Models a Finance Ministry Can Audit

The Climate Policy Assessment Tool, rebuilt and validated: section 5 in ten slides

A model a ministry can use must be small enough to state on a few pages, built at least twice, and tested against data and other models. CPAT now is: fourteen equations, two rebuilds, a public register of its errors, and a check of its elasticities.

Stephen J. Stretton · writing in a personal capacity

Ministries rely on CPAT, but its method was hard to check

180+

countries covered by the Climate Policy Assessment Tool

57,000

words in the mitigation documentation, which disagreed with the model in places

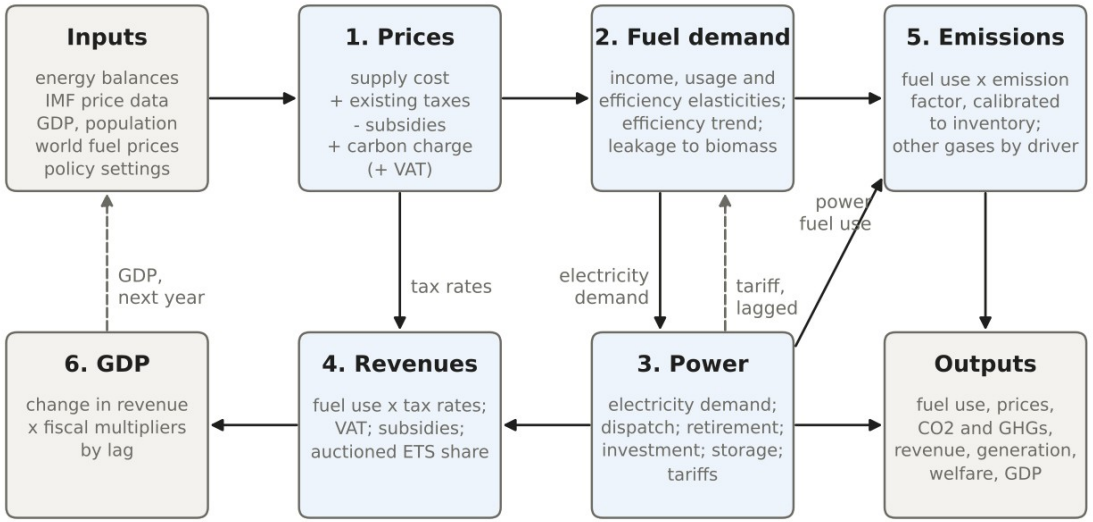
231

legacy bugs in the public register (release 1.1.0); 24 of the 228 at release 1.0.2 were high priority

A recommendation from a model nobody can audit is a recommendation taken on trust.

Source: CPAT Documentation (2023); papers 5.1 and 5.2; CPAT-AI register.

The mitigation module is a one-way chain of six blocks and fourteen equations



Prices, demand, power, emissions, revenue, GDP: each a few equations long.

$50/t in the example: industrial coal use falls about a third, road fuel about 6%.

Two judgment calls drive power results: an unestimated logit and caps on build rates.

Source: Paper 5.1, Figure 1 and Table 1.

Power-sector results rest on two judgment calls, and one has a fixable flaw

The logit is a penalized cost minimum: its curvature is a penalty weight in $/MWh.

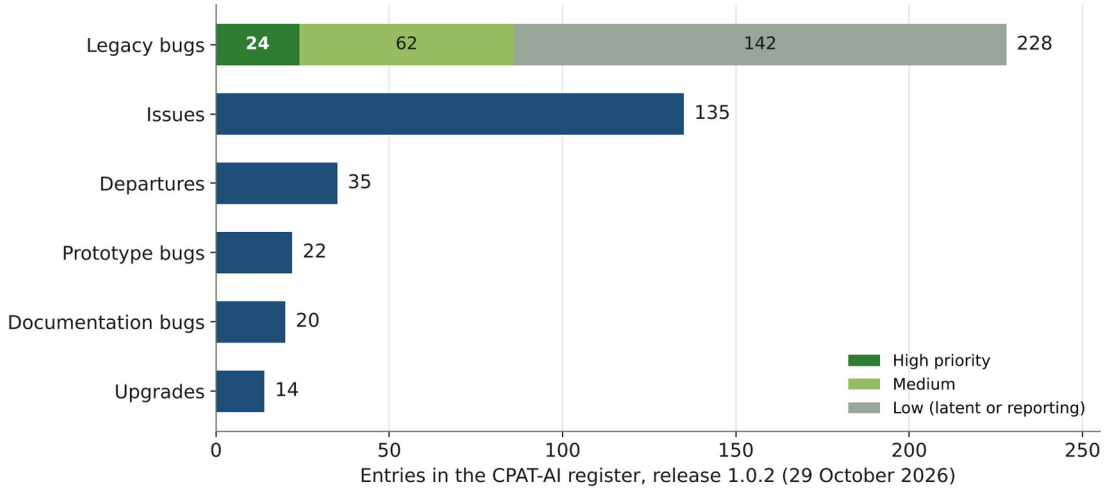
Its flaw: listing a technology twice changes the shares; solar listed twice rises from 36% to 58% of new investment.

