

# The Barrier to Human Understanding of MARS

*Reading a MARS model as rates, contributions and adjustments*

*Foundations • The Valuation Engineer*

William Bert Craytor • October 1, 2026

Version 1.2.0

DOI: [10.67330/qmnctv78](https://doi.org/10.67330/qmnctv78)

---

This is a supplement to the Foundations department, not a concept entry. Its title names a barrier the editor has come to recognize over years of indirect observation: appraisers who read a regression coefficient without difficulty stop at the list of hinge terms a MARS program prints, because nothing in that list looks like a rate. The four statistical entries that follow this supplement (*Conditional Expectation*, *Omitted Variable Bias*, *Endogeneity* and *Parameter Identification*) reason about the rate a regression assigns to a feature. The Journal’s own modeling uses multivariate adaptive regression splines (MARS; Friedman 1991)<sup>1</sup>, which report not one rate per feature but one per range, and the path from what a MARS program prints to the rate table an appraiser can defend has one step at which hand-built tables go wrong. This supplement gives the path, the step, and a program that walks it and checks its own work.

**Formal definition.** In a MARS model the fitted value is an intercept plus a sum of terms, each a coefficient times a product of hinge functions of the features. The fitted value is the model’s estimate of price. For one feature, with the others held fixed, it is a series of straight segments. The points where the segments meet are the knots. The graph of the feature’s value contribution over its range is the *variable graph*. Its slope in each range is the *rate*: the change in fitted value per unit of the feature in that range; the *value contribution* of a quantity of the feature is the sum of the model’s terms in that feature at that quantity; in the example all those terms are zero at 1,200 square feet, so the contribution is measured from there and is negative below it; a model need not have such a quantity, and only differences between contributions, which are adjustments, carry meaning; and the *adjustment* from a comparable to a subject is the amount added to the comparable’s contribution to make it equal to the subject’s, which is the sum, over the ranges between the two quantities, of rate times width. The rate so defined is the model’s; it is read as the rate the market pays only once the model has been accepted as a fair account of the sales, and the entries that follow this supplement are about when that acceptance is warranted.

## Why this supplement exists

An appraiser who fits a MARS model, in earthUI (Craytor 2026) or directly in the R package earth (Milborrow 2024a), is handed a list of terms like these:

|             |         |
|-------------|---------|
| (Intercept) | 1520000 |
| h(1200-gla) | -600    |
| h(gla-1200) | 1000    |

---

<sup>1</sup>MARS is a registered trademark of Minitab, LLC. It is used in this Journal as the name of the statistical method introduced by Friedman (1991), and the software referred to is the R package earth, whose notation this Journal follows: hinges are written  $h(x - t)$  and  $h(t - x)$  and models are shown as earth prints them. Salford’s MARS writes the same functions as numbered basis functions,  $BF1 = \max(0, X - t)$  and so on, and differs from earth in such details of notation, not in the method.

|                      |      |
|----------------------|------|
| $h(\text{gla-2000})$ | -250 |
| $h(\text{gla-2800})$ | -650 |

Nobody reports that. What goes into the sales grid is the adjustment: the subject's value contribution less the comparable's. The rates per square foot are discussed in the report, as a step in justifying the MARS model. The path from the printed terms to those numbers is short, but it has one step where hand-built tables go wrong, and the whole point of using a model is lost if the numbers it produces cannot be explained to a reviewer.

## The plain version: four ranges, four rates

A linear regression that enters living area as one linear term gives one rate for it across the whole market; this is what regressions in traditional appraisal do, and it is what "ordinary regression" means in this supplement. MARS lets the rate change where the sales show it changing. In the coastal market of the example, living area may be worth \$600 a square foot up to 1,200 square feet, \$1,000 a square foot from 1,200 to 2,000, \$750 from 2,000 to 2,800, and only \$100 above that. That is the whole difference, and it is one an agent understands at once: the market pays most for the space that turns a small house into a family house, less for the next room, and almost nothing for space beyond what the neighborhood's buyers want.

Given the rates, everything else is arithmetic an appraiser already does. A comparable with 1,800 square feet and a subject with 2,400 are 600 square feet apart, and the knot at 2,000 lies between them. The adjustment added to the comparable is the subject's fitted value less the comparable's,  $\$2,620,000 - \$2,120,000 = \$500,000$ , which is also the subject's living-area contribution less the comparable's,  $\$1,100,000 - \$600,000$ . The rates explain it: the move crosses 200 square feet at \$1,000, then 400 square feet at \$750, which is \$500,000 in all. Figure 3 shows it, on a fitted value (the price the model gives a house with those features) that runs from \$1,280,000 at 800 square feet, the smallest house in the data, to about \$3 million at 3,500, the largest, most of the rise coming before 2,800.

The only rate a straight line through the two points could have used for that move is the one that connects them, \$500,000 over 600 square feet, about \$833 a square foot: the pair's *secant rate*, the effective, or blended, rate for that one move: the average of the model's rates weighted by the width of each range crossed. It is not one of the model's rates, and it would be a different average for the next pair. An ordinary regression, with one rate for living area, has to serve every pair with that one number, so it cannot reproduce these pair-by-pair averages; that is what a single rate does: it averages over a change the model says the market makes.

A reader who stops here has the method: take the difference of the two contributions, then explain it by reading the rate on each range, walking the comparable to the subject at those rates, and adding. What follows is for the reader who wants to know why the rates are what they are, where the one trap lies, what changes when features interact, and how a program checks the arithmetic; the rate table and the adjustment do not change.

## The refined version: hinges, and the sign that gets missed

Call the graph of a variable's value contribution over its range the *variable graph*. It is not a curve in the everyday sense of something that bends smoothly; it is a segmented linear graph, straight between the knots and bending only at them. MARS builds it out of *hinges*; Friedman (1991) wrote them as truncated power functions, and "hinge" is the name the later literature settled on

(Milborrow, n.d.); Hastie, Tibshirani and Friedman (2009, §9.4) call the two at one knot a “reflected pair”.

*Definition (hinge).* A hinge is a function of one variable  $x$  with a single break, called its *knot*  $t$ , whose value is zero on one side of the knot and, on the other side, equal to the distance from the knot. There are two of them at any knot, one for each side, drawn in Figure 1:

$$h(x - t) = \max(0, x - t) = \begin{cases} 0 & x \leq t, \\ x - t & x > t, \end{cases} \quad h(t - x) = \max(0, t - x) = \begin{cases} t - x & x < t, \\ 0 & x \geq t. \end{cases}$$

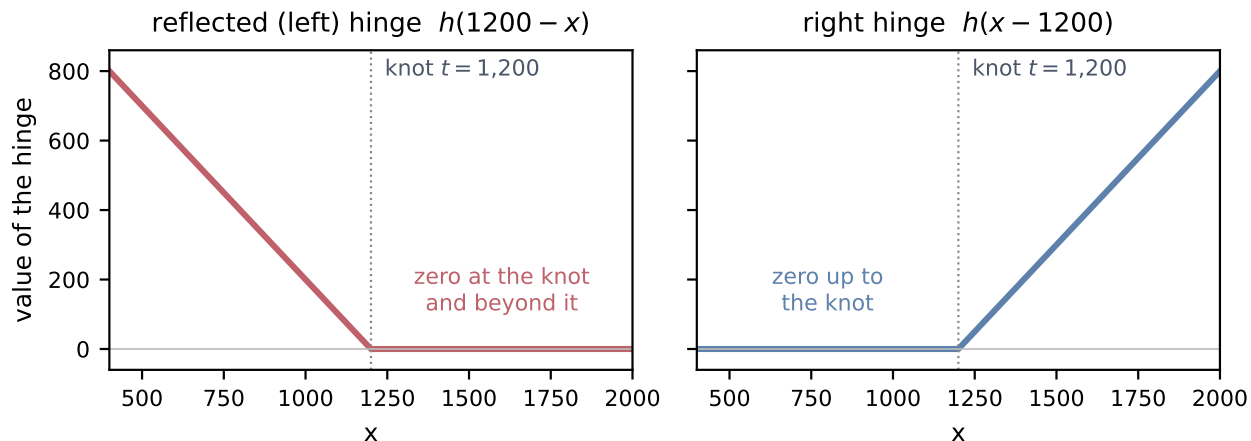


Figure 1: The two hinge functions with the knot at  $t = 1,200$ : the reflected (left) hinge  $h(1200 - x)$ , decreasing from left to right to zero at the knot and zero beyond it; the right hinge  $h(x - 1200)$ , zero up to the knot and increasing from left to right beyond it. These are the bare hinges; in a model each is multiplied by a coefficient, which may reverse the direction.

The first,  $h(x - t)$ , is the *right* hinge: zero up to the knot and, beyond it, increasing from left to right. The second,  $h(t - x)$ , is the *reflected* (or left) hinge: decreasing from left to right up to the knot, and zero at the knot and beyond it. Both are hinges in the same sense; they differ only in which side of the knot they are non-zero on. These directions describe the bare hinges, drawn from the  $x$  axis in Figure 1, and nothing more. In a model a hinge is multiplied by a coefficient, which may be negative, and added to other terms, so a term built on a right hinge can decrease from left to right and one built on a reflected hinge can increase, and the variable graph meets a knot at whatever height the other terms give it. Every direction stated in this supplement is read from left to right. Hinges are not used on their own. In a MARS model each hinge is multiplied by a coefficient and the terms are added, and it is from those coefficients, not from any single hinge, that the slopes of the model, the rates, are calculated: accumulatively, range by range, each range’s rate being the previous range’s rate adjusted by the coefficients of the hinges whose knot lies between them. The hinge is the only building block MARS has for a single variable; every bend in a variable graph is a knot of some hinge, and every straight stretch is a sum of terms that are either zero or linear there. **The point to carry through everything that follows: the coefficient of a hinge is not a slope. It is an adjustment to the slope, the amount added, at that hinge’s knot, to the slope arriving from the left.** Read from left to right, a variable graph starts at the rate of its first segment and each knot adjusts that rate by the coefficients sitting

there. That is the whole reading rule, and Theorems 1 and 2 below are its proof. The model is the intercept plus each coefficient times its hinge, and the rate in any range is the slope of that sum.

Take the printed terms above; Figure 2 draws each of them with their sum and, beneath, the rate on each range, and Figure 3 shows how those rates explain the adjustment between a comparable and a subject. Knots at 1,200, 2,000 and 2,800 make four ranges. In each range, a right hinge is *active* if its knot lies at or to the left of the range, and a reflected hinge if its knot lies at or to the right of it — a right hinge acts in every range beyond its knot, a reflected hinge in every range before it. The rate is calculated from the coefficients of the active hinges: add the coefficients of the active right hinges and subtract the coefficients of the active reflected hinges. The table that results is the model’s rate on each range, the other features held at one set of values (the program uses the subject’s); in a model without interactions that qualification changes nothing, and in Case 2 below it changes everything:

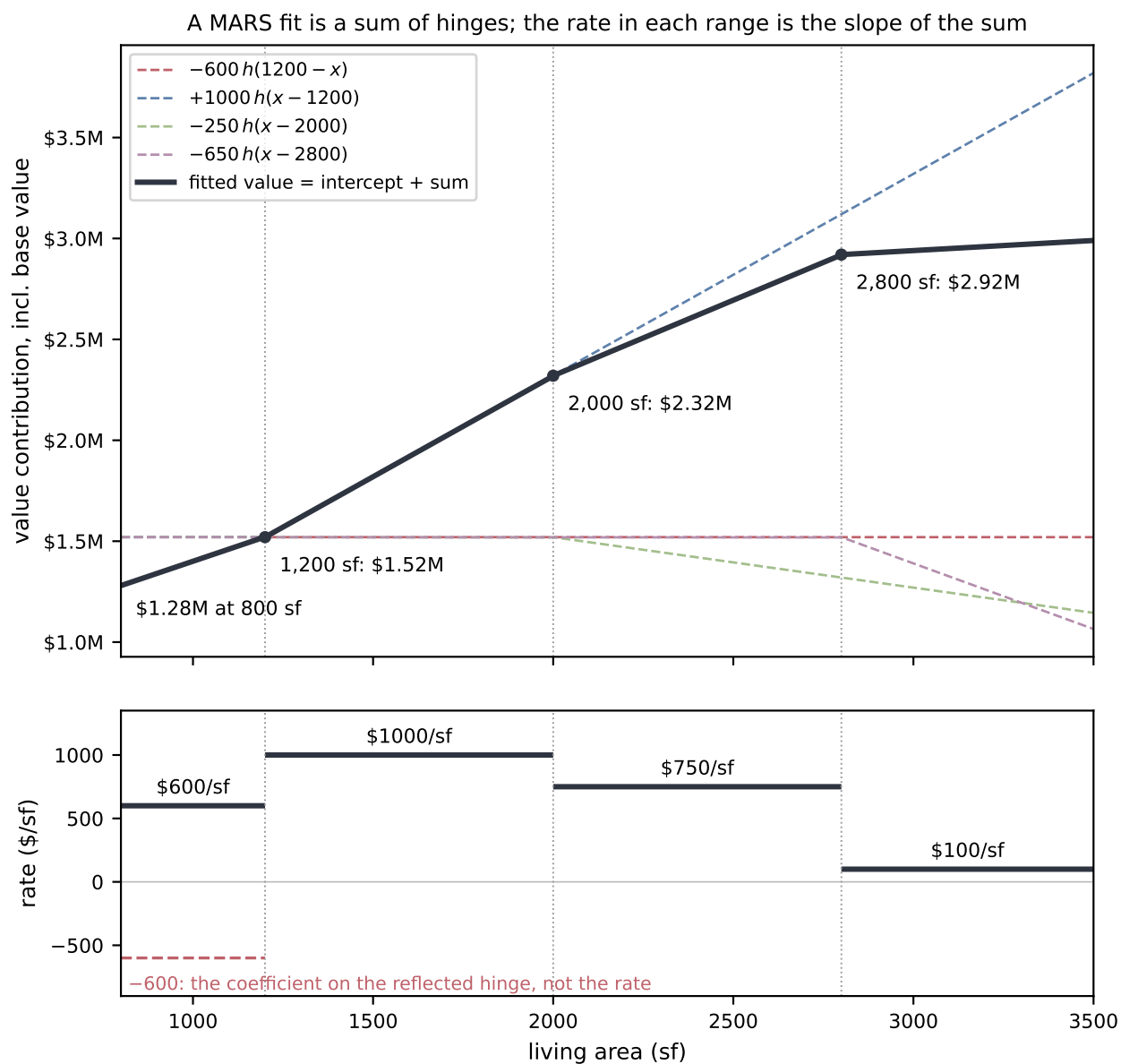
| range (sf)     | active hinges                                | rate (\$/sf)                      |
|----------------|----------------------------------------------|-----------------------------------|
| below 1,200    | $h(1200 - \text{gla})$ only                  | $-(-600) = +\mathbf{600}$         |
| 1,200 to 2,000 | $h(\text{gla} - 1200)$                       | $\mathbf{1000}$                   |
| 2,000 to 2,800 | $h(\text{gla} - 1200), h(\text{gla} - 2000)$ | $1000 - 250 = \mathbf{750}$       |
| above 2,800    | all three right hinges                       | $1000 - 250 - 650 = \mathbf{100}$ |

The first row is the step that gets missed. A coefficient of  $-600$  on a reflected hinge is not a rate of  $-\$600$  a square foot. The hinge  $h(1200 - \text{gla})$  measures how far a house falls short of 1,200 square feet, and  $-600$  times that shortfall is a discount:  $\$240,000$  for an 800-square-foot house, shrinking to nothing at 1,200. Each added square foot removes  $\$600$  of the discount, so the rate is  $+\$600$ . (With a positive coefficient the same hinge adds value for being smaller, and its rate is negative.) The rule, stated once: in a range, the rate is the sum of the coefficients of the active right hinges *minus* the sum of the coefficients of the active reflected hinges, plus the coefficient of a linear term (a feature entered as a straight line with no knot), if the model has one. One convention goes with it: exactly at a knot the rate is taken from the right, so that a right hinge counts from its knot onward and a reflected hinge only strictly below its knot; the rate at 1,200 square feet is therefore  $\$1,000$ , the rate of the range that begins there, and the program reads it the same way.

