

# Engineering Complexity and Community Friction

## Joint Predictors of High-Speed Rail Cost

*A multivariate reference-class analysis of 69 international high-speed lines, scored blind to cost, with cost predictions for the ALTO corridor. Revised edition: project costs from the NYU Transit Costs Project data used by the Parliamentary Budget Officer.*

---

### KEY FINDINGS

Scored blind to cost on 69 high-speed lines, engineering complexity and community friction both predict HSR cost per kilometre: **together they explain 42 % of cross-project variance in log-cost, and both are statistically significant (ECIw  $p < 0.001$ ; CFI  $p = 0.011$ )**. Engineering is the stronger: alone it explains 36 %, friction 10 %.

Friction adds to PBO's own cost model: **added to PBO's variables (tunnel and elevated shares, start year, region), each 10 CFI points is associated with about +13 % cost ( $p = 0.002$ )**. The effect holds without HS2 and California, without China and Japan, and within Europe.

ALTO reference-class prediction (CRI multivariate model): **\$191 M / km at the projected CFI of 65, or \$163 billion on PBO's 850 km route. The probability of delivery at or below the Government's \$90 billion is about 17 %**. PBO's model (CRI replicated) gives \$73 billion at European cost efficiency; with friction added it gives \$115 billion.

Uncertainty is narrower than in the 16-project analysis: **the 95 % prediction interval spans a factor of 11 in the bivariate model and 7 with PBO's variables, against 19 on 16 projects**.

Friction is the part of ALTO's cost still in play: **if ALTO's friction stays at today's 43 rather than rising to the projected 65, the CRI multivariate prediction falls from \$163 to \$129 billion — about \$34 billion**.

**Research Report · October 2026 (revises April 2026 edition)**

[altohsrcitizenresearch.ca](http://altohsrcitizenresearch.ca) · [citizenresearch.ca](http://citizenresearch.ca)

# Executive Summary

---

This report presents a joint multivariate analysis of engineering complexity and community friction as predictors of high-speed rail (HSR) construction cost per kilometre, and applies the resulting models to the ALTO corridor. It is a revised edition of the Initiative's April 2026 report of the same name.

## About this edition

This edition replaces both. Costs come from the NYU Transit Costs Project high-speed rail dataset, the source the Parliamentary Budget Officer (PBO) used in its October 2026 cost analysis (RP-2627-009-C), converted to 2024 Canadian dollars on PBO's basis. The reference class is expanded to the 69 lines in that dataset with complete engineering data, plus Rail Baltica. Every line, including the original 16, has been scored afresh on both indices under a protocol that excluded cost information (section 1.4). The comparison with the April analysis is in Appendix D.

Two analytical instruments developed by the ALTO HSR Citizen Research Initiative provide the input variables. The Engineering Complexity Index (ECI) scores HSR projects on six physical engineering dimensions — geotechnical conditions, major structures, topography and alignment, hydrology and drainage, infrastructure interfaces, and climate and environmental loading — yielding a weighted composite score (ECI<sub>w</sub>) on a 20-to-100 scale. The Community Friction Index (CFI) scores projects on five sociopolitical components — public opposition intensity, legal challenges filed, forced route changes, environmental and NGO opposition, and political controversy — yielding an aggregate score on a 0-to-100 scale.

The ALTO corridor scores ECI<sub>w</sub> = 77 (engineering), CFI = 43 today and a projected CFI = 65 by 2030 (friction). These scores are carried over from the Initiative's earlier work and were not re-scored under the blind protocol.

## Analytical approach

Ordinary least squares regression of log-cost per kilometre was fitted on the 69 reference lines in the April specifications — each index alone, both together (the bivariate model), and with completion year added (the trivariate model). The indices were also added to a replication of PBO's own model, which uses measured tunnel and elevated shares, construction start year and regional indicators. That combination is the most demanding test of whether the indices add information PBO's model does not already contain.

## Findings

The bivariate ECI<sub>w</sub> + CFI model explains 42 per cent of the cross-project variance in log-cost (adjusted  $R^2 = 0.40$ ), against 90 per cent in the April edition and 66 per cent on the 16 April projects with sourced costs. A larger, more varied sample with less hindsight in its scores explains less, which is expected. Both indices are statistically significant: ECI<sub>w</sub> strongly ( $p < 0.001$ ), CFI more modestly ( $p = 0.011$ ). Engineering carries the larger effect (standardised  $\beta = 0.57$  against 0.25). The two indices are almost uncorrelated ( $r = 0.13$ ).

Added to PBO's model, friction is the more useful index. It raises  $R^2$  from 0.66 to 0.71, with each 10 CFI points associated with about 13 per cent higher cost ( $p = 0.002$ ), and the effect holds across every subset tested. Engineering complexity also adds to PBO's model on its own ( $p = 0.009$ ), largely because it captures ground and terrain conditions PBO's variables do not; once friction is also included it is no longer individually significant ( $p = 0.11$ ). Completion year remains significant in every specification, at about +2 per cent a year.

Applied to the ALTO corridor at its projected CFI of 65, the CRI multivariate model (ECIw + CFI) returns \$191 million per kilometre, or \$163 billion on PBO's 850 km baseline route; at today's CFI of 43 it returns \$151 million (\$129 billion). PBO's model (CRI replicated) returns \$86 million (\$73 billion), close to PBO's published \$75 billion; with friction added it returns \$136 million (\$115 billion).

### Implications

Three implications follow. First, the stated ALTO budget of \$60–90 billion sits below the central prediction of every model that includes either index. Under the CRI multivariate model, the probability of delivery at or below \$90 billion on the 850 km route is about 17 per cent at the projected CFI of 65 and 27 per cent at today's 43; at or below PBO's upper bound of \$113 billion it is about 28 per cent.

Second, community friction is a measurable cost driver on this evidence, although a smaller one than the April edition reported. Its effect survives control for measured engineering, region and start year, and does not depend on the U.K. and U.S. projects. Because the friction scores exclude cost-driven controversy, the association is less likely to run from cost to friction than in the April analysis.

Third, part of ALTO's cost is still in play. Engineering complexity is largely fixed by the route; friction is not. If ALTO's friction stays at today's level rather than rising to the projected 65, the CRI multivariate prediction is about \$34 billion lower.

***Scored without reference to cost, on 69 lines, both indices predict HSR cost. The CRI multivariate model places ALTO at about \$163 billion on PBO's 850 km route — about twice the Government's upper figure and above PBO's \$75–113 billion range.***

# 1. Introduction

---

## 1.1 Background and motivation

Canada's proposed ALTO high-speed rail corridor carries a stated capital cost of \$60–90 billion. On 1 October 2026 the Parliamentary Budget Officer published an independent estimate of \$75–113 billion for an 850 km baseline route, built on the NYU Transit Costs Project dataset of international HSR construction costs (Parliamentary Budget Officer, 2026). PBO's model explains cost with measured engineering variables and regional indicators; it does not include the sociopolitical conditions under which a line is built, and PBO notes that institutional and process risks could push costs beyond its range. Oxford reference-class data on more than 16,000 infrastructure projects show rail megaprojects overrunning their approved budgets by a mean of 44.7 per cent (Flyvbjerg et al., 2002; Flyvbjerg & Gardner, 2023).

The Initiative has developed two instruments to convert international evidence into project-specific cost predictions: the Engineering Complexity Index (ECI) and the Community Friction Index (CFI). The April 2026 edition of this report used both as predictors of cost across 16 projects and concluded that they explained about 90 per cent of cross-project variance. This edition tests that conclusion on the dataset PBO used, with a larger reference class and scores assigned without reference to cost.

## 1.2 Methodological scope

The report keeps the April structure. Section 2 summarises the Engineering Complexity Index and its univariate regression; section 3 does the same for the Community Friction Index. Section 4 presents the joint analysis: correlations, the bivariate and trivariate models, influence diagnostics, the indices added to PBO's model, robustness tests and an out-of-sample test. Section 5 applies the fitted models to ALTO. Sections 6 to 8 set out implications, limitations and conclusions. Appendix B lists every score; Appendix D compares the results with the April analysis.

All cost figures are taken from the Transit Costs dataset, which converts each project's cost to U.S. dollars at purchasing-power parity and 2023 prices. Multiplying by 1.40 reproduces PBO's figures in 2024 Canadian dollars. Costs cover civil works, systems, stations and overheads and exclude rolling stock and financing. Rail Baltica is not in the dataset; the Initiative converted its €23.8 billion full-scope estimate on the dataset's basis (\$54 million per kilometre). Regressions use the base-10 logarithm of cost per kilometre, as in April.

## 1.3 Disclosure of analytical framework

Both indices are analytical instruments developed by the Initiative. Neither has been externally peer-reviewed. The scores in this edition were assigned in October 2026 by AI research agents working from the published rubrics and the public record, one pass per line, and have not yet been checked by a second scorer. They are reproduced in full, with the evidence and sources for each line, so that any score can be checked and revised.

This report does not represent the position of the ALTO Infrastructure Corporation, Transport Canada, or any order of government.

## 1.4 Scoring protocol

The April scores were assigned by a scorer who already knew each project's cost and history, which tends to make indices track outcomes and overstate their predictive power. This edition uses a blind protocol (Appendix E). Scorers were instructed not to look up or use any information on cost, budget,

cost per kilometre or overruns. Engineering was scored on the physical characteristics of the costed section. Friction was scored from the documentary record of opposition, litigation, route changes, environmental challenge and political controversy, excluding any event whose subject was the project's cost. Where evidence was equivocal, the lower score was used. Each line carries a confidence rating for each index.

Blindness cannot be complete: a scorer may know a famous project's reputation. The protocol removes cost from the evidence and from the friction record, which addresses the main channel by which cost could feed into the scores.

## 2. The Engineering Complexity Index (ECI)

### 2.1 Rubric

The Engineering Complexity Index scores HSR projects on six physical engineering dimensions. Each dimension is scored 1 to 5 against published anchor descriptions; scores are weighted and summed to produce a composite index on a 20-to-100 scale. The rubric is unchanged from April.

| Dimension                         | Weight | What it captures                                                                                                                         |
|-----------------------------------|--------|------------------------------------------------------------------------------------------------------------------------------------------|
| Geotechnical conditions           | 20 %   | Bedrock quality; soft-soil extent; karst, permafrost, liquefiable or swelling clay presence; groundwater regime.                         |
| Major structures                  | 20 %   | Tunnel-kilometres; viaduct-kilometres; major bridge count; span complexity.                                                              |
| Topography and alignment          | 20 %   | Vertical relief over corridor; horizontal curve radii required; grade and cant-deficiency constraints at design speed.                   |
| Hydrology and drainage            | 15 %   | Major watercourse crossings; floodplain exposure; drainage structure count and design return period; wetland displacement.               |
| Infrastructure interfaces         | 15 %   | Urban alignment density; existing-rail, road, and utility crossings; station integration complexity; staging constraints.                |
| Climate and environmental loading | 10 %   | Winter design temperature; freeze-thaw cycle count; snow, ice, and wind design loads; seismic design category; extreme-weather exposure. |

Table 1. The Engineering Complexity Index rubric. Full scoring descriptors are set out in Appendix A.

