WiNDC: **Wi**sconsin **N**ational **D**ata **C**onsortium

Thomas F. Rutherford[♯⊥] and Andrew Schreiber[♭]

[♭]National Center for Environmental Economics, Environmental Protection Agency, US

[♯]Department of Agricultural and Applied Economics, University of Wisconsin-Madison, US

[⊥]Optimization Group, Wisconsin Institute for Discovery, US

October 16, 2018

Abstract

This paper illustrates the Wisconsin National Data Consortium (WiNDC) build routine for producing sub-national economic accounts for the United States used in economic equilibrium models. The intent of this effort is to provide a completely transparent set of tools for facilitating evidence based sub-national economic research. We describe the computer programs used to generate regional social accounting matrices and a calibrated static multi-regional, multi-sectoral computable general equilibrium model which complements the constructed set of data. The modeling framework is intended to be used as a foundational structure from which an empirical model for policy analysis can be based upon. We apply this framework using blueNOTE, the energy-environment module of WiNDC, to assess carbon leakage rates under different configurations of prospective state level climate action. Calculated leakage rates depend significantly on the representation of trade in the model. Including gravity based state level bilateral trade flows produce rates up to 50% smaller relative to a model calibrated with a pooled national market illustrating the importance of bottom up decisions when constructing a dataset for policy analysis.

JEL Codes: C6, C8, D5, Q5, R1.

Keywords: Computable general equilibrium models; Applied economic analysis; Multi-regional models; Small open economy models; Calibration; Air Pollution; Regional economies.

The views expressed in this research are of the authors and not necessarily those of the U.S Environmental Protection Agency or the University of Wisconsin-Madison. All errors are our own. Support is gratefully acknowledged by the Environmental Defense Fund. For helpful comments, we would like to thank Christoph Böhringer, Gökçe Akin-Olçum, Ed Balistreri, Alan Fox, and participants of a workshop hosted by the Environmental Defense Fund in Washington D.C., USA.

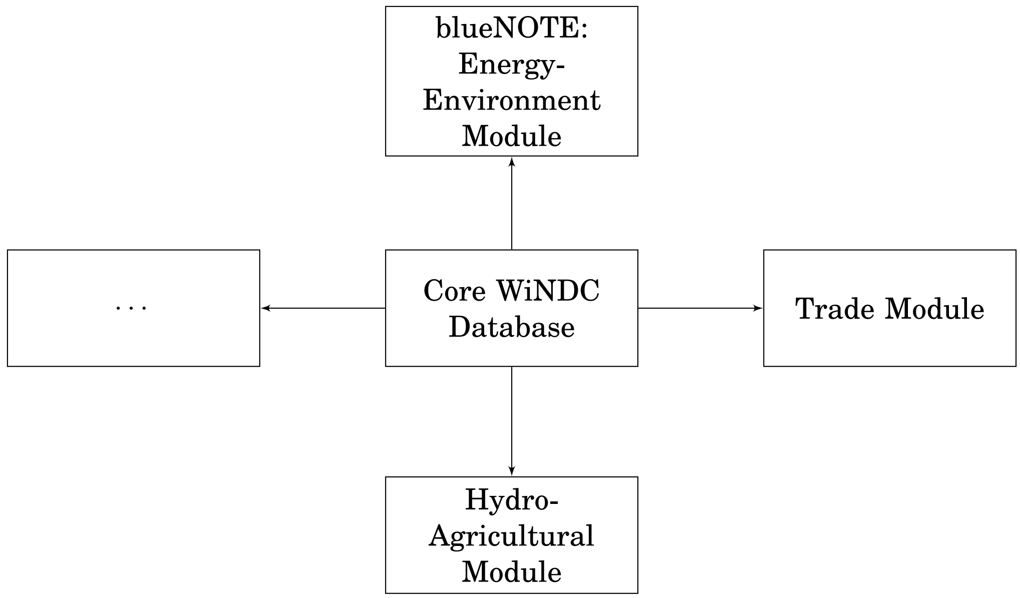
Email addresses: schreiber.andrew@epa.gov (A. Schreiber), rutherford@aae.wisc.edu (T.F. Rutherford).

1 Introduction

This paper describes a build routine and modeling framework for facilitating the use of computable general equilibrium (CGE) modeling as a tool for sub-national economic research in the United States. CGE models provide an ex-ante framework for policy analysis which align an Arrow-Debreu general equilibrium representation of the economy with empirical data. The modeling environment has been used widely across the economics literature; for instance, in trade (e.g. Harrison et al. (1997); Kitwiwattanachai et al. (2010)), energy (e.g. Böhringer and Löschel (2006); Böhringer and Rutherford (2008)), environment (e.g. Böhringer et al. (2016, 2018)), and agriculture (e.g. Calzadilla et al. (2010)). Equilibrium models rely on constructed data sets called input output tables which describe all market transactions between producers and consumers across a given region. The Bureau of Economic Analysis (BEA) reports annual national level input output data, though publicly available regionalized accounts are not available limiting the scope of equilibrium analysis as a tool for decomposing aggregate economic impacts into its distributional components. Economists interested in sub-national policy research are forced to look for outside options (e.g. see Rutherford and Schreiber (2016)) which can be expensive and inflexible in construction.

The process of data construction often relies on assumptions where empirical data is unavailable. We introduce the WiNDC (Wisconsin National Data Consortium) build stream and an extension for specific policy analysis on energy-environment topics in this paper as a system for understanding where the needed assumptions lie along the construction path. The build stream is composed of a set of sub-routines written in GAMS (General Algebraic Modeling System) which produce micro-consistent regionalized economic accounts encompassing the entire United States for 1997-2014 (at the time of writing). Micro-consistent economic data satisfy needed accounting identities to represent an initial equilibrium point in a CGE model. We also presents a canonical modeling framework which complements the constructed set of regional economic accounts. The model, a calibrated multi-regional, multi-sector computable general equilibrium model, provides a foundational structure for specific empirical applications. The aim of WiNDC is to make evidence based regional economic research more accessible and provide a transparent and basic structure of analysis from which to compare

Figure 1: WiNDC Modules



Notes: Included data modules are examples of possible development directions for the WiNDC.

the impacts of policy.

The WiNDC package provides a core database of regionalized accounts with additional modules for specific topical applications as illustrated in figure 1. Such modules rely on re-calibration techniques to enforce consistency between WiNDC accounts and external data sources. This paper introduces blueNOTE as one such module with a focus on energy-environment applications. In blueNOTE, we re-calibrate the core database to match State Energy Data System (SEDS) energy demands and prices and supplement market goods with carbon emissions. Furthermore, we provide a gravity based method for generating state level bilateral trade flows for all sectors and regions in the blueNOTE aggregation.

This module is used to assess the effectiveness of state wide climate policy in an atmosphere of limited federal interest for pollution regulation. We calculate leakage rates of a 20% reduction in statewide emissions for a variety of configurations of state participation depending on the likelihood of climate action. These configurations are based on existing policy frameworks (California's AB-32 and the Regional Greenhouse Gas Initiative), historical attempts for climate action or interest in post-Paris agreement coalitions of state governors and city mayors. This calculation has two intentions. We first use it to benchmark our computed leakage rates with existing studies and second, as an assessment of the importance of bottom up data related decisions in the final results. We find that including explicit bilateral trade flows rather than approximate sub-national trade with a national goods market reduces the calculated leakage rate by up to 50%.

We begin by describing the sub-routines of the WiNDC build stream.[1] A batch file is provided which serves as a launching program for constructing the WiNDC database locally. If explicit code references are necessary, we introduce only fractions of each sub-routine (eliciting the help of "..." to indicate portions of code). We then describe the canonical calibrated multi-regional, multi-sector computable general equilibrium model we use for verifying benchmark consistency. The model is described as a mixed complementarity problem (MCP) with provided code using both MPSGE (Mathematical Programming Software for General Equilibrium) and MCP. We conclude the paper with an application of the blueNOTE module which explores the effectiveness of state level climate policies.

2 Data Reconciliation

The build routine structure is summarized in figure 2. Sub-routines are listed in the order they appear in the figure.

1. `readbea.gms`: routine which collapses and converts downloaded BEA summary input output files (supply and use tables) into a singular file in a GAMS readable format, called GDX (GAMS Data Exchange).

