset Prepayment Models Framework

January 26, 2026

Contents

1 Introduction

This guide explains how to construct custom prepayment models in the platform using a small family of mathematically well-defined building blocks. The key idea is that you assemble a real-valued score function from a set of basis functions, and the engine maps that score into a conditional prepayment rate (CPR) between zero and one.

A user interface (UI) is provided where each row represents one multiplicative term in the score. By combining simple transformations—such as indicator conditions and variable factors—you can express a wide range of behaviors: near-zero prepayment when refinancing is unattractive, near-certain prepayment when refinancing is attractive, burnout effects that depend on the pool state, and regime-like behavior that varies across pool bands or time-to-maturity.

For each explanatory variable there is exactly one associated basis-function choice in a given row. The only exception is the pool state, which appears in two separate columns (POOL and POOL2). This design is intentional: it allows you to use the pool state both as a continuous factor (for example through a power term capturing burnout) and as a regime selector (for example by gating different parameter sets to different pool segments). The worked examples include a mixture-style construction showing how separate sub-models can be targeted to different pool regimes while remaining within a single, transparent specification.

The Terms table maps directly to the mathematical model: the dropdowns fully determine the functions applied to each input, and the platform can display the resulting expression. This makes the specification auditable and supports piecewise/regime behavior via indicator gating, while the logistic map keeps CPR smooth.

Throughout the guide, we refer to the platform's built-in starting specification as the standard model.

The guide is organized as follows. Section 2 defines the score-and-logistic model used by the engine. Section 3 describes the input variables and their units. Section 4 explains how the Terms table encodes the model in the UI. Section 5 gives the exact mapping from UI Kind choices to mathematical basis functions. Section 6 presents worked examples, including pool-regime mixtures and near-step prepayment behavior. Finally, Section 7 summarizes practical implementation notes and common pitfalls.

2 Model definition

The modeled quantity is the conditional prepayment rate (CPR), denoted by λ, which lies in $[0, 1]$. The platform computes a real-valued score f from the Terms rows and maps that score to CPR via a logistic function:

$$\lambda = \frac{1}{1 + e^{-k(f-i)}}. \tag{1}$$

Here $k > 0$ controls how sharply CPR transitions from low to high as the score crosses the inflection level i. At $f = i$ the model yields $\lambda = 0.5$. For fixed i, increasing k makes the transition more step-like, while smaller k makes it smoother.

The score f is defined as a bias plus a sum of row contributions:

$$f = a_0 + \sum_{j=1}^{n} \text{row}_j. \tag{2}$$

The parameter a_0 is the bias. Each row j corresponds to one row in the Terms table and is a coefficient multiplied by the selected per-variable factors:

$$\text{row}_j = a_j \, T_{SR,j}(SR) \, T_{GAIN,j}(GAIN) \, T_{POOL,j}(POOL) \, T_{POOL2,j}(POOL) \, T_{TTM,j}(TTM). \tag{3}$$

Here $POOL$ and $POOL2$ are two independent basis-function columns applied to the same pool-state input, allowing the pool to act both as a continuous factor and as a regime selector.

In the Terms table, a_j is the value in the leftmost column **a** on row j.

Spread parameter. The specification includes a spread field `spread_dec` stored as a decimal. For readability, the platform may display spread in basis points (bps), using the conversion

$$\text{bps} = 10^4 \cdot \texttt{spread\_dec}. \tag{4}$$

For example, 40bps corresponds to `spread_dec` = 0.004. Spread is part of the model specification and may be shown in tooltips or summaries; it does not change the structural form (2)–(3).

3 Input variables and units

The engine provides a fixed input vector. The builder uses four variables in the score: SR, $GAIN$, $POOL$, and TTM. You do not need to supply them manually in the Terms table, but it is important to understand their units and conventions because thresholds are applied to these raw values.

Index	Variable	Description
0	SR	Interest-rate proxy used by the engine (for example a par/swap rate at a fixed tenor).
1	$GAIN$	Refinancing incentive (defined as Coupon − SR), in decimal units (e.g. -0.005 means -0.5%).
3	$POOL$	Pool/burnout state (path-dependent), defined to lie in $[0, 1]$.
4	TTM	Time-to-maturity in years. Thresholds θ specified in the Terms table are also in years.