The check that a table was built right is the fitted value at the knots. At 1,200 square feet every hinge is zero and the value is the intercept,  $\$1,520,000$ ; at 2,000 it is  $\$1,520,000 + 1,000 \times 800 = \$2,320,000$ ; at 2,800,  $\$2,320,000 + 750 \times 800 = \$2,920,000$ ; and at 800 square feet, the smallest house in the data, it is  $\$1,520,000 - 600 \times 400 = \$1,280,000$ . From 800 to 1,200 the value rises  $\$240,000$ , a rate of  $\$600$ , which is the first row; a table that read the coefficient as the rate would have the value decreasing from left to right up to the knot at 1,200, and the values show it does not. Figure 2 draws the four hinges, their sum, and the rate in each range, with the coefficient  $-600$  marked where a reader might mistake it for the rate.

Two further complications the program handles and a hand table must:

- **Interactions.** earth, in its forward pass (the stage that adds terms one at a time; Friedman 1991, §3), prints a product like  $h(\text{gla} - 1500) \times \text{lot}$  when the rate for one feature depends on another: an *interaction*. The rate for lot then differs above and below 1,500 square feet of living area, and the rate table for one feature is always the table *given* the others, held at the subject’s values. The program takes a point, a value for every other feature, and computes every rate at it.

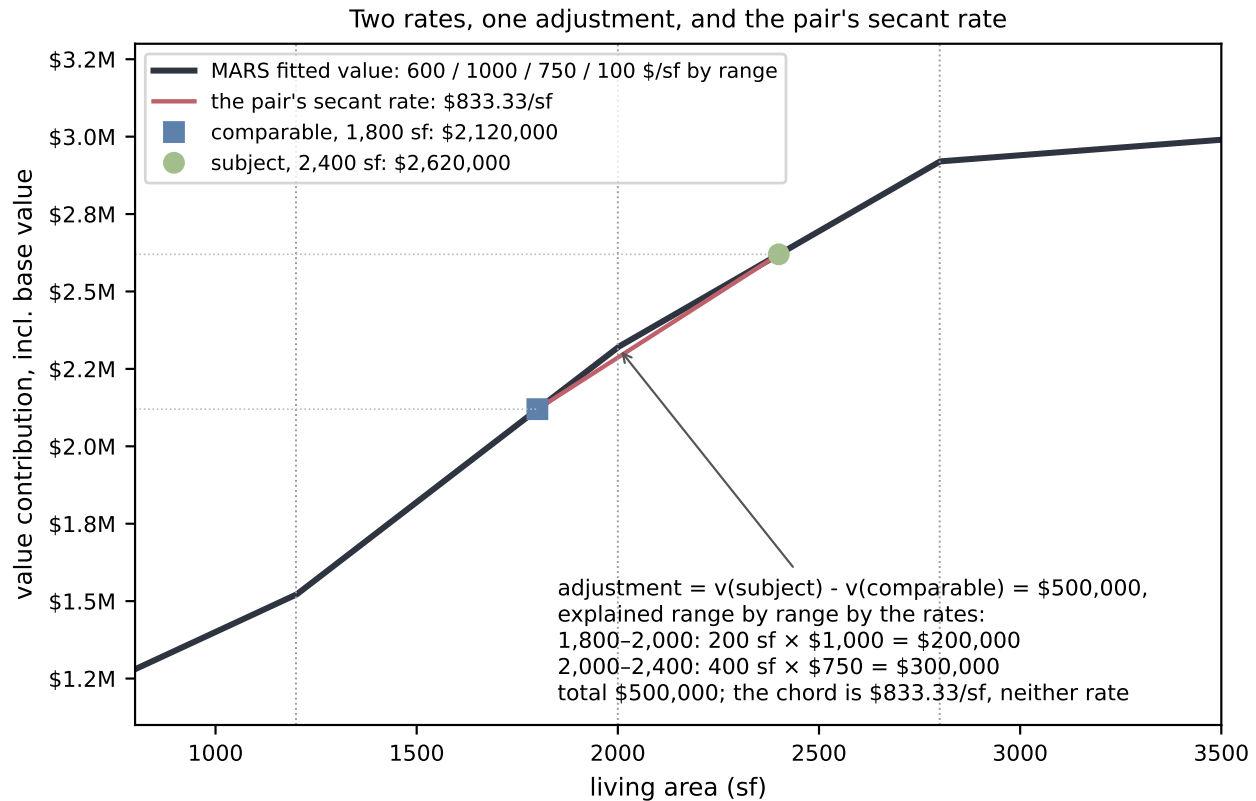


```

Model (v = value contribution, gla = living area in sf):
v = 1,520,000 - 600·h(1200-gla) + 1000·h(gla-1200) - 250·h(gla-2000) - 650·h(gla-2800)
as earth prints it:
(Intercept) 1520000 h(1200-gla) -600 h(gla-1200) 1000 h(gla-2000) -250 h(gla-2800) -650
Rate on a segment = (coefficients of the active right hinges) - (coefficients of the active reflected hinges):
800-1200 sf: -(-600) = $600/sf only the reflected hinge at 1200 is active
1200-2000 sf: 1000 = $1000/sf
2000-2800 sf: 1000 - 250 = $750/sf
2800-3500 sf: 1000 - 250 - 650 = $100/sf all three right hinges active

```

Figure 2: The four hinge terms of the example, their sum, and the rate in each range. The coefficient  $-600$  on the reflected hinge  $h(1200 - \text{gla})$  is marked where a reader might mistake it for the rate; the rate below 1,200 square feet is  $+600$ .



```

Model (v = value contribution, gla = living area in sf):
v = 1,520,000 - 600·h(1200-gla) + 1000·h(gla-1200) - 250·h(gla-2000) - 650·h(gla-2800)
as earth prints it:
(Intercept) 1520000 h(1200-gla) -600 h(gla-1200) 1000 h(gla-2000) -250 h(gla-2800) -650
Adjustment from the comparable (1,800 sf) to the subject (2,400 sf) =  $v(2,400) - v(1,800) = \$500,000$ ;
explained range by range by the rates:
1800-2000 sf: 200 sf  $\times$  $1000/sf = $200,000
2000-2400 sf: 400 sf  $\times$  $750/sf = $300,000
total = $500,000 secant rate for the pair:  $500,000 / 600 = \$833.33/\text{sf}$ 

```

Figure 3: The fitted value with the comparable at 1,800 square feet and the subject at 2,400. The adjustment, the difference of the two contributions, is \$500,000, and the rates in force explain it range by range; the chord joining the two points, about \$833 a square foot, is the pair's secant rate, the single rate a straight line would need for this pair and no other; an ordinary regression must use one rate for every pair.

- **Level versus adjustment.** The intercept and the terms in other features set the level the contribution is measured from; they cancel in every adjustment. The program reports contributions as the terms in the feature alone, so that two contributions differ by exactly the adjustment.

## Two cases: models without and with interactions

The rule of the previous section can be stated as two theorems, proved below, and they give a clean answer to a question every reader of a MARS model asks sooner or later: how many of the printed terms does one rate actually depend on? The answer splits into two cases.

### Case 1: no interactions

**Setting.** Let  $h(u) = \max(0, u)$  for real  $u$ . A MARS model without interactions on the features  $x, z_1, \dots, z_q$  has fitted value

$$f(x, z_1, \dots, z_q) = \beta_0 + g(x) + \sum_{k=1}^q g_k(z_k),$$

where  $\beta_0$  is the intercept, each  $g_k$  is a function of  $z_k$  alone, and  $g$  collects every term in  $x$ :

$$g(x) = c \cdot x + \sum_{j=1}^K a_j h(x - t_j) + \sum_{j=1}^K b_j h(t_j - x).$$

Here  $t_1 < t_2 < \dots < t_K$  are the distinct knots of the hinges in  $x$ ;  $a_j$  is the coefficient of the right hinge at  $t_j$  and  $b_j$  that of the reflected hinge at  $t_j$ , either being 0 when the model has no such term; and  $c$  is the coefficient of a bare linear term  $c \cdot x$ , a feature entered as a straight line with no knot, which earth produces only for features named in `linpreds` or, for any feature, when `Auto.linpreds` is on (see the remark after Theorem 1);  $c = 0$  when there is no such term, as in every model of this supplement. Put  $t_0 = -\infty$  and  $t_{K+1} = +\infty$  and call  $S_i = (t_i, t_{i+1})$ , for  $i = 0, 1, \dots, K$ , the *segments* of  $x$ ;  $t_i$  is the left knot of  $S_i$  and  $t_{i+1}$  its right knot. On each  $S_i$  every summand of  $g$  is either identically zero or linear, so  $g$  is linear on  $S_i$ ; its slope there, which is also  $\partial f / \partial x$  on  $S_i$  because no other summand of  $f$  depends on  $x$ , is the *rate* on  $S_i$  and is written  $r_i$ . The rate table of  $x$  is the list  $r_0, r_1, \dots, r_K$ . (If  $K = 0$  there is one segment,  $(-\infty, +\infty)$ , and the table is the single rate  $r_0 = c$ .) At a knot itself  $g$  has no derivative: its left and right derivatives are  $r_{i-1}$  and  $r_i$ , which differ by  $a_i + b_i$  (Theorem 2 below), so the value a rate table shows at  $t_i$  is a convention, taken here from the right, and not a value of  $g'$ .

**Theorem 1 (the rate on a segment).** With the notation of the setting, for every  $i \in \{0, 1, \dots, K\}$ ,

$$r_i = c + \sum_{j=1}^i a_j - \sum_{j=i+1}^K b_j,$$

an empty sum being 0. In particular  $r_i$  depends on at most  $K$  of the  $2K$  hinge coefficients, one per knot, together with  $c$ .

**Proof of Theorem 1.** Fix  $i$  and a point  $x \in S_i$ , so  $t_i < x < t_{i+1}$ . Consider each summand of  $g$  on a neighborhood of  $x$  small enough to lie inside  $S_i$ . The bare term  $c \cdot x$  has derivative  $c$ . For the right hinge at  $t_j$ : if  $j \leq i$  then  $t_j \leq t_i < x$ , so  $h(x - t_j) = x - t_j$  on the neighborhood and its

derivative is 1; if  $j \geq i + 1$  then  $x < t_{i+1} \leq t_j$ , so  $h(x - t_j) = 0$  there and its derivative is 0. For the reflected hinge at  $t_j$ : if  $j \geq i + 1$  then  $x < t_j$ , so  $h(t_j - x) = t_j - x$  and its derivative is  $-1$ ; if  $j \leq i$  then  $t_j < x$ , so  $h(t_j - x) = 0$  and its derivative is 0. A finite sum of functions differentiable at  $x$  is differentiable there with the sum of the derivatives, which gives

$$g'(x) = c + \sum_{j \leq i} a_j \cdot 1 + \sum_{j \geq i+1} a_j \cdot 0 + \sum_{j \leq i} b_j \cdot 0 + \sum_{j \geq i+1} b_j \cdot (-1),$$

the stated formula. For the count: the index sets  $\{j : j \leq i\}$  and  $\{j : j \geq i + 1\}$  are disjoint and together exhaust  $\{1, \dots, K\}$ , so each knot  $t_j$  contributes exactly one of  $a_j$  and  $b_j$  and never both; hence at most  $K$  hinge coefficients enter, and  $r_i$  is the same at every point of  $S_i$ . ■

The bound is attained: if  $b_j = 0$  for every  $j$ , the rightmost segment  $S_K$  uses all  $K$  of the  $a_j$ .

**Remark (when  $c$  is not zero).** earth can express a feature as a single straight line with no knot, the term  $c \cdot x$  of the setting, in two different ways, and the difference matters. `linpreds`, a list of named features such as `linpreds = c("GLA")`, *forces* each feature on the list to enter linearly, even when a fit with knots would be better; it is the analyst's decision, and it can cost fit. `Auto.linpreds`, a single switch that is on by default, forces nothing: it only changes how earth writes a feature whose fit has turned out to be truly linear, a single hinge with its knot at the smallest observed value, by printing it as a coefficient and the variable name instead of as a hinge; the fit is the same. In either case earth prints the term as the bare variable name with its coefficient, `221.4 * GLA` in earth's summary form, as in the example below. Theorem 1 handles it through  $c$ , which adds to the rate on every segment. For example, the call

```
mod_linear <- earth(
  SalePrice ~ GLA + LotSize + BathrmCount + BedrmCount,
  data = sales,
  degree = 1,
  linpreds = c("GLA", "LotSize", "BathrmCount", "BedrmCount")
)
```

enters every feature linearly (with `degree = 1`, so that no products of the linear terms are fitted either) and prints, in earth's summary form (illustrative output, not a fitted model),

```
SalePrice =
  179843
  + 221.4 * GLA
  + 11.8 * LotSize
  + 34872 * BathrmCount
  + 17641 * BedrmCount
```

For living area this is the setting with  $c = 221.4$  and no hinges,  $K = 0$ : there is one segment, and Theorem 1 gives the single rate  $r_0 = c = \$221.40$  a square foot over the whole range of the data; lot size, bathrooms and bedrooms read the same way, at \$11.80 a square foot of lot, \$34,872 a bathroom and \$17,641 a bedroom. Such a model is an ordinary regression in the sense of this supplement, one rate per feature. Over the data such a term is the same line as a single right hinge with its knot at the smallest observed value, and the Journal runs earth with `Auto.linpreds = FALSE`, so in its models a feature whose variable graph is straight over the data appears as that single hinge, `h(gla-800) 700` for instance, the knot at the smallest house in the data, with  $c = 0$ ,  $K = 1$  and  $a_1 = 700$ : the rate is 0 below the knot, on a stretch that holds no sales, and 700 above

it, so the rate table over the data has one row. One corollary follows here; the change of rate at a knot is Theorem 2 below.