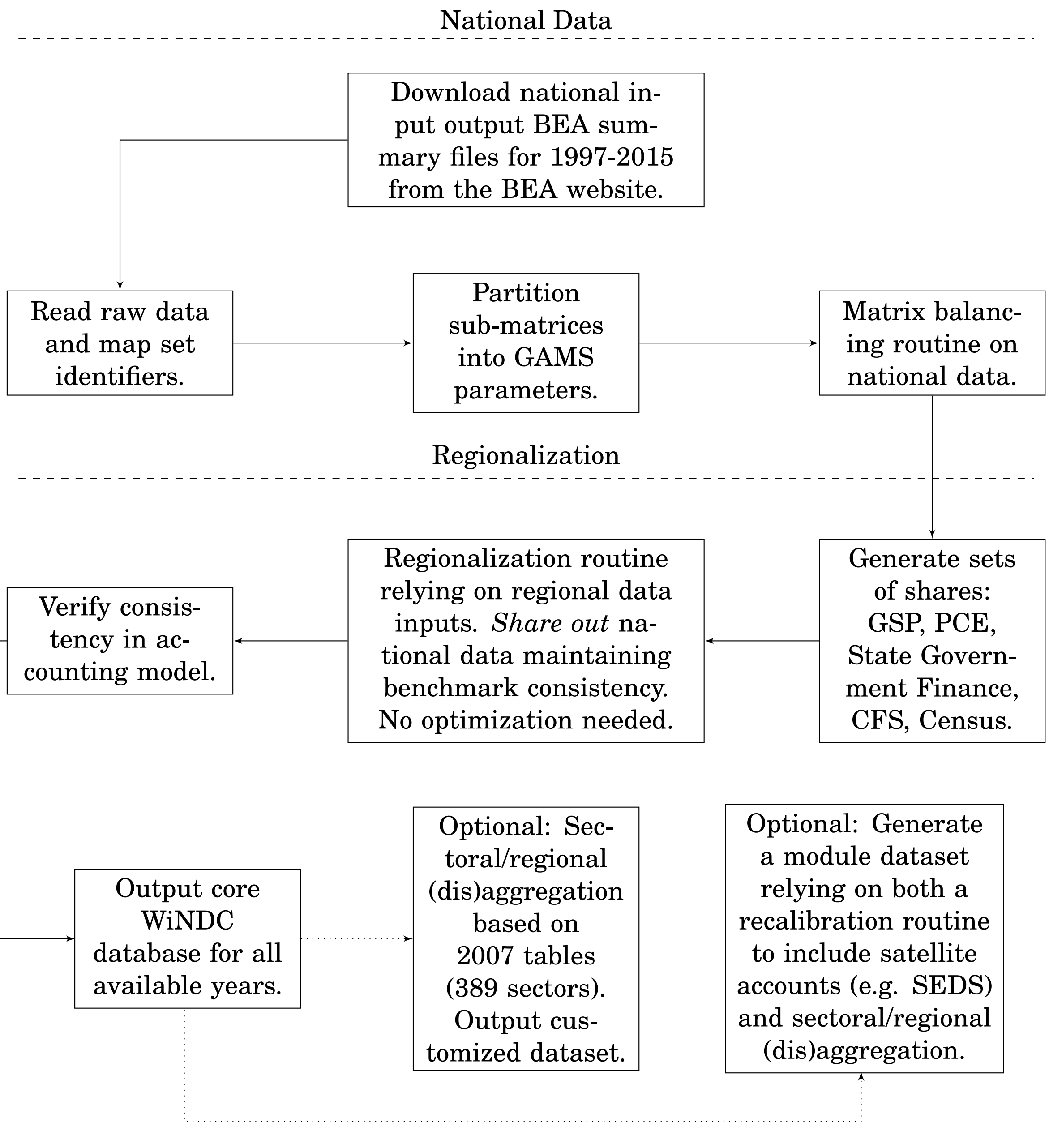
2. `mapbea.gms`: mapping program that re-labels raw input output data sectoring schemes to non-numeric indices.

3. `partitionbea.gms`: matrix partitioning routine which allocates portions of the national input output table to associated CGE based parameters (e.g. intermediate input demand).

4. `calibrate.gms`: optimization based matrix balancing scripts enforces accounting identities in the data.

[1]For the most updated version of WiNDC, see our webpage (`https://windc.wisc.edu`).

Figure 2: Build Stream Process



Notes: Dotted arrowed lines indicate optional portions of the build routine. The core routine ends after outputting a full set of computable general equilibrium parameters for every available year in the source data sets. Recalibrating and/or (dis)aggregating these accounts based on additional information is optional.

5. `read*.gms`: data sources used for regionalization (see table 1) are first read into GAMS and reconciled to match the blueNOTE sectoring scheme. The "*" indicates multiple

programs with names that match the particular source of data that is being introduced into the routine.

6. `*share.gms`: regional shares based on "*" data are generated. Shares based on commodity flow survey data are calculated as *regional purchase coefficients*, or the share of nationally traded goods that remain within the region of production.

7. `statedisagg.gms`: state level regional disaggration routine. All shares are read into this program and used to disaggregate reconciled national data parameters.

8. `statemodel.gms`: state level accounting computable general equilibrium model used to verify benchmark consistency.

9. `sectordisagg.gms`: depending on the scope of the analysis, this routine disaggregates chosen aggregate sectors in the core 71 sectoring scheme from the BEA based on 2007 tables (389 sectors).

10. `aggregate.gms`: sectoral and/or regional aggregation routine.

11. `aggchk.gms`: consistency check on (dis)aggregation.

12. `bluenote.gms`: optional routine which recalibrates the WiNDC dataset to match additional satellite information (e.g. energy demands and supplies and emissions levels).

13. `enforcechk.gms`: computable general equilibrium model used to verify benchmark consistency of the recalibrated dataset.

14. `mcp.gms`: provides a regional-level accounting computable general equilibrium model in both MPSGE and MCP used to verify benchmark consistency.

All source data required for producing sub-national economic input output tables are included in table 1. The first column indicates one of three broader governmental organizations who provide the needed data: the Bureau of Economic Analysis (BEA), the Census Bureau, and the Energy Information Administration (EIA). Specific datasets are indicated along with links to the website for access. The column entitled "Years" indicates the intersection of available data across data sources. For instance, more data is available from the State Energy

Table 1: WiNDC Data Sources

Source	Description	ID	URL	Years
Bureau of Economic Analysis	Supply and Use Tables	*BEA*	https://www.bea.gov/industry/io_annual.htm	1997-2014
	Gross State Product	*GSP*	https://www.bea.gov/newsreleases/regional/gdp_state/qgsp_newsrelease.htm	1997-2014
	Personal Cosumer Expenditures	*PCE*	https://www.bea.gov/newsreleases/regional/pce/pce_newsrelease.htm	1997-2014
Census Bureau	Commodity Flow Survey	*CFS*	https://www.census.gov/econ/cfs/	2012
	State Government Finance	*SGF*	https://www.census.gov/programs-surveys/state/data/tables.All.html	1997-2014
	State Exports	*UTD*	https://usatrade.census.gov	2002-2014
Energy Information Administration	State Energy Data System	*SEDS*	https://www.eia.gov/state/seds/	1997-2014

Notes: Years indicates the intersection of available data across data sources. For instance, more data is available for SEDS on the EIA website but the listed years of data indicate information *used* in the routine.

Data System (SEDS) provided by the EIA, but most of which are not applicable to the data build.

2.1 Using the Batch File

We use a batch file to direct the build routine. The only code that need be altered to generate a local version of the database is contained within `run.bat`.[2] The batch file begins by declaring environment variables for both generating the directory structure for the local copy and options for (dis)aggregation. Options for the level of (dis)aggregation include both regions and sectors (the 2007 BEA tables allows for up to 389 sectors) along with the generation of preset modules. The following section of code is the only instance in the batch file that needs customization. The remainder of the batch file uses these inputs and generates the WiNDC database, providing a scripted version of figure 2.