#### A note on notation

ECI refers to the instrument; ECI<sub>w</sub> denotes the weighted composite score on a 20-to-100 scale. Individual dimension scores, each on a 1-to-5 scale, are referred to by their dimension name (Geo, Str, Top, Hyd, Int, Cli; Appendix B). The CFI aggregate is an unweighted sum, and the same abbreviation serves for the instrument and the score.

### 2.2 Reference class

The reference class comprises the 68 Transit Costs lines with complete tunnel, elevated-structure and start-year data — the set on which PBO's model can be replicated — plus Rail Baltica. It covers 21 countries. Table 2 summarises it by region; Appendix B lists every line.

| Region        | Lines     | Mean ECI <sub>w</sub> | Mean CFI  | Median \$M / km |
|---------------|-----------|-----------------------|-----------|-----------------|
| Europe        | 42        | 60                    | 30        | \$64            |
| Asia          | 19        | 73                    | 14        | \$115           |
| U.K./U.S.     | 3         | 58                    | 56        | \$216           |
| Rest of world | 5         | 59                    | 14        | \$63            |
| <b>All</b>    | <b>69</b> | <b>63</b>             | <b>26</b> | <b>\$82</b>     |

Table 2. Reference class by region. Europe is the EU members, Switzerland, Norway and Rail Baltica; rest of world is Turkey and Morocco; Asia is China, Japan, Korea and Taiwan. Costs in 2024 CAD, PBO basis.

## 2.3 Univariate regression

Ordinary least squares regression of the base-10 logarithm of cost per kilometre on the weighted complexity index, across the 69 lines, gives the following fit:

$$\log_{10} (\text{cost CAD M / km}) = 0.0152 \cdot \text{ECIw} + 0.944$$

$$R^2 = 0.36 \quad \cdot \quad \text{residual } \sigma = 0.26 \text{ log-units} \quad \cdot \quad n = 69$$

The fit implies approximately 42 per cent cost growth per 10 ECIw points, close to the 49 per cent reported in April, with a lower  $R^2$  (0.36 against 0.63). Applied to the ALTO corridor (ECIw = 77), the univariate prediction is \$129 million per kilometre. The U.K. and U.S. projects sit above the line, as in April, and several Spanish and French lines below it.

## 3. The Community Friction Index (CFI)

### 3.1 Rationale

The Community Friction Index measures the sociopolitical resistance a linear infrastructure project generates in the communities along its corridor. Sustained opposition can produce design changes — tunnel extensions, alignment shifts, elevated sections — that impose direct structural cost; it can extend programme duration; litigation imposes legal and administrative cost; and contingencies held against litigation risk add a further cost line (Flyvbjerg, 2014; Priemus, Flyvbjerg & van Wee, 2008). The theoretical foundations draw on work on place attachment and local opposition (Dear, 1992; Devine-Wright, 2009).

### 3.2 Framework

This edition uses the framework of the Initiative's friction note (October 2026). The CFI comprises five components, each scored 0 to 20 from the documentary record against a 20-indicator rubric; the components are summed without weighting to an aggregate on a 0-to-100 scale. For this edition, events whose subject was the project's cost were excluded from every component.

| Component                      | Range | What it captures                                                                                                                                                                      |
|--------------------------------|-------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Public opposition intensity    | 0–20  | Scale, organisation and reach of grassroots opposition: objections, community groups, coalitions, media coverage, mobilisation.                                                       |
| Legal challenges filed         | 0–20  | Formal legal activity against the project or its approvals: pre-action letters, tribunal applications, judicial review, injunctions, constitutional and Indigenous-rights litigation. |
| Forced route changes           | 0–20  | Alignment, station or scope changes made under community, regulatory or political pressure.                                                                                           |
| Environmental / NGO opposition | 0–20  | NGO engagement from routine submissions through species-at-risk triggers, environmental litigation and regulatory intervention.                                                       |
| Political controversy          | 0–20  | Inter-party, inter-governmental and intra-caucus contest over the project, other than over its cost.                                                                                  |

Table 3. The Community Friction Index components (friction note, October 2026). Each is scored against 20 cumulative indicators; the aggregate is the unweighted sum.

### 3.3 Band classification

| CFI range | Classification            | Reference lines in this band (blind scores)                                                                                                  |
|-----------|---------------------------|----------------------------------------------------------------------------------------------------------------------------------------------|
| 0–25      | <b>Low friction</b>       | Most Chinese, Japanese, Turkish and Spanish lines; LGV Sud-Est (11), Madrid–Seville (11), Taiwan HSR (17).                                   |
| 25–45     | <b>Moderate friction</b>  | LGV Nord (29), LGV Atlantique (32), Florence–Bologna (38), Madrid–Barcelona (42), Gardermobanen (42); ALTO today (43).                       |
| 45–65     | <b>High friction</b>      | HS1 (50), California IOS (50), VDE 8 (55), HSL-Zuid (56), Cologne–Frankfurt (57), Rail Baltica (59), Lyon–Turin (64), Gyeongbu Phase 2 (64). |
| 65–85     | <b>Very high friction</b> | HS2 Phase 1 (68); ALTO projected (65).                                                                                                       |
| 85–100    | <b>Extreme friction</b>   | No reference line once cost-driven controversy is excluded.                                                                                  |

Table 4. CFI band classification (friction note, October 2026), with reference lines placed by their blind scores. Bands are interpretive; the regressions use the continuous score.

### 3.4 Univariate regression

Ordinary least squares regression of log-cost on the CFI aggregate gives the following fit:

$$\log_{10}(\text{cost CAD M / km}) = 0.0060 \cdot \text{CFI} + 1.746$$

$$R^2 = 0.10 \quad \cdot \quad \text{residual } \sigma = 0.31 \text{ log-units} \quad \cdot \quad n = 69$$

The fit implies approximately 15 per cent cost growth per 10 CFI points, against 38 per cent in April, and friction alone explains only 10 per cent of variance (April: 81 per cent). Two things account for most of the change. Removing cost-driven controversy lowers the scores of the costliest projects most: HS2 falls from 92 to 68 and California from 78 to 50. And the expanded class includes many expensive Asian lines with little documented friction, which weakens the simple association. Applied to ALTO, the univariate prediction is \$137 million per kilometre at the projected CFI of 65 and \$101 million at today's 43.

***Once cost-driven controversy is excluded from the record, friction alone explains 10 per cent of cross-project cost variance, not 81 per cent. Its contribution appears once engineering and region are controlled for.***

## 4. Joint Multivariate Analysis

### 4.1 Dataset

The merged dataset has 69 lines, each with a Transit Costs cost per kilometre, an ECIw score and a CFI score assigned under the blind protocol, completion and start years, and tunnel and elevated shares. Engineering scores carry medium or high confidence for all but one line. Friction scores are rated low confidence for 34 lines, mostly older French, Italian, Spanish, Japanese and Chinese lines with little documented opposition; section 4.6 tests whether they drive the results.

### 4.2 Pairwise relationships

Pearson correlations across the four principal variables:

|                          | ECIw   | CFI    | Year   | $\log_{10}(\text{cost})$ |
|--------------------------|--------|--------|--------|--------------------------|
| ECIw                     | 1.000  | +0.128 | +0.165 | +0.602                   |
| CFI                      | +0.128 | 1.000  | +0.169 | +0.320                   |
| Year                     | +0.165 | +0.169 | 1.000  | +0.431                   |
| $\log_{10}(\text{cost})$ | +0.602 | +0.320 | +0.431 | 1.000                    |

Table 6. Pearson correlation matrix ( $n = 69$ ).

Engineering complexity has the strongest correlation with log-cost ( $r = +0.60$ ), followed by completion year ( $r = +0.43$ ) and friction ( $r = +0.32$ ). The three predictors are nearly independent of each other ( $r$  between 0.13 and 0.17), so their separate effects can be estimated cleanly — an improvement on the 16-project analysis, where friction and year moved together.

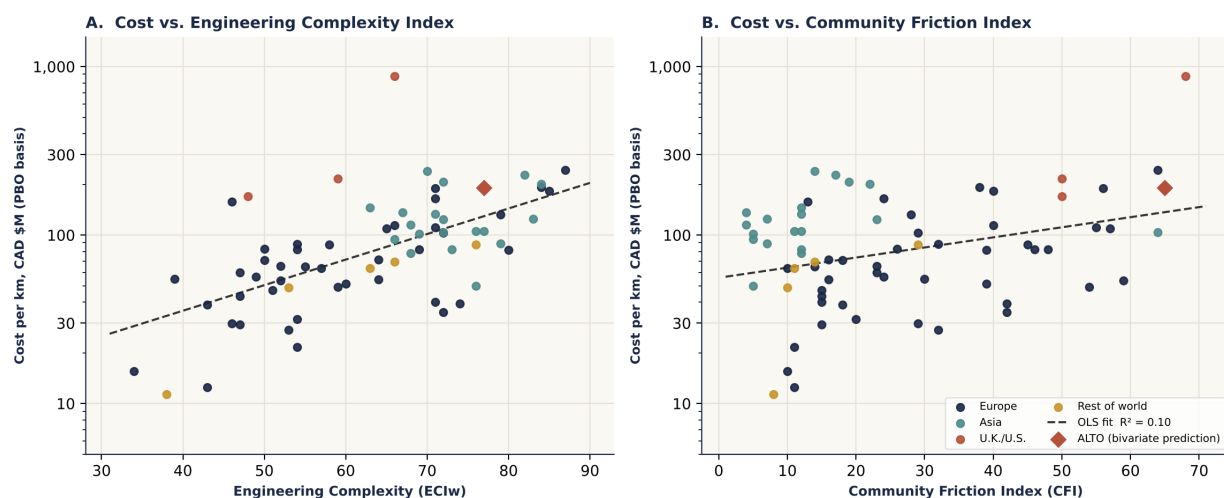


Figure 1. Univariate relationships between cost per kilometre and each index, coloured by region. Log-y axis. Panel A: cost against ECIw; OLS fit  $R^2 = 0.36$ . Panel B: cost against CFI; OLS fit  $R^2 = 0.10$ . Diamond: ALTO bivariate prediction (CFI 65).