Table 1: Input variable structure used by the builder and their unit conventions.

A key unit convention is that $GAIN$ thresholds are in decimals. For example, $\theta = -0.005$ corresponds to a threshold of -0.5%. The platform may display the same number as a percentage for readability, but the model always evaluates thresholds in decimals.

Note. When using Power with a non-integer exponent, ensure the underlying variable is non-negative in the region where the term is active.

4 The Terms table

Figure 1 illustrates how a prepayment-model specification (here the platform's standard built-in model) is represented in the Terms table. Each row defines one multiplicative term (one row_j) in the score (2). Rows are added, while the selected functions within a row are multiplied, exactly as in (3).

Row layout. Each row contains:

- a coefficient a_j in the leftmost column **a**, and
- one basis-function column for each of SR, $GAIN$, $POOL$, $POOL2$, and TTM.

Each basis-function column has the same UI structure: a dropdown (called **Kind** in this guide) and two numeric parameter fields θ and p. The mapping from Kind/θ/p to the mathematical factor $T(\cdot)$ is defined in Section 5.

#	a	SR	θ	p	GAIN	θ	p	POOL	θ	p	POOL2	θ	p	TTM	θ	p
1	-12.8718	line	0	1	Indicator	-0.005	0	Non	0	0	None	0	0	Nor	0	0
2	18.2932	Nor	0	0	linear	0	1	pow	0	0.3198	None	0	0	Nor	0	0
3	0.6727	Nor	0	0	Indicator	-0.005	0	Non	0	0	None	0	0	line	0	1

Figure 1: Terms table UI. Each row specifies a coefficient a_j and one basis-function choice per column. Each variable column (SR, GAIN, POOL, POOL2, TTM) provides a Kind dropdown and parameters (θ, p).

POOL vs. POOL2. The inputs contain a single pool-state variable $POOL$, but the Terms table provides two independent pool columns, $POOL$ and $POOL2$. Both columns are applied to the same underlying pool-state input, and their effects multiply in each row. This allows the pool to act both as a continuous factor (for example via a Power term) and as a regime selector (for example via an Indicator term).

Reading a row. A row with coefficient a_j and column selections defines one multiplicative contribution row_j as in (3). The score is the bias plus the sum of all row contributions (2). If multiple columns in the same row use indicators, their product behaves like a logical "AND".

5 Basis functions and Terms table mapping

In the Terms table, each variable column (SR, $GAIN$, $POOL$, $POOL2$, TTM) has the same structure described in Section 4: a dropdown (Kind) and two numeric parameters θ and p. This section gives the exact mapping from those column settings to the factor $T_{v,j}(\cdot)$ used in (3).

Write $T(x)$ for the single-variable factor selected in one column, where x is one of SR, $GAIN$, $POOL$, or TTM. Kind and (θ, p) define T. The pool variable is available twice ($POOL$ and $POOL2$), meaning two independent factors may be applied to the same underlying input. We denote indicators by $\mathbf{1}_{\{\cdot\}}$.

Strict thresholds for indicators. Indicator+ is on only when $x > \theta$ and Indicator− is on only when $x < \theta$. If $x = \theta$ exactly, the indicator is off.

None. When Kind is **None**, the variable is not used and the factor is constant:

$$T(x) = 1. \tag{5}$$

In this case the displayed θ and p have no effect on the model.

Indicator+. When Kind is **Indicator+**, the factor is an activation condition above a threshold:

$$T(x) = \mathbf{1}_{\{x>\theta\}}. \tag{6}$$

Indicator−. When Kind is **Indicator−**, the factor is the analogous condition below a threshold:

$$T(x) = \mathbf{1}_{\{x<\theta\}}. \tag{7}$$

For Indicator kinds, the threshold θ is taken directly from the UI field, while p is ignored (it may be shown as 0 for consistency, but it does not change (6)–(7)).