Declared Environment Variables (`run.bat`)

```
: -------------------------------------------------------------------------------

: Environment variables:

@echo off

: Set working and data source directories:

set   wdir=c:\work\projects\windc\build\
set   ddir=c:\work\data\windc\datasources\

: Output dataset directory:

set   dsdir=c:\work\data\windc\datasets\

: Declare temporary directory which stores all intermediate .gdx files and .lst
: files.
```

[2]This can be easily be done in any text editor. Double clicking on the batch file will run the routine with pre-selected options as included in this paper.

```
set   temp=c:\work\data\windc\temp\

: Level of regionalization.

set   reg=State

: ------------------------------------------------------------------------------
:           Sectoral\Regional (dis)aggregation options
: ------------------------------------------------------------------------------

                                    ...

: Example sectoral and\or regional (dis)aggregations are provided in two files:
: %wdir%defines\secaggregations\example.map
: %wdir%defines\aggregations\example.map

set   sdisagg=yes
set   aggr=example

: Recalibration routine matching satellite data sources. Currently available:
: blueNOTE (SEDS data - 1997-2014), nass (NASS data - 2012), gtap (consistent
: with GTAP), none (no recalibration -- core accounts).

set   year=2014
set   satdata=none
if not %satdata%==none set sdisagg=yes
```

2.2 National Data Reconciliation

The WiNDC build routine begins by reconciling multi-year national level use and supply tables from the Bureau of Economic Analysis. The GAMS programming language relies on declared parameters (data) and variables indexed over sets. Raw BEA data read using supply and use matrices as singular data structures, which is subsequently partitioned into sub-components based on specific row and column subsets. Table 2 reports all sets used within the WiNDC build routine (note that the regional set, r, is used in subsequent sections). The listed sets correspond to row or column sub-indices of the supply and use tables. Sets s and g denote aliased sets for the 71 sectors and commodities in the summary BEA data.[3] Value added, referred to by set va, contains wage payments, other taxes on production and gross operating surplus. Final demand accounts in the national data, denoted by fd, correspond to personal consumption expenditures, investment categories, and government (both federal and state levels) payments. Set m indicates margin adjustments for both trade and transport.

Using these subsets, we partition the annual supply and use tables into CGE parameters. National level parameters are described in table 3. All data parameters are scaled to be *10s*

[3] See table 12 for a listing of core economic sectors in Appendix A.

Table 2: Set Notation in blueNOTE

Type	Item	Description
Sets:	*yr*	Years
	s,g	Sectors/Goods
	m	Margin type
	va	Value added components
	fd	Final demand accounts
	r	Regions

of billions of dollars. Sectoral supply is characterized by the first quadrant of the supply table which provides data on byproducts. For a given sector s and goods g, the parameter $\tilde{ys}_{yr,g,s}$ describes a matrix of annual output vectors. Intermediate demand describing the material inputs needed to produce sectoral level output is defined by the first quadrant of the use table, final demand payments and exports of goods g are captured by the second quadrant of the use table, and value added by component va for sectoral production s and tax payments are partitioned from the third quadrant. Imports, margins (supply of margins are negative in the data and positive for margin demands) and tax payments for each commodity are partitioned from the second quadrant of the supply table. Average tax (duty) rates are subsequently defined based on overall tax payments relative to total input (import) demand.

Given the raw data parameters, we formulate other CGE parameters based on accounting identities needed in a computable general equilibrium model. Gross output is defined as total production (household production is defined as negative payments) net of margin supplies.

$$\tilde{y}_{yr,g} = \sum_{s} \tilde{ys}_{yr,s,g} + \tilde{fs}_{yr,g} - \sum_{m} \tilde{ms}_{yr,g,m} \quad \forall \quad (yr, g) \tag{1}$$

The Armington supply parameter, $\tilde{a}_{yr,g}$, is defined as the total value of goods purchased as both final input and consumption demand.

$$\tilde{a}_{yr,g} = \sum_{s} \tilde{id}_{yr,g,s} + \sum_{fd} \tilde{fd}_{yr,g,fd} \quad \forall \quad (yr, g) \tag{2}$$

The balance of payments, $\tilde{bop}_{yr}$, is defined by the overall difference in value between total

Table 3: Annual National Level Partitioned Parameters From BEA Data

Parameter	GAMS Code	Description
$\tilde{ys}_{yr,s,g}$	`ys0(yr,s,g)`	Sectoral supply (with byproducts)
$\tilde{y}_{yr,s}$	`y0(yr,s)`	Gross output (net margin supply)
$\tilde{fs}_{yr,g}$	`fs0(yr,g)`	Household production
$\tilde{id}_{yr,g,s}$	`id0(yr,g,s)`	Intermediate input demand
$\tilde{va}_{yr,va,s}$	`va0(yr,va,s)`	Value added factor demand
$\tilde{x}_{yr,g}$	`x0(yr,g)`	Foreign exports
$\tilde{m}_{yr,g}$	`m0(yr,g)`	Imports
$\tilde{ms}_{yr,g,m}$	`ms0(yr,g,m)`	Margin supply
$\tilde{md}_{yr,m,g}$	`md0(yr,m,g)`	Margin demand
$\tilde{a}_{yr,g}$	`a0(yr,g)`	Armington supply
$\tilde{ta}_{yr,g}$	`ta0(yr,g)`	Tax (subsidy) rate on intermediate demand
$\tilde{tm}_{yr,g}$	`tm0(yr,g)`	Import tariff rate
$\tilde{fd}_{yr,g,fd}$	`fd0(yr,g,fd)`	Final demand payments
$\tilde{bop}_{yr}$	`bopdef0(yr)`	Balance of payments

Notes: We use ˜ to indicate national data parameters. In subsequent sections, ¯ is used to indicate associated regionalized parameters.

imports and total exports.

$$\tilde{bop}_{yr} = \sum_{g} (\tilde{m}_{yr,g} - \tilde{x}_{yr,g}) \quad \forall \quad (yr, g) \tag{3}$$

The core WiNDC database is constructed using the 71 sector configuration based on BEA summary files. The reference set of partitioned data parameters is assumed to reflect a benchmark equilibrium. An equilibrium is characterized by three sets of conditions in a canonical competitive general equilibrium framework: profits cannot be greater than zero, markets must clear and incomes must balance with expenditures. Given these requirements, certain accounting identities must hold in the data to properly calibrate a model. The subroutine, `calibrate.gms`, ensures that such conditions hold using optimization based matrix balancing techniques. This routine works using *savepoint* and *loadpoint*. Following the initial solve of each problem, the solution is stored in a temporary directory. This allows subsequent runs of the problem to start from a previous solution point to minimize the time of computations.

Zero profits must hold in sectoral production and the absorption and allocation of goods in the national economy. Zero profit in sectoral production requires that the value of supply equals the total value of the cost of production for each sector s.

$$\sum_{g} \tilde{ys}_{yr,s,g} = \sum_{g} \tilde{id}_{yr,g,s} + \sum_{va} \tilde{va}_{yr,va,s} \quad \forall \quad (yr, s) \tag{4}$$

The composition of the goods market must also satisfy a zero profit condition. The total value of goods demanded both via intermediate inputs and final demand along with those exported out of the domestic economy must equal the value of domestically produced goods, imported goods and trade margins.

$$\left(1 - \tilde{ta}_{yr,g}\right) \tilde{a}_{yr,g} + \tilde{x}_{yr,g} = \tilde{y}_{yr,g} + \left(1 + \tilde{tm}_{yr,g}\right) \tilde{m}_{yr,g} + \sum_{m} \tilde{md}_{yr,m,g} \quad \forall \quad (yr, g) \tag{5}$$

Market clearance conditions must hold in factor and goods markets. Multiple goods markets exist based on the location of production, the absorption into the Armington supply and margin use. Because factor payments are not differentiated between agent types in the national data, we refrain from explicit representation. Domestically produced goods both through household and sectoral production must equal gross output demand in the absorption of goods and supply of margins.

$$\sum_{s} \tilde{ys}_{yr,s,g} + \tilde{fs}_{yr,g} = \tilde{y}_{yr,g} + \sum_{m} \tilde{ms}_{yr,g,m} \quad \forall \quad (yr, g) \tag{6}$$

Trade and transport margins are generated through retail and transport sectors and demanded through the absorption of imported and domestically produced goods.

$$\sum_{g} \tilde{ms}_{yr,g,m} = \sum_{g} \tilde{md}_{yr,m,g} \quad \forall \quad (yr, m) \tag{7}$$

Foreign exchange is characterized through imports, exports and the balance of payments. Given the construction of $\tilde{bop}_{yr}$ in equation 3, this holds automatically.

$$\sum_{g} \tilde{x}_{yr,g} + \tilde{bop}_{yr} = \sum_{g} \tilde{m}_{yr,g} \quad \forall \quad yr \tag{8}$$

Finally, the total use of goods via intermediate input demand or final consumption must equal the amount supplied via the Armington supply. This holds through equation 2.