### 4.3 Bivariate model

Ordinary least squares regression of log-cost jointly on ECIw and CFI — referred to in this report as the CRI multivariate model — gives the following fit:

$$\log_{10}(\text{cost CAD M / km}) = 0.0144 \cdot \text{ECIw} + 0.0046 \cdot \text{CFI} + 0.874$$

$$R^2 = 0.422 \quad \cdot \quad \text{adjusted } R^2 = 0.405 \quad \cdot \quad F(2, 66) = 24.1 \quad \cdot \quad p < 0.001 \quad \cdot \quad n = 69$$

Both coefficients are statistically significant: ECIw ( $t = 6.04$ ,  $p < 0.001$ ) and CFI ( $t = 2.62$ ,  $p = 0.011$ ). Each 10 ECIw points is associated with about 39 per cent higher cost, and each 10 CFI points with about 11 per cent. The standardised coefficients are  $\beta(\text{ECIw}) = 0.57$  and  $\beta(\text{CFI}) = 0.25$ : engineering has a little more than twice friction's effect per standard deviation, reversing the April result (0.35 and 0.70).

#### Collinearity diagnostics

Variance-inflation factors are 1.02 for both predictors, against 4.95 in April. The condition number of the design matrix is 355. Collinearity is not a concern.

#### Incremental explanatory power

| Contribution             | $\Delta R^2$ | Partial $r$ | Partial $F(1, 66)$ | p-value |
|--------------------------|--------------|-------------|--------------------|---------|
| ECIw added over CFI-only | +0.320       | +0.60       | 36.53              | < 0.001 |
| CFI added over ECIw-only | +0.060       | +0.31       | 6.88               | 0.011   |

Table 7. Incremental contribution of each predictor above and beyond the other.

Engineering complexity adds 32 points of  $R^2$  over friction; friction adds 6 points over engineering. Both are significant. The April conclusion that both indices are independent contributors is confirmed, but the balance between them is reversed.

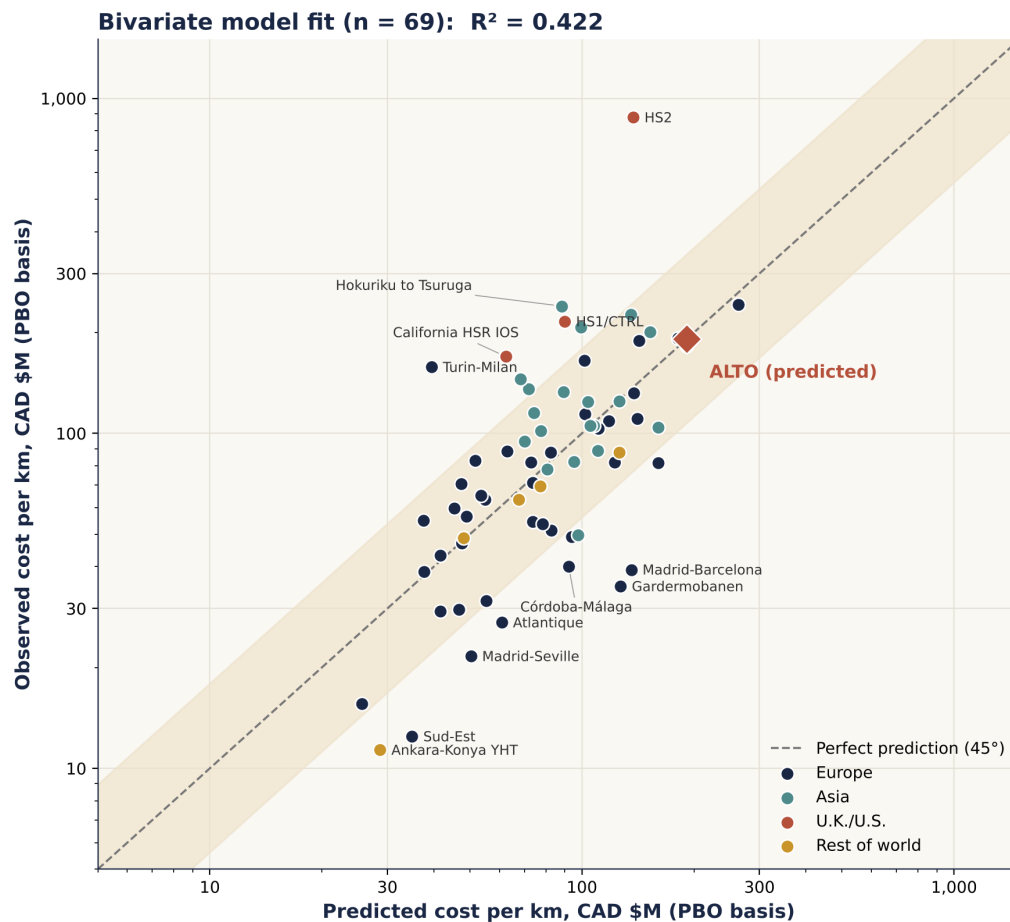


Figure 2. Bivariate model fit: observed cost per kilometre against the model's predicted value for each of the 69 reference lines, coloured by region; the twelve largest residuals are labelled. The 45° dashed line marks perfect prediction; the shaded band marks  $\pm 0.25$  log-units (about  $\times 1.8$ ). ALTO plots on the 45° line by construction.

#### 4.4 Trivariate test: does completion year add anything?

The April edition found that completion year added nothing once complexity and friction were controlled for. Adding completion year (centred on 2000) as a third predictor gives:

$$\log_{10}(\text{cost}) = 0.0132 \cdot \text{ECIw} + 0.0037 \cdot \text{CFI} + 0.0085 \cdot (\text{Year} - 2000) + 0.868$$

$$R^2 = 0.514 \quad \text{adjusted } R^2 = 0.492 \quad F(3, 65) = 22.9 \quad n = 69$$

The year coefficient is 0.0085 log-units per year ( $p = 0.001$ ), equivalent to about +2.0 per cent a year in real construction cost per kilometre, close to the start-year effect in PBO's model. The partial F-test for year over the bivariate model gives  $F(1, 65) = 12.30$ ,  $p = 0.0008$ , and adjusted  $R^2$  rises from 0.405 to 0.492. Both indices remain significant (ECIw  $p < 0.001$ ; CFI  $p = 0.026$ ). Real HSR construction cost

has risen over time, independently of complexity and friction; the April conclusion that it has not does not hold.

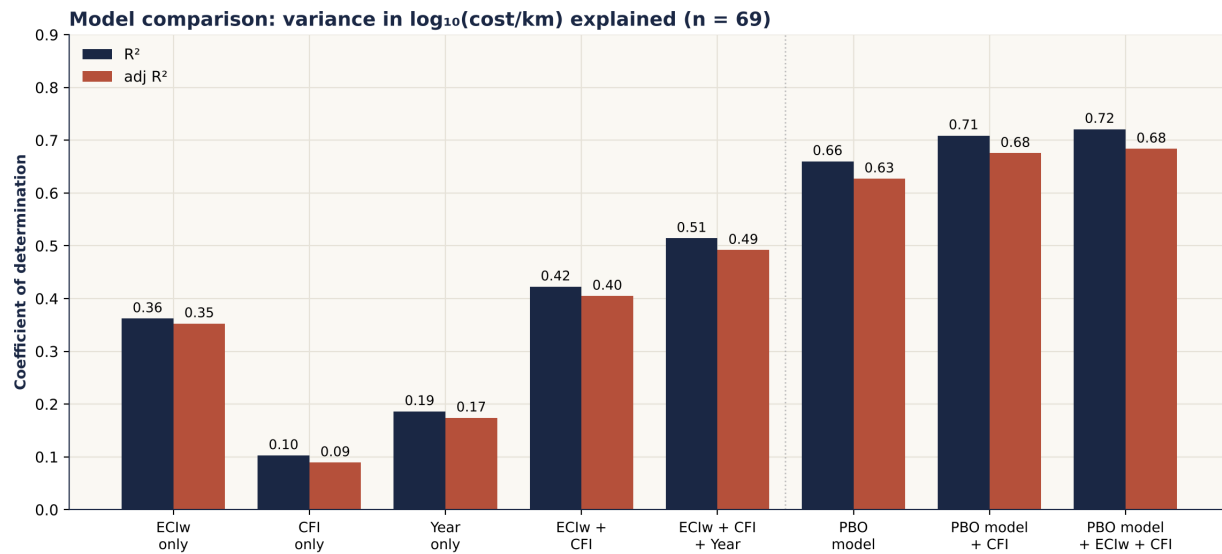


Figure 3. Model comparison (n = 69). Left of the dotted line, the index specifications; right, PBO's model (tunnel and elevated shares, start year, regional indicators) alone and with the indices added (section 4.6).

## 4.5 Influence diagnostics

With n = 69 the conventional Cook's distance threshold of  $4/n$  is 0.058. 5 lines exceed it: HS2 (D = 0.42) by a wide margin, and Turin–Milan, California IOS, Ankara–Konya and LGV Sud-Est marginally. HS2 sits far above the fitted line. Excluding HS2 and California together lowers the friction coefficient from 0.0046 to 0.0020, and it loses significance in the bivariate model ( $p = 0.23$ ); engineering is unaffected ( $p < 0.001$ ). In the bivariate model, then, the friction effect still leans on the two U.K./U.S. outliers. Section 4.6 shows that it does not once the regional cost level is controlled for.

## 4.6 The indices added to PBO's model

PBO's model explains log-cost with tunnel share, elevated share, construction start year and indicators for Asia, the U.K./U.S. and the rest of the world, with Europe as the reference. Fitted to these 69 lines it reproduces PBO's published results closely (Table 9) and explains 66 per cent of variance. Because its engineering variables are measured rather than scored, and it controls for the regional cost level, it is the most demanding test of the indices.

| Model                                            | R <sup>2</sup> | adj R <sup>2</sup> | Key result                                                    |
|--------------------------------------------------|----------------|--------------------|---------------------------------------------------------------|
| PBO model (tunnel, elevated, start year, region) | 0.660          | 0.627              | All engineering variables significant; U.K./U.S. $\times 4.8$ |
| + CFI                                            | 0.709          | 0.675              | CFI +13 % per 10 points; $p = 0.002$                          |
| + ECIw                                           | 0.696          | 0.661              | ECIw +20 % per 10 points; $p = 0.009$                         |
| + ECIw + CFI                                     | 0.721          | 0.684              | CFI $p = 0.024$ ; ECIw $p = 0.11$                             |

Table 8. The indices added to PBO's model (n = 69).