Linear. When Kind is **Linear**, the factor is the identity map:

$$T(x) = x. \tag{8}$$

Operationally the UI typically shows $\theta = 0$ and $p = 1$ for Linear. The mathematical definition of Linear is always (8), independent of the displayed parameters: θ is ignored, and p is treated as implicitly equal to 1 (changing the UI value of p does not change the Linear factor).

Power. When Kind is **Power**, the factor is a raw power of the variable:

$$T(x) = x^p. \tag{9}$$

The exponent p is taken from the UI field p. For Power, the displayed θ is not used and is conventionally shown as 0. Because fractional powers are not real-valued for negative bases, Power with non-integer p should be used only on variables that are nonnegative in the intended operating region (POOL is the usual candidate).

Notes on Power. Variables such as $GAIN$ may be negative. This does not prevent using Power when p is an integer (e.g. $p = 2, 3, 4$), since x^p is then well-defined for negative x. The main restriction is for non-integer exponents: if p is not an integer, then x^p is not real-valued for negative x, so such choices should be used only when the variable is nonnegative in the operating region (commonly $POOL$). Also avoid $p < 0$ unless you are sure the variable is strictly nonzero (and typically strictly positive), since x^p diverges at $x = 0$ for negative p.

Row multiplication and gating. Within a row, the selected single-variable factors multiply. This is how gating and regime logic are expressed: if one column contributes an indicator (for example $\mathbf{1}_{\{GAIN>\theta_g\}}$) and another column contributes a nontrivial factor (for example $POOL^p$), then the row contribution becomes $a_j\, \mathbf{1}_{\{GAIN>\theta_g\}}\, POOL^p$. If multiple columns in the same row use indicators, the product of indicators behaves like a logical "AND".

Displayed expression. For readability, the platform may apply harmless algebraic simplifications in the human-readable expression. In particular, if Power is selected with $p = 0$, then $T(x) = x^0 = 1$ and the factor is neutral, so it may be omitted from the displayed expression. If Power is selected with $p = 1$, then $T(x) = x^1 = x$ and the display may show the base variable without an explicit exponent. These display choices do not change the underlying model; they only simplify its printed form.

The table below summarizes all basis-function choices available in the Kind dropdown (including **None**) and shows, for each choice, the corresponding factor $T(x)$ together with whether θ and p are used or ignored.

Kind (UI)	**Factor** $T(x)$	**Meaning / notes**	$\boldsymbol{\theta}$	$\boldsymbol{p}$
None	1	Variable not used in this row (neutral multiplicative factor).	UI-only	UI-only
Indicator+	$\mathbf{1}_{\{x>\theta\}}$	Activates row only when x is above the threshold.	used	UI-only
Indicator−	$\mathbf{1}_{\{x<\theta\}}$	Activates row only when x is below the threshold.	used	UI-only
Linear	x	Identity map. Displayed UI values (often $\theta = 0$, $p = 1$) are cosmetic.	UI-only	UI-only (implicit $p = 1$)
Power	x^p	Raw power. Use non-integer p only where $x \geq 0$ in the operating region (commonly $POOL$). Avoid $p < 0$ unless $x > 0$ always holds.	UI-only	used

Table 2: Summary of basis-function choices in a single variable column. The Kind dropdown selects the functional form of $T(x)$; the parameters θ and p are used only where indicated.

6 Worked examples

This section shows a baseline model and several modifications. Each example is written as a score function f and accompanied by a Terms-table-style configuration table that mirrors exactly what the Terms table displays: for each variable column, a Kind plus the two numeric fields θ and p. The intent is that the algebra in each example matches the Terms-table construction literally: rows add, factors within a row multiply, and regime logic is introduced by Indicator selections in the relevant variable columns.