The remaining equilibrium assumption requires that endowment income must satisfy the total value of demand expenditures. Notably, the national level data does not provide information on transfers between agents, so we represent this condition through an aggregated expression.

$$\sum_{fd,g} \tilde{fd}_{yr,g,fd} = \sum_{g} \tilde{fs}_{yr,g} + \tilde{bop}_{yr} + \sum_{va,s} \tilde{va}_{yr,va,s} + \sum_{g} \left(\tilde{ta}_{yr,g} \tilde{a}_{yr,g} + \tilde{tm}_{yr,g} \tilde{m}_{yr,g} \right) \quad \forall \quad yr \tag{9}$$

These identities are enforced through matrix balancing techniques. Optimization based matrix balancing problems are formulated using objective functions which penalize deviations away from available data while satisfying accounting identity constraints. We provide two techniques in the WiNDC build stream to test the sensitivity of different objective functions in the resulting micro-consistent dataset. The first option is formulated as a hybrid between a linear and quadratic objective function. For index pairings which have positive data entries, a weighted least squares penalty is used which minimizes the percent difference between observable data. In instances where parameters are zero, we impose a linear zero penalty weight to limit any changes from zero in the resulting solution. This problem is written generally in square format, using set indices (r, c) to denote a particular row and column element.[4] Let Φ_{rc} denote the subset of (r, c) with non-zero elements, Φ_{rc}^c as its converse, γ represents a positive penalty on zero elements, $\tilde{a}_{rc}$ as available data and A_{rc} as its corresponding variable. The quadratic program is characterized as follows, with $F_i(\cdot) = 0$ denoting the set of accounting identity constraints.

$$\min_{A_{rc}} \quad \sum_{\Phi_{rc}} |\tilde{a_{rc}}| \left(\frac{A_{rc}}{\tilde{a_{rc}}} - 1 \right)^2 + \gamma \sum_{\Phi_{rc}^c} A_{rc}$$

$$\text{s.t} \quad F_i(A, \tilde{a}) = 0 \quad \forall i$$

The second option is a hybrid approach based on Huber (1964). Because the least squares objective function is sensitive to outliers, we formulate an objective function which is piecewise

[4] In practice, however, we have separate terms for each endogenized parameter value.

and depend on the deviation away from the target value. This piecewise objective function adds a log term to penalize values which go to zero.

$$\min_{A_{rc}} \quad \sum_{rc} L(A_{rc}, \tilde{a_{rc}})$$

$$\text{s.t} \quad F_i(A, \tilde{a}) = 0 \quad \forall i$$

$$\text{where } L(A_{rc}, \tilde{a_{rc}}) = \begin{cases} \tilde{a_{rc}}\theta\left(\frac{A_{rc}}{\tilde{a_{rc}}} - 1\right) & \frac{\tilde{A_{rc}}}{\tilde{a}_{rc}} - 1 \geq \theta \\ \tilde{a_{rc}}\left(\frac{A_{rc}}{\tilde{a_{rc}}} - 1\right)^2 & \theta \geq \frac{\tilde{A_{rc}}}{\tilde{a}_{rc}} - 1 \geq -\gamma \\ \tilde{a_{rc}}\gamma(1-\gamma)\log\left(\frac{A_{rc}}{\tilde{a_{rc}}}\right) & \frac{\tilde{A_{rc}}}{\tilde{a}_{rc}} - 1 \leq -\gamma \end{cases}$$

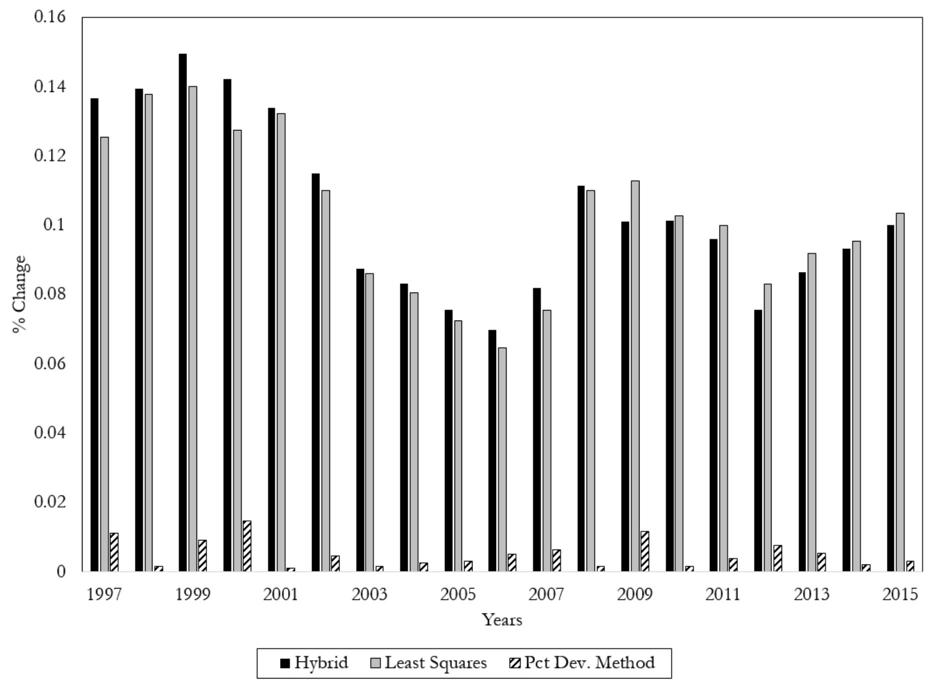
The differences in objective functions between methods produces slight differences in the resulting micro-consistent datasets. Figure 3 illustrates the annual aggregate deviations from the benchmark data by optimization method (least squares or the hybrid method) and the percent deviation between both optimization frameworks. The datasets are balanced in each year. The total percent deviation from the original data needed to satisfy all accounting identities ranges from 0.07-0.15%. By the percent deviation metric, the hybrid method tends to outperform least squares in more recent years, though changes are otherwise small.

2.3 Sub-National Regionalization

The method of regionalization used in the WiNDC build routine relies on *sharing*, rather than optimization. Given data on the composition of the regional economy, we generate regional shares for associated CGE parameters to disaggregate the set of fully micro-consistent national accounts. This method has a number of advantages. The goal of the WiNDC is to provide a comprehensive dataset of the United States with specific representation of sub-national regions. The sharing method provides the routine a handshake on national totals (summing across regions) and perhaps most importantly, a measure of simplicity and sense of transparency.

The build stream uses data for generating regional shares as outlined in table 1. `read*.gms` and `*share.gms` correspond to subroutines which read a particular dataset (where the asterisk indicates the dataset) and produce regional shares. In cases aside from the trade data, all

Figure 3: Total Matrix Balancing Deviations



shares are generated such that summing across regions equals unity. Because there is typically not a 1-1 mapping between regional sector level data and national data (regional data tends to be more aggregated), sector mappings are created to associate disaggregate sector national level data with the regional economic composition of aggregated accounts.

Let $\alpha^*_{yr,r,s}$ be the regional sharing parameter associated with the $\star$ dataset, such that $\sum_r \alpha^*_{yr,r,s} = 1$. Table 4 provides a listing of all regional CGE parameters generated through the regional disaggregation process in `statedisagg.gms`. Gross State Product (GSP) data is used to disaggregate sectoral production accounts. We also use information in this dataset to calculate the regional share of labor relative to capital to distinguish regional production technologies. Let $\theta^{ls}_{yr,r,s} \in (0,1)$ denote the share of labor in total value added.[5] Using gross state product shares maintains the zero-profit condition by region.

$$\bar{ys}_{yr,r,s,g} = \alpha^{gsp}_{yr,r,s} \tilde{ys}_{yr,s,g} \quad \forall \quad (yr,r,s,g) \tag{10}$$

$$\bar{id}_{yr,r,g,s} = \alpha^{gsp}_{yr,r,s} \tilde{id}_{yr,g,s} \quad \forall \quad (yr,r,g,s) \tag{11}$$

[5] We formulate labor shares according to averages (across years) in years with negative capital payments to enforce shares of less than 1.

