Referee review of PC National modelling NHVDCF licensing scenarios

May 2026 — Tim Murray (Founding Partner, Subject-Matter Expert)

Key Points

1. The PC's approach is fit for purpose.
2. PC National (a comparative static, CGE model) is appropriate for assessing the macroeconomic impacts of the licensing scenarios, taking the inputs from BITRE work, scaling it for the PC model database, and adjusting (shocking) appropriate variables and industries in the PC model.
3. The exposition is highly detailed and transparent. Sources, shocks and adjustments are well documented, properties of the model and closure are clear, and results illustrated and explained clearly. PC staff were very responsive and forthcoming with additional explanations when asked.
4. There are a few areas of the PC work that could be examined in a little more detail, mostly via sensitivity analysis. These are unlikely to impact the broad magnitude of the GDP result for this application.

The PC's approach to modelling the NHVDCF licensing scenarios (PC scenarios 3L and 3U, which adopt BITRE derived input shocks, scaling them for the PC's model database, and running them in the PC's own national comparative static model with a long-run

closure) is broadly defensible, well-executed, and documented in a clear and transparent way.

The strongest positive point is that PC National is a reasonable vehicle for a long-run, indicative productivity-reform exercise of this nature to estimate broad magnitudes of macroeconomic impact (assuming the impacts are not particularly different between states and territories, or that at the least the top-down national result is unlikely to be changed by state-specific nuances which would need to be modelled bottom up), and the qualitative CGE mechanisms are well explained: lower road freight costs flow through downstream sectors, real wages adjust under fixed employment, and capital adjusts under the long-run closure. These are intuitive results for a straight-forward simulation, reflecting the model structure. Overall the results for the scenarios examined (in particular 3L) are small.

There are several areas of the PC work and appendix that are worth highlighting. These are not wrong per se, but are assumptions (or include assumption driven results) and are therefore important to unpack/discuss in this report.

1. The PC relies on the BITRE inputs to derive a shock, and largely takes this as given. That's fine to translate to an approximate national shock, but the mechanisms used to derive the shock are not in the PC National model
2. The PC rightly flags that the industry in PC National is an imperfect representation of the targeted industry (it includes buses, taxis and others, which will have different cost and input mixes, and will have different levels of margin and regular activity/contribution).
3. The L and U scenarios apply two 'sensitivities' to the input shock base (a direct effect, and a coverage effect). These are just 2 sources of input uncertainty among many possible sources. Both of those are valid, but combining them invites the question of why just combine these two to get an upper and lower bound? Why not others? It could be preferable to present these two separately as illustrations of result sensitivities to each, even around a central estimate. Or just include a broad, 'rounded' upper and low scaling to avoid false precision.
4. There are some relationships in the macroeconomic results that look very similar. This could be sensible and correct, but is driven by closure decisions and attributes of the model. It could be worth unpacking them, or running closure sensitivity tests, since closure and specification can have a material impact on the

distribution of the elements of GDP for any CGE simulation. GDP is unlikely to be impacted, but the macroeconomic compositional story would be.

5. There are a few presentational points. In particular, that results are presented with a large number of decimal places giving a false sense of precision, and that caution needs to be taken in taking an arithmetic sum of different CGE simulations. CGE models don't show 'the real world' in a precise way, and are highly stylised (especially national comparative static models, with a broad-brush productivity shock), so avoiding false precision is important (especially if simulations are summed).

1 BITRE licensing work as a core input

The core input (shock) — translating BITRE's licensing reform effect into a Road Transport TFP shock — is good as a reduced-form approximation to be fed to the national CGE model. But the mechanisms of action could be explained a bit more. The mechanism is not that licensing reform raises productivity across all Road Transport inputs per se; it is that faster progression to HC/MC licences may allow some freight to shift/substitute to larger vehicles/adjust fleet composition etc, reducing cost per tonne-kilometre (no doubt discussed in great detail in the report proper). The appendix should therefore describe the shock as an equivalent cost-reducing technical change or fleet-composition/logistics-efficiency proxy, rather than a literal TFP improvement from licensing reform.

The biggest upstream uncertainty is BITRE's behavioural chain (which isn't the PC's work). The results rest on assumptions that shorter licence progression increases qualified drivers, that this relaxes a potentially binding driver-availability constraint, and that it induces some of the freight task to shift to larger vehicles. That might be plausible, but it is still assumption driven in the BITRE work. The appendix could specifically say that the CGE results are conditional on BITRE's substitution assumption, and ideally include or reference sensitivity around that assumption beyond the 'one number multiplication' mentioned below.

