

Comparing Subnational Social Accounting Matrices for use in the SAGE Model*

Andrew Schreiber and Joseph M. Perla

National Center for Environmental Economics, Environmental Protection Agency, USA

October 1, 2025

1 Introduction

Access to high quality data has been a major obstacle to producing economy-wide analysis using computable general equilibrium (CGE) models in the United States. This class of models (as well as other types of economy-wide frameworks) use social accounting matrices as its primary data source. These data are constructed to capture a static representation of the inter-industry transactions and flows of incomes and expenditures in an economy (Miller and Blair, 2009). They are created by combining several distinct data sources to reflect a simplified view of the economy in which markets clear and the value of production inputs equals the value of its outputs. In the United States, the Bureau of Economic Analysis (BEA) produces supply and use tables at the national level based on information from the quinquennial economic census and several additional annual survey-based data sources (Horowitz and Planting, 2009). They do not, however, produce a unifying dataset that links characteristics of subnational economies in comparable ways.

For models with subnational detail, the choice set for subregional social accounting matrices is sparse, and had, until recently, been dominated by proprietary options such as IMPLAN (IMPLAN Group LLC., 2016). Though widely used in the CGE modeling literature (e.g., Marten et al., 2024; Yuan et al., 2019; Ross, 2008; Webster et al., 2022; Wing et al., 2025),¹ use of this data source can represent a significant cost burden for existing users looking to update their model or for new users interested in learning about the modeling framework. Furthermore, because of its proprietary nature, it is not possible to observe or modify the data generating process and underlying assumptions needed to produce the IMPLAN dataset nor is it possible to distribute the source data for reproducing analyses.

***Disclaimer:** The views expressed in this paper are those of the author(s) and do not necessarily represent those of the U.S. Environmental Protection Agency (EPA). No official Agency endorsement should be inferred. All errors are our own. Declarations of interest: none.

¹This data source is also used often in regional economic impact assessments (e.g., Bae and Dall’erba (2016); Parajuli et al. (2018); Deller and Hadachek (2024)).

In its advice on using economy-wide models to evaluate environmental policies, the Science Advisory Board (SAB) recommended that EPA collaborate with external researchers to establish an open-source dataset for conducting economy-wide analysis to reduce barriers to modeling and improve the transparency of modeling results (SAB, 2017). The Wisconsin National Data Consortium (WiNDC) has since been developed and has largely achieved these goals. Using the BEA national input output data as its starting point, WiNDC produces both an open-source state-level database based on publicly available source datasets (many of which are similarly used in the creation of IMPLAN) and a fully transparent build routine or set of computer code that can be used to reproduce the creation of the database (Rutherford and Schreiber, 2019). Relative to its first release, the WiNDC database has expanded available resolutions with comparable sectoral, regional, and household-level detail to IMPLAN.² WiNDC data has also been used in several peer reviewed studies (Akin-Olcum et al., 2021; Bistline et al., 2024; Brown et al., 2024; Becker et al., 2025; Brown et al., 2025).

The SAB’s review of EPA’s SAGE model recommended that the agency “assess the feasibility” of transitioning toward the open-source dataset for use in the model (SAB, 2020). To support that transition, this paper documents the similarities and differences across comparable IMPLAN and WiNDC datasets and assesses how observed discrepancies translate into modeled differences when evaluating hypothetical environmental regulatory standards using a static version of the SAGE model. Illustrative sector-specific environmental regulatory scenarios are based on scenarios explored in Marten et al. (2019). We find that while there exists some differences in the data, they produce relatively small effects on aggregate social cost estimates. Furthermore, when the data are aggregated to the SAGE level of resolution (four regions, twenty-three sectors), the data differences attenuate because both datasets are largely based on the BEA national tables. In general, the open-source alternative (WiNDC) has the benefit of allowing data users to assess underlying assumptions in the creation of the dataset that may affect model results. Doing so using IMPLAN is not possible.

The rest of this paper is organized as follows. In Section 2, we provide an overview of each database and offer a quantitative comparison across the components of each database that characterizes production, trade, and consumption. We then describe the outcomes of running illustrative environmental regulatory scenarios using a static version of the SAGE model in Section 3 and conclude in Section 4. Appendix A includes several additional quantitative dataset comparisons that extend the discussion from Section 2 for the interested reader.

2 Data Overview

We conduct a systematic evaluation of the key similarities and differences between WiNDC version 4.0.0 and IMPLAN data for 2016. The purpose of this evaluation is not necessarily to suggest one

²SAB (2017) also notes that “building such datasets is resource-intensive and time-consuming to do well, so this should be regarded as an iterative task to be completed over time with contributions made by EPA and nongovernmental entities.” Indeed the version of WiNDC that we use in this paper is version 4.0.0, which includes several improvements relative to the initial release of the database including heterogeneous household accounts denominated by income level.

data source is more correct than the other because it is challenging to ground truth several aspects of these data. Rather, we document where differences exist and explore the underlying assumptions in the WiNDC framework that may lead to these difference.

The evaluation is done in two steps. First, we consider the direct outputs of the WiNDC build routine and IMPLAN data with minimal post processing of the data (e.g., we do not impose any other matrix balancing constraints in this first step outside what is already incorporated in the respective build routines). In the second step, we run each data source through the build routine that produces the SAGE model benchmark dataset, integrating additional data sources from the Energy Information Administration (EIA), Bureau of Labor Statistics (BLS), National Bureau of Economic Research (NBER), Census, and BEA, and impose further balancing constraints on the data to align the capital and investment accounts with steady state growth assumptions in a Ramsey model framework (Marten et al., 2025).

The WiNDC data are constructed using both the “core” sub-module and the “household” sub-module in the WiNDC build routine retaining the most disaggregated characterization that the package offers without additional post processing. In the first step, the sectoral resolution retains the summary sector definitions from BEA (73 sectors). Households are constructed by approximating pure income quintiles using the Current Population Survey. We process the IMPLAN data using tools that accompanied past versions of SAGE as described in Marten et al. (2024), which are based on methods developed in Rausch and Rutherford (2008). Because the stock IMPLAN data has a higher level of sectoral detail and fixed household income bins, we aggregate the IMPLAN data to the summary sector level and five household groups which approximates, to the degree feasible with IMPLAN, the quintile approach in WiNDC.

Two systematic differences are worth mentioning qualitatively before presenting quantitative findings. First, the characterization of margins differs between the two datasets. Intermediate demand in the IMPLAN data assign transportation services and retail markups as additional inputs to production. In WiNDC, intermediate demands reflect delivered costs that include an upstream margin adjustment when characterizing aggregate demand. Second, the base WiNDC data has a more disaggregate representation of taxes on products. IMPLAN has a single indirect business tax assigned to production outputs which is a composite tax that includes sales and excise taxes and production taxes. WiNDC separately specifies production taxes and sales/excise taxes. Furthermore, WiNDC characterizes imports as net of import duties by explicitly identifying tariff rates. Import duties in IMPLAN are not explicitly identified suggesting imports are gross of tariffs.

2.1 State-Level Data, BEA Summary Sectors

In the first step, we calculate the percent difference between WiNDC and IMPLAN data for CGE parameters constructed in the respective build routines. The percent change is calculated with the IMPLAN value in the denominator. Figure 1 reports the aggregate percent difference for parameters related to production and consumption across states, census regions, and nationally. `ys0` denotes sectoral production, `ld0` is labor demand, `kd0` is capital demand, `id0` is intermediate demand, `x0` is

international exports, $m0$ is international imports, $i0$ is investment demand, and $g0$ is government purchases. In general, as the data are aggregated to the national level, differences between data sources tend toward zero since both data sources are largely based on the same national data from BEA. In the state-level data, there exist relatively small deviations in the aggregate production data across datasets. The WiNDC data retain intermediate demand shares from the national data in the subnational data, constructs sector by region capital and labor shares based on BEA data on the components of gross state product, and shares sectoral output to states based on aggregate gross state product. The relatively small difference between data sets suggests that IMPLAN may use similar assumptions.

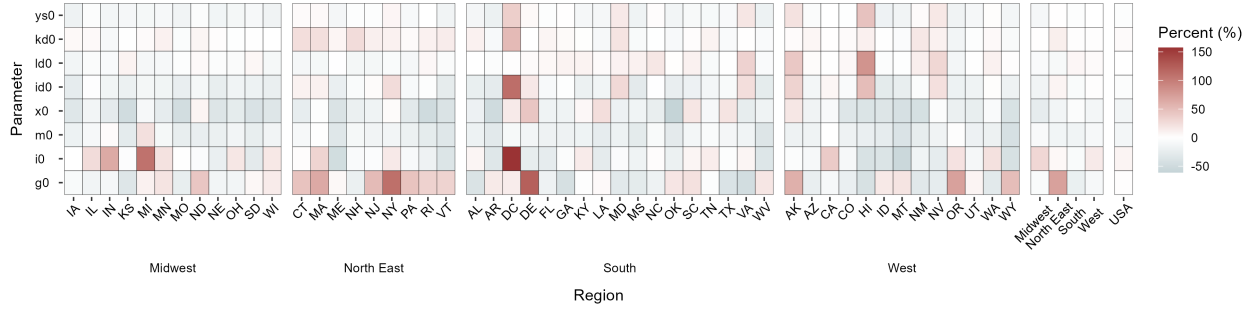


Figure 1: Percentage Difference in Select Production and Consumption Data Between WiNDC and IMPLAN

On the non-household consumption side of the dataset, we observe the largest percent differences in investment and government demand. WiNDC relies on estimates for state gross state product from BEA and state government finance data from Census to inform these distributions. Given the large observable differences across datasets, this may be a place where further research is needed to explore these sets of assumptions. However, a direct comparison is challenging in this case because as explored more in the next section using a more aggregate dataset, the composition of investment and government demand differs across datasets. In both cases, IMPLAN deviates from totals provided by the national level BEA use table. International trade follows similar patterns across datasets, though magnitudes are systematically smaller in WiNDC. This is presumably due to characterizing imports as net of import duties (ad-valorem tariffs are explicitly captured in the database) relative to IMPLAN.

Notably, Figure 1 characterized the distribution of differences for data elements that have relatively robust supporting data sources. Figure 2 reports the percent difference for “intermediate” data elements related to intranational trade that are not often observed, but are potentially important for characterizing regional incidence in subnational CGE modeling. In general, both datasets incorporate exogenous assumptions about the sourcing of supply and demand. They both characterize the proportion of supply that is allocated to the region in which it was produced ($xd0$) or the rest of the country ($xn0$) and the demand for commodities produced within the same region ($dd0$) or from the rest of the country ($nd0$). $a0$ represents total Armington demand for commodities which equals demand from regional, national, and international markets. The datasets show relatively large aggregate differences when decomposing the allocation of production and consumption

to place of origin and destination.