Example 1: Default model (platform baseline). The platform's standard model is a four-factor score with two incentive-gated terms and one pool-power interaction. The overall structure follows the modeling approach developed in [1]. The incentive threshold θ_g is expressed in decimal units (for example, $\theta_g = -0.005$ corresponds to -0.5%). The score can be written as

$$f_{\text{def}}(SR, GAIN, POOL, TTM) = \beta_0 + \beta_1\, SR\, \mathbf{1}_{\{GAIN>\theta_g\}} + \beta_2\, GAIN\, POOL^{\beta_3} + \beta_4\, TTM\, \mathbf{1}_{\{GAIN>\theta_g\}}. \tag{10}$$

CPR is obtained by mapping the score f_{def} through the logistic function (1) with standards

$$\begin{aligned} \theta_g &= -0.005, \\ (\beta_0, \beta_1, \beta_2, \beta_3, \beta_4) &= (0.1400645,\ -12.87179006,\ 18.2931884,\ 0.31978514,\ 0.67270253), \\ k &= 5.09871788, \qquad i = 0.85225578. \end{aligned} \tag{11}$$

A typical standard spec also includes a spread value `spread_dec` $= 0.004$ (i.e. 40bp), stored as a decimal.

Row		SR			GAIN			POOL			POOL2			TTM		
Row	a_j	Kind	θ	p	Kind	θ	p	Kind	θ	p	Kind	θ	p	Kind	θ	p
1	β_1	Linear	0	1	Indicator+	θ_g	0	None	0	0	None	0	0	None	0	0
2	β_2	None	0	0	Linear	0	1	Power	0	β_3	None	0	0	None	0	0
3	β_4	None	0	0	Indicator+	θ_g	0	None	0	0	None	0	0	Linear	0	1

Table 3: Terms table configuration for the standard model. For Indicator kinds, p is ignored (shown as 0 for consistency).

Example 2: Standard model with pool-kill (non-callable below a pool threshold). In some portfolios you may want the model to behave as if the bond is effectively "non-callable" (or operationally inactive for refinancing) in a low-pool/burnout regime. A simple and transparent way to enforce this is to add a strongly negative "kill" contribution that activates only when $POOL$ is below a chosen threshold. Because the CPR is produced by a logistic map, pushing the score f far into the negative region makes λ extremely close to zero, without changing the functional form of the model elsewhere. Concretely, to force CPR to be essentially zero whenever $POOL \le 0.30$ while leaving the standard model unchanged for $POOL > 0.30$, add one extra row:

$$f_{\text{kill}} = f_{\text{def}} + a_{\text{kill}}\, \mathbf{1}_{\{POOL<0.30\}}, \qquad a_{\text{kill}} \ll 0. \tag{12}$$

A practical starting value is $a_{\text{kill}} = -20$, which typically makes CPR negligible when the indicator is active.

Row		SR			GAIN			POOL			POOL2			TTM		
Row	a_j	Kind	θ	p	Kind	θ	p	Kind	θ	p	Kind	θ	p	Kind	θ	p
4	a_{kill}	None	0	0	None	0	0	Indicator−	0.30	0	None	0	0	None	0	0

Table 4: Standard model plus a kill row implementing (12). This table shows the additional kill row; the baseline rows are given in Table 3.

Example 3: Pool-regime variants of the standard model (three pool levels). If a single specification is too restrictive—for example because borrower behavior is highly heterogeneous—you can split the pool into segments using the POOL factor and use a separate set of rows for each segment. This regime/mixture construction allows different parameterizations in different segments, which can amplify burnout effects beyond what a single $POOL^p$ term provides and produce more realistic variation in prepayment behavior across mortgagors.

Choose two pool cutoffs $0 < p_1 < p_2 < 1$ (for concreteness you can think of $p_1 = 0.25$ and $p_2 = 0.50$). Define the three regime indicators

$$\begin{aligned} I_L &:= \mathbf{1}_{\{POOL<p_1\}}, \\ I_H &:= \mathbf{1}_{\{POOL>p_2\}}, \\ I_M &:= \mathbf{1}_{\{POOL>p_1\}} - \mathbf{1}_{\{POOL>p_2\}}. \end{aligned} \tag{13}$$

Then I_L selects the low-pool regime, I_H selects the high-pool regime, and I_M selects the middle band $p_1 < POOL \le p_2$.

