WORKING PAPER | VERSION 0.3 | OCTOBER 2026

Models a Finance Ministry Can Audit

The Climate Policy Assessment Tool, rebuilt and validated: an overview of section 5

Stephen J. Stretton*

ABSTRACT

Finance ministries decide on carbon taxes, fuel subsidy reform and climate investment with the help of models they cannot inspect. This section is the paper-length companion to the CPAT-AI project, which rebuilt the Climate Policy Assessment Tool so that every step can be read and tested. It states the mitigation module's method in fourteen equations; describes the two AI-assisted rebuilds and the legacy bugs they found (231 in the register, 24 of them high priority); gathers the published validation, which supports the model for ranges and policy comparisons but not point forecasts; and checks the price elasticities that drive every result, finding most defaults within the published range and the timing of the response the larger problem. The through-line is that 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. The extensions ministries ask for are written as modular upgrades in section 6.

Overview of section 5, version 0.3, October 2026. Stephen J. Stretton, writing in a personal capacity. The views expressed are the author's own and do not represent those of any institution with which he is or has been affiliated. Companion to the public CPAT-AI repository, https://github.com/cpmodelai/cpat_ai. Each paper below comes as a short policy brief and a full working paper.

*Writing in a personal capacity.

A model nobody can audit is a recommendation taken on trust

Ministries set carbon prices and sequence reforms with the help of a model few can check. The Climate Policy Assessment Tool (CPAT), built by the World Bank and the International Monetary Fund, estimates what carbon pricing and related policies do to energy use, emissions, revenue, growth and households for more than 180 countries (CPAT Team 2023). Until recently almost nobody could check its numbers: the legacy model was a very large Excel workbook, and its 57,000-word mitigation documentation disagreed with the model in places (Stretton 2026c). A corrected edition of the documentation, fixing the twenty errors the rebuild found, was published with the CPAT-AI project on 31 October 2026.

The papers in this section make the model legible and check it; section 6 extends it. Each is written so that an economist can understand it in an afternoon, and each points to the code and documentation in the public CPAT-AI repository for the detail.

What the model computes (paper 5.1)

The mitigation module is a one-way chain of six blocks and fourteen equations (paper 5.1). Retail prices are built up from supply cost, existing taxes, subsidies and the carbon charge. Fuel use responds through income, usage and efficiency elasticities. An engineering power model dispatches, retires and invests. Emissions and revenue follow by accounting identities, and revenue moves GDP a year later through fiscal multipliers. In the paper's stylized example a $50 per tonne carbon price cuts industrial coal use by about a third and road fuel use by about 6 percent (Stretton 2026c).

Power-sector results rest on two judgment calls. New investment is shared out by an unestimated logit rule, and build rates are capped. The logit turns out to be an entropy-penalized cost minimum, so its curvature is a penalty weight in dollars per megawatt-hour, and its real flaw is that listing a technology twice changes the answer: solar listed twice rises from 36 to 58 percent of new investment under the alternative velocity rule. The paper proposes a quadratic-penalty rule that lets uncompetitive technologies fall to zero. Investment is priced at a country cost of capital plus a technology premium by default; the income-group rates of 7.5–15 percent that the documentation describes are an option (Stretton 2026c).

Road transport is the module's weak joint. CPAT describes it three times, in a fuel equation, an engineering fleet model and a co-benefits module, and the three are not consistent: the co-benefits module always takes vehicle-kilometers from the fleet model, so the change in driving it values need not match the modelled change in fuel use (CPAT-AI register issue I100).

Rebuild it twice (paper 5.2)

The model was rebuilt twice in October 2026, with AI coding agents doing most of the construction (paper 5.2): once as a readable Excel workbook generated by code (Excel-AI) and once as a Python model (Coded CPAT) that reproduces it number for number. One rule governed both: errors in the legacy model are fixed and never copied; questionable but deliberate choices are copied and logged.