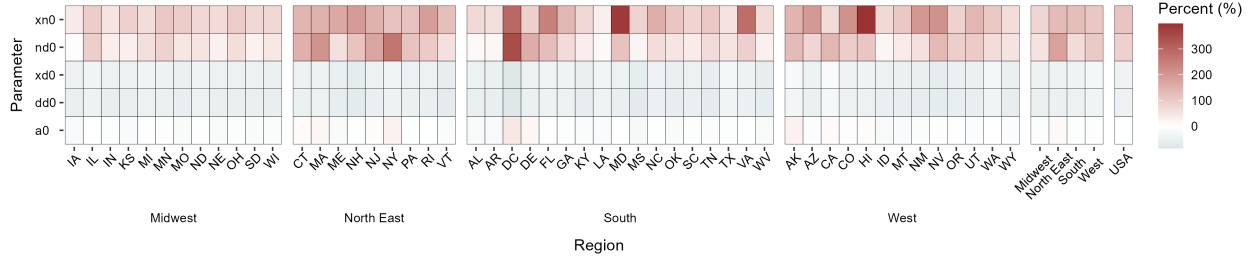


Figure 2: Percentage Difference in Select Trade Parameters Between WiNDC and IMPLAN

These data rely on exogenously specified regional purchase coefficients to characterize the source and destination for supply. In the United States, there is limited information on the trade of goods between states that characterizes the origin of production and the final destination of demand (either final or intermediate). Data sources like the Commodity Flow Survey and the Freight Analysis Framework (FAF) do exist but contains transshipments (or flows of goods that pass through a state on its way to its place of final or intermediate demand) which makes using this information a challenging attribution exercise. In WiNDC, regional purchase coefficients are based on the FAF data with the implicit assumption that transshipments follow the distribution of origin and destination pairings of supply and demand. However, those data only supply estimates for commodities that can be shipped as freight, not service based flows. Therefore, in the absence of information, the build routine makes an assumptions regarding the regional purchase coefficient. In sectors other than utilities and construction, the build routine assumes an average economy-wide regional purchase coefficient from the FAF data. IMPLAN, conversely, relies on a gravity model framework that uses estimates for where production and consumption occur to estimate a trade matrix through a bi-proportional balancing procedure based on a decaying distance function (Lindall et al., 2006). The comparison shows that the WiNDC approach reports relatively larger flows of goods and services outside of the region in which it was produced. Both approaches are challenging to validate. Taken together, however, both datasets show similar aggregate state level demand, as reported by small differences in $a0$.

Figure 3 reports the percent difference in household consumption ($cd0$), savings ($sav0$), labor income ($le0$), and capital income ($ke0$). Like the aggregate production and consumption data, state-level differences across states attenuate at the national level. In WiNDC, the distribution of household consumption to states is based on personal consumer expenditure data from the BEA. Here there is only a small degree of differences across the WiNDC and IMPLAN datasets. For other parameters, the construction of these datasets relies on satisfying a series of “accounting identities,” meaning here that savings must equal total investment demand ($i0$), labor income must equal labor demands, and capital income must equal capital demands. Therefore, the datasets should be quite close if the national estimates are similar by construction. The degree of difference in regional markets stems from what is assumed about market closures in the databases. For labor, WiNDC does allow for inter-state commuting based on data from the American Community Survey (e.g.,

labor endowments are smaller in DC in WiNDC relative to IMPLAN due to commuters from Maryland and Virginia), but this represents a relatively small fraction of the total labor demand in a given state which is assumed to clear regionally. Therefore, both datasets show similar estimates. Closure assumptions in capital and investment markets are assumed to clear at the national level which means that any given subregion or state may have discrepancies between the ownership of capital or savings and the demand for investment and capital. WiNDC bases its decomposition of household income on both the Current Population Survey (CPS) and the National Income and Product Accounts (NIPA) for capital interest income and retirement contributions. This national closure assumption was chosen to better match the underlying income distribution data used to calibrate household accounts. In IMPLAN, household estimates for net household investments and capital rental income are scaled to match aggregate demands in the data. Therefore, differences in the data in these markets stem from assumptions underpinning the household decomposition in each dataset.

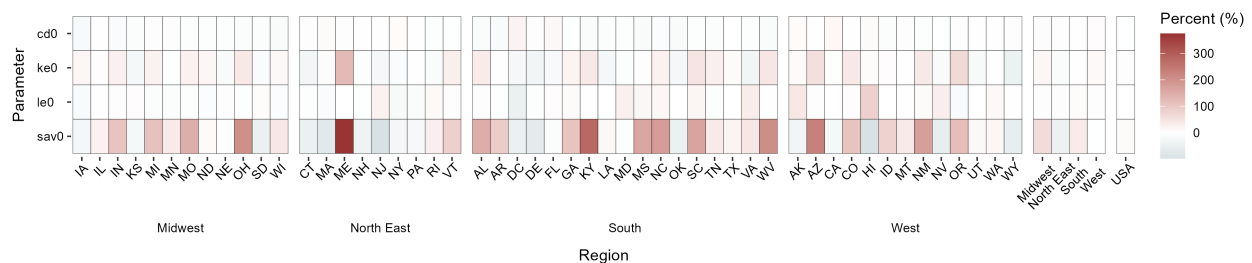


Figure 3: Percentage Difference in Select Household Parameters Between WiNDC and IMPLAN

Both datasets provide information across heterogeneous households denominated by income. Household income thresholds are reported in Table 1. As previously mentioned, WiNDC households are defined based on pure quintiles where income thresholds are determined based on the number of households in the economy (even bins) whereas the base IMPLAN data provides information on nine separate household types denominated by income. We’ve aggregated the IMPLAN households to be as comparable to WiNDC as possible.

Figure 4 reports the percent difference in aggregate household consumption by region and household type. The top panel reports the percent difference across the entire income category. Because income categories represent different types of households across the datasets, we also report a “scaled” version where we calculate the percent change in per household estimates (e.g., dividing consumption by the number of households in a particular income bin). Dataset difference decrease when the data are scaled. WiNDC has slightly higher consumption (and therefore income) accruing to the top income quintile relative to IMPLAN regardless of the regional aggregation. Total household consumption is estimated in the WiNDC framework by enforcing a budget balance closure condition that simultaneously targets components of income based on the underlying Census and NIPA data as well as the distribution of consumption across income quintiles, using data from the BLS’s Consumer Expenditure Survey.³

³Consumption by household quintile is approximated using [average annual expenditures by income quintile](#) from

Table 1: SAGE Representative Households Denominated by Income

Index	WiNDC		IMPLAN	
	Income Bins (2016\$)	# Households (millions)	Income Bins (2016\$)	# Households (millions)
hh1	$\leq \$21,219$	25.3	$\leq \$30,000$	33.0
hh2	\$21,219-\$43,016	25.3	\$30,000-\$50,000	22.1
hh3	\$43,016-\$72,015	25.3	\$50,000-\$70,000	17.5
hh4	\$72,015-\$119,148	25.3	\$70,000-\$150,000	36.9
hh5	$\geq \$119,148$	25.3	$\geq \$150,000$	17.0

In WiNDC, income thresholds are calculated using the CBO definition of income before taxes and transfers composed of market income plus social insurance benefits.

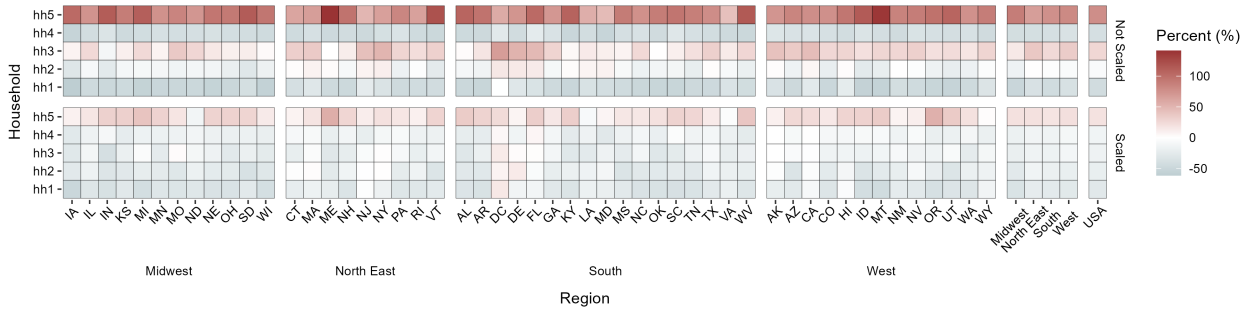


Figure 4: Percentage Difference in Representative Household Level Expenditures (cd0) Between WiNDC and IMPLAN

2.2 Processed Data at the SAGE Resolution

In the second step, we compare more aggregated datasets that have been post processed using the SAGE build routine. Benchmark data refinements include capturing marginal and average income tax rates, disaggregation of production between crude oil and natural gas, in the case of WiNDC, disaggregation of some summary sectors, and including information on government finances and income from the rest of the world. Table 2 reports the sector, region, and household aggregations in the SAGE model. For WiNDC data, while a majority of the summary sectors are easily aggregated into the SAGE sectoring scheme, some need to be further disaggregated. See Section 3.1.1 in Marten et al. (2025) for an overview of this process. In general, we leverage data from the detailed supply and use tables from BEA, labor compensation data from the BLS, and electricity generation data from EIA to inform the disaggregation for most sectors. Notably, for both IMPLAN and WiNDC, the crude oil and natural gas extraction sector is disaggregated equivalently using information from EIA and Census.

Like in the previous section, we report the percent difference between datasets, where the IMPLAN value is in the denominator. Figure 5 reports the percent difference in data across the Consumer Expenditure Survey.

Table 2: Static SAGE Dimensions

Sectors (abbreviations)	Census Regions	Households (income based)
Agriculture, forestry, fishing and hunting (agf)	Northeast	hh1
Crude oil (cru)	South	hh2
Coal mining (col)	Midwest	hh3
Metal ore and nonmetallic mineral mining (min)	West	hh4
Electric power (ele)		hh5
Natural gas (gas)		
Water, sewage, and other utilities (wsu)		
Construction (con)		
Food and beverage manufacturing (fbm)		
Wood product manufacturing (wpm)		
Petroleum refineries (ref)		
Chemical manufacturing (chm)		
Plastics and rubber products manufacturing (prm)		
Cement manufacturing (cem)		
Primary metal manufacturing (pmm)		
Fabricated metal product manufacturing (fmm)		
Electronics and technology manufacturing (cpu)		
Transportation equipment manufacturing (tem)		
Other manufacturing (bom)		
Non-Truck Transportation (trn)		
Truck transportation (ttn)		
Services (srv)		
Healthcare services (hlt)		

production and consumption parameters in the post processed SAGE datasets. In general, trends in aggregate production, non-household demand, and international trade follows similar patterns as the disaggregate data. Relatively small differences exist in production parameters (**ys0**, **kd0**, **ld0**, and **id0**) and differences in aggregate international trade likely stem in part from the treatment of import duties in WiNDC. As before, the largest differences in the data are in investment demand and public purchases. Figure 14 in Appendix A reports the absolute difference between investment and government demands across the datasets (gray tiles indicate places in which data do not exist in one of the databases). As alluded to in the previous section, the composition of both data elements is significantly different between data sources. For instance, WiNDC follows BEA conventions and allocates a significant share of government demand to government enterprise sectors, most of which are aggregated into the broader service sector. For investments, WiNDC has larger demands in sectors like electronics manufacturing or balance of manufacturing and relatively smaller demands in construction. This is potentially significant because the composition of investment governs the formation of capital in dynamic CGE models.