Table 4: Parameters in the Regional CGE Model

Parameter	**GAMS Code**	**Description**
$\bar{ys}_{yr,r,s,g}$	`ys0(yr,r,s,g)`	Sectoral supply (with byproducts)
$\bar{id}_{yr,r,g,s}$	`id0(yr,r,g,s)`	Intermediate demand
$\bar{ld}_{yr,r,s}$	`ld0(yr,r,s)`	Labor demand
$\bar{kd}_{yr,r,s}$	`kd0(yr,r,s)`	Capital demand
$\bar{cd}_{yr,r,g}$	`cd0(yr,r,g)`	Final demand
$\bar{yh}_{yr,r,g}$	`yh0(yr,r,g)`	Household production
$\bar{g}_{yr,r,g}$	`g0(yr,r,g)`	Government demand
$\bar{i}_{yr,r,g}$	`i0(yr,r,g)`	Investment demand
$\bar{s}_{yr,r,g}$	`s0(yr,r,g)`	Aggregate supply
$\bar{xn}_{yr,r,g}$	`xn0(yr,r,g)`	National supply
$\bar{xd}_{yr,r,g}$	`xd0(yr,r,g)`	State level supply
$\bar{x}_{yr,r,g}$	`x0(yr,r,g)`	Foreign exports
$\bar{a}_{yr,r,g}$	`a0(yr,r,g)`	Armington supply
$\bar{m}_{yr,r,g}$	`m0(yr,r,g)`	Imports
$\bar{nd}_{yr,r,g}$	`nd0(yr,r,g)`	National demand
$\bar{dd}_{yr,r,g}$	`dd0(yr,r,g)`	State level demand
$\bar{bop}_{yr,r}$	`bopdef0(yr,r)`	Balance of payments
$\bar{ta}_{yr,r,g}$	`ta0(yr,r,g)`	Tax net subsidy rate on intermediate demand
$\bar{tm}_{yr,r,g}$	`tm0(yr,r,g)`	Import tariff
$\bar{md}_{yr,r,m,g}$	`md0(yr,r,m,g)`	Margin demand
$\bar{nm}_{yr,r,g,m}$	`nm0(yr,r,g,m)`	National margin supply
$\bar{dm}_{yr,r,g,m}$	`dm0(yr,r,g,m)`	State level margin supply

$$\bar{ld}_{yr,r,s} = \theta^{ls}_{yr,r,s}\alpha^{gsp}_{yr,r,s}\sum_{va}\tilde{va}_{yr,va,s} \quad \forall \quad (yr,r,s) \tag{12}$$

$$\bar{ld}_{yr,r,s} = \left(1-\theta^{ls}_{yr,r,s}\right)\alpha^{gsp}_{yr,r,s}\sum_{va}\tilde{va}_{yr,va,s} \quad \forall \quad (yr,r,s) \tag{13}$$

Final demand categories are aggregated into either aggregate household consumption (C), investment (I) or government expenditures (G). The latter is a combination of both federal and state and local levels encompasses spending on defense, infrastructure, education, and equipment. Regional shares used for household final demand come from the Personal Consumption Expenditure (PCE) dataset by the BEA and shares for government purchases come from the State Government Finance (SGF) tables from the Census Bureau. Investment demand and household production are regionally disaggregated according to gross state product data.

$$\bar{cd}_{yr,r,g} = \alpha^{pce}_{yr,r,g}\sum_{C\in fd}\tilde{fd}_{yr,g,fd} \quad \forall \quad (yr,r,g) \tag{14}$$

$$\bar{g}_{yr,r,g} = \alpha^{sgf}_{yr,r,g}\sum_{G\in fd}\tilde{fd}_{yr,g,fd} \quad \forall \quad (yr,r,g) \tag{15}$$

$$\bar{i}_{yr,r,g} = \alpha^{gsp}_{yr,r,g} \sum_{I \in fd} \tilde{fd}_{yr,g,fd} \quad \forall \quad (yr, r, g) \tag{16}$$

$$\bar{yh}_{yr,r,g} = \alpha^{gsp}_{yr,r,g} \tilde{fs}_{yr,g} \quad \forall \quad (yr, r, g) \tag{17}$$

These disaggregate parameters are used to calculate the regionalized Armington supply parameter similarly to the national accounting identity.

$$\bar{a}_{yr,r,g} = \bar{cd}_{yr,r,g} + \bar{g}_{yr,r,g} + \bar{i}_{yr,r,g} + \sum_{s} \bar{id}_{yr,r,g,s} \quad \forall \quad (yr, r, g) \tag{18}$$

Gross output can also be computed from regionalized parameters. Notably, we include margin supply in gross output (in the national parameter, gross output was net of margins), to distinguish margins generated through state level or national level demands.

$$\bar{s}_{yr,r,g} = \sum_{s} \bar{ys}_{yr,r,s,g} + \bar{yh}_{yr,r,g} \quad \forall \quad (yr, r, g) \tag{19}$$

Given a degree of freedom in pinning down trade totals, we use a mixture of state level export data from USA Trade Online (or UTD, from the Census Bureau) and gross state product to regionally disaggregate export totals.[6]

$$\bar{x}_{yr,r,g} = \alpha^{utd}_{yr,r,g} \tilde{x}_{yr,g} \quad \forall \quad (yr, r, g) \tag{20}$$

Without information on region tax rates, we assume that: $\bar{ta}_{yr,r,g} = \tilde{ta}_{yr,g}$ and $\bar{tm}_{yr,r,g} = \tilde{tm}_{yr,g}$. By pinning down the Armington supply, we calculate implicit shares based on regional absorption (which embed the differences in shares across data sources).

$$\alpha^{abs}_{yr,r,g} = \bar{a}_{yr,r,g} / \sum_{rr} \bar{a}_{yr,r,g} \quad \forall \quad (yr, r, g) \tag{21}$$

The regional absorption of goods are generated through imports, local and national supply and margins. Therefore, regional imports and margin demand are computed using the implicit

[6]USA Trade Online does not include a comprehensive list of sectors to cover all of those from the national input output data. For this reason, we assume exports follow gross state product for sectors not included.

share of regional absorption.

$$\bar{m}_{yr,r,g} = \alpha^{abs}_{yr,r,g}\tilde{m}_{yr,g} \tag{22}$$

$$\bar{md}_{yr,r,m,g} = \alpha^{abs}_{yr,r,g}\tilde{md}_{yr,m,g} \tag{23}$$

Given this regional configuration, there are instances where $\bar{s}_{yr,r,g} - \bar{x}_{yr,r,g} < 0$ predominantly in the "Noncomparable imports and rest-of-the-world adjustment" sector market. We create a parameter, $\bar{rx}_{yr,r,g}$, to indicate *re-exports* for this reason, where $\bar{rx}_{yr,r,g} = \bar{x}_{yr,r,g} - \bar{s}_{yr,r,g}$ when the total regional supply is smaller than regional exports for a given good. The regional balance of payments is calculated based on imports and exports (re-exports cancel out).

$$\bar{bop}_{yr,r} = \sum_g \left(\bar{m}_{yr,r,g} - \bar{x}_{yr,r,g}\right) \quad \forall \quad (yr, r) \tag{24}$$

From the *demand* side of the market, the difference between the value of the total regional absorption $((1 - \bar{ta})\bar{a})$ net of re-exports and imports plus margin demand must equal the maximum level of state (or nationally) sourced goods demand. Conversely, the maximum level of state (or national) supply is governed by the supply side of the market, namely, the difference between total supply and foreign exports net of re-exports. We let the supply and demand side of the market dictate the possible levels of state and nationally produced supply and demand. Let $\hat{dd}_{yr,r,g}$ denote the maximum *possible* level of state level goods demands.