The rebuild found what a single audit would not. It registered 228 legacy bugs at release 1.0.2, 24 of them high priority because they change standard results, together with 135 issues and 20 documentation errors; the register lists 231 legacy bugs at release 1.1.0 (Stretton 2026b). Writing the model a second time found seven bugs in the first rebuild. The lesson follows the long experience with independently written software: specify a policy model by shared variable codes, build it at least twice and test every output at every change, and remember that verification is not validation.

Check it against data and other models (papers 5.3 and 5.4)

The published validation supports CPAT for ranges and policy comparisons, not point forecasts (paper 5.3). It reproduces its base year by construction. In back-casts from 2000 its error in the level of emissions is 5 percent of mean emissions for Australia and 92 percent for China. A $75 per tonne tax cuts G20 CO_2 emissions by 29 percent in 2030 by default, and by 7–39 percent as the price elasticities move within their uncertainty. On a common growth-rate basis CPAT's response to a $10 carbon price, about 0.7–0.8 percent a year, is above the 0.5 percent of the most comparable study of real carbon prices (the author's calculation). Every published check predates the bug fixes and should be rerun (Stretton 2026d).

Most price elasticities pass; the timing of the response does not (paper 5.4). Once CPAT's own calibration targets are removed from the evidence, eight of ten main fuel-sector defaults lie within the published range and six are on the cautious side; diesel in industry and gasoline in services are too high, and road gasoline (−0.64) is twice the estimate corrected for publication bias (−0.31). The larger problem is structural: CPAT applies the full long-run response within one year, which overstates five-year responses by 45–72 percent in the central case, and it has almost no cross-price terms (Stretton 2026a).

Extend it where policy needs it (section 6)

Extensions belong outside the core, so that the core stays small enough to audit. Section 6 adds five modular upgrades, each of which reads CPAT's prices, fuel use and emissions and adds one piece of economics: the Total Effective Carbon Price, which measures any policy mix as one carbon tax (paper 6.1); the CBAM Policy Tool, which measures exporters' exposure to the EU carbon border adjustment (paper 6.2); a public investment extension that prices the clean capital and tests whether the budget should pay for it (paper 6.3); a common activity layer that links CPAT to SISEPUEDE, the CBAM tool and the Future Technology Transformations models (paper 6.4); and the Macro-Fiscal-Energy Scenario Tool, which carries a climate program through to growth, revenue and debt (paper 6.5).

The papers in this section

No.	Paper	Status
5.1	The CPAT Mitigation Module: Methods in Brief	Working paper v0.2, brief and deck (31 October 2026)
5.2	CPAT-AI: Rebuilding a Policy Model as a Transparent, Testable Tool	Working paper v0.2, brief and deck (31 October 2026)
5.3	Validating CPAT: Baselines, Back-Casts and Model Comparisons	Working paper v0.2, brief and deck (31 October 2026)
5.4	Are CPAT's Price Elasticities Right? A Check Against the Evidence	Working paper v0.2, brief and deck (31 October 2026)

The instruments these models compare are in section 1. The modular upgrades, including the Total Effective Carbon Price (paper 6.1) that gives the common scale, are in section 6, and the diagnostics that apply the same standard country by country are in section 7.

References

CPAT Team. 2023. *CPAT Documentation: WB-IMF Climate Policy Assessment Tool*. World Bank; International Monetary Fund. https://cpmodelai.github.io/cpat_ai/.

Stretton, Stephen J. 2026a. “Are CPAT’s Price Elasticities Right? A Check Against the Evidence.” Unpublished manuscript.

Stretton, Stephen J. 2026b. “CPAT-AI: Rebuilding a Policy Model as a Transparent, Testable Tool.” Unpublished manuscript.

Stretton, Stephen J. 2026c. “The CPAT Mitigation Module: Methods in Brief.” Unpublished manuscript.

Stretton, Stephen J. 2026d. “Validating CPAT: Baselines, Back-Casts and Model Comparisons.” Unpublished manuscript.