While the percent difference in aggregate regional production is relatively close across datasets,

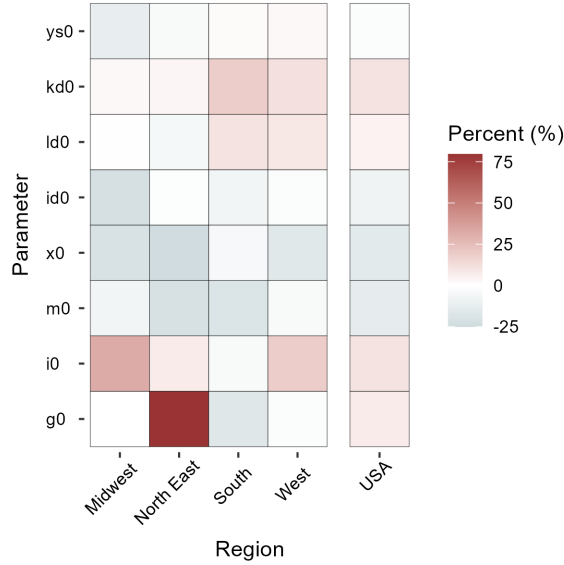


Figure 5: Percentage Difference in Select Production and Consumption Parameters Between WiNDC and IMPLAN

we verify that the distribution of production across sectors is also similar, as reported in Figure 6. Figure 6a reports the percent difference for each sector and region pairing and Figure 6b reports the absolute difference by sector and region. In terms of percent difference, the natural gas and other mining sector in the Southern census region represents the largest change. Interestingly, the natural gas extraction and distribution sector is similarly calibrated across datasets, so this difference is somewhat surprising. However, whenever the difference is denominated in dollar terms, differences in the aggregate service sector dominate other sectors because it represents a disproportionate share of the economy. This is also true of intermediate demands, labor demands, and capital demands as reported in Figure 15 in Appendix A.

We also confirm that trends in subnational trade in the aggregate processed data align with the disaggregate unprocessed data in the previous section. Figures 7 report the percent change in commodity supply allocated to the region of origin or the rest of the country. As noted earlier, alternative approaches are taken across datasets in how they estimate regional purchase coefficients. Interestingly, the decomposition shows that the differences largely stem from assumptions on trade from service-based sectors. The previous section described the different assumption involved in estimating regional purchase coefficients across datasets. In WiNDC, the FAF data is used to characterize subnational trade, but it is only representative of commodities that can be shipped as freight and not service based flows. Therefore, in the absence of data, the build routine assigns trade based on average regional purchase coefficients (coefficients of 0.9 are assigned to construction and electricity). IMPLAN's estimates using the gravity model produce similar regional purchase coefficients for commodities where FAF data are available. In service based sectors, the datasets diverge. It appears that the distance decay function used in IMPLAN's gravity model is sufficiently calibrated to prevent trade flows from traveling far distances in these cases.⁴

⁴gas is not included in Figure 7 because in percentage terms, it obscures the visualization for $xd0$. We choose to

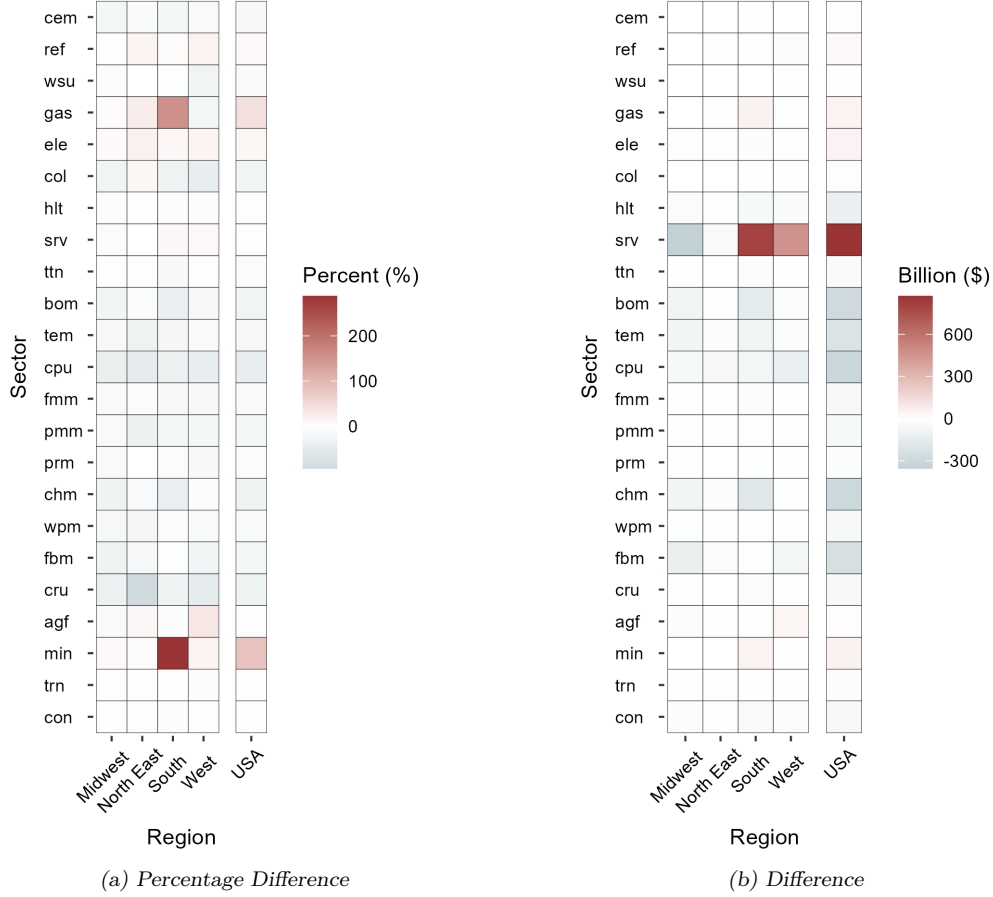


Figure 6: Percentage Difference in Sectoral Output Between WiNDC and IMPLAN

Figures 8 reports the percent difference between WiNDC and IMPLAN for income endowments, and household level consumer demand. The figure reports scaled data which accounts for different household bin sizes across datasets. For household specific parameters, the variation in the datasets decreases when normalizing the data by the respective populations within each household bin. In general, the top income quintile has a larger share of consumption and labor income in the WiNDC data than it does in the IMPLAN data. The distribution of household consumption in WiNDC is informed by aggregate data from BLS's Consumer Expenditure Survey. As discussed in the previous section, labor endowments are calibrated to income shares from the Current Population Survey which, in aggregate, closely approximates the value of employee compensation in the National Income and Product Accounts. WiNDC also allows for commuting based on commuting flows data. That information, however, is not denominated by household income, and therefore the distribution of commuting across households is estimated in the matrix balancing framework that calibrates the household accounts in WiNDC. For capital, WiNDC has relatively larger capital endowments in the Midwest than the North East and Western regions. The Current Population Survey includes capital income from a variety of asset holdings. However, there is a significant

drop the sector because reporting absolute differences for these plots is dominated by the aggregate services sector.

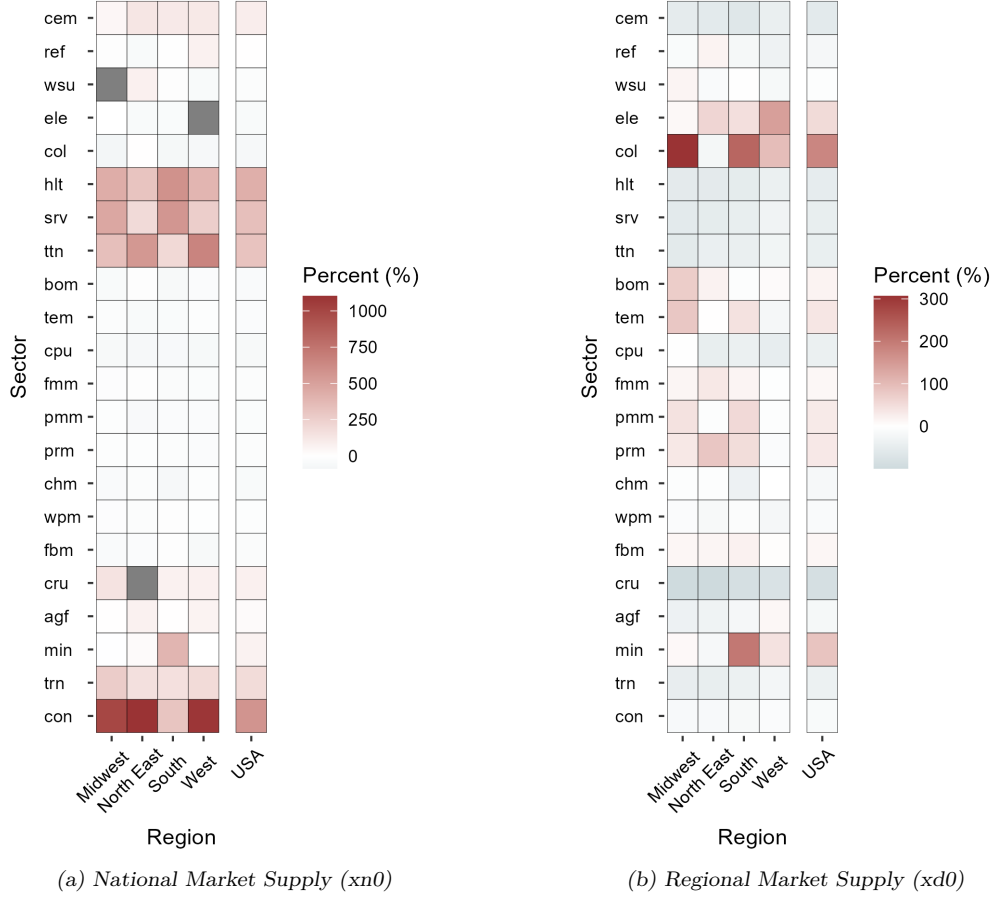


Figure 7: Change Between WiNDC and IMPLAN in Market Supply

discrepancy between the total value of capital income in the CPS relative to the NIPA tables. Furthermore, the issue of capital ownership is complicated by foreign capital owners. In both datasets, this latter issue is assumed away by requiring that the total capital demands in the domestic economy equal the total capital rental income. However, because this national closure is assigned to both datasets, regional variation exists. In general, limited data exist for attributing capital ownership and this area is ripe for additional research. The distribution of household savings also differs across databases. Across households, WiNDC allocates savings based on estimates for distribution of wealth across income quintile reported by the Congressional Budget Office.⁵ In the normalized data, this leads to slightly less savings in the top household income quintile and slightly more in hh4 in WiNDC relative to IMPLAN.

We also report household final demand consumption shares by good, region, and household type in Figure 9. Given the aggregation in SAGE, household consumption shares are largely made up of service and health care expenditures (services include housing costs) across the income distribution. Both datasets have virtually identical consumption shares across Census regions.

⁵The distribution of wealth across income quintiles accounts for home equity, non-retirement and retirement assets, and other assets (see the section on the composition of wealth by income group).