$$\begin{aligned}\hat{dd}_{yr,r,g} = \min\{&(1 - \bar{ta}_{yr,r,g})\bar{a}_{yr,r,g} + \bar{rx}_{yr,r,g} - (1 + \bar{tm}_{yr,r,g})\bar{m}_{yr,r,g} - \sum_m \bar{md}_{yr,r,m,g},\\ &\bar{s}_{yr,r,g} - (\bar{x}_{yr,r,g} - \bar{rx}_{yr,r,g})\} \quad \forall \quad (yr, r, g)\end{aligned} \tag{25}$$

In order to determine the share of the maximum which is represented by the parameter $\bar{dd}_{yr,r,g}$, we use data from the Commodity Flow Survey (CFS) for 2012. The CFS catalogs all commodity shipments between and within states.[7] Using these data, we characterize *regional purchase coefficients*, $\rho^{cfs}_{r,g}$, as the share of a given commodity's state demand relative to total national demand (from all states, included itself). For service sectors not included in the CFS, averages are used. Furthermore, we set the RPC for utilities near unity. State level

[7]Notably, these indicate *all* shipments, not origin and final destination. As such, the data includes state pairings for transshipments.

demand is determined by these regional purchase coefficients.

$$\bar{dd}_{yr,r,g} = \rho^{cfs}_{r,g} \hat{dd}_{yr,r,g} \quad \forall \quad (yr, r, g) \tag{26}$$

Goods demand from other national markets must satisfy the following closure condition, or the difference between the Armington supply (net of re-exports) and local and foreign demand for goods (net of margins).

$$\begin{aligned} \bar{nd}_{yr,r,g} =& (1 - \bar{ta}_{yr,r,g})\bar{a}_{yr,r,g} + \bar{rx}_{yr,r,g} - \bar{dd}_{yr,r,g} \\ & - (1 + \bar{tm}_{yr,r,g})\bar{m}_{yr,r,g} - \sum_m \bar{md}_{yr,r,m,g} \quad \forall \quad (yr, r, g) \end{aligned} \tag{27}$$

Margins are supplied either through the state or national supply of goods. Total margin supply, $\hat{ms}_{yr,r,g,m}$ can be characterized through shares generated from $\bar{md}_{yr,r,m,g}$.

$$\hat{ms}_{yr,r,g,m} = \frac{\sum_{g'} \bar{md}_{yr,r,m,g'}}{\sum_{r'g'} \bar{md}_{yr,r',m,g'}} \tilde{ms}_{yr,g,m} \tag{28}$$

The share of trade and transport margins can calculated as: $\beta^{mar}_{yr,r,m,g} = \hat{ms}_{yr,r,g,m} / \sum_{m'} \hat{ms}_{yr,r,g,m'}$. Using these parameters, we characterize the share of total margin supply coming from the state supply of goods relative to the national supply similarly using the supply and demand side of the market along with information on the state level demand relative to national demands using the regional purchase coefficient.

$$\begin{aligned} \bar{dm}_{yr,r,g,m} = \min\{&\rho^{cfs}_{r,g} \hat{ms}_{yr,r,g,m}, \\ & \beta^{mar}_{yr,r,m,g}(\bar{s}_{yr,r,g} - \bar{x}_{yr,r,g} + \bar{rx}_{yr,r,g} - \bar{dd}_{yr,r,g})\} \quad \forall \quad (yr, r, m, g) \end{aligned} \tag{29}$$

Margins from the national supply follow directly.

$$\bar{nm}_{yr,r,g,m} = \hat{ms}_{yr,r,g,m} - \bar{dm}_{yr,r,g,m} \quad \forall \quad (yr, r, m, g) \tag{30}$$

The regional and national supply must then be tied down using margin supply and goods demand parameters previously computed.

$$\bar{xd}_{yr,r,g} = \sum_{m} \bar{dm}_{yr,r,g,m} + \bar{dd}_{yr,r,g} \quad \forall \quad (yr, r, g) \tag{31}$$

The national supply of goods follows from the difference in total supply and state and foreign supply (net of re-exports).

$$\bar{xn}_{yr,r,g} = \bar{s}_{yr,r,g} + \bar{rx}_{yr,r,g} - \bar{xd}_{yr,r,g} - \bar{x}_{yr,r,g} \quad \forall \quad (yr, r, g) \tag{32}$$

The remaining code in `statedisagg.gms` serves to verify equilibrium accounting identities both through reports and the reference state level computable general equilibrium model found in `statemodel.gms`. Once all regionalized parameters pass a series of consistency checks, the core WiNDC database is constructed including all regional data for all available years.

2.4 Sectoral/Regional Customizations

Basic sector and regional customizations are possible once the core WiNDC database is constructed using 51 regions (in a state level database) and 71 sectors. Often times, a CGE analysis requires specific attention to particular region-sector pairings and an aggregated treatment of other sectors and regions for better numerical precision and faster computational speeds. The subroutine, `sectordisagg.gms`, provides the ability to partially (or fully) disaggregate any of the 71 sectors based on the 2007 BEA make and use input output tables with 389 sectors. For instance, in energy applications, splitting the aggregated utilities sector into its electricity, gas and water subcomponents may provide a level of detail necessary to model the policy. The routine, `aggregate.gms`, allows the user to subsequently aggregate sectors and regions fitting a given policy application.

3 Canonical Static Multi-Regional Multi-Sectoral CGE Model

The canonical, multi-regional, multi-sectoral computable general equilibrium model which complements the constructed set of regional economic parameters is formulated in a static

Table 5: Nomenclature in the Regional CGE Model

Type	Item	Description
Activity Levels:	$Y_{r,s}$	Sectoral output
	$A_{r,g}$	Armington composite
	$X_{r,g}$	Supply allocation
	$MS_{r,m}$	Margin supply
Prices:	$p^Y_{r,g}$	Output market price
	$p^A_{r,g}$	Armington composite price index
	$p^D_{r,g}$	State level market price for goods
	p^N_g	National market price for goods
	p^{FX}	Foreign exchange rate
	p^L_r	Wage rates
	$p^K_{r,s}$	Capital rental rates
	$p^M_{r,m}$	Margins markup
Agents:	RA_r	Representative household income

framework. For this reason, a single year of data must be chosen for model calibration (the year index will be subsequently suppressed). The variables required for the canonical model are described in table 5. $Y_{r,s}$ is total production by sector s in region r, $X_{r,g}$ denotes the allocation of good g in region r to either the state, national or foreign market, $A_{r,g}$ represents the absorption of goods in region r for good g and $MS_{r,m}$ is margin supply of margin type m in region r. Each activity level variable is associated with zero profit accounting identities in the data.

$$Y_{r,s}: \quad \sum_g \bar{ys}_{r,s,g} - \sum_g \bar{id}_{r,g,s} - \bar{ld}_{r,s} - \bar{kd}_{r,s} = 0 \quad \forall \quad (r,s) \tag{33}$$

$$X_{r,g}: \quad \bar{x}_{r,g} - \bar{rx}_{r,g} + \bar{xn}_{r,g} + \bar{xd}_{r,g} - \bar{s}_{r,g} = 0 \quad \forall \quad (r,g) \tag{34}$$

$$A_{r,g}: \quad (1-\bar{ta}_{r,g})\bar{a}_{r,g} + \bar{rx}_{r,g} - \bar{nd}_{r,g} - \bar{dd}_{r,g} - (1+\bar{tm}_{r,g})\bar{m}_{r,g} - \sum_m \bar{md}_{r,m,g} = 0 \quad \forall \quad (r,g) \tag{35}$$

$$MS_{r,m}: \quad \sum_g \bar{md}_{r,m,g} - \sum_g \left(\bar{nm}_{r,g,m} + \bar{dm}_{r,g,m}\right) = 0 \quad \forall \quad (r,m) \tag{36}$$

Price variables are associated with market clearance accounting identities. We distinguish between eight price categories. p^L_r denotes the regional wage rate, $p^K_{r,s}$ is the sector specific capital rental rate in region r for sector s, $p^Y_{r,g}$ denotes the output market price, $p^D_{r,g}$ is the state level market price for good g, p_g^N represents the price for good g from the national market, p^{FX} is the price of foreign exchange and p_{rm}^M denotes the price of margin type m in region r. Notably, some of these accounting conditions hold trivially, though are explicitly represented for completeness.