**Corollary 1 (the reflected sign).** A reflected hinge with coefficient  $b_j$  contributes  $-b_j$  to the rate on every segment to the left of  $t_j$  and 0 to the right. In the running example,  $b_1 = -600$  at  $t_1 = 1,200$  contributes  $+600$  to  $r_0$ , which is the whole of the trap this supplement is built around.

**Proof.** Read off the second sum of Theorem 1. ■

### A seven-knot example, and the assumption that defeats readers

The count matters because of an assumption almost everyone brings to a printed MARS model: that the coefficient printed beside a knot is the slope of the segment whose left knot it is. It is not. **The coefficient of a hinge is an adjustment to the slope coming in from the left, not the slope itself**, and the slope of a segment is therefore the slope of the first segment plus the running total of every such adjustment at its left knot and at every knot further left. A reader who holds the wrong assumption tries the arithmetic in their head, or in a small program or a spreadsheet formula, finds that the numbers do not reproduce the fitted values, and gives up, concluding that MARS is opaque. It is not opaque; it is cumulative.

**Theorem 2 (the running total).** With the notation of the setting of Case 1, for every  $i \in \{1, \dots, K\}$ ,

$$r_i = r_{i-1} + a_i + b_i.$$

Consequently  $r_i = r_0 + \sum_{j=1}^i (a_j + b_j)$  for every  $i$ , where  $r_0 = c - \sum_{j=1}^K b_j$  is the rate on the leftmost segment and  $r_K = c + \sum_{j=1}^K a_j$  is the rate on the rightmost.

**Proof of Theorem 2.** Fix  $i \in \{1, \dots, K\}$ . By Theorem 1,

$$r_i = c + \sum_{j=1}^i a_j - \sum_{j=i+1}^K b_j \quad \text{and} \quad r_{i-1} = c + \sum_{j=1}^{i-1} a_j - \sum_{j=i}^K b_j.$$

Subtracting, the  $c$  cancels; the two sums of  $a$ 's differ by the single term  $a_i$ ; and since  $\sum_{j=i}^K b_j = b_i + \sum_{j=i+1}^K b_j$ , the two sums of  $b$ 's differ by the single term  $b_i$ , which is subtracted in  $r_{i-1}$  and not in  $r_i$ . Hence  $r_i - r_{i-1} = a_i + b_i$ . Adding this identity for  $j = 1, \dots, i$  in place of  $i$ , the left sides telescope:  $(r_1 - r_0) + (r_2 - r_1) + \dots + (r_i - r_{i-1}) = r_i - r_0$ , which gives the sum formula. The values of  $r_0$  and  $r_K$  are Theorem 1 at  $i = 0$ , where the sum of  $a$ 's is empty, and at  $i = K$ , where the sum of  $b$ 's is empty. ■

**Corollary 2 (the reading rule).** In the setting of Case 1, for every  $i \in \{1, \dots, K\}$  the rate on the segment whose left knot is  $t_i$  is  $r_i = r_{i-1} + a_i + b_i$ : the coefficients printed at a knot are added to the rate arriving from the left, and are not themselves the rate of the segment that begins there. The two coincide,  $r_i = a_i + b_i$ , exactly when  $r_{i-1} = 0$ ; in particular, for a model whose only hinges are right hinges and which has no linear term, this happens at  $i = 1$  and, by Theorem 2, at no later knot unless the coefficients at the knots before it sum to zero.

**Proof.** The first equality is Theorem 2.  $r_i = a_i + b_i$  if and only if  $r_{i-1} = 0$  follows from it. If  $b_j = 0$  for all  $j$  and  $c = 0$ , Theorem 2 gives  $r_0 = 0$ , so  $r_1 = a_1$ , and  $r_{i-1} = \sum_{j < i} a_j$ , which is 0 only when those coefficients sum to zero. ■

**Corollary 3 (the adjustment as an integral).** In the setting of Case 1,  $g$  is continuous at every knot, because each hinge is, and so is  $f(\cdot, z)$  for every  $z$ ; that is what makes the value at a knot

the same whether it is reached from the left or the right, and what lets the rates be walked from one quantity to another. Let  $x_c$  and  $x_s$  be the quantities of  $x$  for a comparable and a subject, with the other features held at the same values for both, so that  $g(x)$ , the terms of  $f$  that involve  $x$ , is the contribution of  $x$  at quantity  $x$ . Write  $u_1 = \min(x_c, x_s)$  and  $u_2 = \max(x_c, x_s)$ , and  $|I|$  for the length of an interval  $I$ . Then the adjustment  $g(x_s) - g(x_c)$ , the amount added to the comparable's contribution to make it equal to the subject's, is

$$g(x_s) - g(x_c) = \int_{x_c}^{x_s} g'(u) du = \text{sign}(x_s - x_c) \sum_{i=0}^K r_i \cdot |S_i \cap [u_1, u_2]|.$$

In words: the adjustment is the integral of the rate between the two quantities, and because the rate is constant on each segment the integral is the sum, over the segments met on the way, of rate times width.

**Proof.**  $g$  is continuous and piecewise linear with finitely many pieces, hence absolutely continuous, so the fundamental theorem of calculus gives  $g(x_s) - g(x_c) = \int_{x_c}^{x_s} g'(u) du$ , the derivative existing everywhere except at the knots, a finite set. On each  $S_i$  the integrand equals  $r_i$  by Theorem 1, so  $\int_{u_1}^{u_2} g'(u) du$  is the sum over  $i$  of  $r_i$  times the length of  $S_i \cap [u_1, u_2]$ , segments not met contributing length 0; and  $\int_{x_c}^{x_s} = \text{sign}(x_s - x_c) \int_{u_1}^{u_2}$  by the orientation of the integral. ■

The rate table is  $g'$ ; the adjustment, computed directly as  $g(x_s) - g(x_c)$ , equals its integral between the two quantities; and Theorems 1 and 2 say what  $g'$  is made of. In practice the adjustment is always computed as that difference of contributions, which needs no rates at all; the corollary says why the rates, read range by range, reproduce it, which is what makes them an explanation of the adjustment and a check on it, not a way of computing it.

Take living area in a dataset whose smallest house has 800 square feet, with seven knots, at 1,200, 1,600, 2,000, 2,400, 2,800, 3,200 and 3,600 square feet, and nine terms in the model, two of them reflected hinges, each paired with a right hinge at the same knot: at 1,200 and at 2,400:

|             |      |
|-------------|------|
| h(1200-gla) | -500 |
| h(gla-1200) | 700  |
| h(gla-1600) | -100 |
| h(gla-2000) | 150  |
| h(2400-gla) | -200 |
| h(gla-2400) | -250 |
| h(gla-2800) | -300 |
| h(gla-3200) | -150 |
| h(gla-3600) | 50   |

Seven knots make eight segments. The rate on each is the sum of the coefficients of the active right hinges minus the sum of the coefficients of the active reflected hinges:

| segment (sf)   | active terms                                           | rate (\$/sf)                  |
|----------------|--------------------------------------------------------|-------------------------------|
| 800 to 1,200   | $h(1200 - \text{gla}), h(2400 - \text{gla})$           | $500 + 200 = 700$             |
| 1,200 to 1,600 | $h(\text{gla} - 1200), h(2400 - \text{gla})$           | $700 + 200 = 900$             |
| 1,600 to 2,000 | right hinges at 1,200, 1,600; $h(2400 - \text{gla})$   | $700 - 100 + 200 = 800$       |
| 2,000 to 2,400 | right hinges at 1,200 to 2,000; $h(2400 - \text{gla})$ | $700 - 100 + 150 + 200 = 950$ |
| 2,400 to 2,800 | right hinges at 1,200 to 2,400                         | $700 - 100 + 150 - 250 = 500$ |
| 2,800 to 3,200 | right hinges at 1,200 to 2,800                         | $500 - 300 = 200$             |
| 3,200 to 3,600 | right hinges at 1,200 to 3,200                         | $200 - 150 = 50$              |
| above 3,600    | all seven right hinges                                 | $50 + 50 = 100$               |

Three things to read off the table. The rightmost segment's rate,  $700 - 100 + 150 - 250 - 300 - 150 + 50 = 100$ , combines the coefficients of all seven right hinges, one per knot, which is the bound of Theorem 1 with  $K = 7$ ; the two reflected hinges play no part there, because a reflected hinge is flat to the right of its own knot and the rightmost segment lies to the right of every knot. The leftmost segment, from 800 up to its right knot at 1,200, is the mirror image: only the reflected hinges count, with their signs reversed, because every knot lies at or to the right of its right knot. And at each knot that carries a pair the step is the sum of the two coefficients: at 1,200,  $-500 + 700 = +200$ , from 700 to 900; at 2,400 the rate drops from 950 to 500, a change of  $-450$ , which is exactly the sum of the two coefficients at that knot,  $-250$  and  $-200$ : crossing the knot switches the right hinge on and the reflected hinge off, and both changes push the same way.

The same rate on the rightmost segment can be reached by walking: start at 700 on the first segment, 800 to 1,200, and add, at each knot crossed, the sum of the coefficients sitting there,  $+200, -100, +150, -450, -300, -150, +50$ . That route touches all eight segments and all nine terms and lands on the same 100. The direct sum is shorter and is what the program does; the walk is what a reader should do once, by hand, to see that the coefficient beside a knot is a step and not a level. The  $-450$  at 2,400 is two half-steps: crossing the knot leaves the reflected hinge  $h(2400 - \text{gla})$  behind, and with it the  $+\$200$  it added below the knot; and it brings in the right hinge  $h(\text{gla} - 2400)$ , whose  $-250$  starts there.  $-200 - 250 = -450$ , and the rate falls from \$950 to \$500.

Figure 4 draws the same lesson for a five-knot model, small enough to check by eye. The bars are the coefficients printed at the knots; the staircase is the rate on each segment; and the staircase starts at the first segment's rate, \$900, set by the reflected hinge, and each step is the sum of the bars at that knot. The rate on the rightmost segment,  $650 - 250 + 150 - 200 - 150 = 200$ , uses all five right hinges; the reflected hinge at 1,200, printed as  $-900$ , sets the first segment to  $+900$  and touches nothing else.

Two things about the shape of that graph are worth saying, because they are about houses rather than about MARS. First, the axis starts at 800 square feet and not at zero, and the graph starts at the axis: every graph of a fitted model in this supplement, like the plots the modeling software draws, runs from the smallest house in the data to the largest and draws nothing outside that range. Every market area has a smallest house, below which it holds no sales, and the variable graph shows nothing below it: not zero, but no value the sales support; the formula would still return a number there, and that number is extrapolation. The figures use a floor of 800 square feet, but the floor is a characteristic of the market area, not a constant: different market areas have different minimum sizes of properties that show up in the MLS, and the dataset, not the appraiser, says where it lies. Second, the shape varies from market to market: in the coastal example the steepest range is 1,200 to 2,000 square feet, and in Figure 4 it is the first segment. In a high-priced area such as Atherton or Hillsborough the steepest segment of the living area (GLA) curve may lie past 4,000 or 5,000 square feet, where the buyers are. And a neighborhood in which a builder put up houses of one size with problems in one period could show a sharp drop in the rate over an intermediate range, as the constructed graph does between 1,600 and 2,000. Over a larger market area the variable graphs are varied, and the rate table is how the appraiser reads what a particular one says. Figures 5 and 6 draw two of those shapes as earth would print them, so that a reader can build each rate table by the rule and check it against the staircase.

## Case 2: interactions

An interaction term is a coefficient times a product of factors in different variables, each factor a hinge or the variable itself, for instance  $\gamma \cdot h(\text{gla} - 1500) \cdot \text{lot}$ . Everything in Case 1 survives except

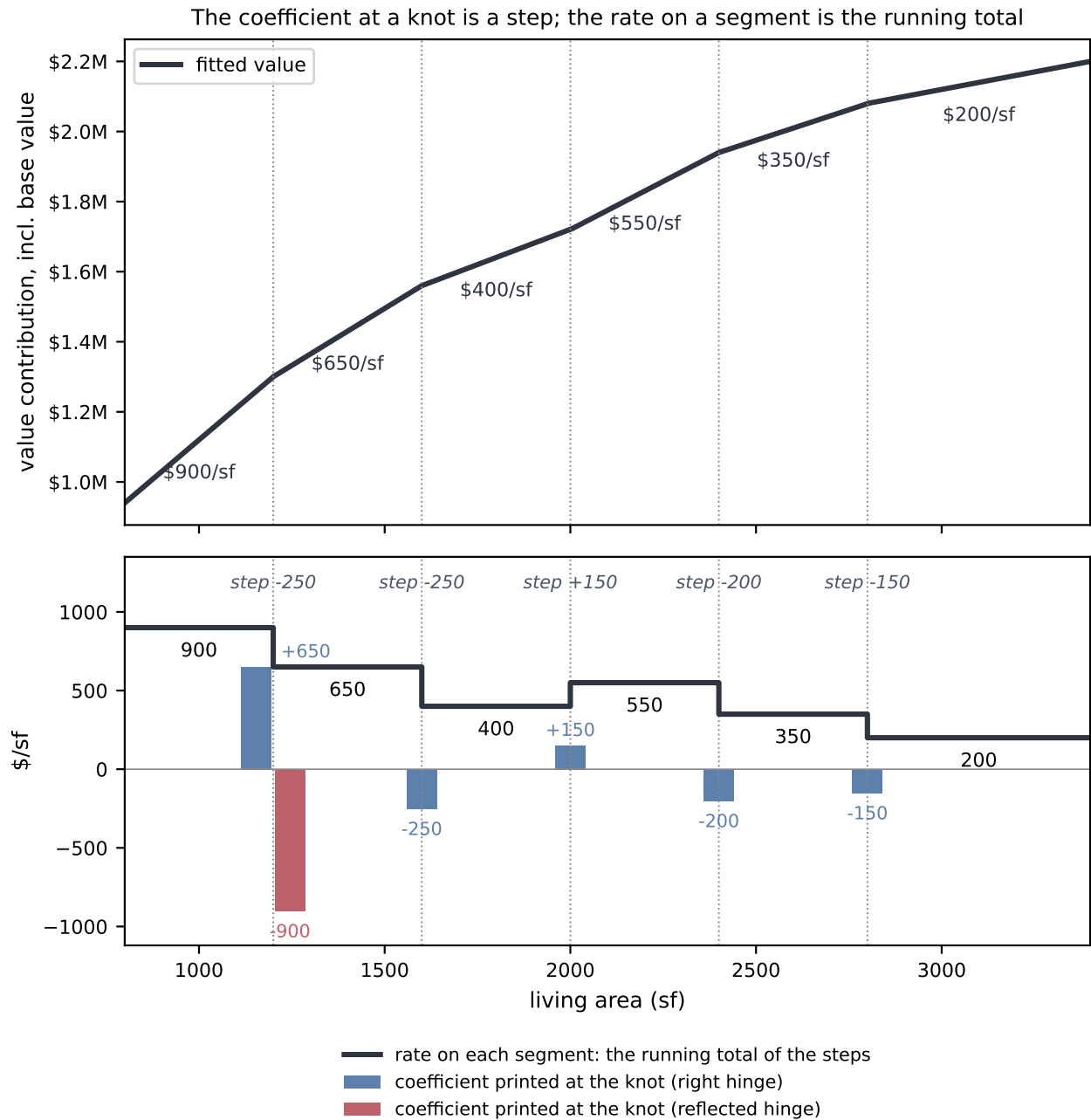


Figure 4: A five-knot model on data that start at 800 square feet:  $-900 h(1200 - \text{gla}) + 650 h(\text{gla} - 1200) - 250 h(\text{gla} - 1600) + 150 h(\text{gla} - 2000) - 200 h(\text{gla} - 2400) - 150 h(\text{gla} - 2800)$ . Above, the fitted value with the rate written on each segment; there is no value below 800 square feet because there are no such houses. Below, the coefficient printed at each knot as a bar and the rate on each segment as a staircase: each step of the staircase is the sum of the bars at that knot (one bar here, two where a knot carries a pair, Theorem 2), so the rate on a segment is the first segment's rate plus the running total of the steps to its left, not the coefficient beside it: each coefficient adjusts the rate coming in from the left.