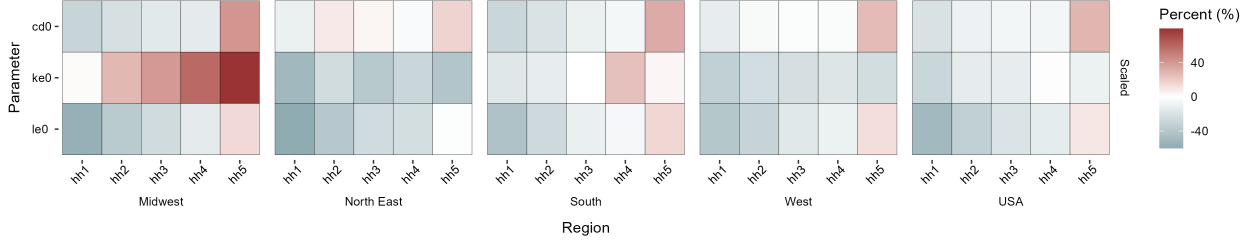


Figure 8: Percentage Difference in Select Representative Household Parameters Between WiNDC and IMPLAN

However, WiNDC demand accounts are calibrated to empirically estimated income elasticities based on the methodology in Aguiar and Bils (2015) using data from the Consumer Expenditure Survey from BLS. Therefore, consumption shares shift across the income distribution in WiNDC depending on the income elasticity associated with a particular commodity.

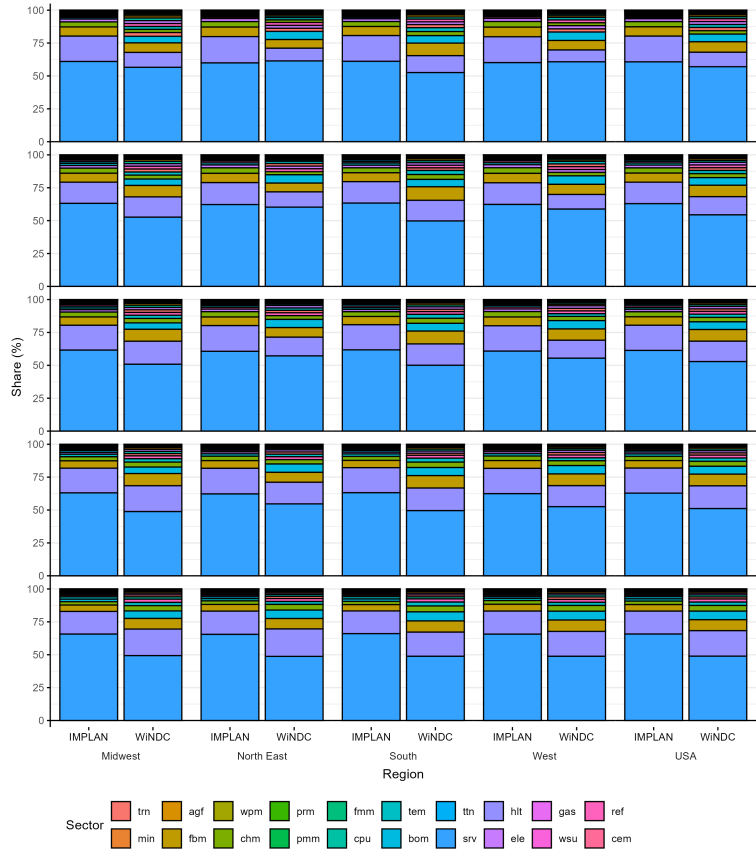


Figure 9: Share of Final Demand ($cd0$) for Goods (g) from Representative Household Bins by Data Source

3 Modeling Application

In this section, we explore the sensitivity of modeling outcomes to the assumed social accounting matrix using a static version of the SAGE model. In this model comparison exercise, we use an

otherwise identical modeling framework (same structural assumptions and elasticities) that differs only by the assumed benchmark data (e.g., cost shares in the calibrated model). We use the static version of the model as an initial diagnostic tool to simplify the comparison between datasets. Per [Marten et al. \(2019\)](#) and [Marten et al. \(2025\)](#), environmental regulations in the United States typically take the form of source-level technology standards, performance based emissions rate limits, or workplace standards. These types of actions often mandate additional inputs to comply with the regulatory standard to produce the same amount of output. Therefore, in this paper, we assess model outcomes to changes in sector specific productivity in which the directly affected sector requires an additional \$100 million of labor and capital (evenly split) to comply with the hypothetical standard. We assume that the shock is distributed across model regions according to reference production shares from the WinDC dataset. Doing so ensures that the models are equivalently shocked.

3.1 Static SAGE Model

The static version of SAGE assumes away the intertemporal dynamics in the full version of the model with perfect foresight, which is described in detail in [Marten et al. \(2025\)](#). The model is generally structured similarly to other CGE models designed to assess environmental and energy policies (e.g., [Yuan et al., 2019](#)). The model features four sub-regions (defined by Census Region), 23 sectors, five household quintiles, and is calibrated and parameterized using several data sources. Model dimensions are further described in Tables 1 and 2. Assumed substitution elasticities are either estimated or reflect the best available estimates from the empirical literature.

Firms are assumed to be perfectly competitive and maximize profits subject to technological constraints. Production is generally characterized by nested constant elasticity of substitution (CES) functions that differentiate substitution possibilities between factors, energy inputs, and other material inputs. CES functions are used widely in general equilibrium modeling for their parsimony and global regularity properties ([Brockway et al., 2017](#)). For non-resource sectors, the top level nest combines non-energy intermediate inputs with a value added-energy composite. The model assumes a capital-labor-energy composite structure that is consistent with empirical data ([Van der Werf, 2008](#)). For resource extraction (crude oil, natural gas, coal, and other mining) and agriculture sectors, the model assumes an additional factor input to represent finite natural resources that is calibrated to supply elasticities from the literature.

On the demand side of the model, consumers maximize their utility over consumption and leisure subject to a budget constraint composed of endowment income from capital returns, natural resources, and time. The preference structure is characterized by a nested CES-LES (Linear Expenditure System) function with a CES aggregator in the top level nest between leisure and aggregate commodity consumption and an LES expenditure function for commodity consumption in the lower nest. We calibrate the leisure-composite material goods substitution elasticity to existing estimates for labor supply income and substitution elasticities. The LES sub-nest is calibrated to empirically estimated income elasticities by separating subsistence and discretionary spending

(Shojaeddini et al., 2024). There is a single representative government representing all jurisdictions. The government raises revenue through ad-valorem taxes on capital, labor, production, commodity purchases, and imports and balances its budget through lump sum transfers.

The United States is treated as a large open economy. While the rest of the world is not explicitly represented, the model provides a reduced form mechanism to influence world commodity prices based on Schreiber et al. (2024), Lanz and Rutherford (2016), and Aguiar et al. (2022). Export supply and import demand are governed by the Armington assumption, where goods are differentiated by place of production (Armington, 1969). The import aggregation function distinguishes imported varieties from the regional (defined at the Census Region), national, and international markets through a nested CES function. Conversely, production is allocated to the regional, national, and international markets through a constant elasticity of transformation function.

The model is written as a mixed complementarity problem in GAMS and solved using PATH. The model is solved as a square system in which market clearance conditions determine equilibrium price levels, zero profit conditions determine activity levels, and budget balance conditions determine income levels. Labor and resource markets are assumed to clear regionally while capital rental markets clear nationally.

3.2 Results

Figure 10 reports aggregate equivalent variation (also referred to here as social costs) from our set of stylized scenarios. The percent difference in aggregate social costs as estimated by the WiNDC dataset relative to the IMPLAN dataset range from -2.6% in the coal mining sector to 9.4% in the electronics and technology manufacturing sector. Across models calibrated to the alternative datasets, the majority of sector specific shocks produce impacts within 3% of each other, with exception of balance of manufacturing (or **bom**, 3.2%), non-truck transportation (or **trn**, 4.6%), truck transportation (or **ttn**, 5.3%), transportation equipment manufacturing (or **tem**, 8.4%), and the electronics and technology manufacturing (or **cpu**, 9.4%) sectors. We focus the discussion on these five sectors to understand where data differences seem to drive the largest differences in model results.

In non-transportation sectors, several differences in the datasets are key contributors to model output differences. First, sectoral production in WiNDC is smaller in **bom**, **tem**, and significantly so in **cpu**. The value of production in the latter sector in the WiNDC dataset is approximately half the size in the IMPLAN dataset (see Figure 18 in Appendix A). Notably, WiNDC levels are closer to the requisite values in the national BEA supply and use tables. Therefore the scenarios lead to larger price effects because the assumed shock is a larger share of production costs in these sectors. Secondly, while the value of output is smaller in these three sectors, the values of labor demand are comparable across datasets which means that the labor value share of production is larger in WiNDC (see Figure 19 in Appendix A). In the **cpu** sector, the labor value share in WiNDC is roughly double the share in IMPLAN. Therefore, the scale effect, or the reduction in output in response to higher production costs, leads to a larger labor demand response in the

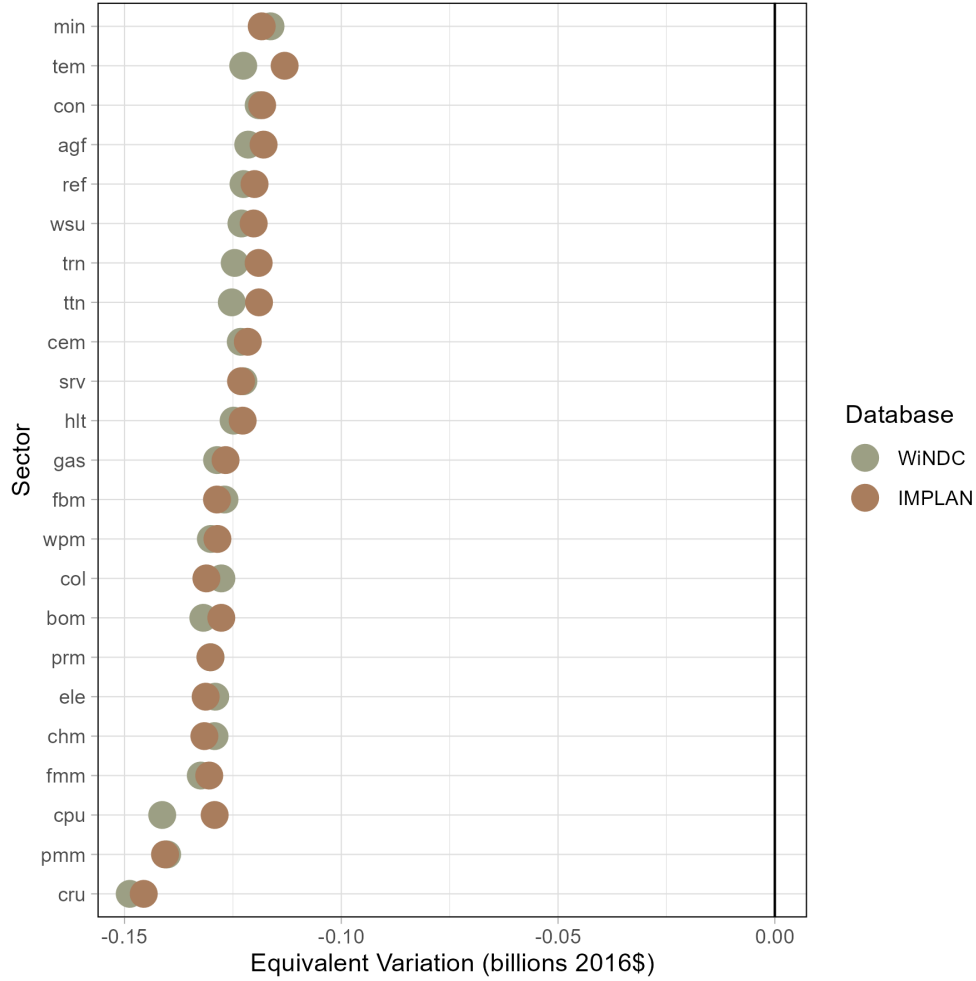


Figure 10: Aggregate Equivalent Variation from Sector-Specific Illustrative Scenarios

directly regulated sector in WiNDC. Given the marginal and average labor income tax structure in the model, reductions in labor demand in an already distorted labor market can cause additional welfare effects from the tax interaction effect.