$$p_{r,g}^Y: \quad \sum_s \bar{ys}_{r,s,g} + \bar{yh}_{r,g} - \bar{s}_{r,g} = 0 \quad \forall \quad (r,g) \tag{37}$$

$$p_{r,g}^A: \quad \bar{a}_{r,g} - \sum_s \bar{id}_{r,g,s} - \bar{cd}_{r,g} - \bar{g}_{r,g} - \bar{i}_{r,g} = 0 \quad \forall \quad (r,g) \tag{38}$$

$$p_{r,g}^D: \quad \bar{xd}_{r,g} - \sum_m \bar{dm}_{r,g,m} - \bar{dd}_{r,g} = 0 \quad \forall \quad (r,g) \tag{39}$$

$$p_g^N: \quad \sum_r \left(\bar{xn}_{r,g} - \sum_m \bar{nm}_{r,g,m} - \bar{nd}_{r,g} \right) = 0 \quad \forall \quad (g) \tag{40}$$

$$p^{FX}: \quad \sum_{r,g} (\bar{x}_{r,g} - \bar{m}_{r,g}) - \sum_r \bar{bop}_r = 0 \tag{41}$$

$$p_r^L: \quad \sum_s \bar{ld}_{r,s} - \sum_s \bar{ld}_{r,s} = 0 \quad \forall \quad (r) \tag{42}$$

$$p_{r,s}^K: \quad \bar{kd}_{r,s} - \bar{kd}_{r,s} = 0 \quad \forall \quad (r,s) \tag{43}$$

$$p_{r,m}^M: \quad \sum_g \bar{md}_{r,m,g} - \sum_g \bar{md}_{r,m,g} = 0 \quad \forall \quad (r,m) \tag{44}$$

Income variables are associated with income balance accounting identities. In our canonical framework, we explicitly represent a representative regional household (RA_r denotes such agent's income level), holding fixed government and investment demand. This is something that can be extended if revenue recycling policy mechanisms is of interest. Given the regionalization routine outline in this paper, regional household income $(\bar{adj}_r^{hh})$ adjustment parameters are necessary for redistributing incomes across regions in order to equate endowment incomes with reference demands. Note that: $\sum_r \bar{adj}^{hh}{}_r = 0$.

$$\begin{aligned} RA_r: \quad & \sum_g (\bar{cd}_{r,g} + \bar{g}_{r,g} + \bar{i}_{r,g}) - \sum_g \bar{yh}_{r,g} - \sum_s (\bar{ld}_{r,s} + \bar{kd}_{r,s}) \\ & - \bar{bop}_r - \bar{adj}_r^{hh} - \sum_g (\bar{ta}_{r,g}\bar{a}_{r,g} + \bar{tm}_{r,g}\bar{m}_{r,g}) = 0 \quad \forall \quad (r) \end{aligned} \tag{45}$$

Table 6: Regional Social Accounting Matrix

		Production	**Exports**	**Absorption Composite**	**Margins**	**Output Market**	**Regional Market**	**National Market**	**Domestic Composite**	**Factors**	**Margins Market**	**Trade**	**Agents**
		$Y_{r,s}$	$X_{r,g}$	$A_{r,g}$	$M_{r,m}$	$p^Y_{r,g}$	$p^D_{r,g}$	p^N_s	$p^A_{r,g}$	$p^L_r, p^K_{r,s}$	$p^M_{r,m}$	P^{FX}	
Production	$Y_{r,s}$					$\bar{ys}_{r,s,g}$							
Exports	$X_{r,g}$						$\bar{xd}_{r,g}$	$\bar{xn}_{r,g}$				$\bar{x}_{r,g}$	
Absorption Composite	$A_{r,g}$								$\bar{a}_{r,g}$				
Margins	$M_{r,m}$										$\bar{md}_{r,m,g}$		
Output Market	$p^Y_{r,g}$		$\bar{s}_{r,g}$										
Regional Market	$p^D_{r,g}$			$\bar{dd}_{r,g}$	$\bar{dm}_{r,g,m}$								
National Market	p^N_g			$\bar{nd}_{r,g}$	$\bar{nm}_{r,g,m}$								
Domestic Composite	$p^A_{r,g}$	$\bar{id}_{r,g,s}$											$\bar{cd}_{r,g}$, $\bar{g}_{r,g}$,$\bar{i}_{r,g}$
Factors	$p^L_r, p^K_{r,s}$	$\bar{ld}_{r,s}, \bar{kd}_{r,s}$											
Margins Market	p^M_m			$\bar{md}_{r,m,g}$									
Trade	p^{FX}			$\bar{m}_{r,g}$									
Agents						$\bar{yh}_{r,g}$				$\bar{ld}_{r,s}, \bar{kd}_{r,s}$		$\bar{bop}_r$	

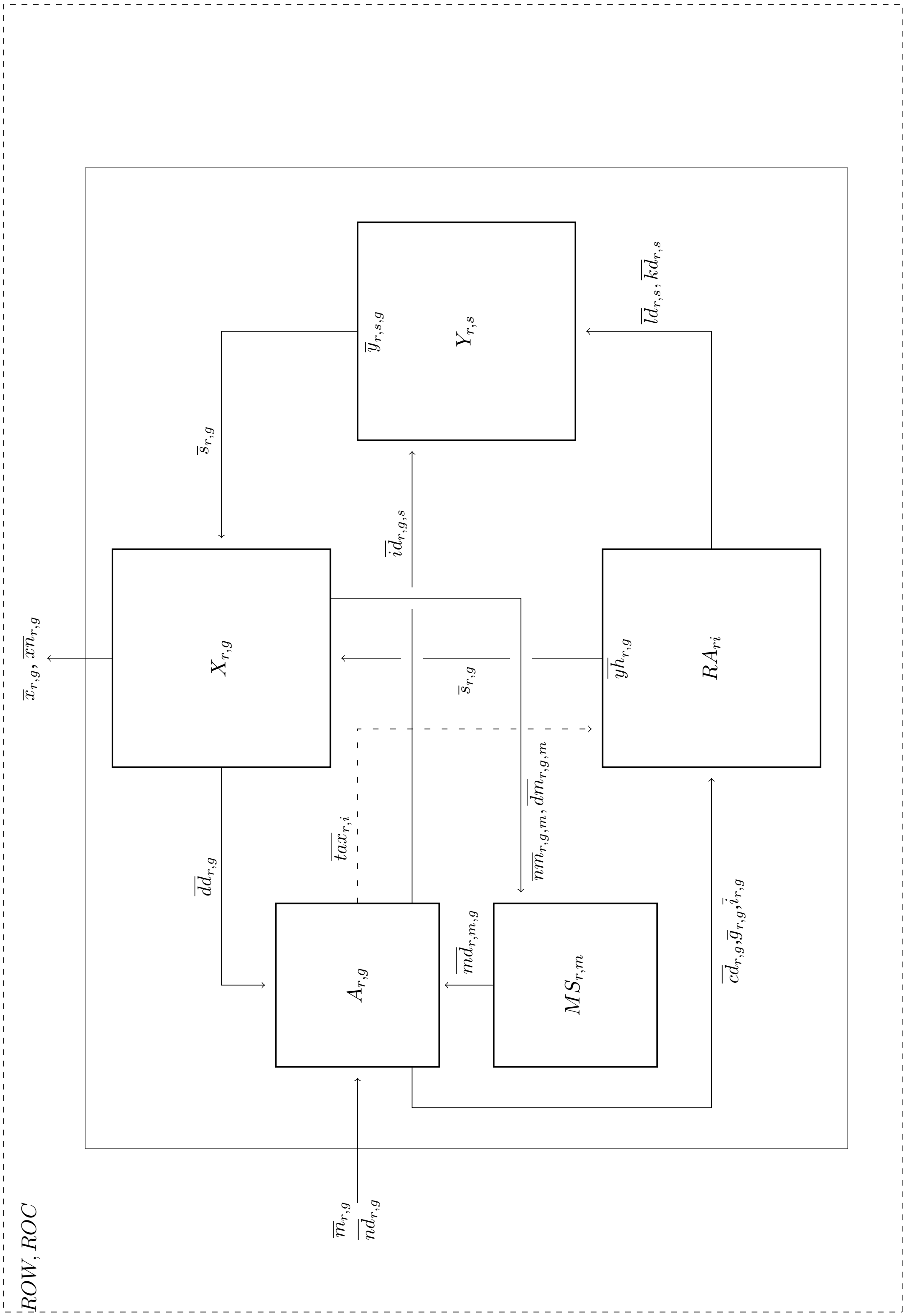
These accounting identities can be visually represented in the context of a regional social accounting matrix in table 6. Rows indicate outputs or endowments for the associated condition (zero profit, market clearance or income balance) and columns denote inputs or demands. Rows/columns which are labeled with activity levels represent zero-profit accounting identities, those labeled with prices represent market clearance conditions and income balance identities can be characterized in the Agent column/row. The gray cell in the bottom right corner of the matrix characterizes the "fourth quadrant" which typically includes transfers between agents. The core WiNDC build refrains from constructing a set of transfers between agents, though may be added as satellite information if distributional impacts of policy is of interest.