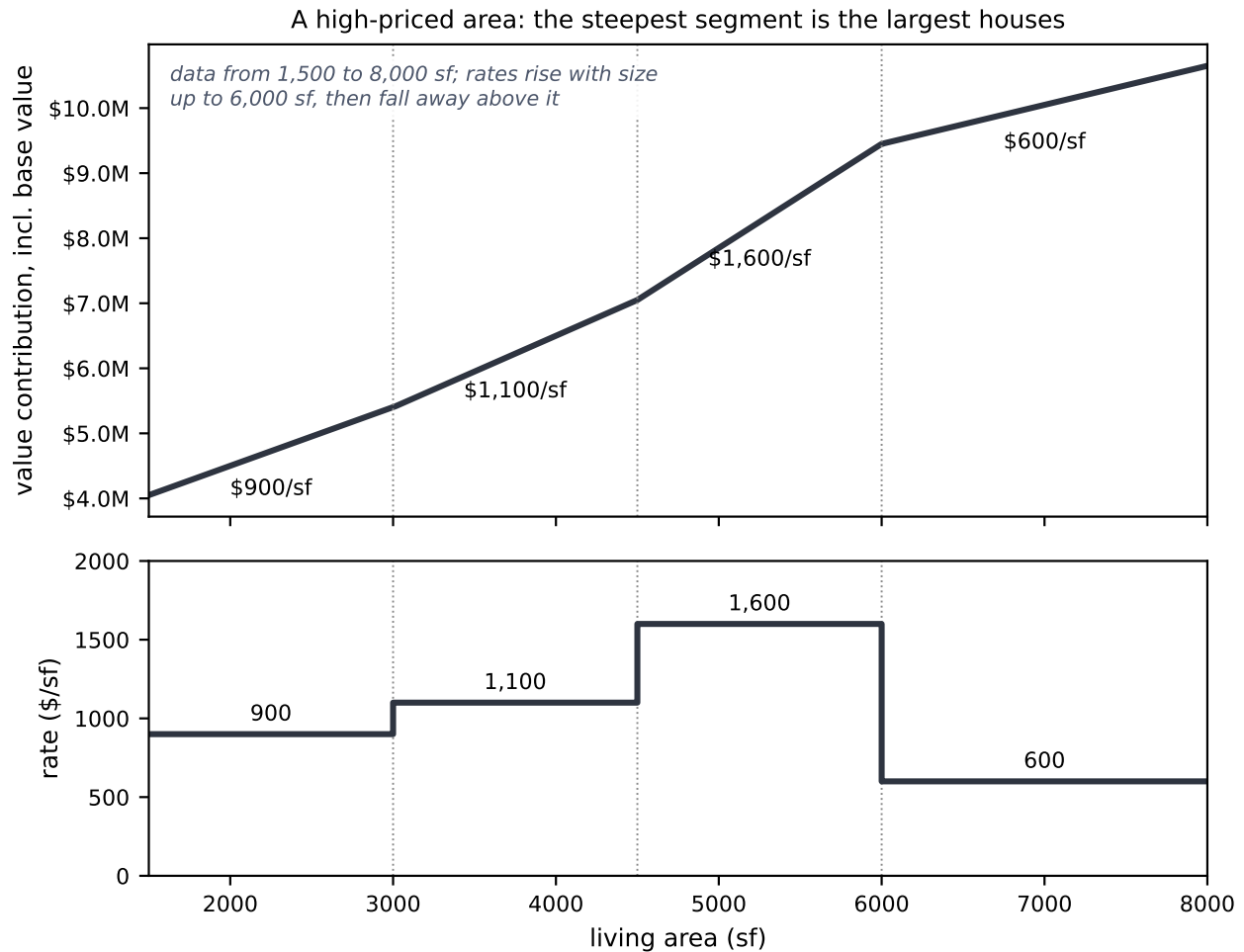


Figure 5: A high-priced area, on data from 1,500 to 8,000 square feet:  $-900h(3000 - gla) + 1100h(gla - 3000) + 500h(gla - 4500) - 1000h(gla - 6000)$ , intercept \$5,400,000 at the first knot. Rates \$900, \$1,100, \$1,600 and \$600 a square foot: the steepest segment is 4,500 to 6,000 square feet, where the buyers are, the first segment is shallower, and above 6,000 the rate falls away. The rule is unchanged; only the market is.

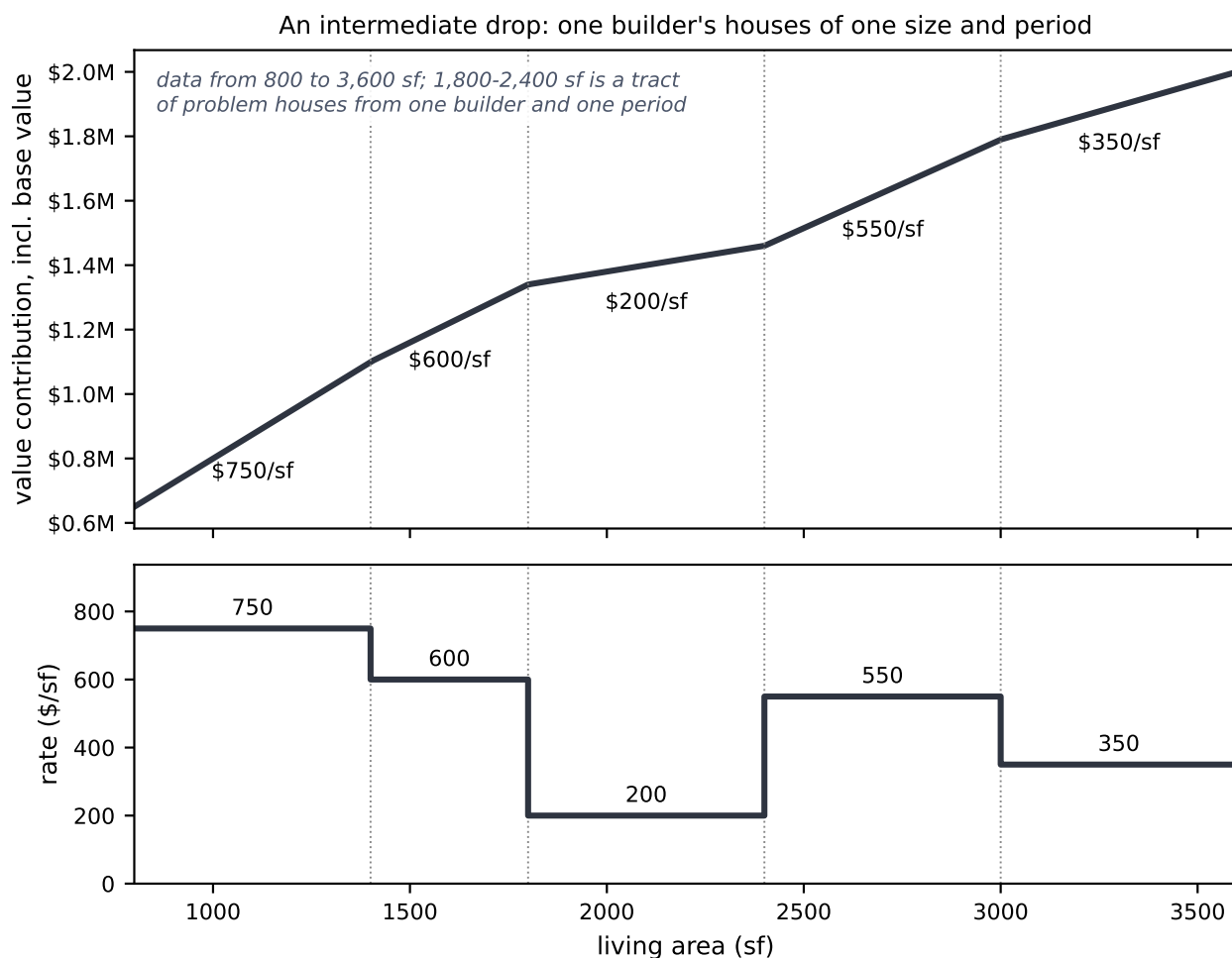


Figure 6: An intermediate drop, on data from 800 to 3,600 square feet:  $-750 h(1400 - \text{gla}) + 600 h(\text{gla} - 1400) - 400 h(\text{gla} - 1800) + 350 h(\text{gla} - 2400) - 200 h(\text{gla} - 3000)$ , intercept \$1,100,000 at the first knot. Rates \$750, \$600, \$200, \$550 and \$350 a square foot. Suppose the segment from 1,800 to 2,400 is a tract of houses of that size, built by one builder in one period, with problems the market prices (a constructed illustration); the rate falls to \$200 there and recovers above it. A reader who takes the coefficient  $-400$  at 1,800 for the rate of that segment, or who assumes rates only fall with size, gets the tract wrong.

one word: the rate on a segment is now a function of the other variables, not a constant. The statement and its proof follow, and then what it does to an appraiser's arithmetic.

**Setting.** A MARS model with interactions on the features  $x, z_1, \dots, z_q$  has fitted value

$$f(x, z) = \beta_0 + \sum_{m=1}^M \gamma_m \prod_{k \in F_m} B_{m,k}(z_k),$$

where  $M$  is the number of terms other than the intercept,  $\beta_0$  is the intercept, and the model is the intercept plus the sum of its terms, term  $m$  being  $\gamma_m \prod_{k \in F_m} B_{m,k}(z_k)$ : its coefficient  $\gamma_m$  times one factor for each feature in the term. Here  $z = (z_1, \dots, z_q)$  and, so that one index runs over every feature,  $z_0$  stands for  $x$ . The features appearing in term  $m$  form the set  $F_m \subseteq \{0, 1, \dots, q\}$ , no feature appears twice in one term, and each factor  $B_{m,k}$  is a function of the one feature  $z_k$ , of one of three kinds: the feature itself,  $z_k$ ; a right hinge  $h(z_k - t)$ ; or a reflected hinge  $h(t - z_k)$ . Let  $M_x = \{m : 0 \in F_m\}$  be the terms that contain  $x$ . For  $m \in M_x$ ,  $B_{m,0}(x)$  is the term's factor in  $x$ , and the term's other features, those  $z_k$  with  $k \in F_m$ ,  $k \neq 0$ , are the *partners* of  $x$  in that term. Write

$$P_m(z) = \prod_{k \in F_m, k \neq 0} B_{m,k}(z_k)$$

for the product of the partner factors, an empty product being 1. In words: once the subject's other features are filled in, the partners of a term multiply out to one ordinary number; that number times the printed coefficient is the coefficient of a plain living-area term, and the term reads as in Case 1. The partner product  $P_m(z)$  is what remains of term  $m$  when its factor in  $x$  is taken out: term  $m$  is  $\gamma_m B_{m,0}(x) P_m(z)$ . At a given point  $z$  it is a number, so on a segment the term acts on  $x$  as a single hinge with coefficient  $\gamma_m P_m(z)$ , and that product is how the partners enter the rate. A term with no partner has  $P_m(z) = 1$  and is a term of Case 1.

For example, in the term  $0.03 h(\text{gla} - 1200) h(\text{lot} - 6000)$  of the worked example below, with  $x$  the living area, the factor in  $x$  is  $h(\text{gla} - 1200)$ , the one partner is the lot, and  $P_m(z) = h(\text{lot} - 6000)$ . At a 5,000 square foot lot  $P_m(z) = 0$  and the term adds nothing to the rate; at a 10,000 square foot lot  $P_m(z) = 4,000$  and the term acts as a hinge at 1,200 with coefficient  $0.03 \times 4,000 = 120$ , adding \$120 a square foot to the rate above 1,200 square feet. Nothing here limits how many features a term may contain; earth's **degree** argument sets that maximum, and the theorems below hold for any degree. A term in  $d$  features has a graph in  $d + 1$  dimensions, which no one can draw once  $d$  exceeds two; the software shows only slices, one or two features at a time with the others held fixed, and Theorem 3 is the same slicing done at the subject's own point.  $P_m(z)$  is where that slicing happens: however many partners a term has, at a point their product is one number. In the four-feature term  $\gamma h(\text{gla} - 1200) h(\text{lot} - 6000) h(\text{age} - 10) \text{baths}$ , with  $\gamma = 0.001$ , whose graph has five dimensions,  $P_m(z) = h(\text{lot} - 6000) h(\text{age} - 10) \text{baths}$ ; at a 10,000 square foot lot, an age of 20 years and 3 baths it is  $4,000 \times 10 \times 3 = 120,000$ , and along living area the whole term is a hinge at 1,200 with coefficient  $\gamma \times 120,000 = 0.001 \times 120,000 = 120$ , adding \$120 a square foot: the same as the two-feature term above, from a coefficient thirty times smaller. That is why the coefficient of a high-degree term often prints as a very small number when its partners are measured in large units: it multiplies a product of several partner values. And if any partner is on the zero side of its knot, a lot under 6,000 square feet or an age under 10 years, then  $P_m(z) = 0$  and the term contributes nothing, whatever the other partners are; the higher the degree, the more ways a term can be switched off at a given property. The examples in this supplement use two-way interactions (**degree** = 2).