On the demand side of the accounts, **bom**, **tem**, **cpu**, all represent a larger share of total commodity consumption in WiNDC relative to IMPLAN, and significantly so in the case of **bom** (the consumption share is approximately five times larger in WiNDC). Therefore, rising prices in the regulated sector in response to the regulatory scenario lead to a larger reduction in the purchasing power of the modeled households at benchmark consumption patterns. These three sectors also represent a larger portion of the composition of aggregate investment in WiNDC relative to IMPLAN. Because our modeling is static and fixes investment to benchmark levels, price increases in the regulated sector lead to a larger reduction in purchasing power for households because the portion of real income directed toward savings increases. Finally, differences in inter-sectoral linkages also likely contribute to the welfare effect (which may be particularly true for **tem** since a substantial share of downstream demand is due to transportation service sectors that are differentially

represented across datasets).

For scenarios directly affecting transportation sectors (**trn** and **ttn**), the WiNDC-based model estimates larger impacts than the IMPLAN-based model. WiNDC represents the demand for transportation services through margins, or price markups from shipping costs. The characterization of margins in WiNDC is also driven by the assumption that transportation margins are supplied by the regional or national markets. Furthermore, the WiNDC data is associated with much smaller exports than IMPLAN in these two sectors (see Figure 20 in Appendix A). Therefore, more of the burden of the regulation falls on the domestic economy.

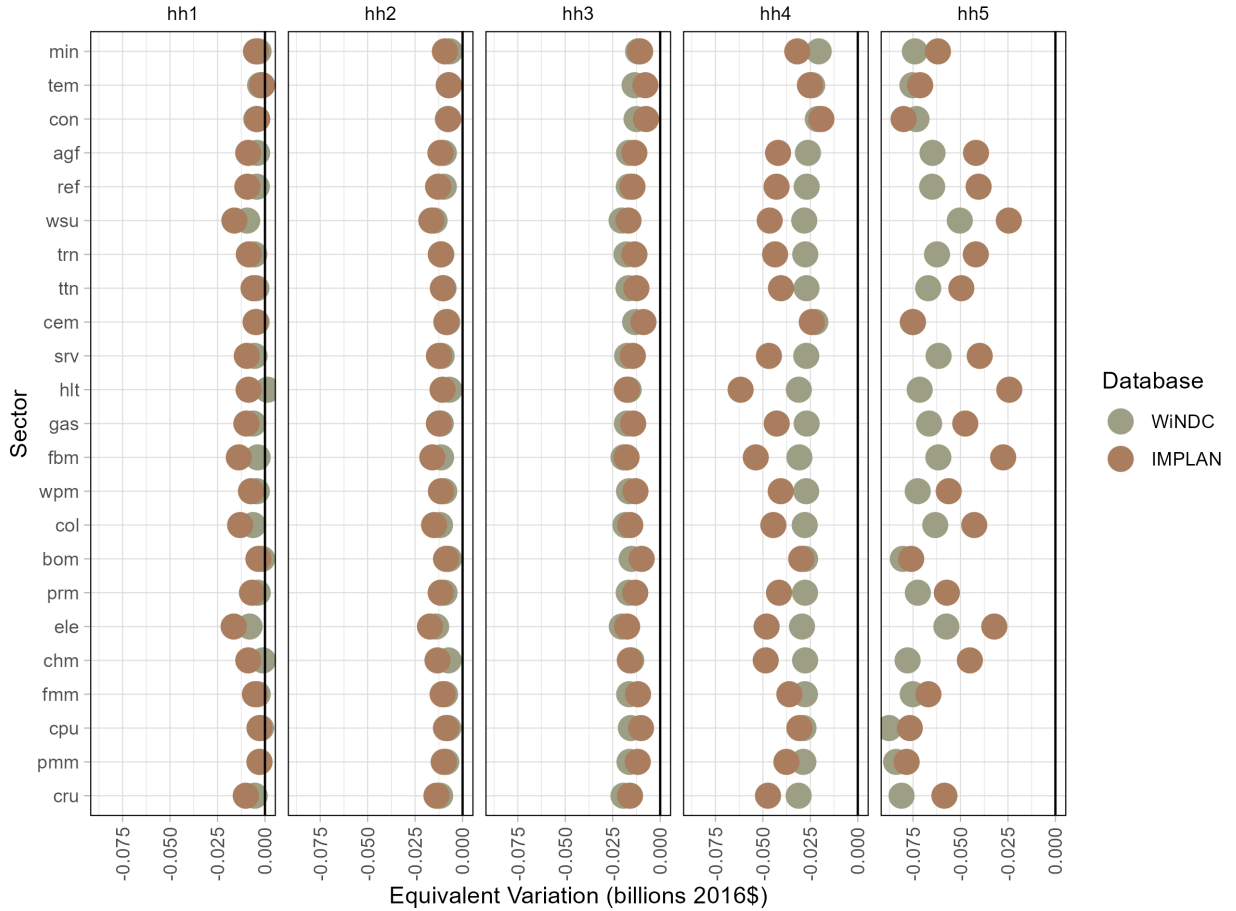


Figure 11: Aggregate Household-Level Equivalent Variation from Sector-Specific Illustrative Scenarios

Figure 11 reports the distribution of social costs across households in the model for our set of stylized scenarios. Generally across models based on either dataset, more of the burden of the stylized regulatory scenarios falls on primary factor owners (hh4 and hh5) and less so on lower income households who derive a larger share of their income through transfer payments. Reported results are for aggregate equivalent variation for each household bin which represents a different number of households across datasets (see Table 1 for the distribution). Most notable differences are present in the dataset for the fourth and fifth income quintile, where in the former, WiNDC

produces systematically smaller social costs than IMPLAN; the converse is true in the upper income quintile.

To make a more consistent comparison, we normalize the distributional results by dividing the estimated social cost for each household bin by the number households within each income bin. The outcomes of this calculation are reported in Figure 12 (shifting the units to dollars per household). Generally, impacts look much more similar, particularly for hh1-hh4. Impacts in the upper income quintile are relatively more “stable” under the WiNDC dataset across regulatory scenarios. Many of the larger differences in results appear to occur for sector-specific regulatory scenarios that directly affect sectors that are larger components of aggregate investment, which disproportionately affects upper income households. The distribution of savings is slightly different between the datasets (where WiNDC has a relatively higher proportion of savings in hh4 than in IMPLAN and a relatively lower proportion in hh5, as reported in Figure 9) which explains the shift in welfare costs from hh5 to hh4 in the WiNDC-based model relative to the IMPLAN-based model in those cases.

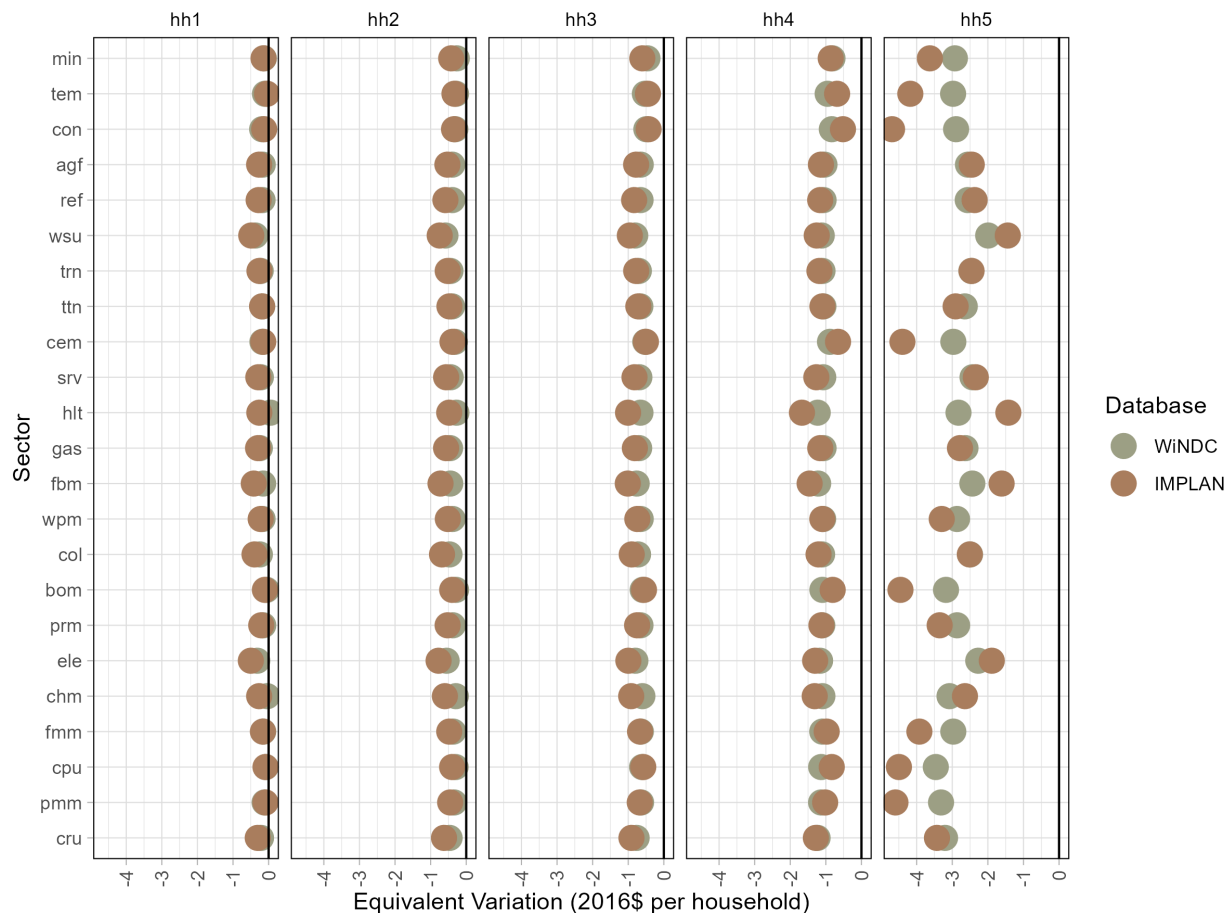


Figure 12: Per Household Equivalent Variation from Sector-Specific Illustrative Scenarios

Finally, we report the regional incidence of the suite of stylized scenarios in Figure 13. In

general, the results are mostly similar. Modeled differences do exist which likely corresponds to the significant difference in implicit regional purchase coefficients assumed by either dataset. Furthermore, those differences may be more pronounced in the results if the assumed shock is not economy-wide, or follows a distribution that differs from the underlying production data in the database.