Friction adds significant explanatory power to PBO's model: R<sup>2</sup> rises from 0.66 to 0.71, and each 10 CFI points is associated with about 13 per cent higher cost ( $p = 0.002$ ). The estimated U.K./U.S. premium falls from  $\times 4.8$  to  $\times 3.4$ , so friction accounts for part of what PBO's model attributes to region. Engineering complexity also adds to PBO's model on its own ( $p = 0.009$ ), consistent with the ECI

capturing ground and terrain conditions that tunnel and elevated shares miss — PBO itself notes that its model does not capture the Canadian Shield. With both indices included, friction remains significant and engineering does not ( $p = 0.11$ ), in part because the ECI's structures dimension overlaps PBO's tunnel and elevated shares.

### Robustness of the friction effect

| Sample                          | n  | Bivariate: per 10 CFI | p     | PBO model + CFI: per 10 CFI | p     |
|---------------------------------|----|-----------------------|-------|-----------------------------|-------|
| All 69                          | 69 | +11 %                 | 0.011 | +13 %                       | 0.002 |
| Excluding HS2 and California    | 67 | +5 %                  | 0.230 | +13 %                       | 0.004 |
| Excluding China and Japan       | 54 | +17 %                 | 0.005 | +12 %                       | 0.012 |
| Europe only                     | 42 | +10 %                 | 0.141 | +12 %                       | 0.017 |
| Friction confidence H or M only | 35 | +9 %                  | 0.186 | +15 %                       | 0.014 |

Table 8b. Friction effect across subsets. “Europe only” omits the regional indicators. Low-confidence friction scores are mostly lines with little documented opposition.

In PBO's model the friction effect is stable — between +12 and +15 per cent per 10 points — and significant in every subset: without HS2 and California, without China and Japan (where documented opposition may be under-recorded), within Europe alone, and using only lines whose friction scores carry medium or high confidence. In the bivariate model the point estimates are similar, but significance depends on the U.K./U.S. projects and on sample size. The friction effect is best estimated alongside a control for the regional cost level.

## 4.7 Out-of-sample test: Rail Baltica

Rail Baltica, an 870 km standard-gauge line through the Baltic states, was not in the April reference class and is not in the Transit Costs dataset. Its blind scores are ECI<sub>w</sub> = 52 and CFI = 59; its cost on the dataset's basis is \$54 million per kilometre (€23.8 billion over 870 km, European Court of Auditors, 2026). Fitted on the other 68 lines, the bivariate model predicts \$81 million (95 per cent interval \$24–274 million), PBO's model \$55 million, and PBO's model with friction \$90 million (\$35–231 million). The actual cost lies inside every interval, nearest PBO's model. Adding friction moves Rail Baltica's prediction up, away from its cost: it is a high-friction line that has so far been relatively cheap per kilometre. Its estimate covers a route only about a third designed, and may rise.

## 5. Predictions for ALTO

### 5.1 Scoring

The ALTO corridor scores  $EClw = 77$ : 5 on geotechnical conditions (Leda clay, karst, Frontenac Arch bedrock exposure, freeze-thaw regime), 4 on major structures, 3 on topography and alignment, 4 on hydrology and drainage (Moir, Napanee and Salmon River crossings), 3 on infrastructure interfaces, and 4 on climate and environmental loading.

The friction scores follow the Initiative's friction note. ALTO scores  $CFI = 43$  at April 2026 (public opposition 16, legal challenges 5, forced route changes 2, environmental and NGO opposition 9, political controversy 11), in the upper Moderate band. The note's mid-energy trajectory projects  $CFI = 65$  by 2030.

These scores were not re-assigned under the blind protocol. Two adjustments would bring them closer to it. On the blind scorers' conventions, ALTO's structures score would be 3 rather than 4 (about 13 per cent of the route on tunnel or elevated structure), giving  $EClw = 73$ ; the bivariate prediction would fall from \$191 to \$168 million per kilometre. And for the 16 April projects, blind friction scores averaged 9 points lower than the April scores, mainly because cost-driven controversy was excluded; ALTO's current score of 43 includes little such controversy, but the projected 65 may be high on the blind scale. The predictions at  $CFI 43$  bracket that case.

The trivariate model uses a completion year of 2037 for ALTO. PBO's model uses PBO's assumptions for ALTO: about 15 km of tunnel and 100 km of elevated-equivalent structure on 850 km, with construction from 2029, at the European cost level.

### 5.2 Cost predictions

| Model                                                   | ALTO \$M / km<br>(CFI 65,<br>projected) | Total, 850 km<br>(CFI 65) | ALTO \$M / km<br>(CFI 43, today) | Total, 850 km<br>(CFI 43) |
|---------------------------------------------------------|-----------------------------------------|---------------------------|----------------------------------|---------------------------|
| <b>CRI multivariate model (<math>EClw + CFI</math>)</b> | <b>\$191</b>                            | <b>\$163B</b>             | <b>\$151</b>                     | <b>\$129B</b>             |
| CRI multivariate model + completion year                | \$278                                   | \$237B                    | \$230                            | \$196B                    |
| $EClw$ only                                             | \$129                                   | \$110B                    | \$129                            | \$110B                    |
| CFI only                                                | \$137                                   | \$116B                    | \$101                            | \$86B                     |
| PBO model, CRI replicated (Europe cost level)           | \$86                                    | \$73B                     | \$86                             | \$73B                     |
| PBO model + CFI                                         | \$136                                   | \$115B                    | \$103                            | \$88B                     |
| PBO model + $EClw + CFI$                                | \$159                                   | \$135B                    | \$128                            | \$109B                    |
| PBO model, CRI replicated (U.K./U.S. cost level)        | \$411                                   | \$349B                    | \$411                            | \$349B                    |
| PBO published (RP-2627-009-C)                           | \$88                                    | \$75–113B <sup>1</sup>    | \$88                             | \$75–113B <sup>1</sup>    |

Table 9. Cost predictions for ALTO ( $n = 69$ ). Medians on the log scale; totals apply the median to PBO's 850 km baseline route. The CRI multivariate model, the subject of this report, is highlighted. Models without CFI give the same value at both friction levels. <sup>1</sup> PBO's range runs from its central estimate to a one-sided upper bound expected to be exceeded one time in six.

The CRI multivariate model — engineering complexity and community friction together — places ALTO at \$191 million per kilometre at the projected  $CFI$  of 65, or \$163 billion on PBO's 850 km route, and at \$151 million (\$129 billion) at today's  $CFI$  of 43. Its 95 per cent prediction interval is wide, \$57–644 million per kilometre (\$48 – \$547 billion), reflecting the scatter of real projects around the line. Allowing

for that scatter, the expected (mean) cost is about \$228 million per kilometre. Adding completion year raises the prediction to \$278 million (\$237 billion), because ALTO would be built later than most of the reference class.

The CRI replicated PBO model gives \$86 million per kilometre (\$73 billion), against PBO's published \$88 million (\$75 billion); its 84th-percentile value is \$141 million (\$119 billion), against PBO's upper bound of \$113 billion. This confirms that the dataset and specification match PBO's. Adding friction to PBO's model, the most demanding test of the index (section 4.6), raises its prediction to \$136 million (\$115 billion).

The spread between models reflects the cost level each assumes. PBO's model prices ALTO at European efficiency, with no allowance for friction. The CRI multivariate model is fitted across all regions and, for a corridor scored as complex and contested, predicts more. With the U.K./U.S. indicator switched on, PBO's own model gives \$411 million per kilometre (\$349 billion). The CRI multivariate prediction sits between PBO's European reference and that adverse case.

***The CRI multivariate model places ALTO at \$163 billion on PBO's 850 km route at its projected level of friction, and \$129 billion if friction is held at today's level — above the Government's \$60–90 billion in both cases.***

## 6. Implications

### 6.1 The stated ALTO budget

The ALTO proposal carries a stated capital cost of \$60–90 billion, or about \$71–106 million per kilometre on PBO's 850 km route (PBO's own \$75–113 billion is \$88–133 million per kilometre). The CRI multivariate model places ALTO well above that range, at \$191 million per kilometre. PBO's model (CRI replicated) places its central estimate inside the Government's range at \$86 million, on the assumption of European cost efficiency and no friction; with friction added it rises to \$136 million.

| Model                                         | ≤ \$60 billion | ≤ \$75 billion | ≤ \$90 billion | ≤ \$113 billion |
|-----------------------------------------------|----------------|----------------|----------------|-----------------|
| CRI multivariate model, CFI 65                | 5 %            | 10 %           | 17 %           | 28 %            |
| CRI multivariate model, CFI 43                | 10 %           | 18 %           | 27 %           | 41 %            |
| PBO model + CFI, CFI 65                       | 9 %            | 19 %           | 30 %           | 48 %            |
| PBO model, CRI replicated (Europe cost level) | 35 %           | 52 %           | 66 %           | 81 %            |

Table 10. Probability that ALTO's construction cost on the 850 km route falls at or below each threshold, from each model's prediction distribution.

Under the CRI multivariate model, the Government's upper figure of \$90 billion has about a 17 per cent chance of being met at the projected level of friction, and PBO's upper bound of \$113 billion about 28 per cent. On PBO's model alone, \$90 billion is more likely than not to be met. Friction and engineering complexity, which PBO's model does not score, account for the difference.

### 6.2 Community friction as a cost driver

The April edition concluded that community friction is a measurable, independent cost driver. On blind scores and sourced costs that conclusion holds, with a smaller effect: about 13 per cent per 10 CFI points once measured engineering, region and start year are controlled for, robust across subsets. Because cost-driven controversy was excluded, the association is less likely to reflect expensive projects attracting opposition. Some reverse causation may remain — a line that takes longer can accumulate more friction — and the Initiative's main report found that, among friction measures, litigation exposure scored before construction is the clearest forecasting counterpart.

For ALTO the practical point is that friction is not yet fixed. Each 10 points of friction avoided — through consultation that changes the route rather than defends it, settled land acquisition, and early resolution of environmental objections — is worth roughly 13 per cent of construction cost on this evidence.