Let  $t_1 < \dots < t_K$  be the distinct knots of those  $B_{m,0}$  that are hinges, with  $t_0 = -\infty$ ,  $t_{K+1} = +\infty$  and segments  $S_i = (t_i, t_{i+1})$  as in Case 1. For  $m \in M_x$  and  $i \in \{0, \dots, K\}$  define

$$s_{m,i} = \begin{cases} +1 & \text{if } B_{m,0} \text{ is the feature } x \text{ itself, or a right hinge } h(x - t_j) \text{ with } t_j \leq t_i, \\ -1 & \text{if } B_{m,0} \text{ is a reflected hinge } h(t_j - x) \text{ with } t_j \geq t_{i+1}, \\ 0 & \text{otherwise.} \end{cases}$$

A *point* is a value  $z$  of the other features. Since every  $B_{m,0}$  is linear or identically zero on  $S_i$ ,  $f$  is linear in  $x$  on  $S_i$  for fixed  $z$ ; its slope there is the *rate* on  $S_i$  at the point  $z$ , written  $r_i(z) = \partial f / \partial x$  for  $x \in S_i$ .

To see every symbol at once, take the model of the worked example below as earth prints it: intercept 900,000,  $600 h(\text{gla} - 1200)$ ,  $-350 h(\text{gla} - 2400)$ ,  $0.03 h(\text{gla} - 1200) h(\text{lot} - 6000)$  and  $15 \text{ lot}$ . Its features are living area,  $x = z_0 = \text{gla}$ , and lot size,  $z_1 = \text{lot}$ , so  $q = 1$ ; the intercept is  $\beta_0 = 900,000$  and there are  $M = 4$  terms. At a property with 2,000 square feet of living area on a 10,000 square foot lot:

| $m$ | $\gamma_m$ | $F_m$      | $B_{m,0}(x)$           | $B_{m,1}(z_1)$         | $P_m(z)$               | term at the property                    |
|-----|------------|------------|------------------------|------------------------|------------------------|-----------------------------------------|
| 1   | 600        | $\{0\}$    | $h(\text{gla} - 1200)$ | none                   | 1                      | $600 \times 800 = 480,000$              |
| 2   | -350       | $\{0\}$    | $h(\text{gla} - 2400)$ | none                   | 1                      | $-350 \times 0 = 0$                     |
| 3   | 0.03       | $\{0, 1\}$ | $h(\text{gla} - 1200)$ | $h(\text{lot} - 6000)$ | $h(\text{lot} - 6000)$ | $0.03 \times 800 \times 4,000 = 96,000$ |
| 4   | 15         | $\{1\}$    | none                   | lot                    | (not in $M_x$ )        | $15 \times 10,000 = 150,000$            |

So  $M_x = \{1, 2, 3\}$ : term 4 has no factor in  $x$ , adds the same amount whatever the living area, and does not enter the rate of  $x$ . The fitted value is  $f = 900,000 + 480,000 + 0 + 96,000 + 150,000 = \$1,626,000$ . Terms 1 and 2 have no partner, so  $P_1(z) = P_2(z) = 1$ ; term 3 has one, and  $P_3(z) = h(10,000 - 6,000) = 4,000$ , the multiplier of the example above. Theorem 3, which follows, turns these into the rate. Living area has knots  $t_1 = 1,200$  and  $t_2 = 2,400$ , and the property is on the segment  $S_1$  between them, where terms 1 and 3 are active and term 2 is not ( $s_{1,1} = s_{3,1} = 1$ ,  $s_{2,1} = 0$ ); so  $r_1(z) = 600 \times 1 + 0.03 \times 4,000 = 720$ , the rate at a 10,000 square foot lot in the worked example's table.

**Theorem 3 (the rate on a segment, with interactions).** With this notation, for every  $i$  and every point  $z$ ,

$$r_i(z) = \sum_{m \in M_x} \gamma_m s_{m,i} P_m(z).$$

**Proof of Theorem 3.** Fix  $i$ , a point  $z$ , and  $x \in S_i$ . A term with  $m \notin M_x$  contains no factor in  $x$  and is constant in  $x$ ; its derivative in  $x$  is 0. A term with  $m \in M_x$  equals  $\gamma_m B_{m,0}(x) P_m(z)$ , in which  $P_m(z)$  does not depend on  $x$ . By the case analysis in the proof of Theorem 1, applied to the single factor  $B_{m,0}$ , the derivative of  $B_{m,0}$  at  $x$  is 1 for the feature itself, 1 or 0 for a right hinge according as its knot is at or left of  $t_i$  or at or right of  $t_{i+1}$ , and  $-1$  or 0 for a reflected hinge according as its knot is at or right of  $t_{i+1}$  or at or left of  $t_i$ ; that is, it equals  $s_{m,i}$ . So the derivative of the term is  $\gamma_m s_{m,i} P_m(z)$ , and the derivative of  $f$ , a finite sum, is the sum of these. The value does not depend on which  $x \in S_i$  was chosen. ■

Theorem 1 is the special case in which no term in  $M_x$  has a partner, so that every  $P_m(z) = 1$ ; grouping the terms of the sum by their knot gives the formula proved there. Three corollaries say what the partners change; the last of them is the failure of the one-per-knot bound, and the “term can vanish” case that the worked example below shows.

**Corollary 4 (the rate is a number only at a point).** In the setting of Case 2, for each  $i$  the rate  $r_i$  is a function of the point  $z$ . It is constant in  $z$  whenever every term  $m \in M_x$  with  $s_{m,i} \neq 0$  has no partner. When some term with  $s_{m,i} \neq 0$  has a partner,  $r_i(z)$  is the function  $z \mapsto \sum_m \gamma_m s_{m,i} P_m(z)$ , which changes with the partners' values except where the terms cancel exactly; two properties with the same  $x$  on the same segment then have different rates for  $x$  whenever their partner products differ and the differences do not cancel. This is why the program takes a point, and why no rate table for one feature exists on its own.

**Proof.** By Theorem 3,  $r_i(z) = \sum_{m \in M_x} \gamma_m s_{m,i} P_m(z)$ . If every term with  $s_{m,i} \neq 0$  has no partner, each such  $P_m$  is the empty product 1 and the sum is a constant. Otherwise the sum contains at least one summand  $\gamma_m s_{m,i} P_m(z)$  with  $\gamma_m s_{m,i} \neq 0$  and  $P_m$  a product of one or more factors each of which is a non-constant function of a partner, so the sum is constant in  $z$  only if those summands cancel identically. ■

**Corollary 5 (on a segment of  $x$ , the rate is a variable graph in each partner).** In the setting of Case 2 fix  $i$ , a partner  $z_p$  ( $p \geq 1$ ), and the values of all other features. Then  $z_p \mapsto r_i(z)$  is a continuous piecewise-linear function of  $z_p$  whose breaks lie among the knots of the factors  $B_{m,p}$  of the terms in  $M_x$ . Along  $x$  with every partner held fixed, the rate changes only at the knots  $t_1, \dots, t_K$  of  $x$ ; it is in the joint space of  $x$  and its partners that the rate surface is cut both by the knots of  $x$  and by the knots of every partner.

**Proof.** With the other features fixed, each  $P_m(z)$  is a constant times either 1 (if  $p \notin F_m$ ) or the single factor  $B_{m,p}(z_p)$ , which is  $z_p$  itself,  $h(z_p - t)$  or  $h(t - z_p)$ : a continuous piecewise-linear function of  $z_p$  with at most one break, at its knot. A finite linear combination of such functions, which is what Theorem 3 gives for  $r_i(z)$ , is continuous and piecewise linear with breaks only where some summand breaks. The statement along  $x$  is Theorem 3 itself: for fixed  $z$ ,  $r_i(z)$  is one number on each  $S_i$  and the segments meet only at the knots of  $x$ . ■

**Corollary 6 (the bound of Theorem 1 fails).** In the setting of Case 2 the number of terms with a non-zero summand in Theorem 3's formula for  $r_i(z)$  is bounded only by  $|M_x|$ , the number of terms containing  $x$ , and may exceed  $K$ : a knot of  $x$  may appear in several terms of  $M_x$  with different partners, each a distinct term. Conversely a term with  $s_{m,i} \neq 0$  contributes nothing at a point where its partner product is zero.

**Proof.** The terms  $\gamma_1 h(x - 1500) \text{lot}$  and  $\gamma_2 h(x - 1500) h(\text{age} - 10)$  are distinct terms of  $M_x$  with the same knot, and on the segment above 1,500 both have  $s_{m,i} = 1$ , so both contribute, and  $r_i(z) = \gamma_1 \text{lot} + \gamma_2 h(\text{age} - 10)$  has two summands from one knot; adding further partners adds further terms at the same knot without bound. For the converse, take a term whose partner is a hinge  $h(\text{lot} - 6000)$  at a point with  $\text{lot} < 6000$ : then  $P_m(z) = 0$  and the summand vanishes although  $s_{m,i} \neq 0$ ; at a point with  $\text{lot} > 6000$  it does not. ■

A practical observation from the editor's own appraisals belongs here. Once interactions are allowed, the living-area graph the software draws for the variable on its own often comes out straight, a single rate, while the bends that a model without interactions would have put into living area appear instead in the interaction terms, typically living area with sale age (the time between the sale and the effective date) or living area with bathrooms. Theorem 3 says why the straight line is not the rate. The per-variable plot in earthUI draws the terms with no partner; plotmo's degree-one plot instead draws a slice of the whole model with the other variables at their medians. The interaction terms are drawn as surfaces elsewhere (the degree-one and degree-two plots of plotmo, Milborrow 2024b, which earthUI shows side by side). At any actual property the rate for living area is the straight line's slope plus, for every active interaction term, its coefficient times its partner factor's value; it changes at the living-area knots of those terms, and by how much depends on the

partners: a sale six years ago and a sale last month, or a two-bath house and a four-bath one, sit on different rate tables for living area even though the living-area plot shows one line. A reader who takes the line for the rate has met the barrier again, one level up; the rate table at the subject's point is the way through it. Figure 7 draws one such model the way the software does, as a surface over living area and sale age, and then reads the living-area variable graph off it at two sale ages.

Figure 8 runs the same idea in the other direction, starting from the market of Figure 2. A model is built for that market with a living-area by sale-age interaction, so that the living-area main effect is one straight line, \$500 a square foot, and the four interaction terms carry the bends. The plot the software draws for living area on its own is the straight line. Read at a sale four years old, the living-area rates are Figure 2's to the dollar, \$600, \$1,000, \$750 and \$100; read at a new sale it is the straight line; read at two years it is between. The model is constructed to make the point, not fitted to sales, and a real refit would not land on Figure 2 so neatly; but the mechanism is the one a real refit uses, and it is why an appraiser who sees the living-area plot straighten after adding interactions has not lost the bends. The bends have moved into the interaction terms, where only a reading at a point recovers them. The main-effect plot of a model with interactions draws the variable's terms alone, which is the slice at a setting of the partners where every interaction term is zero (here, a new sale), not the market's shape; read the variable graph at the subject.

The last consequence is about the adjustment rather than the rate, and it is the one that reaches the sales grid.

**Theorem 4 (the split of an adjustment is order-dependent).** In the setting of Case 2 let  $y = z_p$  ( $p \geq 1$ ) be one of the other features and let  $w$  stand for the remaining ones,  $z_k$  for  $k \notin \{0, p\}$ ; write  $f(x, y, w)$  for  $f$  with its arguments so grouped. Let a comparable be at  $(x_c, y_c, w)$  and a subject at  $(x_s, y_s, w)$ , the same  $w$  for both. For a value  $y^*$  of  $y$  define the *adjustment for  $x$  with  $y$  held at  $y^*$*  as

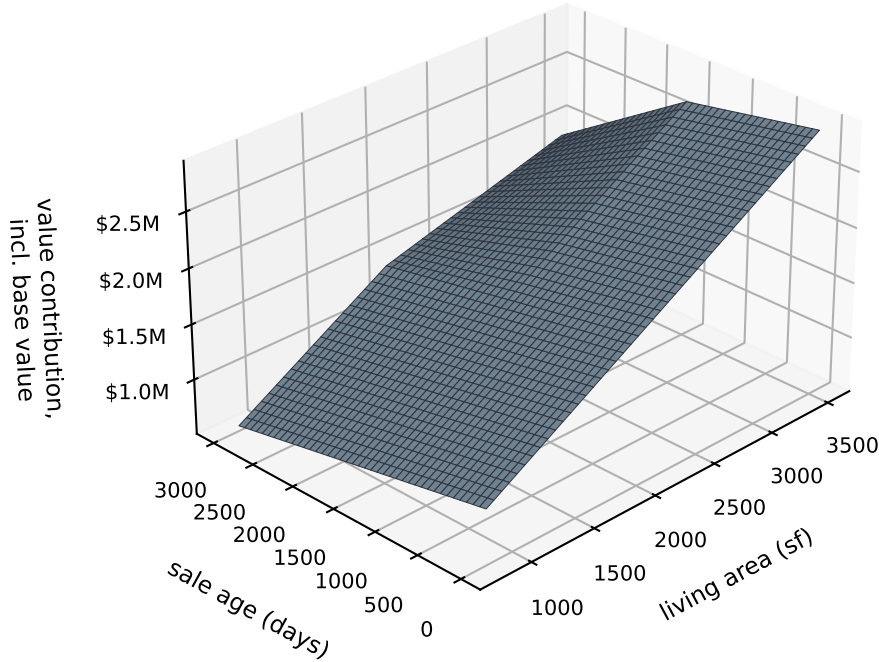
$$\Delta_x(y^*) = f(x_s, y^*, w) - f(x_c, y^*, w).$$