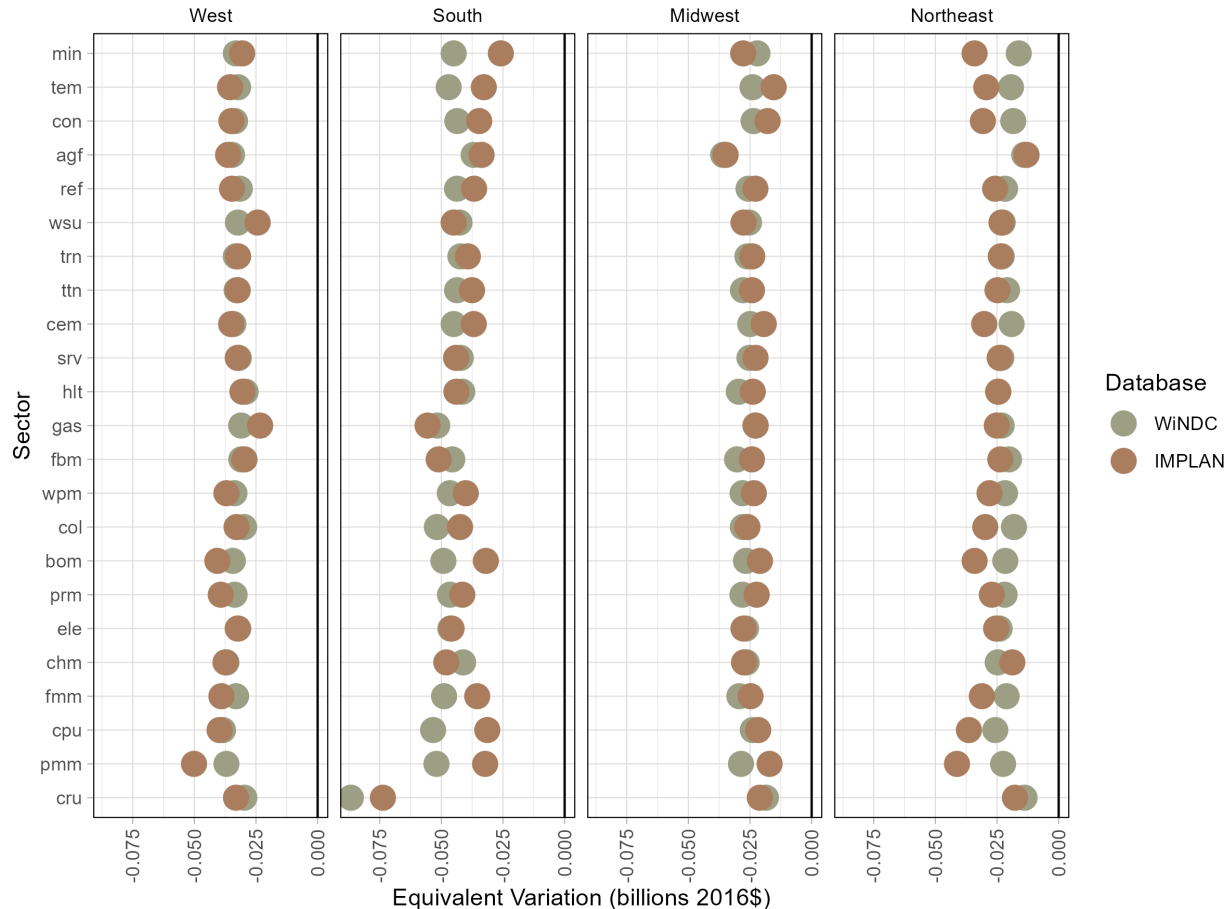


Figure 13: Aggregate Regional Equivalent Variation from Sector-Specific Regulatory Scenarios

4 Conclusion

This paper quantifies underlying differences in competing data sources for subnational CGE modeling in the United States. We find that there are structural differences in how the data characterizes production, trade, and consumption, but those differences produce mostly small impacts on aggregate welfare based on a static version of EPA’s SAGE computable general equilibrium model. We verify this outcome for a range of sector-specific scenarios that approximate the types of environmental regulatory standards most often considered by EPA using a modeling framework that can attribute differences in results to the underlying benchmark dataset. In instances where results

differ somewhat notably between models calibrated to the alternative datasets, the difference is due to the underlying assumptions in the data generating process in WiNDC. Note, however, that these outcomes may be specific to the types of modeling scenarios explored in this paper. We have not explored their generalizability across a wider range of policy scenarios.

The purpose of this paper is not to validate either the WiNDC or IMPLAN data. Both WiNDC and IMPLAN datasets are constructed datasets, meaning they impose a series of assumptions (sometimes based on best professional judgment and other times based on underlying source datasets) while simultaneously satisfying accounting constraints implicit in the social accounting matrix structure. However, because WiNDC is open source with a publicly available build routine that allows us to investigate the data sources or assumptions used in the construction of the database, we can observe where WiNDC and IMPLAN differ highlighting places where IMPLAN may take an alternative approach. It is not possible to verify how differences manifest from specific approaches used in IMPLAN due to its proprietary status.

In general, we find that the WiNDC dataset is a suitable replacement to the IMPLAN dataset for modeling the economy-wide impacts of environmental regulations. Furthermore, use of WiNDC is responsive to SAB recommendations to move to open-source datasets to improve the transparency and accessibility of conducting economy-wide modeling in regulatory analysis.

Acknowledgements

The authors thank Ann Wolverton, Alex Marten, Dave Evans, and Wade Davis for helpful comments in the preparation of this draft.

References

- Aguiar, A., M. Chepeliev, E. Corong, and D. Van Der Mensbrugghe (2022). The global trade analysis project (GTAP) data base: Version 11. *Journal of Global Economic Analysis* 7(2).
- Aguiar, M. and M. Bils (2015). Has consumption inequality mirrored income inequality? *American Economic Review* 105(9), 2725–2756.
- Akin-Olcum, G., C. Böhringer, T. Rutherford, and A. Schreiber (2021). Economic and environmental impacts of a proposed ‘carbon adder’ on New York’s energy market. *Climate Policy* 21(6), 823–842.
- Armington, P. S. (1969). A theory of demand for products distinguished by place of production. *Staff Papers (International Monetary Fund)* 16(1), 159–178.
- Bae, J. and S. Dall’erba (2016). The economic impact of a new solar power plant in Arizona: Comparing the input-output results generated by JEDI vs. IMPLAN. *Regional Science Policy & Practice* 8(1-2), 61–74.
- Becker, J., M. Brown, M. Browning, Y. Cai, D. Goldfinger, J. McFarland, S. Paltsev, S. Weisberg, and M. Yuan (2025). U.S. economy-wide decarbonization: Sectoral and distributional impacts. *Energy and Climate Change* 6, 100181.
- Bistline, J., C. Onda, M. Browning, J. Emmerling, G. Iyer, M. Mahajan, J. McFarland, H. McJeon, R. Orvis, F. R. Fonseca, et al. (2024). Equity implications of net-zero emissions: A multi-model analysis of energy expenditures across income classes under economy-wide deep decarbonization policies. *Energy and Climate Change* 5, 100118.
- Brockway, P. E., M. K. Heun, J. Santos, and J. R. Barrett (2017). Energy-extended CES aggregate production: Current aspects of their specification and econometric estimation. *Energies* 10(2), 202.
- Brown, M., S. Awara, J. Ntiamoah, and M. Yuan (2025). Distributional impacts of technology pathways for us power sector transformation. *Applied Energy* 388, 125500.
- Brown, M., J. M. Becker, J. C. Carbone, T. Goforth, J. McFarland, D. Nock, K. Pitman, and D. Steinberg (2024). Tax credits for clean electricity: the distributional impacts of supply-push policies in the power sector. *Journal of the Association of Environmental and Resource Economists* 11(S1), S199–S229.
- Deller, S. and J. Hadachek (2024). The contribution of agriculture to the wisconsin economy: An update for 2022. <https://aae.wisc.edu/wp-content/uploads/2024/11/The-Contribution-of-Agriculture-to-the-Wisconsin-Economy.pdf>.

-
- Horowitz, K. J. and M. A. Planting (2009). Concepts and methods of the us input-output accounts. Bureau of Economic Analysis: https://www.bea.gov/sites/default/files/methodologies/IOmanual_092906.pdf.
- IMPLAN Group LLC. (2016). IMPLAN 2016. Huntersville, NC.: <https://www.implan.com>.
- Lanz, B. and T. F. Rutherford (2016). GTAPinGAMS: Multiregional and small open economy models. *Journal of Global Economic Analysis* 1(2), 1–77.
- Lindall, S. A., D. C. Olson, and G. S. Alward (2006). Deriving multi-regional models using the implan national trade flows model. *Journal of Regional Analysis and Policy* 36(1).
- Marten, A., J. Perla, A. Schreiber, and A. Wolverton (2025). SAGE model documentation (3.0.0). U.S. Environmental Protection Agency: <https://www.epa.gov/environmental-economics/cge-modeling-regulatory-analysis>.
- Marten, A., A. Schreiber, and A. Wolverton (2024). SAGE model documentation (2.1.1). U.S. Environmental Protection Agency: <https://www.epa.gov/environmental-economics/cge-modeling-regulatory-analysis>.
- Marten, A. L., R. Garbaccio, and A. Wolverton (2019). Exploring the general equilibrium costs of sector-specific environmental regulations. *Journal of the Association of Environmental and Resource Economists* 6(6), 1065–1104.
- Miller, R. E. and P. D. Blair (2009). *Input-output analysis: Foundations and extensions*. Cambridge university press.
- Parajuli, R., J. E. Henderson, S. Tanger, O. Joshi, and R. Dahal (2018). Economic contribution analysis of the forest-product industry: A comparison of the two methods for multisector contribution analysis using IMPLAN. *Journal of Forestry* 116(6), 513–519.
- Rausch, S. and T. F. Rutherford (2008). Tools for building national economic models using state-level IMPLAN social accounts. <https://www.mpsge.org/IMPLAN2006inGAMS/IMPLAN2006inGAMS.pdf>.
- Ross, M. T. (2008). Documentation of the applied dynamic analysis of the global economy (ADAGE) model. RTI Working Paper (05-01), https://www.rti.org/sites/default/files/resources/adage-model-doc_ross_sep08.pdf.
- Rutherford, T. F. and A. Schreiber (2019). Tools for open source, subnational CGE modeling with an illustrative analysis of carbon leakage. *Journal of Global Economic Analysis* 4(2), 10–21642.
- SAB (2017). SAB advice on the use of economy-wide models in evaluating the social costs, benefits, and economic impacts of air regulations. https://www.epa.gov/system/files/documents/2023-02/EPA-SAB-17-012_1.pdf.