2 Important acknowledgement of the industry coverage

The most important PC National-specific limitation is the Road Transport mapping, and the margin/non-margin separation. The appendix is right to acknowledge that the IO road transport industry is both too narrow (excluding substantial in-house freight in retail, construction, mining, agriculture, etc) and too broad (including non-freight road transport

such as taxis, buses and other activities). Shocking only Road Transport (and factor augmenting technical change in particular) may understate the effect on the full heavy-vehicle freight footprint, while the current scaling also tries to remove over-included non-freight activity. A sensitivity analysis (illustrating the extent of the assumption on macro aggregates) would improve credibility.

The PC correctly scales their shock for that (further mentioned below), but it's important to note that the compositional story between margins and regular activity is different, and scaling alone doesn't address that (since heavy vehicles will have a different composition in regular activity than they do in the margin service).

Ideally, the heavy vehicle transport sector would be split out separately from the database, reflecting the margin and regular industry differences as well.

3 Uncertainty bounds on shock inputs

The main scenarios for 3* define lower and upper case shocks which actually combine two different uncertainties: BITRE's direct productivity estimate and the affected share of Road Transport. In 3L the shock uses 0.11% * 96.1% * 64.3%; in 3U it uses 0.32% * 96.1% * 85.9%. That combines a lower direct effect (0.11) with a smaller coverage (64.3), and an upper direct effect (0.32) with a broader coverage (85.9). This is not necessarily wrong, but it should be explicit, because there are many other factors that COULD have been included, and each of the two that are included have their own impact. It isn't a global 'lower' or 'upper' bound because it's not including or testing a wide range of factors.

An alternative would be to present each of the 2 separately, to highlight the impact of each. Or try to include a larger number of sensitivity factors, and run several simulations with combinations of inputs to illustrate the broader distribution of potential impacts around some central estimate. A cruder, more broad brush alternative (also avoiding false precision) would be to put rounded 'whole number' upper and lower bounds which include those effects and potentially others.

4 Investment and GDP (and other GDP components)

There's a high degree of similarity between the macroeconomic GDP result and the investment result (as well as the consumption result), likely driven by the model closure. Under the long-run closure, as well as more specific PC choices (fixed employment, fixed rate of return, capital adjustment, and fixed savings behaviour) near one-for-one movements in investment and GDP are close to being a structural property of the model (independent of policy, scenario, simulation, database, etc). The appendix could explain

this so readers do not interpret investment as an independent corroborating result driven from the bottom up. Put differently, the scenario isn't driving foreign investors to want more stake in Australian businesses based on RoR or driving Australian households to save more which is then transmitted to investment, but is a necessary consequence of "activity expands", therefore capital must expand (since labour is fixed), therefore investment expands. Similarly, real wages rising by more than GDP is consistent with fixed employment and a higher capital-labour ratio (so different closures, and demand theory — see below — could change that result), but could be signposted for non-CGE readers.

This need not be addressed, it's just flagged for awareness. But there are two ways it could be addressed: the first would be to document and describe the closure attributes driving those properties of the results; the second would be to run sensitivity tests of an identical model with alternative fiscal and macroeconomic closures.

It's also noteworthy that consumption moves pretty closely with GDP. This is a property of the settings and theory in the current version of PC National: households' consumption is Cobb-Douglas (a utility specification that could be made much better in PC National, since it implies fixed expenditure shares, unit own-price elasticity, no cross-price substitution except via income effects and consequent expenditure shares), and other closure settings effectively leave real consumption 1-to-1 linked back to GDP via income. Large scale, applied CGE models typically do not use Cobb-Douglas utility functions, and testing the sensitivity of this assumption on results would be very valuable.

WIth the investment binding discussed above, that consumer theory and closure (along with a more-or-less fixed government), means that the trade balance is forced as a residual to adjust to preserve the balance. In that sense, the composition of GDP is pretty shackled. That said, for this exercise, these model and closure properties are unlikely to material change the magnitude of the GDP result.

5 Precision and reporting

Results are presented to a lot of decimal places. While the results are very small, being as small as they are it makes sense to only present 1 (or maybe 2) significant figures (so probably 2 and at most 3 decimal places). The 3L result is so small that it is best read as "small and positive", not as a precise point estimate at all (especially in light of the points covered above).

It’s also important to be careful when adding scenarios together unless a combined simulation is run. CGE models are used precisely because wages adjust with capacity constraints, so the scenarios run jointly would potentially be more muted (since the models are linearisations of non-linear models, and if run with a multistep method, should approximate the true non-linear solution of the model theory).