Let  $M_{x,y}$  be the terms containing both  $x$  and  $y$ . Each such term is its coefficient times three parts,  $\gamma_m A_m(x) C_m(y) Q_m(w)$ :  $A_m(x) = B_{m,0}(x)$  is its factor in  $x$ , which is  $x$  itself or a hinge in  $x$ ;  $C_m(y) = B_{m,p}(y)$  is its factor in  $y$ , which is  $y$  itself or a hinge in  $y$ ; and  $Q_m(w)$  is the product of its factors in the remaining features, each such a feature itself or a hinge in it, and is 1 when the term has no other factor. In the worked example below the one such term is  $0.03 h(\text{gla} - 1200) h(\text{lot} - 6000)$ , with  $A_m(x) = h(\text{gla} - 1200)$ ,  $C_m(y) = h(\text{lot} - 6000)$  and  $Q_m(w) = 1$ . Write  $\Delta_x(y_s)$  and  $\Delta_x(y_c)$  for the adjustment for  $x$  with  $y$  held at the subject's value  $y_s$  and at the comparable's value  $y_c$ , that is,  $\Delta_x(y^*)$  at  $y^* = y_s$  and at  $y^* = y_c$ ; in the worked example, with  $x$  the living area and  $y$  the lot,  $\Delta_x(y_c) = \$240,000$  at the comparable's 5,000 square foot lot and  $\Delta_x(y_s) = \$288,000$  at the subject's 10,000. Then

$$\Delta_x(y_s) - \Delta_x(y_c) = \sum_{m \in M_{x,y}} \gamma_m [A_m(x_s) - A_m(x_c)] [C_m(y_s) - C_m(y_c)] Q_m(w),$$

while the total difference  $f(x_s, y_s, w) - f(x_c, y_c, w)$  involves no choice of  $y^*$ . The statement is about this two-feature slice, one partner  $y$  with everything else held at the same  $w$  for both properties; when three or more features differ between the two, the split of the total among them is path-dependent in the same way along a longer chain of one-feature moves, and the theorem applies to each step of the chain, with the features not yet moved held at their current values. In particular, with a single such term  $\gamma A(x) C(y)$  and no factors in  $w$ , the two adjustments for  $x$  differ by  $\gamma [A(x_s) - A(x_c)] [C(y_s) - C(y_c)]$ , which is not 0 whenever the move in  $x$  changes  $A$  and the two properties differ in  $C$ . For  $A = h(x - t)$  and  $C = y$  this is  $\gamma [h(x_s - t) - h(x_c - t)] (y_s - y_c)$ ; for

Living area by sale age: the surface the software draws



The living-area variable graph, read at two points

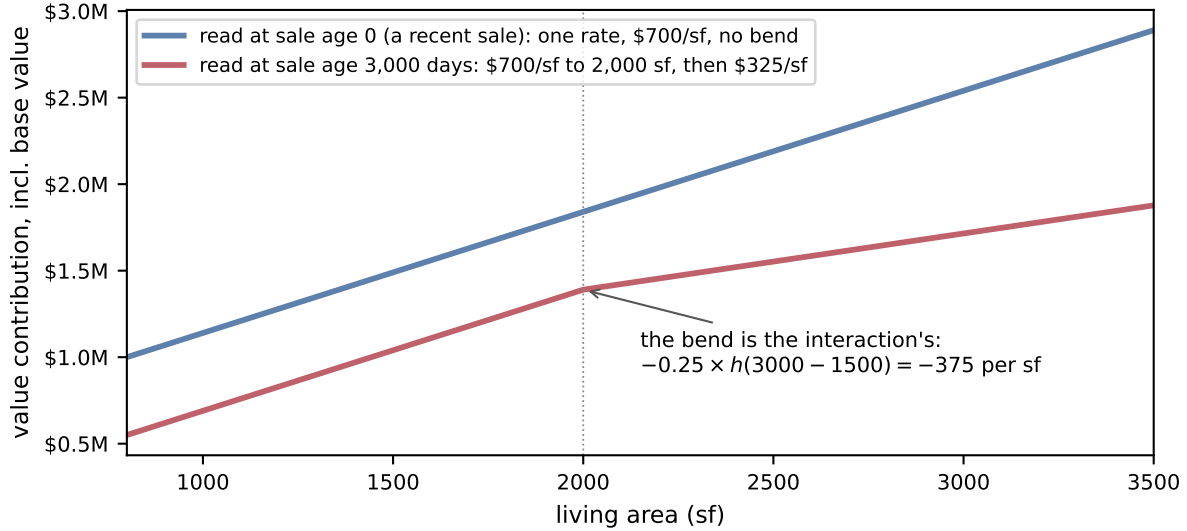


Figure 7: An interaction between living area and sale age:  $1,000,000 + 700h(\text{gla} - 800) - 150\text{saleage} - 0.25h(\text{gla} - 2000)h(\text{saleage} - 1500)$ , on data from 800 to 3,500 square feet and sales up to 3,000 days old. Above, the surface the modeling software draws for the pair. Below, the living-area variable graph read at two points, as Theorem 3 requires: at any sale up to 1,500 days old the partner factor  $h(\text{saleage} - 1500)$  is zero, the interaction contributes nothing and the graph is one straight line at \$700 a square foot; at a sale 3,000 days old the partner is 1,500, the interaction subtracts  $0.25 \times 1,500 = 375$  a square foot above 2,000 square feet, and the rate there is \$325. The living-area main effect (the terms in living area alone, with no partner) is the same straight line in both; the bend belongs to the interaction, and appears only where the partner is active.

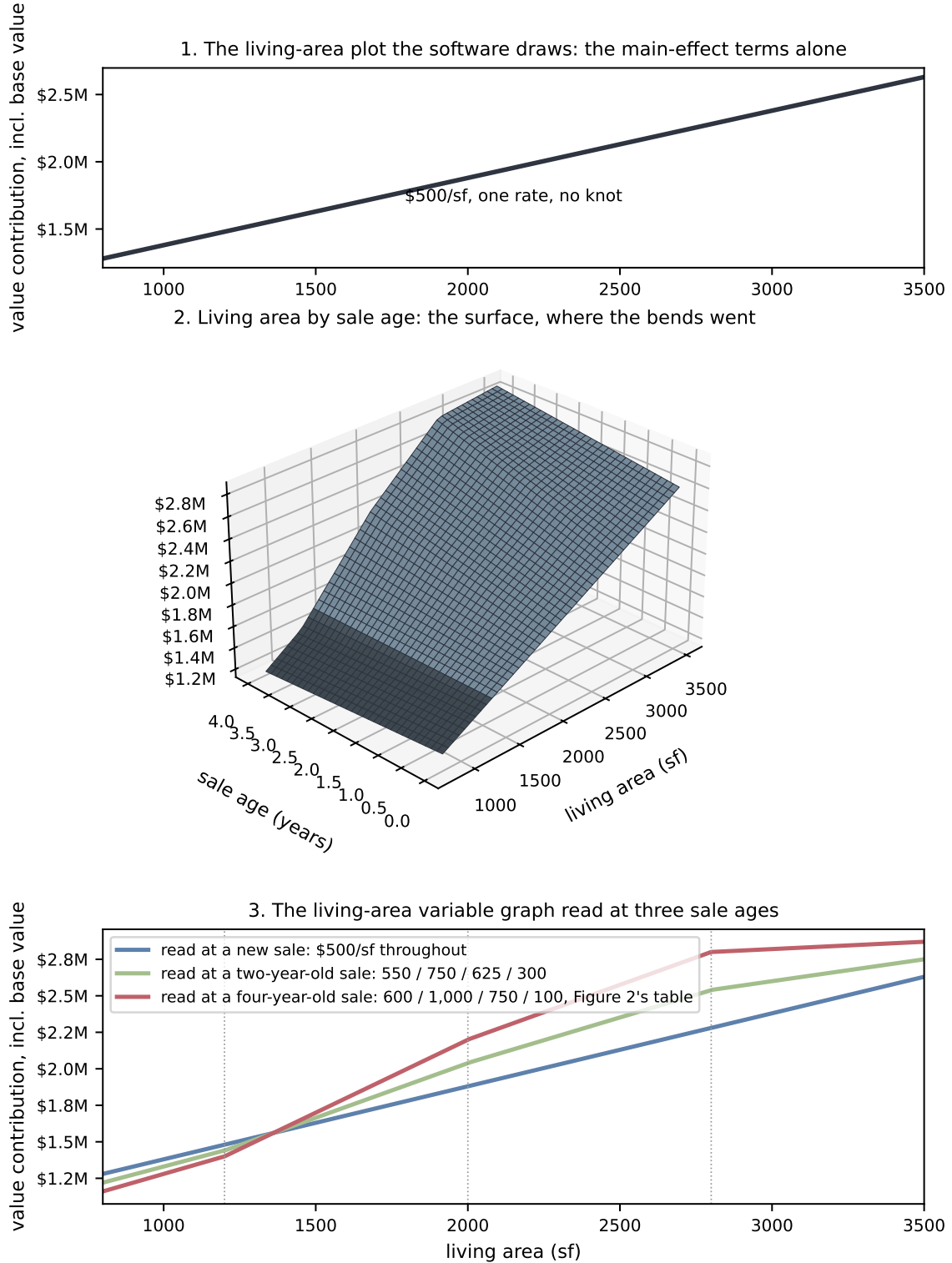


Figure 8: The market of Figure 2, with a living-area by sale-age interaction:  $1,280,000 + 500 h(\text{gla} - 800) - 20,000 \text{saleyrs} + [-25 h(1200 - \text{gla}) + 125 h(\text{gla} - 1200) - 62.5 h(\text{gla} - 2000) - 162.5 h(\text{gla} - 2800)] \times \text{saleyrs}$ , sale age in years from 0 to 4. The panels are described in their titles; at four years the bottom graph has exactly the rates of Figure 2, \$120,000 lower in level. Built for the demonstration, not fitted.

$A = h(x - t)$  and  $C = h(y - s)$ , the form of the worked example below, it is  $\gamma [h(x_s - t) - h(x_c - t)] [h(y_s - s) - h(y_c - s)]$ .

**Proof of Theorem 4.** Group the terms of  $f$  by which of  $x, y$  they contain. A term containing neither  $x$  nor  $y$  takes the same value at  $(x_s, y^*, w)$  and  $(x_c, y^*, w)$  and cancels in  $\Delta_x(y^*)$ . A term containing  $y$  but not  $x$  likewise cancels. A term containing  $x$  but not  $y$  contributes to  $\Delta_x(y^*)$  an amount that does not depend on  $y^*$ , so it cancels in  $\Delta_x(y_s) - \Delta_x(y_c)$ . A term  $m \in M_{x,y}$  contributes  $\gamma_m [A_m(x_s) - A_m(x_c)] C_m(y^*) Q_m(w)$  to  $\Delta_x(y^*)$ ; setting  $y^* = y_s$  and  $y^* = y_c$  and subtracting gives the stated sum. The total difference is a difference of two values of the one function  $f$  and mentions no  $y^*$ . ■

So the interaction's value belongs to the pair of features, not to either. A grid that adjusts one feature at a time at the comparable's values, and one that adjusts at the subject's values, agree on the total and disagree on the split. Holding the other features at the subject's values is one defensible convention and is the one the program uses; a report has to say which convention it used. A standard residential grid, such as that of the Uniform Residential Appraisal Report (URAR), has no line for an interaction. Done one feature at a time, the two moves must hold the partner at opposite ends — living area adjusted at the subject's lot, then the lot at the comparable's living area, or the reverse — so that the interaction is counted once and the lines sum to the model's total difference. With contributions computed directly, a grid can also carry the interaction as a line of its own, the difference of the interaction terms between subject and comparable, next to the lines for each feature; each line is then one group of terms, and no order has to be chosen. In the worked example below the three lines are living area \$240,000, lot \$75,000, and living area by lot \$120,000, which sum to the same \$435,000. With interactions, the pair's secant rate changes with the pair in two ways: once through the ranges crossed, and once through where the pair sits in the partners.

### A worked example with an interaction

A coastal market again, now with two features, living area and lot size, and one interaction between them. earth prints:

|                         |        |
|-------------------------|--------|
| (Intercept)             | 900000 |
| h(gla-1200)             | 600    |
| h(gla-2400)             | -350   |
| h(gla-1200)*h(lot-6000) | 0.03   |
| lot                     | 15     |

Read as Theorem 3 says. Living area has two knots, 1,200 and 2,400, and three terms in  $M_{\text{gla}}$ : the two plain hinges and the interaction, whose partner is  $h(\text{lot} - 6000)$ . On the segment from 1,200 to 2,400 the plain hinge at 1,200 is active and so is the interaction's hinge at 1,200, so the rate there is  $600 + 0.03 h(\text{lot} - 6000)$ : a variable graph in the lot, constant at 600 for lots up to 6,000 square feet and increasing by 3 cents for each further square foot of lot. Above 2,400 the hinge at 2,400 comes in and the rate is  $250 + 0.03 h(\text{lot} - 6000)$ . The lot's own rate is  $15 + 0.03 h(\text{gla} - 1200)$  above 6,000 square feet of lot and 15 below, so each feature's rate depends on the other's quantity.

Two points, two tables. The program's rate table for living area over the data, 800 to 3,500 square feet, once at a 5,000 square foot lot and once at a 10,000 square foot lot. Each column is the model's slope in living area with the lot held at that one value, and at no other:

| segment of living area (sf) | rate at lot 5,000 (\$/sf) | rate at lot 10,000 (\$/sf) |
|-----------------------------|---------------------------|----------------------------|
| 800 to 1,200                | 0                         | 0                          |
| 1,200 to 2,400              | 600                       | 720                        |
| 2,400 to 3,500              | 250                       | 370                        |

At the smaller lot the partner factor  $h(5000-6000)$  is zero, so the interaction contributes nothing and the table is the one a model without the interaction would give; that is the “term can vanish” case of Corollary 6. At the larger lot the partner is 4,000, the interaction adds  $0.03 \times 4,000 = 120$  to every segment above 1,200, and the same knot, 1,200, is feeding two active terms, which is the failure of the one-per-knot bound.

Now the adjustment. The comparable has 1,800 square feet on a 5,000 square foot lot with a fitted value of \$1,335,000; the subject has 2,200 square feet on a 10,000 square foot lot, fitted \$1,770,000. The total to explain is \$435,000. Adjust living area first, at the comparable’s lot, then the lot at the subject’s living area. Each adjustment is the difference of the two contributions it compares; where the move stays within one segment that difference is width times rate, which is how the table writes it:

|                        | living area first, lot second | lot first, living area second |
|------------------------|-------------------------------|-------------------------------|
| living-area adjustment | $400 \times 600 = 240,000$    | $400 \times 720 = 288,000$    |
| lot adjustment         | 195,000                       | 147,000                       |
| total                  | 435,000                       | 435,000                       |

The lot adjustments are differences of contributions in the same way, and the rates explain them; the lot’s own rate breaks at 6,000 square feet: at the subject’s 2,200 square feet the lot rate is \$15 below the knot and \$45 above it, so the move from 5,000 to 10,000 is  $1,000 \times 15 + 4,000 \times 45 = 195,000$ , not  $5,000 \times 45$ ; at the comparable’s 1,800 square feet it is  $1,000 \times 15 + 4,000 \times 33 = 147,000$ . The totals agree, as Theorem 4 says they must, and the splits differ by \$48,000, which is exactly the theorem’s formula:  $0.03 \times [h(2200 - 1200) - h(1800 - 1200)] \times [h(10000 - 6000) - h(5000 - 6000)] = 0.03 \times 400 \times 4,000$ . That \$48,000 is the value of having the extra 400 square feet of house *and* the extra 4,000 square feet of lot together; it belongs to neither feature, and whichever is adjusted second collects it. The program’s convention, the other features held at the subject’s values, gives the right-hand column for living area, \$288,000 at \$720 a square foot, and its report says at what point the rate was read:

Feature: gla

Rates by range (data from 800 to 3500; nothing is reported outside it):

```
800 .. 1200   rate 0.0
1200 .. 2400   rate 720.0
2400 .. 3500   rate 370.0
```

Fitted value at the knots (intercept and partners included):

```
1200   1050000.0
2400   1914000.0
```

Comparable at 1800: contribution 432000.0

Subject at 2200: contribution 720000.0

Adjustment, range by range:

```
1800 .. 2200   400 at 720.0 = 288000.0
```

Adjustment added to the comparable: 288000.0

Secant rate for the pair (one straight-line rate): 720.0  
Check: segments sum to the adjustment.