***Scored without reference to cost, engineering complexity is the stronger predictor of HSR cost and community friction a smaller but robust independent driver. Together they place ALTO well above its stated budget, and friction is the part of that cost still open to management.***

## 7. Limitations

---

Six limitations merit explicit statement.

### Sample and design

The reference class contains 69 lines from 21 countries, 23 observations per parameter in the bivariate model. It is the set of lines with complete data in one dataset, not a random sample, and conventional standard errors are optimistic for out-of-sample prediction. Residual scatter remains large: even with an unlimited sample, the bivariate interval would span a factor of about ten.

### Single-pass scoring

Each line was scored once, by AI research agents applying the rubrics to the public record, in eleven batches. No second scorer has checked the scores, and batches may have read the rubrics slightly differently — for example, some counted tunnel and viaduct share together for the structures dimension. Inter-rater reliability is untested.

### Documentation bias

Friction is scored from the documentary record. Where opposition is under-reported — notably for Chinese lines, and for older lines generally — scores may understate friction. Excluding China and Japan, or low-confidence scores, does not change the friction result in PBO's model.

### Imperfect blindness

Scorers were told not to use cost information, but may know the reputation of well-known projects. ALTO's scores were not re-assigned under the protocol (section 5.1).

### Cost data

Transit Costs figures follow each owner's reported scope. HS2 and California include land, trains and large contingencies; several lines are estimates for projects still under construction. Rail Baltica is not in the dataset and was converted by the Initiative.

### Governance as an omitted variable

The U.K. and U.S. projects remain far above the engineering line. The Initiative's main report traces most of this premium to delivery: consultant-led owners, contracts let on immature designs, consent-driven scope and city-centre terminals. Friction captures part of this, the regional indicator more.

## 8. Conclusions

---

Re-estimated on the cost data PBO used, with 69 lines scored without reference to cost, the joint analysis supports four conclusions.

**One.** Engineering complexity is the stronger predictor of HSR cost. ECIw alone explains 36 per cent of cross-project variance; each 10 points is associated with about 39 per cent higher cost in the bivariate model; and it adds to PBO's measured engineering variables on its own.

**Two.** Community friction is an independent cost driver, smaller than the April edition reported. It explains 10 per cent of variance alone, is significant alongside engineering ( $p = 0.011$ ), and adds about 13 per cent per 10 points to PBO's model, robustly across subsets.

**Three.** Real construction cost has risen over time, by about 2 per cent a year, independently of complexity and friction.

**Four.** The CRI multivariate model places ALTO at \$191 million per kilometre at the projected CFI of 65, or \$163 billion on PBO's 850 km route, and \$129 billion at today's CFI. The probability of delivery within the Government's \$90 billion is about 17 per cent. PBO's own model, with friction added, gives \$115 billion.

### 8.1 Recommendations

Two recommendations follow, addressed to parliamentary and municipal actors with standing in Canadian infrastructure decisions.

**First, extend PBO's estimate to friction.** PBO's model is sound and replicates closely. It omits the sociopolitical conditions of delivery, which on this evidence add materially to cost. PBO should be asked to test a friction or litigation-exposure measure in its model and to report ALTO's estimate with it.

**Second, manage friction as a cost line.** The project's cost documentation should set out the allowance made for consent-driven scope, litigation and delay, and the consultation and land-acquisition strategy intended to contain them.

***Re-estimated on sourced data and blind scores, the reference class still places ALTO above its stated budget. Engineering complexity carries most of the weight; community friction adds a measurable share that PBO's reference estimate does not include.***

## Appendix A. ECI scoring descriptors

Each dimension of the Engineering Complexity Index is scored 1 to 5 against descriptor-anchored criteria. The descriptors are unchanged from April.

### Geotechnical conditions

| Score | Descriptor                                                                                                                                 |
|-------|--------------------------------------------------------------------------------------------------------------------------------------------|
| 1     | Uniform competent bedrock or consolidated soils; negligible problematic-ground presence.                                                   |
| 2     | Predominantly competent ground with limited soft-soil or expansive-clay segments; routine site investigation sufficient.                   |
| 3     | Mixed ground conditions with localised geotechnical complications; standard mitigation measures required.                                  |
| 4     | Substantial soft-soil, permafrost, karst, or liquefiable-soil exposure requiring specialised treatment across significant corridor length. |
| 5     | Severe geotechnical regime: Leda clay, extensive karst, permafrost, or combined exposures requiring programme-defining ground treatment.   |

### Major structures

| Score | Descriptor                                                                                                                 |
|-------|----------------------------------------------------------------------------------------------------------------------------|
| 1     | Predominantly at-grade alignment; minimal tunnel or viaduct content.                                                       |
| 2     | Routine bridge and short viaduct count; tunnel content below 10 per cent of corridor.                                      |
| 3     | Substantial bridging, moderate viaduct or tunnel content (10–25 per cent of corridor).                                     |
| 4     | Major tunnel or viaduct backbone (25–50 per cent of corridor); significant structure-count driver.                         |
| 5     | Programme-defining tunnelling or viaduct requirement (> 50 per cent of corridor); structures dominate the capital account. |

### Topography and alignment

| Score | Descriptor                                                                                                                       |
|-------|----------------------------------------------------------------------------------------------------------------------------------|
| 1     | Flat or gently rolling terrain; vertical relief below 100 m over corridor; generous curve radii easily achieved.                 |
| 2     | Modest relief; occasional grade and curve constraints at design speed.                                                           |
| 3     | Significant relief or dense urban segments forcing tighter-than-optimal alignment choices; routine cant-deficiency use.          |
| 4     | Substantial mountain crossings or dense urban segments forcing extensive grade and curve constraints; speed-envelope compromise. |
| 5     | Mountainous terrain with relief above 500 m; alignment substantially shaped by topography; multiple speed-envelope compromises.  |

### Hydrology and drainage

| Score | Descriptor                                                                                                            |
|-------|-----------------------------------------------------------------------------------------------------------------------|
| 1     | Minimal watercourse crossings; no significant floodplain or wetland exposure.                                         |
| 2     | Routine drainage structures; occasional medium watercourse crossings; limited wetland displacement.                   |
| 3     | Several major watercourse crossings; localised floodplain and wetland exposure; substantial drainage structure count. |

| Score | Descriptor                                                                                                              |
|-------|-------------------------------------------------------------------------------------------------------------------------|
| 4     | Multiple major watercourse crossings with significant span requirements; programme-affecting floodplain exposure.       |
| 5     | Dominant hydrological constraint: extensive watercourse crossings, karst hydrogeology, or dominant floodplain exposure. |

### Infrastructure interfaces

| Score | Descriptor                                                                                                  |
|-------|-------------------------------------------------------------------------------------------------------------|
| 1     | Rural corridor; minimal existing-infrastructure interaction; greenfield alignment.                          |
| 2     | Modest urban segments; routine road and rail crossings; manageable utility relocation.                      |
| 3     | Significant urban segments; multiple major crossings; non-trivial staging and utility relocation.           |
| 4     | Extensive urban corridor; complex station integration; dense utility and existing-rail interface.           |
| 5     | Dominant urban and interface complexity; integrated major-station construction; severe staging constraints. |

### Climate and environmental loading

| Score | Descriptor                                                                                                                  |
|-------|-----------------------------------------------------------------------------------------------------------------------------|
| 1     | Mild climate; negligible freeze-thaw cycling; low wind and seismic loads.                                                   |
| 2     | Moderate climate; routine winter design; moderate wind and seismic loads.                                                   |
| 3     | Substantial winter design requirement; notable freeze-thaw cycling; elevated wind or seismic loads.                         |
| 4     | Severe climate loading: cold-climate catenary, extensive freeze-thaw, or high seismic design category.                      |
| 5     | Extreme combined loading: deep winter temperatures, heavy snow/ice design, major seismic zone, or extreme-weather exposure. |

## Appendix B. Full project scoring matrix

Blind scores for all 69 reference lines. Confidence ratings (H, M, L) are the scorers' own. Costs in 2024 CAD million per kilometre, PBO basis. The evidence and sources for every score are in the accompanying workbook (CRI\_ECI\_CFI\_blind\_scores\_69.xlsx).

| Line                 | Ctry     | Built   | Geo | Str | Top | Hyd | Int | Cli | ECIw | Conf. |
|----------------------|----------|---------|-----|-----|-----|-----|-----|-----|------|-------|
| Koralm Railway       | AT       | 2001–26 | 4   | 4   | 4   | 3   | 3   | 3   | 72   | M     |
| Lower Inn Valley Ph1 | AT       | 2006–12 | 4   | 5   | 2   | 4   | 3   | 3   | 71   | M     |
| HSL 4                | BE       | 2000–09 | 2   | 2   | 1   | 2   | 3   | 2   | 39   | M     |
| HSL 3                | BE       | 2001–07 | 3   | 3   | 3   | 2   | 3   | 2   | 55   | M     |
| Mattstetten-Rohrist  | CH       | 1996–04 | 3   | 4   | 3   | 3   | 2   | 2   | 59   | M     |
| Beijing-Tianjin Ph1  | CN       | 2005–08 | 4   | 5   | 1   | 3   | 4   | 3   | 67   | M     |
| Wuhan-Guangzhou      | CN       | 2005–08 | 4   | 5   | 4   | 4   | 3   | 3   | 79   | M     |
| Zhengzhou-Xi'an      | CN       | 2005–10 | 4   | 5   | 4   | 3   | 3   | 3   | 76   | M     |
| Beijing-Shanghai     | CN       | 2008–11 | 4   | 5   | 2   | 5   | 4   | 3   | 77   | H     |
| Shanghai-Hangzhou    | CN       | 2008–10 | 4   | 5   | 1   | 4   | 4   | 2   | 68   | M     |
| Hanover-Würzburg     | DE       | 1973–91 | 3   | 4   | 4   | 4   | 3   | 2   | 69   | M     |
| Mannheim-Stuttgart   | DE       | 1983–91 | 3   | 4   | 2   | 3   | 3   | 2   | 58   | H     |
| Cologne-Frankfurt    | DE       | 1995–02 | 3   | 4   | 3   | 3   | 4   | 2   | 65   | M     |
| VDE 8                | DE       | 1996–17 | 3   | 4   | 4   | 4   | 3   | 3   | 71   | H     |
| Nuremberg-Ingolstadt | DE       | 1998–06 | 4   | 4   | 3   | 3   | 3   | 2   | 66   | H     |
| Wendlingen-Ulm       | DE       | 2010–22 | 4   | 5   | 4   | 4   | 3   | 3   | 79   | H     |
| Rail Baltica         | EE/LV/LT | 2019–30 | 4   | 2   | 1   | 3   | 3   | 3   | 52   | M     |
| Madrid-Seville       | ES       | 1988–92 | 3   | 2   | 3   | 3   | 3   | 2   | 54   | M     |
| Madrid-Barcelona     | ES       | 1996–08 | 4   | 3   | 4   | 3   | 5   | 3   | 74   | M     |
| Madrid-Valladolid    | ES       | 1998–07 | 3   | 3   | 4   | 3   | 3   | 3   | 64   | H     |
| Córdoba-Málaga       | ES       | 2001–07 | 4   | 3   | 4   | 4   | 3   | 3   | 71   | H     |
| Figueres-Perpignan   | ES       | 2004–10 | 3   | 3   | 3   | 3   | 2   | 3   | 57   | M     |
| Basque Y             | ES       | 2006–23 | 4   | 5   | 4   | 4   | 4   | 2   | 80   | M     |
| Venta de Baños-León  | ES       | 2008–15 | 2   | 1   | 1   | 2   | 2   | 3   | 34   | M     |
| Helsinki-Turku       | FI       | 2026–34 | 3   | 3   | 2   | 2   | 2   | 4   | 52   | L     |
| Sud-Est              | FR       | 1976–83 | 2   | 1   | 3   | 3   | 2   | 2   | 43   | M     |
| Atlantique           | FR       | 1985–90 | 3   | 2   | 2   | 3   | 4   | 2   | 53   | M     |
| Interconnexion Est   | FR       | 1990–96 | 2   | 3   | 2   | 2   | 4   | 2   | 50   | M     |
| Rhône-Alpes          | FR       | 1991–94 | 2   | 3   | 3   | 3   | 3   | 2   | 54   | M     |
| Nord                 | FR       | 1991–96 | 3   | 2   | 1   | 3   | 3   | 2   | 46   | M     |
| Méditerranée         | FR       | 1996–01 | 3   | 3   | 3   | 4   | 2   | 3   | 60   | H     |
| Est Phase 1          | FR       | 2002–07 | 3   | 2   | 2   | 3   | 2   | 2   | 47   | M     |