-
- SAB (2020). Technical review of EPA’s computable general equilibrium model, SAGE. <https://www.epa.gov/environmental-economics/technical-review-epas-computable-general-equilibrium-model-sage-pdf>.
- Schreiber, A., A. Marten, and A. Wolverton (2024). Approximating terms of trade effects in single country CGE models. *Journal of Global Economic Analysis* 9(2), 70.
- Shojaeddini, E., A. Schreiber, A. Wolverton, and A. Marten (2024). Consumer demand and the economy-wide costs of regulation: Modeling households with empirically estimated flexible functional forms. *Journal of Environmental Economics and Management* 125, 102972.
- Van der Werf, E. (2008). Production functions for climate policy modeling: An empirical analysis. *Energy Economics* 30(6), 2964–2979.
- Webster, M., K. Fisher-Vanden, V. Kumar, R. B. Lammers, and J. Perla (2022). Integrated hydrological, power system and economic modelling of climate impacts on electricity demand and cost. *Nature Energy* 7(2), 163–169.
- Wing, I. S., P. H. Larsen, J. P. Carvallo, A. Sanstad, D. Wei, A. Rose, S. Baik, J. Smith, C. Ramee, and R. Peterson (2025). A method to estimate the economy-wide consequences of widespread, long duration electric power interruptions. *Nature Communications* 16(1), 3335.
- Yuan, M., S. Rausch, J. Caron, S. Paltsev, and J. Reilly (2019). The MIT US Regional Energy Policy (USREP) model: The base model and revisions. *Joint Program Technical Note* 18.

A Appendix

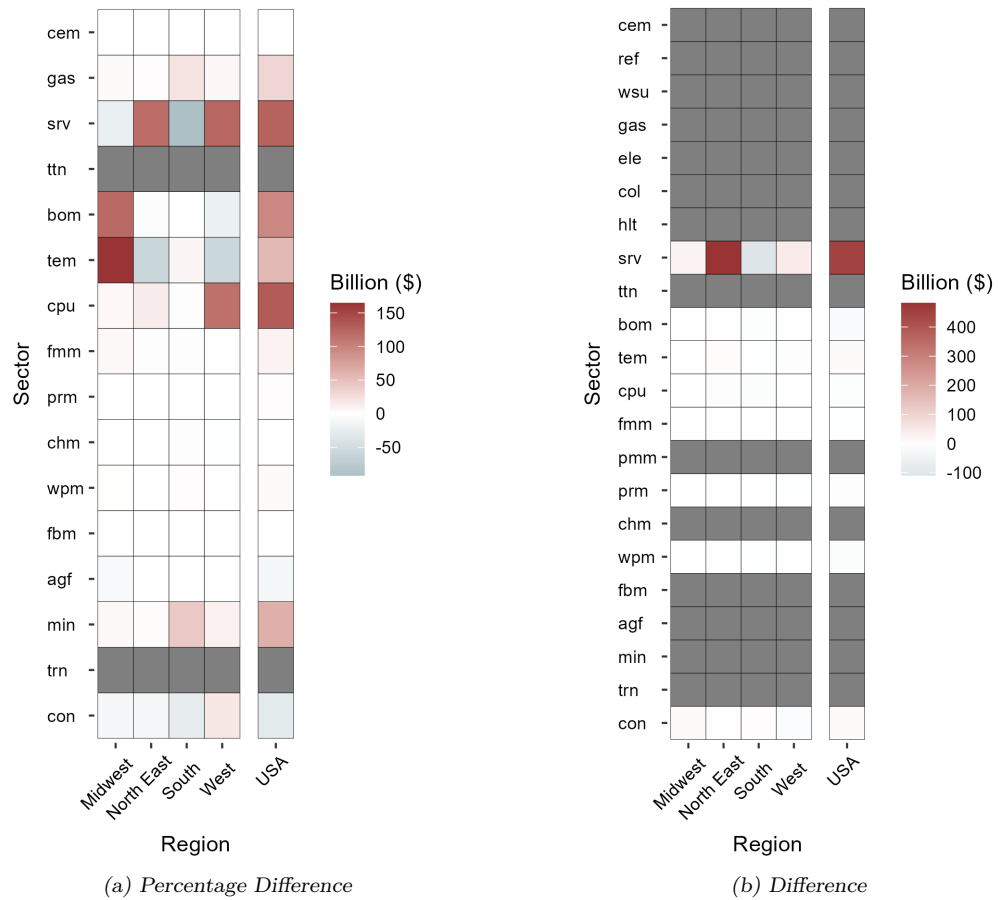
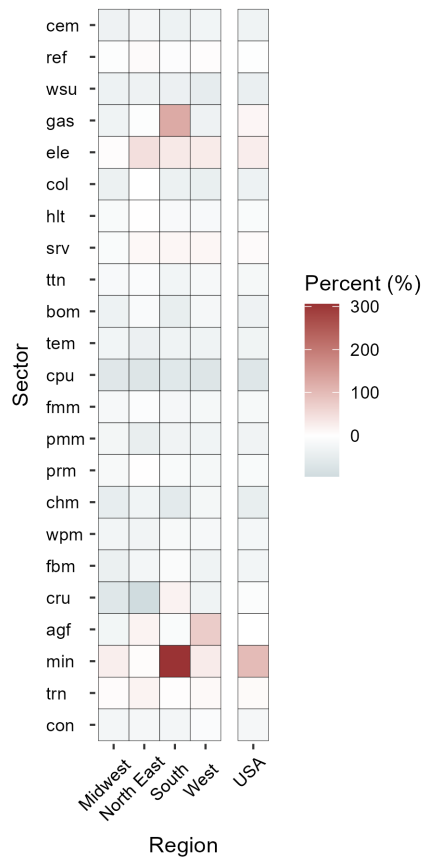
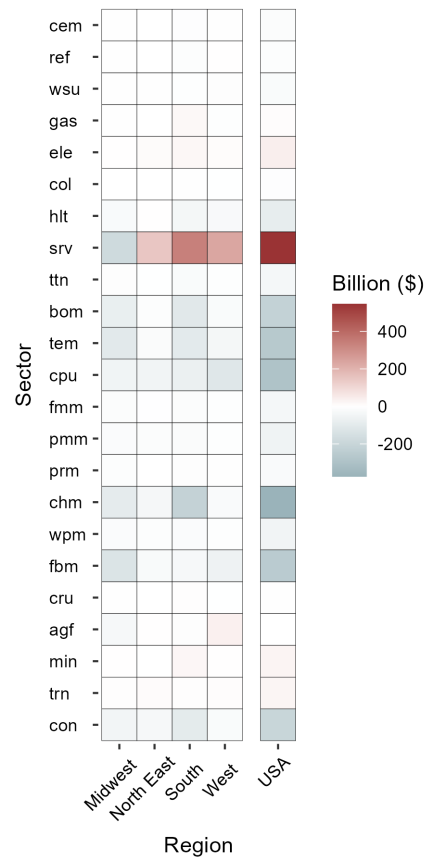


Figure 14: Change from IMPLAN in Investment Demand (i_0) and Government Demand (g_0)

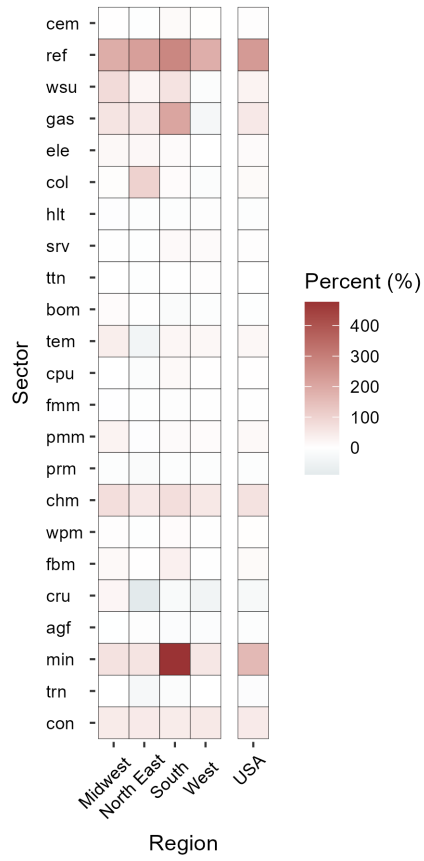


(a) Percentage Difference

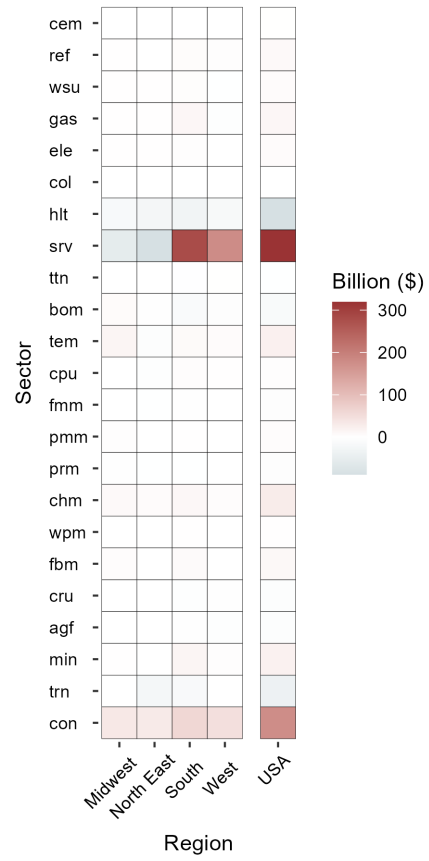


(b) Difference

Figure 15: Change from IMPLAN in Intermediate Demand (id0)

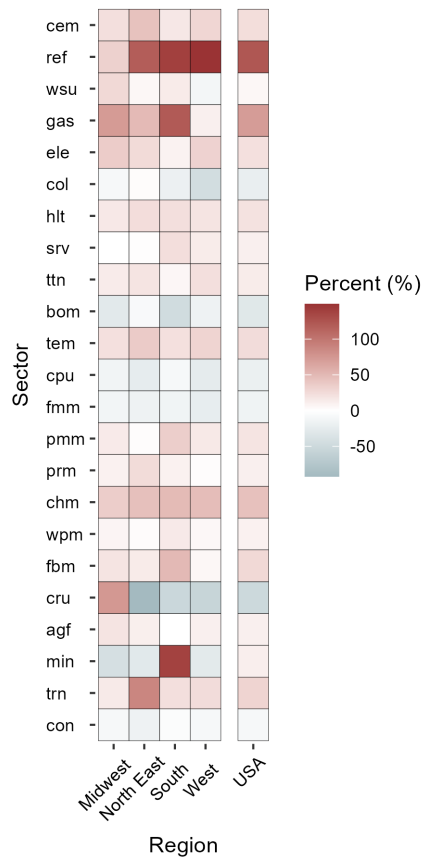


(a) Percentage Difference

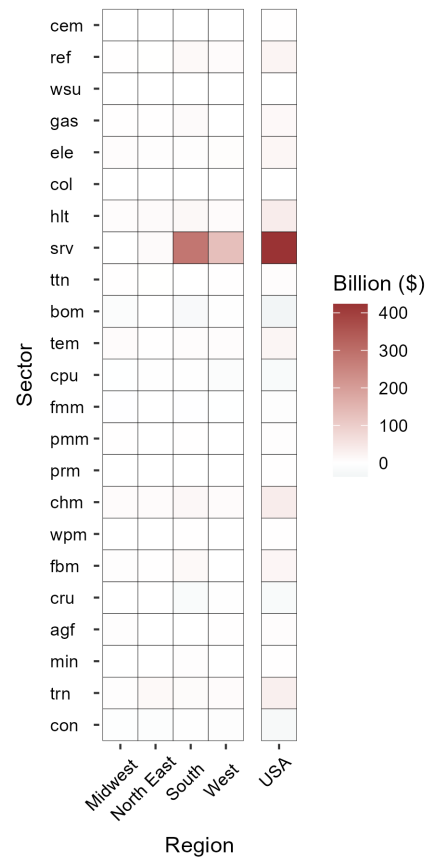


(b) Difference

Figure 16: Change from IMPLAN in Labor Demand (ld_0)