Let $R_{\text{def}} := f_{\text{def}} - \beta_0$ denote the row-sum part of the standard model (the standard score without the bias). Define three pool-specific variants of the standard row-sum by allowing the coefficients to depend on regime:

$$R_{\text{def}}^{(r)} = \beta_1^{(r)}\, SR\, \mathbf{1}_{\{GAIN>\theta_g\}} + \beta_2^{(r)}\, GAIN\, POOL^{\beta_3^{(r)}} + \beta_4^{(r)}\, TTM\, \mathbf{1}_{\{GAIN>\theta_g\}}, \qquad r \in \{L, M, H\}.$$

The combined score is then the mixture

$$f_{\text{pool}} = \beta_0 + I_L\, R_{\text{def}}^{(L)} + I_M\, R_{\text{def}}^{(M)} + I_H\, R_{\text{def}}^{(H)}. \tag{14}$$

In each region of POOL only one of (I_L, I_M, I_H) is active, so the model reduces to the corresponding standard-variant.

How to implement in the Terms table. Duplicate the three standard rows once per regime. Use **POOL** for the Power term (the burnout interaction) and use **POOL2** to gate the regime:

- **Low pool ($POOL < p_1$):** in each duplicated row set POOL2 to Indicator− with $\theta = p_1$.
- **High pool ($POOL > p_2$):** in each duplicated row set POOL2 to Indicator+ with $\theta = p_2$.
- **Middle pool ($p_1 < POOL \le p_2$):** implement I_M as a difference of two "greater-than" indicators using POOL2:

  $$I_M = \mathbf{1}_{\{POOL>p_1\}} - \mathbf{1}_{\{POOL>p_2\}}.$$

  For each middle-regime row, create two copies that are identical except for the POOL2 threshold and the sign of the coefficient: the first copy uses POOL2 Indicator+ with $\theta = p_1$ and coefficient $+\beta_k^{(M)}$; the second uses POOL2 Indicator+ with $\theta = p_2$ and coefficient $-\beta_k^{(M)}$.

If you do not need a strict middle band, you can simplify further by using only one cutoff and its complement, but the three-regime construction above gives a clean, explicit partition while keeping POOL available for the Power term.

Illustrative Terms-table mapping (structure). Table 5 lists the rows explicitly. POOL is reserved for the Power term in the gain–pool interaction row, while POOL2 is used to select the pool regime. The middle band is implemented as the difference $\mathbf{1}_{\{POOL>p_1\}} - \mathbf{1}_{\{POOL>p_2\}}$ by pairing each middle-regime row with a second copy whose coefficient is negated.

Example 4: Standard model for $TTM \le 3$ (implemented with $TTM < 3$), then a less aggressive model for $TTM > 3$. Sometimes you want the standard behavior in the short-horizon regime ($TTM \le 3$), but a less aggressive response for longer remaining time ($TTM > 3$). This is implemented by using TTM to select between two specifications. Since

Row		SR			GAIN			POOL			POOL2			TTM		
Row	a_j	Kind	θ	p	Kind	θ	p	Kind	θ	p	Kind	θ	p	Kind	θ	p
Low pool regime: $POOL < p_1$ (apply coefficients $\beta^{(L)}$)																
1L	$\beta_1^{(L)}$	Linear	0	1	Indicator+	θ_g	0	None	0	0	Indicator−	p_1	0	None	0	0
2L	$\beta_2^{(L)}$	None	0	0	Linear	0	1	Power	0	$\beta_3^{(L)}$	Indicator−	p_1	0	None	0	0
3L	$\beta_4^{(L)}$	None	0	0	Indicator+	θ_g	0	None	0	0	Indicator−	p_1	0	Linear	0	1
Middle pool regime: $p_1 < POOL \leq p_2$ (implement as "$> p_1$" minus "$> p_2$"; use $\beta^{(M)}$)																
1M1	$\beta_1^{(M)}$	Linear	0	1	Indicator+	θ_g	0	None	0	0	Indicator+	p_1	0	None	0	0
1M2	$-\beta_1^{(M)}$	Linear	0	1	Indicator+	θ_g	0	None	0	0	Indicator+	p_2	0	None	0	0
2M1	$\beta_2^{(M)}$	None	0	0	Linear	0	1	Power	0	$\beta_3^{(M)}$	Indicator+	p_1	0	None	0	0
2M2	$-\beta_2^{(M)}$	None	0	0	Linear	0	1	Power	0	$\beta_3^{(M)}$	Indicator+	p_2	0	None	0	0
3M1	$\beta_4^{(M)}$	None	0	0	Indicator+	θ_g	0	None	0	0	Indicator+	p_1	0	Linear	0	1
3M2	$-\beta_4^{(M)}$	None	0	0	Indicator+	θ_g	0	None	0	0	Indicator+	p_2	0	Linear	0	1
High pool regime: $POOL > p_2$ (apply coefficients $\beta^{(H)}$)																
1H	$\beta_1^{(H)}$	Linear	0	1	Indicator+	θ_g	0	None	0	0	Indicator+	p_2	0	None	0	0
2H	$\beta_2^{(H)}$	None	0	0	Linear	0	1	Power	0	$\beta_3^{(H)}$	Indicator+	p_2	0	None	0	0
3H	$\beta_4^{(H)}$	None	0	0	Indicator+	θ_g	0	None	0	0	Indicator+	p_2	0	Linear	0	1