| Line                                 | Ctry | Built   | Geo | Str | Top | Hyd | Int | Cli | ECIw | Conf. |
|--------------------------------------|------|---------|-----|-----|-----|-----|-----|-----|------|-------|
| Rhin-Rhône                           | FR   | 2009–14 | 3   | 2   | 2   | 3   | 2   | 2   | 47   | M     |
| Est Phase 2                          | FR   | 2010–16 | 3   | 2   | 3   | 3   | 2   | 2   | 51   | M     |
| Sud-Europe-Atlantique                | FR   | 2011–17 | 3   | 2   | 2   | 3   | 2   | 2   | 47   | M     |
| Nîmes-Montpellier                    | FR   | 2012–18 | 3   | 2   | 1   | 4   | 3   | 2   | 49   | M     |
| Bretagne                             | FR   | 2012–17 | 2   | 2   | 2   | 3   | 2   | 2   | 43   | M     |
| Lyon-Turin                           | FR   | 2016–30 | 5   | 5   | 5   | 4   | 3   | 3   | 87   | H     |
| Bordeaux-Toulouse                    | FR   | 2024–32 | 3   | 3   | 2   | 3   | 3   | 2   | 54   | M     |
| Rome-Naples                          | IT   | 1994–09 | 3   | 4   | 3   | 3   | 3   | 3   | 64   | M     |
| Florence-Bologna                     | IT   | 1996–09 | 4   | 5   | 5   | 5   | 3   | 2   | 84   | H     |
| Milan-Bologna                        | IT   | 2000–08 | 3   | 3   | 1   | 3   | 3   | 2   | 50   | M     |
| Turin-Milan                          | IT   | 2002–09 | 2   | 3   | 1   | 3   | 3   | 2   | 46   | M     |
| Tortona-Genoa                        | IT   | 2013–23 | 5   | 5   | 5   | 3   | 4   | 2   | 85   | H     |
| Brescia-Verona                       | IT   | 2019–26 | 3   | 3   | 2   | 3   | 3   | 2   | 54   | M     |
| Hokuriku to Nagano                   | JP   | 1989–96 | 3   | 5   | 5   | 3   | 3   | 3   | 76   | M     |
| Kyushu Shin-Yasushiro-Kagoshima Chuo | JP   | 1991–04 | 4   | 5   | 4   | 3   | 2   | 3   | 73   | M     |
| Tohoku Morioka-Hachinohe             | JP   | 1991–02 | 4   | 5   | 3   | 2   | 2   | 4   | 68   | M     |
| Shin-Aomori                          | JP   | 1998–10 | 3   | 5   | 3   | 2   | 2   | 5   | 66   | M     |
| Hokuriku to Kanazawa                 | JP   | 2001–15 | 4   | 5   | 4   | 4   | 3   | 5   | 83   | M     |
| Kyushu Hakata-Shin-Yasushiro         | JP   | 2004–11 | 4   | 5   | 2   | 3   | 4   | 3   | 71   | M     |
| Hokkaido                             | JP   | 2005–16 | 3   | 5   | 3   | 2   | 3   | 5   | 69   | M     |
| West Kyushu Takeo–Nagasaki           | JP   | 2008–22 | 3   | 5   | 3   | 2   | 3   | 2   | 63   | M     |
| Hokuriku to Tsuruga                  | JP   | 2012–24 | 3   | 5   | 3   | 3   | 3   | 4   | 70   | M     |
| Hokkaido Sapporo-Shin-Hakodate       | JP   | 2012–38 | 5   | 5   | 4   | 3   | 3   | 5   | 84   | M     |
| Gyeongbu Phase 1                     | KR   | 1994–04 | 3   | 5   | 4   | 3   | 3   | 3   | 72   | M     |
| Gyeongbu Phase 2                     | KR   | 2002–10 | 3   | 5   | 4   | 3   | 3   | 3   | 72   | M     |
| Suseo/SRT                            | KR   | 2011–16 | 4   | 5   | 3   | 2   | 4   | 3   | 72   | M     |
| Al-Boraq                             | MA   | 2011–18 | 4   | 2   | 2   | 3   | 2   | 3   | 53   | M     |
| HSL Zuid                             | NL   | 2000–09 | 5   | 4   | 1   | 5   | 4   | 2   | 71   | H     |
| Gardermobanen                        | NO   | 1994–98 | 4   | 4   | 2   | 4   | 4   | 4   | 72   | M     |
| Ankara-Istanbul YHT                  | TR   | 2003–30 | 4   | 3   | 4   | 4   | 4   | 4   | 76   | M     |
| Ankara-Konya YHT                     | TR   | 2006–11 | 2   | 1   | 2   | 2   | 2   | 3   | 38   | M     |
| Ankara-Sivas YHT                     | TR   | 2009–23 | 3   | 3   | 4   | 3   | 2   | 4   | 63   | M     |
| Ankara (Polatli)-İzmir YST           | TR   | 2013–25 | 3   | 3   | 4   | 3   | 3   | 4   | 66   | M     |
| THSR                                 | TW   | 2000–07 | 4   | 5   | 3   | 4   | 4   | 5   | 82   | H     |
| HS1/CTRL                             | UK   | 1998–07 | 3   | 4   | 2   | 3   | 4   | 1   | 59   | H     |

| Line               | Ctry | Built   | Geo | Str | Top | Hyd | Int | Cli | ECIw | Conf. |
|--------------------|------|---------|-----|-----|-----|-----|-----|-----|------|-------|
| HS2                | UK   | 2020–36 | 3   | 4   | 3   | 3   | 5   | 1   | 66   | H     |
| California HSR IOS | US   | 2013–33 | 3   | 2   | 1   | 3   | 3   | 3   | 48   | M     |

Table B-1. Engineering Complexity Index: dimension scores and composite (n = 69).

| Line                 | Ctry     | Opp. | Legal | Route | Env. | Pol. | CFI | Conf. | \$M / km |
|----------------------|----------|------|-------|-------|------|------|-----|-------|----------|
| Koralbm Railway      | AT       | 6    | 6     | 4     | 6    | 7    | 29  | L     | \$103    |
| Lower Inn Valley Ph1 | AT       | 6    | 3     | 6     | 4    | 5    | 24  | L     | \$165    |
| HSL 4                | BE       | 7    | 2     | 8     | 5    | 8    | 30  | L     | \$55     |
| HSL 3                | BE       | 4    | 2     | 3     | 3    | 2    | 14  | L     | \$65     |
| Mattstetten-Rohrist  | CH       | 14   | 12    | 9     | 7    | 12   | 54  | M     | \$49     |
| Beijing-Tianjin Ph1  | CN       | 1    | 0     | 1     | 1    | 1    | 4   | L     | \$136    |
| Wuhan-Guangzhou      | CN       | 2    | 0     | 1     | 1    | 3    | 7   | L     | \$89     |
| Zhengzhou-Xi'an      | CN       | 1    | 0     | 1     | 1    | 2    | 5   | L     | \$50     |
| Beijing-Shanghai     | CN       | 3    | 2     | 1     | 1    | 4    | 11  | L     | \$106    |
| Shanghai-Hangzhou    | CN       | 1    | 0     | 1     | 1    | 1    | 4   | L     | \$115    |
| Hanover-Würzburg     | DE       | 13   | 11    | 10    | 7    | 7    | 48  | M     | \$82     |
| Mannheim-Stuttgart   | DE       | 12   | 10    | 9     | 5    | 9    | 45  | M     | \$88     |
| Cologne-Frankfurt    | DE       | 12   | 11    | 10    | 11   | 13   | 57  | M     | \$109    |
| VDE 8                | DE       | 16   | 10    | 6     | 11   | 12   | 55  | M     | \$111    |
| Nuremberg-Ingolstadt | DE       | 10   | 11    | 3     | 10   | 6    | 40  | M     | \$114    |
| Wendlingen-Ulm       | DE       | 7    | 3     | 4     | 8    | 6    | 28  | M     | \$132    |
| Rail Baltica         | EE/LV/LT | 15   | 10    | 5     | 14   | 15   | 59  | M     | \$54     |
| Madrid-Seville       | ES       | 1    | 1     | 1     | 2    | 6    | 11  | M     | \$22     |
| Madrid-Barcelona     | ES       | 13   | 8     | 6     | 6    | 9    | 42  | M     | \$39     |
| Madrid-Valladolid    | ES       | 3    | 1     | 4     | 5    | 3    | 16  | L     | \$54     |
| Córdoba-Málaga       | ES       | 5    | 2     | 1     | 4    | 3    | 15  | M     | \$40     |
| Figueres-Perpignan   | ES       | 2    | 2     | 1     | 2    | 3    | 10  | L     | \$64     |
| Basque Y             | ES       | 16   | 4     | 4     | 6    | 9    | 39  | M     | \$82     |
| Venta de Baños-León  | ES       | 3    | 1     | 1     | 1    | 4    | 10  | L     | \$16     |
| Helsinki-Turku       | FI       | 6    | 6     | 2     | 4    | 5    | 23  | L     | \$65     |
| Sud-Est              | FR       | 3    | 1     | 2     | 2    | 3    | 11  | L     | \$12     |
| Atlantique           | FR       | 9    | 8     | 7     | 4    | 4    | 32  | M     | \$27     |
| Interconnexion Est   | FR       | 5    | 1     | 4     | 5    | 3    | 18  | L     | \$71     |
| Rhône-Alpes          | FR       | 6    | 1     | 5     | 4    | 4    | 20  | L     | \$32     |
| Nord                 | FR       | 10   | 3     | 6     | 4    | 6    | 29  | M     | \$30     |
| Méditerranée         | FR       | 16   | 3     | 9     | 5    | 6    | 39  | M     | \$51     |
| Est Phase 1          | FR       | 4    | 2     | 3     | 3    | 3    | 15  | L     | \$29     |
| Rhin-Rhône           | FR       | 3    | 2     | 3     | 4    | 3    | 15  | L     | \$43     |