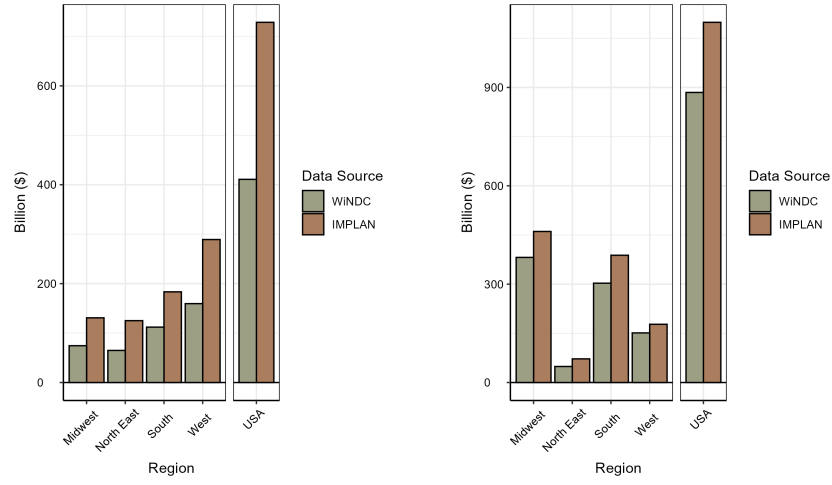


(a) Percentage Difference

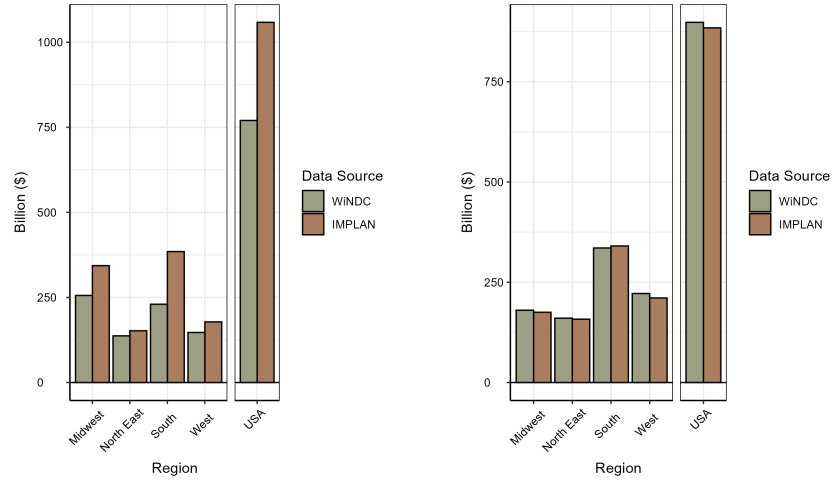


(b) Difference

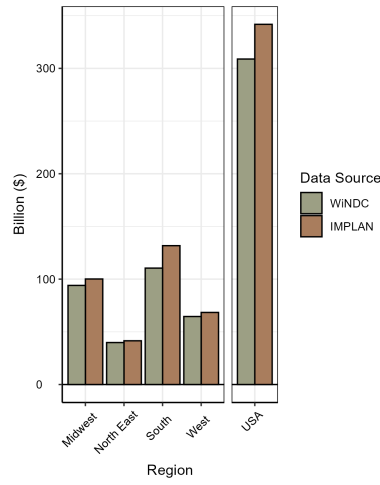
Figure 17: Change from IMPLAN in Capital Demand (kd0)



(a) Electronics and Technology Manufacturing (cpu) (b) Transportation Equipment Manufacturing (tem)

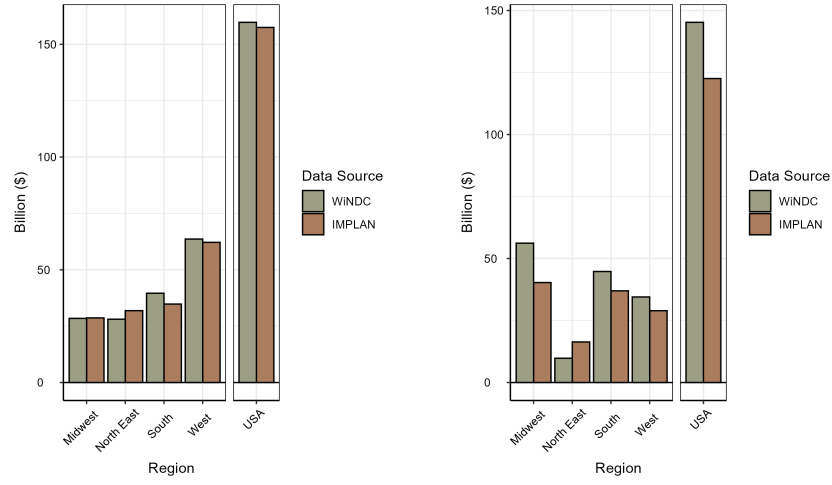


(c) Balance of Manufacturing (bom) (d) Non-truck Transportation (trn)

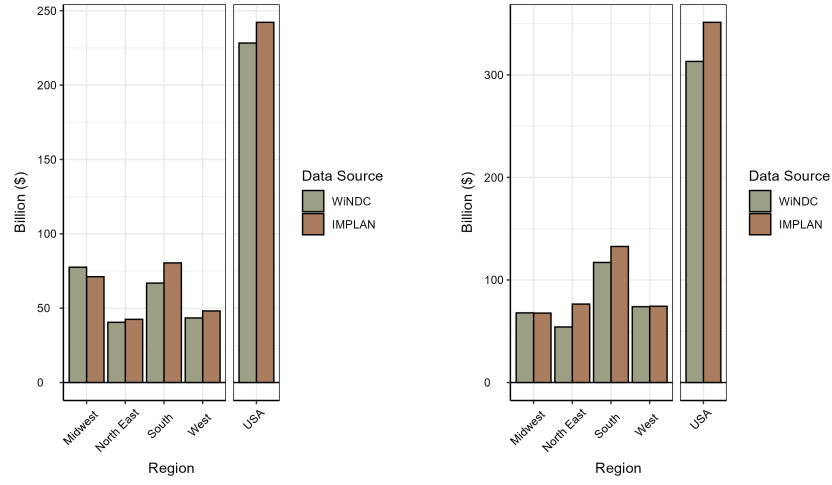


(e) Truck Transportation (ttn)

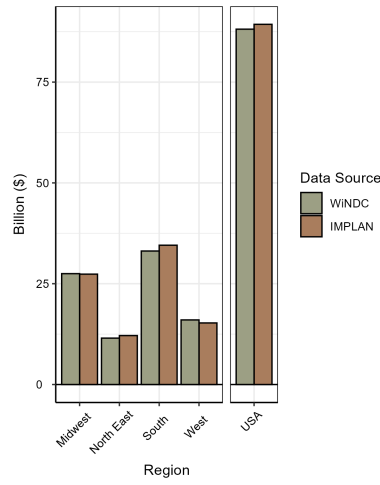
Figure 18: Sectoral Production (ys0) for Select Sectors



(a) Electronics and Technology Manufacturing (cpu) (b) Transportation Equipment Manufacturing (tem)

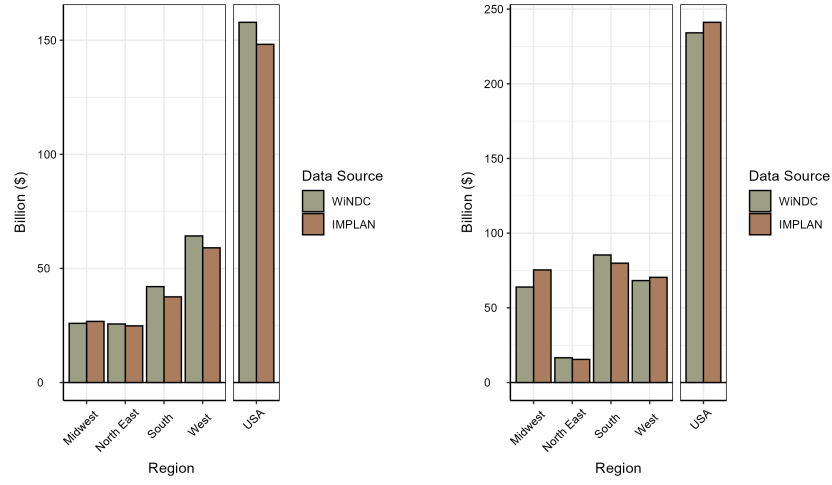


(c) Balance of Manufacturing (bom) (d) Non-truck Transportation (trn)

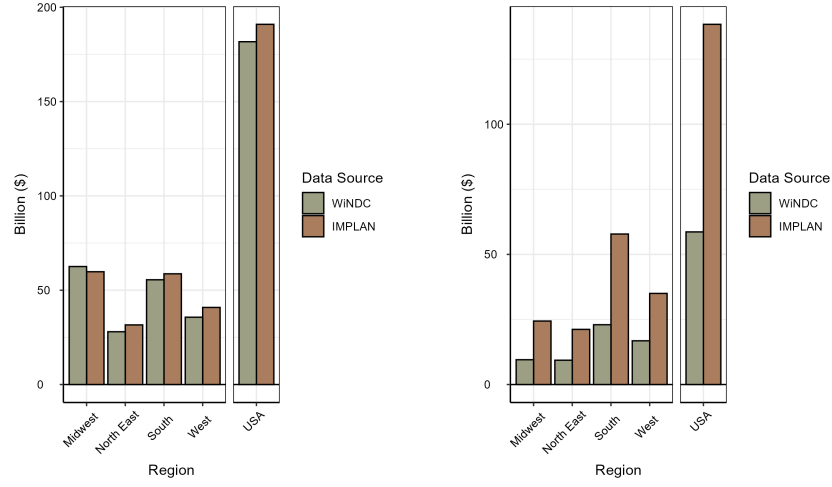


(e) Truck Transportation (ttn)

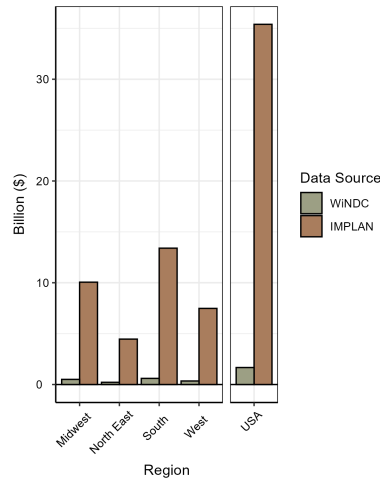
Figure 19: Labor Demand (ld0) for Select Sectors



(a) Electronics and Technology Manufacturing (cpu) (b) Transportation Equipment Manufacturing (tem)



(c) Balance of Manufacturing (bom) (d) Non-truck Transportation (trn)



(e) Truck Transportation (ttn)

Figure 20: Exports (x0) for Select Sectors