Table 5: Three pool-regime variants of the standard model with POOL reserved for the Power term and POOL2 used for regime selection. Low and high regimes are gated by a single POOL2 indicator. The middle band $p_1 < POOL \leq p_2$ is implemented as $\mathbf{1}_{\{POOL>p_1\}} - \mathbf{1}_{\{POOL>p_2\}}$ by pairing each middle-regime row with a second copy whose coefficient is negated.

each row has only one TTM field, TTM cannot simultaneously enter a row as a linear factor and act as an indicator gate. In this example, TTM is used only for regime selection, and the standard model's explicit TTM-linear term is therefore omitted from the per-row factors.

Because the indicator test is strict, we implement the short regime using $TTM < 3$ as the practical equivalent of $TTM \leq 3$. Define

$$\begin{aligned} I_S(TTM) &:= \mathbf{1}_{\{TTM\leq 3\}}, \\ I_L(TTM) &:= \mathbf{1}_{\{TTM>3\}}. \end{aligned} \tag{15}$$

Let

$$R_{\text{def}}^{\neg TTM}(SR, GAIN, POOL) := \beta_1\, SR\, \mathbf{1}_{\{GAIN>\theta_g\}} + \beta_2\, GAIN\, POOL^{\beta_3}$$

denote the row-sum part of the standard specification without the explicit TTM term. Define a less aggressive long-horizon variant by scaling down the incentive-sensitive coefficients and optionally using a smaller pool-power exponent:

$$R_{\text{less}}(SR, GAIN, POOL) = \alpha_1\, SR\, \mathbf{1}_{\{GAIN>\theta_g\}} + \alpha_2\, GAIN\, POOL^{\alpha_3}. \tag{16}$$

A typical choice is $|\alpha_1| < |\beta_1|$, $\alpha_2 < \beta_2$, and often $\alpha_3 \leq \beta_3$ (weaker burnout interaction). The combined score is then

$$f_{\text{switch}} = \beta_0 + I_S(TTM)\, R_{\text{def}}^{\neg TTM}(SR, GAIN, POOL) + I_L(TTM)\, R_{\text{less}}(SR, GAIN, POOL). \tag{17}$$

How to implement in the Terms table. Duplicate the two standard rows that do not require TTM as a factor (the SR-gated term and the gain–pool power term) and gate them by TTM:

- **Short regime ($TTM \leq 3$):** duplicate the rows and set the TTM column to Indicator− with $\theta = 3$ (this multiplies each row by $\mathbf{1}_{\{TTM<3\}}$, used here as the practical equivalent of $\mathbf{1}_{\{TTM\leq 3\}}$).

- **Long regime ($TTM > 3$):** create corresponding rows using (α_1, α_2) and α_3, and set the TTM column to Indicator+ with $\theta = 3$ (this multiplies each row by $\mathbf{1}_{\{TTM>3\}}$).

Illustrative Terms-table mapping. The table below mirrors the Terms table format (Kind, θ, p) and shows the four rows explicitly.