| Line                                 | Ctry | Opp. | Legal | Route | Env. | Pol. | CFI | Conf. | \$M / km |
|--------------------------------------|------|------|-------|-------|------|------|-----|-------|----------|
| Est Phase 2                          | FR   | 3    | 2     | 3     | 4    | 3    | 15  | L     | \$47     |
| Sud-Europe-Atlantique                | FR   | 6    | 6     | 2     | 5    | 4    | 23  | L     | \$60     |
| Nîmes-Montpellier                    | FR   | 5    | 7     | 2     | 6    | 4    | 24  | M     | \$56     |
| Bretagne                             | FR   | 6    | 2     | 2     | 4    | 4    | 18  | L     | \$39     |
| Lyon-Turin                           | FR   | 17   | 10    | 9     | 10   | 18   | 64  | M     | \$242    |
| Bordeaux-Toulouse                    | FR   | 16   | 11    | 3     | 10   | 6    | 46  | H     | \$82     |
| Rome-Naples                          | IT   | 5    | 2     | 2     | 4    | 3    | 16  | L     | \$71     |
| Florence-Bologna                     | IT   | 12   | 10    | 2     | 8    | 6    | 38  | M     | \$193    |
| Milan-Bologna                        | IT   | 6    | 5     | 5     | 5    | 5    | 26  | L     | \$83     |
| Turin-Milan                          | IT   | 3    | 2     | 2     | 4    | 2    | 13  | L     | \$158    |
| Tortona-Genoa                        | IT   | 15   | 6     | 3     | 7    | 9    | 40  | M     | \$182    |
| Brescia-Verona                       | IT   | 8    | 8     | 3     | 8    | 5    | 32  | M     | \$89     |
| Hokuriku to Nagano                   | JP   | 4    | 1     | 2     | 1    | 4    | 12  | L     | \$106    |
| Kyushu Shin-Yasushiro-Kagoshima Chuo | JP   | 4    | 1     | 2     | 1    | 4    | 12  | L     | \$82     |
| Tohoku Morioka-Hachinohe             | JP   | 4    | 1     | 2     | 1    | 4    | 12  | L     | \$78     |
| Shin-Aomori                          | JP   | 1    | 0     | 1     | 1    | 2    | 5   | M     | \$94     |
| Hokuriku to Kanazawa                 | JP   | 2    | 0     | 1     | 1    | 3    | 7   | L     | \$125    |
| Kyushu Hakata-Shin-Yasushiro         | JP   | 3    | 2     | 2     | 2    | 3    | 12  | L     | \$133    |
| Hokkaido                             | JP   | 1    | 0     | 1     | 1    | 2    | 5   | M     | \$102    |
| West Kyushu Takeo–Nagasaki           | JP   | 4    | 0     | 1     | 1    | 6    | 12  | M     | \$145    |
| Hokuriku to Tsuruga                  | JP   | 3    | 0     | 2     | 5    | 4    | 14  | M     | \$240    |
| Hokkaido Sapporo-Shin-Hakodate       | JP   | 7    | 3     | 2     | 6    | 4    | 22  | M     | \$201    |
| Gyeongbu Phase 1                     | KR   | 4    | 2     | 3     | 2    | 8    | 19  | M     | \$207    |
| Gyeongbu Phase 2                     | KR   | 17   | 12    | 9     | 13   | 13   | 64  | H     | \$104    |
| Suseo/SRT                            | KR   | 9    | 2     | 1     | 2    | 9    | 23  | L     | \$124    |
| Al-Boraq                             | MA   | 5    | 0     | 1     | 1    | 3    | 10  | M     | \$49     |
| HSL Zuid                             | NL   | 15   | 10    | 9     | 7    | 15   | 56  | M     | \$189    |
| Gardermobanen                        | NO   | 13   | 7     | 2     | 11   | 9    | 42  | M     | \$35     |
| Ankara-Istanbul YHT                  | TR   | 8    | 6     | 4     | 3    | 8    | 29  | M     | \$88     |
| Ankara-Konya YHT                     | TR   | 2    | 1     | 1     | 1    | 3    | 8   | L     | \$11     |
| Ankara-Sivas YHT                     | TR   | 3    | 1     | 1     | 1    | 5    | 11  | L     | \$63     |
| Ankara (Polatli)-İzmir YST           | TR   | 3    | 1     | 2     | 3    | 5    | 14  | L     | \$70     |
| THSR                                 | TW   | 3    | 2     | 2     | 4    | 6    | 17  | M     | \$227    |
| HS1/CTRL                             | UK   | 16   | 7     | 9     | 6    | 12   | 50  | L     | \$216    |
| HS2                                  | UK   | 19   | 13    | 9     | 12   | 15   | 68  | H     | \$877    |

| Line               | Ctry | Opp. | Legal | Route | Env. | Pol. | CFI | Conf. | \$M / km |
|--------------------|------|------|-------|-------|------|------|-----|-------|----------|
| California HSR IOS | US   | 15   | 11    | 7     | 5    | 12   | 50  | M     | \$170    |

*Table B-2. Community Friction Index: component scores and aggregate, with cost per kilometre (n = 69).*

## Appendix C. Regression output

### C.1 Summary statistics across all specifications

| Model                  | k        | R <sup>2</sup> | adj R <sup>2</sup> | AIC         | BIC          | σ (log)      |
|------------------------|----------|----------------|--------------------|-------------|--------------|--------------|
| ECIw only              | 1        | 0.362          | 0.353              | +13.4       | +17.8        | 0.263        |
| CFI only               | 1        | 0.103          | 0.089              | +36.9       | +41.4        | 0.312        |
| Year only              | 1        | 0.186          | 0.174              | +30.2       | +34.6        | 0.297        |
| <b>ECIw + CFI</b>      | <b>2</b> | <b>0.422</b>   | <b>0.405</b>       | <b>+8.5</b> | <b>+15.2</b> | <b>0.252</b> |
| ECIw + Year            | 2        | 0.476          | 0.460              | +1.8        | +8.5         | 0.240        |
| CFI + Year             | 2        | 0.249          | 0.226              | +26.6       | +33.3        | 0.287        |
| ECIw + CFI + Year      | 3        | 0.514          | 0.492              | -1.4        | +7.5         | 0.233        |
| PBO model              | 6        | 0.660          | 0.627              | -20.1       | -4.4         | 0.199        |
| PBO model + CFI        | 7        | 0.709          | 0.675              | -28.7       | -10.9        | 0.186        |
| PBO model + ECIw       | 7        | 0.696          | 0.661              | -25.8       | -7.9         | 0.190        |
| PBO model + ECIw + CFI | 8        | 0.721          | 0.684              | -29.7       | -9.6         | 0.184        |
| ECIw + CFI + region    | 5        | 0.608          | 0.577              | -12.2       | +1.2         | 0.213        |

Table C-1. Summary statistics for all specifications ( $n = 69$ ).  $k$  is the number of predictors excluding the constant. PBO model = tunnel share, elevated share, start year and indicators for Asia, U.K./U.S. and rest of world.

### C.2 Bivariate model

| Predictor | Coefficient | Std. error | t-statistic | p-value |
|-----------|-------------|------------|-------------|---------|
| Constant  | +0.8741     | 0.1543     | 5.67        | < 0.001 |
| ECIw      | +0.0144     | 0.0024     | 6.04        | < 0.001 |
| CFI       | +0.0046     | 0.0018     | 2.62        | 0.011   |

Table C-2. Bivariate model.  $F(2, 66) = 24.1$ ,  $p < 0.001$ .  $VIF = 1.02$ .

### C.3 Trivariate model

| Predictor   | Coefficient | Std. error | t-statistic | p-value |
|-------------|-------------|------------|-------------|---------|
| Constant    | +0.8682     | 0.1426     | 6.09        | < 0.001 |
| ECIw        | +0.0132     | 0.0022     | 5.96        | < 0.001 |
| CFI         | +0.0037     | 0.0016     | 2.27        | 0.026   |
| Year – 2000 | +0.0085     | 0.0024     | 3.51        | < 0.001 |

Table C-3. Trivariate model.  $F(3, 65) = 22.9$ .  $VIFs$  1.04, 1.04, 1.05.

### C.4 PBO model with friction

| Predictor      | Coefficient | Std. error | t-statistic | p-value |
|----------------|-------------|------------|-------------|---------|
| Constant       | +1.6105     | 0.0795     | 20.26       | < 0.001 |
| Tunnel share   | +0.7718     | 0.1242     | 6.21        | < 0.001 |
| Elevated share | +0.9416     | 0.1912     | 4.92        | < 0.001 |

| Predictor         | Coefficient | Std. error | t-statistic | p-value |
|-------------------|-------------|------------|-------------|---------|
| Start year – 2023 | +0.0073     | 0.0023     | 3.20        | 0.002   |
| Asia              | -0.1981     | 0.1150     | -1.72       | 0.090   |
| Rest of world     | -0.0017     | 0.0938     | -0.02       | 0.986   |
| U.K./U.S.         | +0.5367     | 0.1218     | 4.40        | < 0.001 |
| CFI               | +0.0055     | 0.0017     | 3.20        | 0.002   |

Table C-4. PBO model with CFI. Dependent variable  $\log_{10}$  cost per km; tunnel and elevated as shares of route length; start year relative to 2023.  $R^2 = 0.709$ .

