

Is the migration of the model to WiNDC operationalized in a transparent and defensible way?

Yes, the migration is both transparent and defensible.

The WiNDC database is constructed using an open-source build routine, which is publicly available and documented. This ensures reproducibility and transparency. This version includes disaggregated household data and is built from publicly available sources like BEA and Census data.

The core module of WiNDC builds state-level social accounting matrices. The household module adds income and expenditure detail using CPS data. SAGE integrates WiNDC with additional data from BEA, BLS, EIA, and others to align with its sectoring and modeling framework.

WiNDC's summary sectors are disaggregated to match SAGE's 23-sector scheme using detailed BEA tables, BLS compensation data, and EIA energy data. Crude oil and natural gas extraction are further split using EIA and Census trade data. Matrix balancing routines ensure consistency with steady-state growth assumptions.

Specific GAMS and R scripts are named and described showing how WiNDC data is processed and integrated.

There are, of course, a number of aspects of the calibration that require judgement calls – such as the treatment of intranational trade flows and the treatment of capital accounts. There are multiple ways one could approach these issues, but in each case the authors have described a clear, defensible logic for their approach and presented metrics that would help readers evaluate the calibration in a transparent way.

Is the implementation of WiNDC data within SAGE clearly explained in the documentation?

Yes, the documentation provides a clear and detailed explanation. The documentation outlines the benchmark data and the rationale for using WiNDC. It also describes how WiNDC sectors are disaggregated and calibrated to match SAGE's structure. Appendices and examples show how the data is used in simulations and policy modeling. The documentation includes code-level references and file names for transparency and reproducibility.

Recommendations

- a. Recommendations that should be addressed prior to using SAGE v. 3.0 in a regulatory analysis of an EPA rule
 - None.
- b. Recommendations that should be prioritized in the near-term (1 – 2 years) but do not prevent use of SAGE v. 3.0 in a regulatory analysis
 - There could be greater discussion of the differences in the treatment of intraregional trade and margins in WiNDC versus IMPLAN as well as the consequences of the different assumptions for the policy experiments presented in the comparison report.
- c. Longer-run recommendations (>2 years) for improvements to SAGE and areas where more research is needed.
 - **Develop validation framework for subnational trade and transport assumptions:** Both WiNDC and IMPLAN rely on assumptions due to limited data. WiNDC uses FAF data with transshipment assumptions; IMPLAN uses a gravity model. A research agenda to validate or improve these assumptions would be valuable. The evidence in the trade literature suggests that distance matters a lot for trade in general and subnational trade in the US specifically (see Hillberry and Hummels, 2008). Moreover, the datasets treat transport margins differently. Both aspects of the model calibration have the potential to be important for evaluating the regional incidence of regulations and the impacts of regulations related to the transport sectors, so they seem worth thinking harder about.
 - **Improve characterization of capital ownership and investment flows:** Differences in household savings and capital income across datasets stem from assumptions about national closure and CPS/NIPA alignment. More granular data or modeling of foreign ownership could improve

realism with respect to the regional distribution of capital to the extent that this is possible.

- **Add realism to labor markets.** Currently, the description of labor markets is quite simple – markets clear regionally – but there is no differentiation of workers by skill, occupation, or industry for example. Given how central the evaluation of regulatory incidence on workers is in many policy exercises, it would be worth considering the extent to which more detail could reasonably be added to the model on this front.

Smaller items

- I found two, small syntax errors on lines 148 and 157 of `calibrate_les.gms` in the calls to `csv2gdx`. The lists passed to the value and index options should not be in parentheses.

References

- Russell Hillberry, David Hummels, Trade responses to geographic frictions: A decomposition using micro-data, *European Economic Review*, Volume 52, Issue 3, 2008, Pages 527-550, <https://doi.org/10.1016/j.euroecorev.2007.03.003>.