Both contributions are read at the subject's lot of 10,000 square feet; at its own 5,000-square-foot lot the comparable's living-area contribution is \$360,000.

A reviewer who reruns this at the comparable's lot gets \$240,000 and is not wrong; the reviewer and the program are reading the same surface at two points, and the report has to say which.

In short: without interactions a MARS model is a rate table per feature, and every argument about a rate made for ordinary regression transfers segment by segment. With interactions a rate is a surface, Theorem 3 says exactly how it bends, and the appraiser must say at what point it was read.

## The program

`mars_rates.pl`, printed in full in the addendum, reads earth's printed table, or a model written as a Prolog term, and produces the rate table, the values at the knots, the two contributions and their difference, the adjustment, with its range-by-range explanation, the pair's secant rate, and a check that the ranges sum to the adjustment. Given the data range, the smallest and largest house in the sales, it reports no rate outside it: a hinge's algebra runs to infinity on one side, the sales do not, and a rate beyond them is extrapolation. For the example, on data from 800 to 3,500 square feet:

```
Feature: gla
Rates by range (data from 800 to 3500; nothing is reported outside it):
    800 .. 1200   rate 600
    1200 .. 2000   rate 1000
    2000 .. 2800   rate 750
    2800 .. 3500   rate 100
Fitted value at the knots (intercept and partners included):
    1200   1520000
    2000   2320000
    2800   2920000
Comparable at 1800: contribution 600000
Subject at 2400:   contribution 1100000
Adjustment, range by range:
    1800 .. 2000   200 at 1000 = 200000
    2000 .. 2400   400 at 750 = 300000
Adjustment added to the comparable: 500000
Secant rate for the pair (one straight-line rate): 833.3333333333334
Check: segments sum to the adjustment.
```

Two kinds of number stand in that block. The contributions are the living-area terms alone, with the intercept and every other feature's terms left out: \$600,000 and \$1,100,000 differ by exactly the adjustment. The values at the knots are fitted values, intercept included, and in a model with interactions the partners' terms at the point as well; here \$1,520,000 at 1,200 square feet is the intercept because every hinge is zero there. The fitted values at the comparable and the subject, \$2,120,000 and \$2,620,000, carry the intercept the same way. The report labels the knot values as fitted values so that a reader does not take them for contributions.

That block is what belongs in a report's addendum: the rates, the knots, and the arithmetic from one to the other, in a form a reviewer can redo by hand. The program's 69 tests pin the sign rule, the interaction rule, the consistency of segments with adjustments, and the parser for earth's output, and pass under Trealla, the Journal's Prolog, and under SWI-Prolog and Sryer Prolog. For readers who work in R, the same reading is available without leaving it: `mars_rates.R`, published with the program, takes an `earth` object or its printed table and prints the same block, with the rates clamped to the data range; the Prolog remains the checked reference, and the R file carries its own self-checks against the examples of this supplement. The program, its tests and the figure script are published with this issue as a download beside the PDF, under the GNU Affero General Public License, version 3 or later.

## Why MARS, and what this supplement leaves to other entries

The Journal models with MARS for three reasons, and they are the reasons this supplement can exist. Objectivity: the knots are placed where the sales place them, by one stated procedure with one stated pruning rule (the later stage that removes terms that do not earn their place), and nothing about the shape of a variable graph is decided beforehand beyond those stated settings (the maximum degree of interaction, the pruning penalty, the minimum spacing of knots). Discoverability: the forward pass searches every feature and every candidate knot, interactions included up to the stated degree, and reports what it found, which no method that starts from a chosen shape can do. Interpretability: constant rates over ranges the data chose, and interactions built from the same hinges, are what let the model be written as a rate table and an adjustment be explained as rate times width. Two other methods also produce rates that change across a feature's range, and they are different from each other. The first, segmented ("broken-stick") regression, takes a variable and a number of breaks chosen by the analyst and estimates where the breaks fall, with a standard error (a measure of how far the estimate could move with other sales) on each; between breaks the rate is constant, as in MARS. The second, a shape-constrained additive model, fits a smooth curve instead of straight segments, forced to a shape the analyst imposes — for example, that value never falls as living area grows, or that each added square foot adds less than the one before. Neither is used here. Each answers a question this supplement does not ask, and each gives up at least one of the three properties: segmented regression gives up discoverability, because the variable and the number of breaks are chosen in advance; the shape-constrained model gives up interpretability, because a curve has no constant rates to tabulate, and objectivity, because its shape is imposed. This supplement reads the model MARS produces; the choice of method, and the stability of the knots and rates MARS finds from one set of sales to the next, are subjects for entries of their own.

## What this does not settle

The rates are the model's, and the four statistical entries apply to them unchanged: a rate approximates the conditional expectation only as well as the model's form allows (*Conditional Expectation*); it can carry an omitted feature's value (*Omitted Variable Bias*); it can be distorted when a feature moves with the structural error (*Endogeneity*); and it may not be pinned down by the sales at all, however many there are, when two features always move together (*Parameter Identification*). MARS finds where a rate changes only where the sales contain the contrast to show it; a knot is a claim about the market and is only as good as the sales on either side of it. What this supplement adds is narrower: given the model, here is how to say what it says.

## References.

- Craytor, W. B. (2026). earthUI: An Interactive Interface for the earth (MARS) Package. *Valuation Engineer Journal*, 1(1). <https://doi.org/10.67330/7fj41b65>
- Friedman, J. H. (1991). Multivariate Adaptive Regression Splines. *The Annals of Statistics*, 19(1), 1–67. <https://doi.org/10.1214/aos/1176347963>
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (2nd ed.), §9.4, Multivariate Adaptive Regression Splines. Springer.
- Milborrow, S. (2024a). earth: Multivariate Adaptive Regression Splines. R package. <https://CRAN.R-project.org/package=earth>
- Milborrow, S. (2024b). plotmo: Plot a Model's Residuals, Response, and Partial Dependence Plots. R package. <https://CRAN.R-project.org/package=plotmo>
- Milborrow, S. (n.d.). Notes on the earth Package. <http://www.milbo.org/doc/earth-notes.pdf>

---

*Cross-references:* conditional expectation; omitted variable bias; endogeneity; parameter identification; hedonic regression; implicit price.

---

## Addendum: Prolog Formalization

The following machine-readable companion (`mars_rates.pl`) formalizes this entry in the journal's Prolog knowledge base. It is also published as a browsable galley alongside this article.

```
% SPDX-License-Identifier: AGPL-3.0-or-later
% Copyright (C) 2026 William Bert Craytor. Published by The Valuation Engineer
% Journal (Pacific Vista Net). This program is free software: you can
% redistribute it and/or modify it under the terms of the GNU Affero General
% Public License as published by the Free Software Foundation, either version 3
% of the License, or (at your option) any later version. It is distributed
% WITHOUT ANY WARRANTY, without even the implied warranty of MERCHANTABILITY or
% FITNESS FOR A PARTICULAR PURPOSE; see the COPYING file beside it or
% <https://www.gnu.org/licenses/agpl-3.0.html>.

/* mars_rates.pl -- mars-rates: from a fitted MARS model to the rates, value
   contributions and adjustments an appraiser reports.

   A MARS model, as the R package earth prints it, is an intercept plus a sum
   of terms, each a coefficient times a product of basis functions:

       h(x - t)    a hinge, zero for x below the knot t, rising with slope +1 above
       h(t - x)    a reflected hinge, zero above t, rising to the LEFT: slope -1 in x
       x           a linear term
       h(x-t)*h(y-s) an interaction, the product of two factors

   The appraiser does not report basis functions. The appraiser reports, for a
   feature, the RATE the market pays per unit in each range (the slope), the
   VALUE CONTRIBUTION of the quantity a property has (the level), and the
   ADJUSTMENT: the amount added to a comparable's contribution to make it
   equal to the subject's. This program turns the one into the other, and
   checks that they agree.

   The sign rule that hand-built tables get wrong: in a range, the rate is the
   sum of the coefficients of the active right hinges MINUS the sum of the
   coefficients of the active reflected hinges (plus any linear term's
   coefficient). A negative coefficient on h(5 - x) contributes a POSITIVE
   slope below 5.
```

Representation. A model is

```
model(Intercept, Terms)
```

where each term is `term(Coef, Factors)` and each factor is

```
hinge(Var, Knot)      h(Var - Knot)
```

```
rhinge(Var, Knot)     h(Knot - Var)
```

```
linear(Var)           Var
```

A point is a list of `Var = Value` pairs, e.g. `[gla = 1800, lot = 6000]`.

Entry points:

```
value(+Model, +Point, -Y)           the fitted value at a point
contribution(+Model, +Var, +Point, -C) the part of the value that the
                                     terms involving Var contribute
rate(+Model, +Var, +Point, -R)       slope in Var at the point (other
                                     variables held at the point)
knots(+Model, +Var, -Knots)          the sorted knots on Var
rate_table(+Model, +Var, +Point, -Rows)
                                     rows range(Lo, Hi, Rate), Lo/Hi
                                     numbers or the atoms minus_inf,
                                     plus_inf; other variables fixed
                                     at the point
rate_table(+Model, +Var, +Point, +data(Lo, Hi), -Rows)
                                     the same table clamped to the
                                     data range: the first row starts
                                     at Lo, the last ends at Hi, rows
                                     outside the data are dropped.
                                     Rates outside the data are
                                     extrapolation and are not
                                     reported
knot_values(+Model, +Var, +Point, -Pairs)
                                     Knot-Value at each knot
adjustment(+Model, +Var, +Point, +CompQ, +SubjQ, -A)
                                     amount added to the comparable's
                                     contribution (Var = CompQ) to
                                     equal the subject's (Var = SubjQ)
segments(+Model, +Var, +Point, +CompQ, +SubjQ, -Segs)
                                     the adjustment as the appraiser
                                     explains it: seg(Lo, Hi, Rate,
                                     Width, Amount) per range crossed
pair_rate(+Model, +Var, +Point, +CompQ, +SubjQ, -R)
                                     the pair's secant rate: the one rate a straight line
↪ would
                                     need between the two quantities:
                                     adjustment over width
consistent(+Model, +Var, +Point, +CompQ, +SubjQ)
                                     the segments sum to the adjustment
parse_earth(+Atom, -Model)          read earth's printed coefficient
                                     table (see parse_earth below)
print_report(+Model, +Var, +Point, +CompQ, +SubjQ)
```

Plain ISO Prolog: runs under Trealla, SWI-Prolog and Scryer Prolog. No modules, no `format/2`, no double-quoted strings. Consult this file, then `tests/harness.pl` and `tests/test_mars_rates.pl`, and call `run_tests`. \*/

```
/* ===== */
/* 1. Evaluation */
/* ===== */
```

```
point_value(Var, Point, X) :- mr_member(Var = X, Point), !.
point_value(Var, _, _) :- throw(error(missing_variable(Var), point_value/3)).
```

```

factor_value(hinge(Var, T), Point, V) :- point_value(Var, Point, X), D is X - T, V is max(0, D).
factor_value(rhinge(Var, T), Point, V) :- point_value(Var, Point, X), D is T - X, V is max(0, D).
factor_value(linear(Var), Point, V) :- point_value(Var, Point, V).

term_value(term(Coef, Factors), Point, V) :-
    factors_product(Factors, Point, P), V is Coef * P.

factors_product([], _, 1).
factors_product([F | Fs], Point, P) :- factor_value(F, Point, V), factors_product(Fs, Point, P0), P
    ↪ is V * P0.

value(model(Intercept, Terms), Point, Y) :-
    terms_sum(Terms, Point, S), Y is Intercept + S.

terms_sum([], _, 0).
terms_sum([T | Ts], Point, S) :- term_value(T, Point, V), terms_sum(Ts, Point, S0), S is S0 + V.

/* The contribution of a variable at a point: the terms that involve it. The
   intercept and the terms in other variables are the level the feature is
   measured against; an adjustment never depends on them. */

involves(term(_, Factors), Var) :- mr_member(F, Factors), factor_var(F, Var), !.
factor_var(hinge(V, _), V). factor_var(rhinge(V, _), V). factor_var(linear(V), V).

contribution(model(_, Terms), Var, Point, C) :-
    findall(V, ( mr_member(T, Terms), involves(T, Var), term_value(T, Point, V) ), Vs),
    mr_sum(Vs, C).

/* ===== */
/* 2. Rates (slopes) */
/* ===== */

/* d/dVar of a term = Coef * (slope of the factor in Var) * (product of the
   other factors' values). A hinge is active strictly above its knot, a
   reflected hinge strictly below; at the knot itself the slope is taken from
   the right, which is the convention of the rate table below. */

factor_slope(hinge(Var, T), Var, Point, S) :- !, point_value(Var, Point, X), ( X >= T -> S = 1 ; S
    ↪ = 0 ).
factor_slope(rhinge(Var, T), Var, Point, S) :- !, point_value(Var, Point, X), ( X < T -> S = -1 ; S
    ↪ = 0 ).
factor_slope(linear(Var), Var, _, 1) :- !.
factor_slope(_, _, _, 0).

term_slope(term(Coef, Factors), Var, Point, S) :-
    term_slope_(Factors, Var, Point, [], Coef, S).

term_slope_([], _, _, _, 0).
term_slope_([F | Fs], Var, Point, Seen, Coef, S) :-
    ( factor_var(F, Var) ->
        factor_slope(F, Var, Point, S1),
        mr_append(Seen, Fs, Others), factors_product(Others, Point, P),
        S1 is Coef * S1 * P
    ; S1 = 0 ),
    term_slope_(Fs, Var, Point, [F | Seen], Coef, S0),
    S is S0 + S1.

rate(model(_, Terms), Var, Point, R) :-
    findall(S, ( mr_member(T, Terms), term_slope(T, Var, Point, S) ), Ss),
    mr_sum(Ss, R).