Row		SR			GAIN			POOL			POOL2			TTM		
Row	a_j	Kind	θ	p	Kind	θ	p	Kind	θ	p	Kind	θ	p	Kind	θ	p
Short regime: $TTM \le 3$ (implemented with $TTM < 3$; use standard coefficients)																
1S	β_1	Linear	0	1	Indicator+	θ_g	0	None	0	0	None	0	0	Indicator−	3	0
2S	β_2	None	0	0	Linear	0	1	Power	0	β_3	None	0	0	Indicator−	3	0
Long regime: $TTM > 3$ (less aggressive coefficients)																
1L	α_1	Linear	0	1	Indicator+	θ_g	0	None	0	0	None	0	0	Indicator+	3	0
2L	α_2	None	0	0	Linear	0	1	Power	0	α_3	None	0	0	Indicator+	3	0

Table 6: Time switch using TTM for regime selection. Here TTM does not enter as a separate linear factor; it is used only to gate the short regime (Indicator− at $\theta = 3$) versus the long regime (Indicator+ at $\theta = 3$).

Example 5: Near-optimal (almost certain) prepayment when refinancing is beneficial. To model near-optimal behavior, we force the CPR to be close to 0 when refinancing is not beneficial and close to 1 when it is beneficial. A robust way to do this in the builder is to use a strongly negative bias (baseline score) plus a large positive "call" shift gated by an incentive trigger.

Using the simple trigger $GAIN > 0$, define

$$f_{\text{opt}} = a_0 + a_{\text{call}}\,\mathbf{1}_{\{GAIN>0\}}. \tag{18}$$

Choose $a_0 \ll 0$ and $a_{\text{call}} \gg 0$ so the score is far below i when $GAIN \le 0$ and far above i when $GAIN > 0$. With a sufficiently steep k, CPR becomes very close to a step function.

Illustrative parameters. A practical starting point is

$$k = 25, \qquad i = 0, \qquad a_0 = -10, \qquad a_{\text{call}} = 20,$$

which yields $\lambda \approx 0$ for $GAIN \le 0$ and $\lambda \approx 1$ for $GAIN > 0$.

Row		SR			GAIN			POOL			POOL2			TTM		
Row	a_j	Kind	θ	p	Kind	θ	p	Kind	θ	p	Kind	θ	p	Kind	θ	p
1	a_{call}	None	0	0	Indicator+	0	0	None	0	0	None	0	0	None	0	0

Table 7: An "almost step" CPR using a strongly negative bias a_0 and a large positive shift when $GAIN > 0$. The bias a_0 is not a Terms-table row; it is the model intercept in (2).

7 Practical notes

Because the final CPR is produced by a logistic mapping, exact 0% and 100% are not achieved for finite parameters. This is usually desirable because it keeps the model stable and avoids hard discontinuities in downstream computations. When you do want near-step behavior, a robust approach is to use indicator terms together with a sufficiently negative baseline score and an adequately large steepness k.

When using fractional powers (such as $p = 0.32$), ensure that the powered variable is non-negative in the intended operating region. In typical conventions, $POOL$ is a natural candidate for fractional powers, while variables like $GAIN$ may be negative and therefore are more naturally used with indicators or linear factors. If you need a term to apply only in a particular regime, place the corresponding indicator in the same row as that term so the multiplication makes the regime restriction explicit.

Finally, remember the structural rule: the model *adds* rows and *multiplies* factors within each row. This means conjunctions across variables are easy (use multiple indicator columns in the same row), while two-sided bands on the same variable typically require multiple rows. The expression shown in the platform is intended to match the literal definition of the model, but it may omit factors that are algebraically neutral (for example, it will not display x^0 because $x^0 = 1$, and it will display x^1 as x).

8 Conclusion

The builder provides a direct mapping from Terms rows to a mathematically precise score-and-logistic model. Each row contribution is a coefficient times a product of basis factors, and the overall score is transformed into CPR by the logistic mapping. The symbolic expression shown in the platform is intended to be read as the literal model definition; it should match the equations in this guide term-for-term, up to harmless algebraic simplifications.

References

[1] N. Rom, *Callable Mortgage Bonds: Numerical Methods and Valuation Models for Pricing and Risk Analysis*. Springer Cham, 2025. doi: 10.1007/978-3-031-87889-3. ISBN: 978-3-031-87889-3.