## C.5 Most influential observations (bivariate model)

| Line                | Observed | Predicted | Residual | Leverage | Cook's D     |
|---------------------|----------|-----------|----------|----------|--------------|
| HS2                 | 2.943    | 2.138     | +0.805   | 0.100    | <b>0.420</b> |
| Turin-Milan         | 2.198    | 1.596     | +0.602   | 0.045    | <b>0.094</b> |
| California HSR IOS  | 2.229    | 1.796     | +0.434   | 0.070    | <b>0.079</b> |
| Ankara-Konya YHT    | 1.055    | 1.458     | -0.403   | 0.079    | <b>0.079</b> |
| Sud-Est             | 1.095    | 1.543     | -0.448   | 0.056    | <b>0.067</b> |
| Madrid-Barcelona    | 1.593    | 2.132     | -0.539   | 0.035    | 0.057        |
| Gardermobanen       | 1.543    | 2.104     | -0.560   | 0.032    | 0.056        |
| HS1/CTRL            | 2.334    | 1.954     | +0.380   | 0.046    | 0.038        |
| Venta de Baños-León | 1.193    | 1.409     | -0.216   | 0.095    | 0.028        |
| Hokuriku to Tsuruga | 2.380    | 1.945     | +0.434   | 0.027    | 0.028        |

Table C-5. The ten observations with the largest Cook's distance in the bivariate model, in  $\log_{10}$  units. Threshold  $4/n = 0.058$ .

## Appendix D. Comparison with the April analysis

The April edition analysed 16 projects with costs that were not taken from a dataset and scores assigned with knowledge of each project's outcome. An intermediate version of this edition re-ran the April models on the same projects with Transit Costs data and the April scores. Table D-1 compares the three stages; Table D-2 compares the April and blind scores for the 16 projects.

| Measure                         | April 2026 (16, April costs and scores) | 16 projects, sourced costs, April scores | 69 lines, sourced costs, blind scores |
|---------------------------------|-----------------------------------------|------------------------------------------|---------------------------------------|
| R <sup>2</sup> , ECIw alone     | 0.57                                    | 0.54                                     | 0.36                                  |
| R <sup>2</sup> , CFI alone      | 0.81                                    | 0.40                                     | 0.10                                  |
| R <sup>2</sup> , bivariate      | 0.90                                    | 0.66                                     | 0.42                                  |
| CFI p-value, bivariate          | < 0.001                                 | 0.054                                    | 0.011                                 |
| Completion year                 | Not significant (p = 0.83)              | p = 0.14                                 | p = 0.001                             |
| ALTO bivariate, \$M / km        | \$142                                   | \$229                                    | \$191                                 |
| 95 % interval spans a factor of | 3.5                                     | 19                                       | 11                                    |

Table D-1. Key results at each stage. The April interval was narrow because the April costs lay unusually close to the fitted line.

| Project                      | ECIw April | ECIw blind | CFI April | CFI blind |
|------------------------------|------------|------------|-----------|-----------|
| California HSR               | 87         | 48         | 78        | 50        |
| Germany Stuttgart–Ulm        | 63         | 79         | 75        | 28        |
| Germany Cologne–Frankfurt    | 74         | 65         | 50        | 57        |
| Netherlands HSL–Zuid         | 69         | 71         | 40        | 56        |
| Japan Hokkaido               | 79         | 69         | 25        | 5         |
| Taiwan HSR                   | 81         | 82         | 30        | 17        |
| France TGV SEA               | 51         | 47         | 38        | 23        |
| France TGV Paris–Lyon        | 44         | 43         | 23        | 11        |
| Italy AV Rome–Naples         | 66         | 64         | 28        | 16        |
| Morocco Al Boraq             | 42         | 53         | 15        | 10        |
| UK HS1 (CTRL)                | 68         | 59         | 45        | 50        |
| UK HS2                       | 79         | 66         | 92        | 68        |
| Spain AVE Seville            | 54         | 54         | 18        | 11        |
| Spain AVE Barcelona          | 65         | 74         | 22        | 42        |
| China HSR (Beijing–Shanghai) | 60         | 77         | 12        | 11        |
| Rail Baltica                 | 48         | 52         | 60        | 59        |

Table D-2. April and blind scores for the 16 April projects (Stuttgart–Ulm scored as the Wendlingen–Ulm new line). Correlation between April and blind scores: ECIw 0.47, CFI 0.74.

The blind friction scores track the April scores reasonably closely ( $r = 0.74$ ) but are lower on average, by 9 points, most for projects whose controversy was about cost (HS2, California). The engineering scores agree less well ( $r = 0.47$ ). The largest change is California, whose costed section — the initial operating segment between Merced and Bakersfield — runs across the flat Central Valley: scored on

that section it is a moderate-complexity line (48), not the extreme one (87) the April score implied by including the mountain crossings. Stuttgart–Ulm moves the other way, because the Wendlingen–Ulm line it is costed as is largely in tunnel through karst.

## Appendix E. Blind scoring protocol

---

Scorers received the ECI descriptors (Appendix A), the CFI rubric of the friction note (twenty cumulative indicators per component), the line name, construction dates, length, and tunnel and elevated shares from the Transit Costs dataset. They were instructed as follows.

1. Do not look up, use or record any information about the project's construction cost, budget, cost per kilometre, cost overruns, cost escalation or cost-benefit results. Ignore any part of a source that mentions cost, and do not search with cost-related terms.
2. Score engineering on the physical characteristics of the line as built or planned — ground, structures, terrain, water, interfaces and climate — for the specific costed section, not the wider network.
3. Score friction from the documentary record, excluding any opposition, litigation, route change or political controversy whose subject is the project's cost, budget or overruns. Friction about routes, land, environment, noise, local impacts, need and non-cost governance counts.
4. Where evidence is equivocal, use the lower score. Give a justification for each index, at least one source, and a confidence rating (high, medium or low).

Lines were scored in eleven regional batches. Where the rubric's Canadian indicators (for example, access-to-information requests) had no equivalent, scorers treated them as met when a higher indicator was clearly met. Scores were not adjusted after the regressions were run.

## Appendix F. References

---

References are organised by theme. Items marked with a dagger (†) are primary-source government audit or parliamentary documents. Sources for the scores of individual lines are listed in the accompanying workbook.

### F.1 Reference-class forecasting and megaproject cost overruns

- Flyvbjerg, B. (2008). Curbing optimism bias and strategic misrepresentation in planning: reference class forecasting in practice. *European Planning Studies*, 16(1), 3–21.
- Flyvbjerg, B. (2014). What you should know about megaprojects and why: an overview. *Project Management Journal*, 45(2), 6–19.
- Flyvbjerg, B., & Gardner, D. (2023). *How big things get done*. Currency / Penguin Random House.
- Flyvbjerg, B., Holm, M. K. S., & Buhl, S. L. (2002). Underestimating costs in public works projects: error or lie? *Journal of the American Planning Association*, 68(3), 279–295.
- Priemus, H., Flyvbjerg, B., & van Wee, B. (Eds.). (2008). *Decision-making on mega-projects: cost–benefit analysis, planning and innovation*. Edward Elgar Publishing.

### F.2 High-speed rail cost data

- NYU Marron Institute, Transit Costs Project. High-Speed Rail dataset (spreadsheet supplied October 2026). <https://transitcosts.com/high-speed-rail/>
- Office of the Parliamentary Budget Officer. (2026). *Canada's Eastern Corridor High-Speed Rail: A Cost Analysis using Global Data*. RP-2627-009-C, 1 October 2026. †
- Goldwyn, E., Levy, A., Ensari, E., & Chitti, M. (2024). *How to Improve Domestic High-Speed Rail Project Delivery*. Transit Costs Project.
- Mok, J., Chitti, M., & Shalaby, A. (2025). *Understanding the drivers of transit construction costs in Canada: a comparative study*. University of Toronto School of Cities.

### F.3 Project sources cited in the text

- UK National Audit Office. (2012). *The completion and sale of High Speed 1*. HC 1834. †
- UK National Audit Office. (2020). *High Speed Two: a progress update*. †
- California High-Speed Rail Authority. (2024). *2024 Business Plan and Basis of Estimate*.
- California State Auditor. (2018). *California High-Speed Rail Authority. Report 2018-108*. †
- ERR News. (2026, January 20). Rail Baltica phase one will cost at least €24 billion, EU auditors say [reporting European Court of Auditors findings]. <https://news.err.ee/1609915685/>
- World Bank. PPP conversion factor, GDP (LCU per international \$), indicator PA.NUS.PPP, 2023 values for Estonia, Latvia and Lithuania.

### F.4 Community opposition and infrastructure

- Dear, M. (1992). Understanding and overcoming the NIMBY syndrome. *Journal of the American Planning Association*, 58(3), 288–300.
- Devine-Wright, P. (2009). Rethinking NIMBYism: the role of place attachment and place identity in explaining place-protective action. *Journal of Community and Applied Social Psychology*, 19(6), 426–441.

### F.5 Statistical methodology

- Belsley, D. A., Kuh, E., & Welsch, R. E. (1980). *Regression diagnostics: identifying influential data and sources of collinearity*. New York: Wiley.
- Cook, R. D. (1977). Detection of influential observation in linear regression. *Technometrics*, 19(1), 15–18.
- Fox, J. (2015). *Applied regression analysis and generalized linear models* (3rd ed.). Thousand Oaks, CA: SAGE.
- Seabold, S., & Perktold, J. (2010). statsmodels: econometric and statistical modeling with Python. *Proceedings of the 9th Python in Science Conference*, 92–96.

## About the ALTO HSR Citizen Research Initiative

---

The ALTO HSR Citizen Research Initiative (CRI) is an independent, non-partisan citizen research project critically examining Canada's proposed ALTO high-speed rail corridor.

All Initiative research is published openly. Reference data, scoring tables, and regression outputs are reproducible from the published material. Corrections are published as revised editions, as with this one.

[altohsrcitizenresearch.ca](http://altohsrcitizenresearch.ca) · [citizenresearch.ca](http://citizenresearch.ca)