A quadratic penalty lets uncompetitive technologies fall to zero.

Cost of capital: a country rate plus a technology premium by default; income-group rates are an option.

Source: Paper 5.1, Section on investment.

Building the model twice found errors a single audit would not



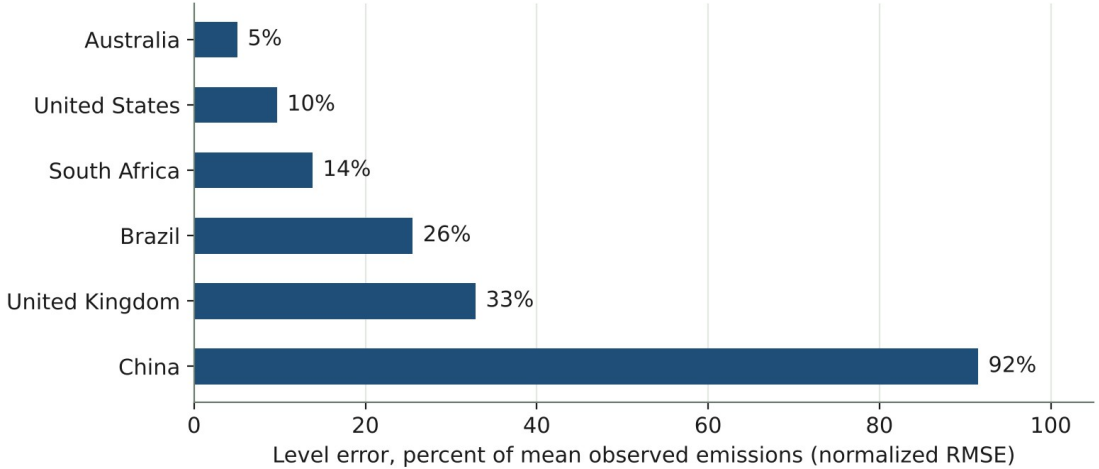
One rule: fix bugs, copy and log deliberate choices.

228 legacy bugs at release 1.0.2 (24 high), 135 issues, 20 documentation errors.

The second build found seven bugs in the first.

Source: Paper 5.2; CPAT-AI register.