```

```

/* ===== */
/* 3. Knots, ranges and the rate table */
/* ===== */

knots(model(_, Terms), Var, Knots) :-
    findall(K, ( mr_member(term(_, Fs), Terms), mr_member(F, Fs), knot_of(F, Var, K) ), Ks),
    sort(Ks, Knots).

knot_of(hinge(Var, K), Var, K).
knot_of(rhinge(Var, K), Var, K).

/* One row per range between consecutive knots, the rate measured at a point
   inside the range (the midpoint, or one unit past the last knot). */

rate_table(Model, Var, Point, Rows) :-
    knots(Model, Var, Knots),
    bounds(Knots, Bounds),
    rows(Bounds, Model, Var, Point, Rows).

bounds([], [minus_inf-plus_inf]).
bounds([K | Ks], [minus_inf-K | Rest]) :- bounds_(K, Ks, Rest).
bounds_(K, [], [K-plus_inf]).
bounds_(K, [K2 | Ks], [K-K2 | Rest]) :- bounds_(K2, Ks, Rest).

/* The table over the observed data only. A hinge's algebra runs to infinity
   on one side; the sales do not. Rates beyond the smallest and largest
   observed quantity are extrapolation, so the report clamps the end rows to
   the data range and drops any row that lies wholly outside it. */

rate_table(Model, Var, Point, data(DLo, DHi), Rows) :-
    rate_table(Model, Var, Point, Rows0),
    clamp_rows(Rows0, DLo, DHi, Rows).

clamp_rows([], _, _, []).
clamp_rows([range(Lo, Hi, R) | Rs], DLo, DHi, Out) :-
    clamp_lo(Lo, DLo, Lo1), clamp_hi(Hi, DHi, Hi1),
    ( Lo1 < Hi1 -> Out = [range(Lo1, Hi1, R) | Rest] ; Out = Rest ),
    clamp_rows(Rs, DLo, DHi, Rest).

clamp_lo(minus_inf, DLo, DLo) :- !.
clamp_lo(Lo, DLo, X) :- ( Lo < DLo -> X = DLo ; X = Lo ).
clamp_hi(plus_inf, DHi, DHi) :- !.
clamp_hi(Hi, DHi, X) :- ( Hi > DHi -> X = DHi ; X = Hi ).

rows([], _, _, _, []).
rows([Lo-Hi | Bs], Model, Var, Point, [range(Lo, Hi, R) | Rows]) :-
    probe(Lo, Hi, X), set_point(Var, X, Point, P), rate(Model, Var, P, R),
    rows(Bs, Model, Var, Point, Rows).

probe(minus_inf, plus_inf, 0) :- !.
probe(minus_inf, Hi, X) :- !, X is Hi - 1.
probe(Lo, plus_inf, X) :- !, X is Lo + 1.
probe(Lo, Hi, X) :- X is (Lo + Hi) / 2.

set_point(Var, X, Point, [Var = X | Rest]) :- mr_delete_var(Point, Var, Rest).

knot_values(Model, Var, Point, Pairs) :-
    knots(Model, Var, Knots),
    findall(K-Y, ( mr_member(K, Knots), set_point(Var, K, Point, P), value(Model, P, Y) ), Pairs).

/* ===== */

```

```

/* 4. Adjustments and their explanation */
/* ===== */

adjustment(Model, Var, Point, CompQ, SubjQ, A) :-
    set_point(Var, CompQ, Point, PC), set_point(Var, SubjQ, Point, PS),
    value(Model, PC, YC), value(Model, PS, YS), A is YS - YC.

/* The path from the comparable's quantity to the subject's, cut at every knot
   between them; each piece is a width at a rate. Going down (subject smaller
   than comparable) the widths are negative, so the amounts are, and the sum
   is still the adjustment as defined: added to the comparable. */

segments(Model, Var, Point, CompQ, SubjQ, Segs) :-
    knots(Model, Var, Knots),
    Lo is min(CompQ, SubjQ), Hi is max(CompQ, SubjQ),
    findall(K, ( mr_member(K, Knots), K > Lo, K < Hi ), Inner),
    mr_append([Lo | Inner], [Hi], Cuts),
    ( SubjQ >= CompQ -> Dir = 1 ; Dir = -1 ),
    pieces(Cuts, Model, Var, Point, Dir, Segs).

pieces([], _, _, _, [], []) :- !.
pieces([A, B | Cs], Model, Var, Point, Dir, [seg(A, B, R, W, Amt) | Segs]) :-
    probe(A, B, X), set_point(Var, X, Point, P), rate(Model, Var, P, R),
    W is Dir * (B - A), Amt is R * W,
    pieces([B | Cs], Model, Var, Point, Dir, Segs).

/* The pair's secant rate. Between two quantities the only single rate
   consistent with the MARS contributions is the chord, adjustment over
   width: the model's rates averaged by the width of each range crossed. An
   ordinary regression, one rate for the feature, must serve every pair with
   one number and so cannot reproduce these. Printed beside the
   MARS rates it shows what a single rate would have had to be for THIS
   comparable and THIS subject, and how far each range's rate departs from
   it. Undefined when the two quantities are equal. */

pair_rate(Model, Var, Point, CompQ, SubjQ, R) :-
    CompQ =\= SubjQ,
    adjustment(Model, Var, Point, CompQ, SubjQ, A), R is A / (SubjQ - CompQ).

consistent(Model, Var, Point, CompQ, SubjQ) :-
    adjustment(Model, Var, Point, CompQ, SubjQ, A),
    segments(Model, Var, Point, CompQ, SubjQ, Segs),
    findall(Amt, mr_member(seg(_, _, _, Amt), Segs), Amts), mr_sum(Amts, S),
    abs(S - A) < 0.000001.

/* ===== */
/* 5. Reading earth's printed coefficients */
/* ===== */

/* earth prints its model as a table, one term per line, e.g.

    (Intercept)      10
    h(x-2)           3
    h(5-x)          -1.5
    h(x-8)           0.8
    h(x-2)*h(y-3)    0.4
    x                1.2

parse_earth/2 takes that table as one atom (lines separated by newline
characters) and returns the model. Whitespace is free; a coefficient is a
decimal number with an optional sign; variable names are letters, digits,

```

```

dots and underscores, as R prints them.                                */

parse_earth(Atom, model(Intercept, Terms)) :-
  atom_chars(Atom, Chars), split_lines(Chars, Lines),
  parse_lines(Lines, 0, Intercept, Terms).

split_lines(Chars, Lines) :- split_lines_(Chars, [], Lines).
split_lines_([], Acc, [Line]) :- mr_reverse(Acc, Line).
split_lines_(['\n' | Cs], Acc, [Line | Lines]) :- !, mr_reverse(Acc, Line), split_lines_(Cs, [],
  ↪ Lines).
split_lines_([C | Cs], Acc, Lines) :- split_lines_(Cs, [C | Acc], Lines).

parse_lines([], I, I, []).
parse_lines([L | Ls], IO, I, Terms) :-
  strip(L, S),
  ( S == [] -> parse_lines(Ls, IO, I, Terms)
  ; split_ws(S, Name, CoefChars), number_chars_safe(CoefChars, Coef),
    ( Name = ['(', 'I', 'n', 't', 'e', 'r', 'c', 'e', 'p', 't', ')'] ->
      parse_lines(Ls, Coef, I, Terms)
    ; parse_factors(Name, Factors),
      Terms = [term(Coef, Factors) | Rest], parse_lines(Ls, IO, I, Rest) ) ).

strip(L, S) :- drop_ws(L, L1), mr_reverse(L1, R1), drop_ws(R1, R2), mr_reverse(R2, S).
drop_ws([C | Cs], S) :- ws(C), !, drop_ws(Cs, S).
drop_ws(S, S).
ws(' '). ws('\t'). ws('\r').

/* the coefficient is the LAST whitespace-separated token; the name is what
   precedes it and may itself contain spaces (earth prints h(x- -3) for a
   negative knot)                                                         */
split_ws(S, Name, Coef) :-
  mr_reverse(S, R), split_ws_(R, [], Coef, NameR),    % consuming the reversed line puts the
  ↪ coefficient back in order
  mr_reverse(NameR, Name0), strip(Name0, Name).
split_ws_([C | Cs], Acc, Coef, Rest) :- ws(C), !, Coef = Acc, Rest = Cs.
split_ws_([C | Cs], Acc, Coef, Rest) :- split_ws_(Cs, [C | Acc], Coef, Rest).
split_ws_([], Acc, Acc, []).

number_chars_safe(Chars, N) :-
  catch(number_chars(N, Chars), _, fail), !.
number_chars_safe(Chars, N) :- % "-.5" and ".5", which earth can print
  ( Chars = ['- ', '.' | R] -> number_chars(N, ['- ', '0', '.' | R])
  ; Chars = ['.' | R] -> number_chars(N, ['0', '.' | R]) ).

/* factors are separated by '*' */
parse_factors(Chars, [F | Fs]) :-
  split_star(Chars, First, Rest), parse_factor(First, F),
  ( Rest == [] -> Fs = [] ; parse_factors(Rest, Fs) ).

split_star(Chars, First, Rest) :- split_star_(Chars, [], First, Rest).
split_star_(['*' | Cs], Acc, First, Cs) :- !, mr_reverse(Acc, First).
split_star_([C | Cs], Acc, First, Rest) :- split_star_(Cs, [C | Acc], First, Rest).
split_star_([], Acc, First, []) :- mr_reverse(Acc, First).

/* h(x-2) / h(2-x) / x -- inside the parentheses, the side that is a number
   is the knot and the other side the variable                           */
parse_factor(['h', '(' | Cs], F) :- !,
  mr_reverse(Cs, [' ' | RevInner]), mr_reverse(RevInner, Inner),
  split_minus(Inner, Left, Right),
  ( number_chars_safe(Right, K) -> atom_chars(Var, Left), F = hinge(Var, K)
  ; number_chars_safe(Left, K) -> atom_chars(Var, Right), F = rhinge(Var, K) ).

```

```

parse_factor(Cs, linear(Var)) :- atom_chars(Var, Cs).

/* the minus that separates the two sides is the last one not at position 0
   and not following an 'e' (so "x-2", "5-x", "gla-1500" and "h(x- -3)" all
   read); a leading '-' belongs to a negative knot */
split_minus(Chars, Left, Right) :-
    mr_length(Chars, N), N1 is N - 1,
    minus_positions(Chars, 0, Ps), mr_member(P, Ps), P > 0, P < N1,
    mr_take(P, Chars, Left0), strip(Left0, Left),
    P2 is P + 1, mr_drop(P2, Chars, Right0), strip(Right0, Right),
    Left \== [], Right \== [], !.

minus_positions([], _, []).
minus_positions(['-' | Cs], I, [I | Ps]) :- !, I1 is I + 1, minus_positions(Cs, I1, Ps).
minus_positions([_ | Cs], I, Ps) :- I1 is I + 1, minus_positions(Cs, I1, Ps).

/* ===== */
/* 6. A report */
/* ===== */

print_report(Model, Var, Point, CompQ, SubjQ) :-
    write('Feature: '), write(Var), nl,
    write('Rates by range:'), nl,
    rate_table(Model, Var, Point, Rows),
    print_report_(Model, Var, Point, CompQ, SubjQ, Rows).

/* The same report over the observed data: print_report/6 takes data(Lo, Hi),
   the smallest and largest observed quantity of Var, and reports no rate
   outside it. */

print_report(Model, Var, Point, data(DLo, DHi), CompQ, SubjQ) :-
    write('Feature: '), write(Var), nl,
    write('Rates by range (data from '), write(DLo), write(' to '), write(DHi), write('; nothing is
    ↪ reported outside it):'), nl,
    rate_table(Model, Var, Point, data(DLo, DHi), Rows),
    print_report_(Model, Var, Point, CompQ, SubjQ, Rows).

print_report_(Model, Var, Point, CompQ, SubjQ, Rows) :-
    ( mr_member(range(Lo, Hi, R), Rows), write(' '), write(Lo), write(' .. '), write(Hi), write('
    ↪ rate '), write(R), nl, fail ; true ),
    write('Fitted value at the knots (intercept and partners included):'), nl,
    knot_values(Model, Var, Point, KVs),
    ( mr_member(K-Y, KVs), write(' '), write(K), write(' '), write(Y), nl, fail ; true ),
    set_point(Var, CompQ, Point, PC), set_point(Var, SubjQ, Point, PS),
    contribution(Model, Var, PC, CC), contribution(Model, Var, PS, CS),
    write('Comparable at '), write(CompQ), write(': contribution '), write(CC), nl,
    write('Subject at '), write(SubjQ), write(': contribution '), write(CS), nl,
    segments(Model, Var, Point, CompQ, SubjQ, Segs),
    write('Adjustment, range by range:'), nl,
    ( mr_member(seg(A, B, R2, W, Amt), Segs), write(' '), write(A), write(' .. '), write(B), write
    ↪ (' '), write(W), write(' at '), write(R2), write(' = '), write(Amt), nl, fail ; true ),
    adjustment(Model, Var, Point, CompQ, SubjQ, Adj),
    write('Adjustment added to the comparable: '), write(Adj), nl,
    ( pair_rate(Model, Var, Point, CompQ, SubjQ, OR) ->
        write('Secant rate for the pair (one straight-line rate): '), write(OR), nl ; true ),
    ( consistent(Model, Var, Point, CompQ, SubjQ) -> write('Check: segments sum to the adjustment
    ↪ .') ; write('CHECK FAILED') ), nl.

/* ===== */
/* 7. List helpers (own names; no library assumed) */
/* ===== */

```

```

mr_member(X, [X | _]).
mr_member(X, [_ | T]) :- mr_member(X, T).

mr_append([], L, L).
mr_append([H | T], L, [H | R]) :- mr_append(T, L, R).

mr_reverse(L, R) :- mr_reverse_(L, [], R).
mr_reverse_([], R, R).
mr_reverse_([H | T], A, R) :- mr_reverse_(T, [H | A], R).

mr_sum([], 0).
mr_sum([X | Xs], S) :- mr_sum(Xs, S0), S is S0 + X.

mr_length([], 0).
mr_length([_ | T], N) :- mr_length(T, N0), N is N0 + 1.

mr_take(0, _, []) :- !.
mr_take(N, [H | T], [H | R]) :- N1 is N - 1, mr_take(N1, T, R).

mr_drop(0, L, L) :- !.
mr_drop(N, [_ | T], R) :- N1 is N - 1, mr_drop(N1, T, R).

mr_delete_var([], _, []).
mr_delete_var([V = _ | T], V, R) :- !, mr_delete_var(T, V, R).
mr_delete_var([H | T], V, [H | R]) :- mr_delete_var(T, V, R).

```