Back-casts show a model good for ranges, not point forecasts

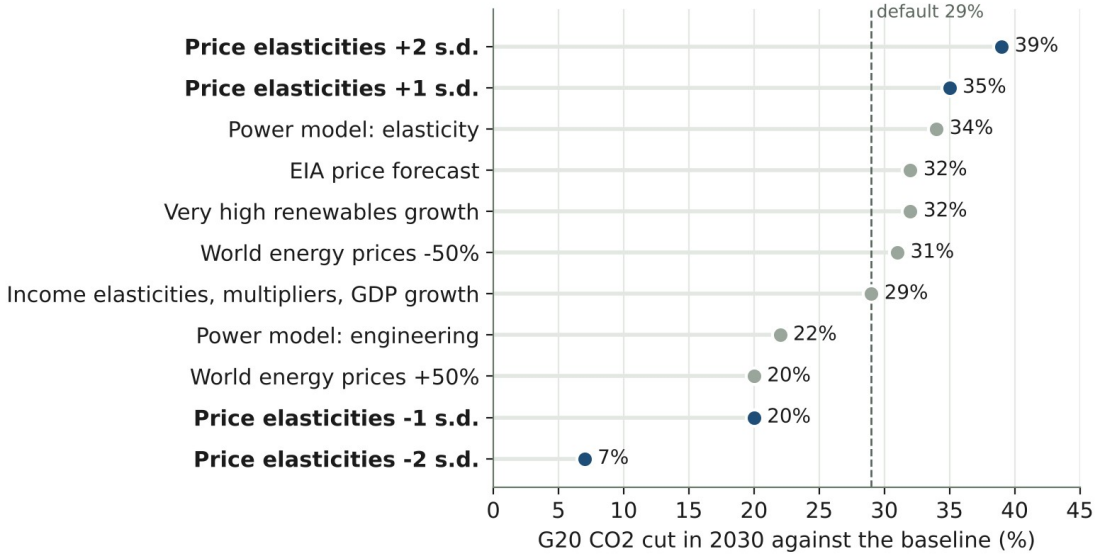


Level error from 2000: 5% of mean emissions for Australia, 92% for China.

Every published check predates the bug fixes and should be rerun.

Source: Paper 5.3, Figure 1.

The price elasticities move the answer more than anything else

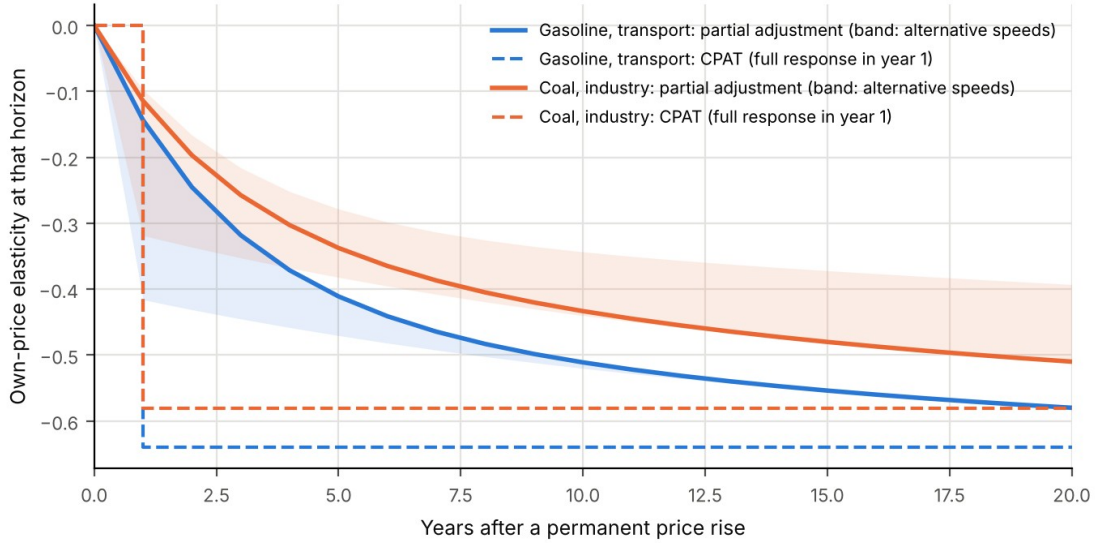


$75/t cuts G20 CO_2 29% by 2030 by default, 7–39% as elasticities move.

Against real carbon prices: about 0.7–0.8% a year at $10/t, against 0.5% in the most comparable study (author's calculation).

Source: Paper 5.3, Figure 2 and Table 4.

Most elasticities pass; applying the long-run response in year one does not



8 of 10 in range, six cautious; diesel in industry and gasoline in services too high.

Road gasoline −0.64: twice the bias-corrected −0.31.

Timing: five-year responses overstated by 45–72%.

Source: Paper 5.4, Tables 1 and 3, Figure 2.

Extend the model outside its core, so the core stays small enough to audit

6.1: the Total Effective Carbon Price, one scale for any policy mix.

6.2: the CBAM Policy Tool, exporters' exposure to the EU border charge.

6.3: public investment, priced at the cost of capital countries face.

6.4 and 6.5: one activity layer across models; growth, revenue and debt.

Source: Section 6.

What to do

What section 5 shows

Legible: fourteen equations, one page each block.

Built twice: 231 legacy bugs registered and fixed.

Good for ranges: not point forecasts.

Timing matters more than any single elasticity.

Actions

Ministries: read CPAT results as ranges; ask for the elasticity case and the error register.

The CPAT team: rerun the published checks on the fixed model; add adjustment speeds and the quadratic investment rule.

Funders: pay for a second build of any model that sets policy.