The regional economic flows are illustrated in figure 4. Sectoral production ($Y_{r,s}$) is supplied with factors of production (both capital and labor) by the representative agent (RA_r) and material inputs from the goods market ($A_{r,g}$). The resulting supply is sent to the allocation market ($X_{r,g}$) where total supply (including household production $\bar{yh}_{r,g}$) is distributed to the regional level market ($\bar{xd}_{r,g}$), the national market ($\bar{xn}_{r,g}$) or the foreign market ($\bar{x}_{r,g}$). Regional level supply is allocated either to margin formulation ($\bar{dm}_{r,m}$) or for regional demand of goods ($\bar{dd}_{r,g}$). The total pool of goods is a blend of regional goods, nationally supplied goods outside of region r ($\bar{nd}_{r,g}$) and imports ($\bar{m}_{r,g}$). Taxes on goods are allocated back to the representative agent. The total pool of goods is allocated to final demanding agents ($\bar{cd}_{r,g}, \bar{g}_{r,g}, \bar{i}_{r,g}$) or back for intermediate demand in sectoral production ($\bar{id}_{r,g,s}$).

3.1 The Primal Formulation

The identities presented in the previous section indicate the conditions that must hold in the *data* to reflect a benchmark equilibrium but do not, however, characterize the behavior of agents in the model. We characterize the decentralized optimization problems of the subcomponents of the model in this section. Producers are assumed to be profit maximizers in a constant returns to scale environment (which we equivalently write as a cost minimization problem). Input choices for sector $Y_{r,s}$ are found by the following minimization problem. Note that these optimization problems solve for quantities in terms of prices. Data notation is used

Figure 4: Regional Economic Structure

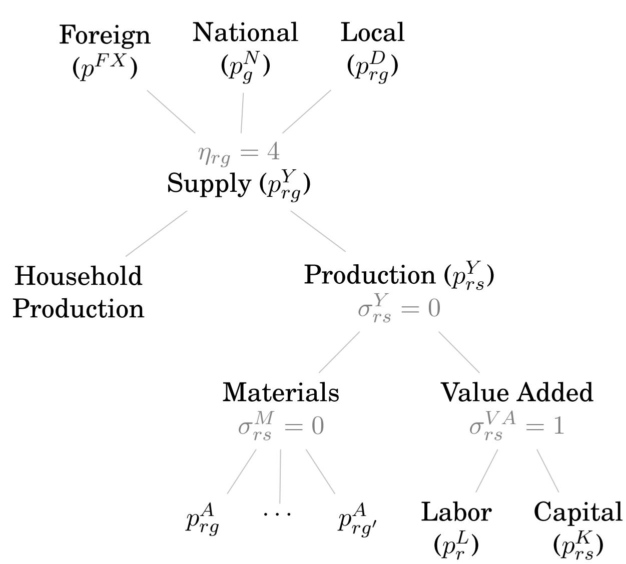


to describe variables without the overline.

$$
\begin{aligned}
\min_{id,ld,kd} \quad & c^M_{r,s} + c^F_{r,s} \\
\text{s.t.} \quad & c^M_{r,s} = \sum_g p^A_{r,g} id_{r,g,s} \\
& c^F_{r,s} = p^L_r ld_{r,s} + p^K_{r,s} kd_{r,s} \\
& F^Y_{r,s}(id_{r,g,s}, ld_{r,s}, kd_{r,s}) = Y_{r,s}
\end{aligned}
$$

The production function $F^Y(\cdot)$ is described by a nested constant elasticity of substitution (CES) form shown in the lower part of the tree diagram in figure 5. We characterize the nesting structure using associated prices of input demands. We let σ denote the constant elasticity of substitution governing input trade offs in production. $\sigma^{VA}_{r,s}$ represents the substitution elasticity between factors of production (assumed to be Cobb-Douglas), $\sigma^M_{r,s}$ is the substitution elasticity between material inputs (assumed to be Leontief) and $\sigma^Y_{r,s}$ denotes the top level substitution elasticity between total value added and total material inputs (also assumed to be Leontief).

Figure 5: Production (CES) and Supply (CET) Structure



The allocation of supply to the foreign, state or national market arise from the following profit maximization problem.[8]

$$\max_{x,xn,xd,rx} \quad p^{FX}(x_{r,g} - rx_{r,g}) + p^N_g xn_{r,g} + p^D_{r,g} xd_g$$
$$\text{s.t.} \quad F^X_{r,g}(x_{r,g}, rx_{r,g}, xn_{r,g}, xd_{r,g}) = X_{r,g}$$

$F^X(\cdot)$ is a constant elasticity of transformation (CET) function with form shown in the upper part of figure 5. $\eta_{r,g}$ denotes the elasticity of transformation governing the supply of output to regional markets. For both production and supply, sample elasticities are chosen, though estimated or calibrated values can be used (for instance, see Lanz and Rutherford (2016)).

The demand for goods across geographically distinct markets (state, national and foreign) follows Armington (1969) where substitution possibilities depend on the location of production. This is reflected by the following cost minimization problem.

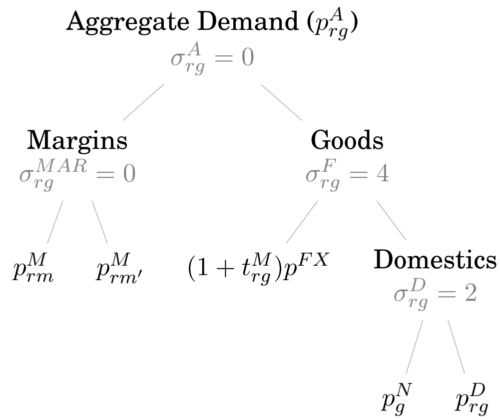
$$\min_{m,nd,dd,md} \quad (1 + t^M_{r,g})p^{FX} m_{r,g} + p^N_g nd_{r,g} + p^D_{r,g} dd_{r,g} + p^M_{r,m} md_{r,m,g}$$
$$\text{s.t.} \quad F^A_{r,g}(m_{r,g}, nd_{r,g}, dd_{r,g}, md_{r,m,g}) = A_{r,g}$$

The import aggregation function given by F^A_g is described by a nested CES function shown in Figure 6. The total pool of goods is determined through foreign imports, national demand, and state level demand (and margin demands). σ^F is the substitution elasticity governing the trade off between foreign and domestic demand, while σ^D denotes the substitution elasticity for intra-national goods demand (between a given state and other states). σ^{MAR} is the substitution elasticity between margin types (trade and transport) and σ^A is the top level elasticity, both of which are set to zero to reflect an assumption on equal proportions.

Margin demands are formed through inputs supplied by the national and regional markets, which is given by a cost minimization problem. We assume that margins are formulated

[8]Notably, the description of the canonical model assumes a pooled national market where explicit bilateral trade flows between regions are suppressed. A description is provided with the blueNOTE module on how the model changes with this additional detail.

Figure 6: Nested CES Import Aggregation Function



through fixed proportions (Leotief production).

$$\begin{aligned} \min_{nm,dm} \quad & \sum_g \left(p^N_g nm_{r,g,m} + p^D_{r,g} dm_{r,g,m}\right) \\ \text{s.t.} \quad & F^M_{r,m}(nm_{r,g,m}, dm_{r,g,m}) = MS_{r,m} \end{aligned}$$

The demand side of the model is characterized by a regional household representative agent. Optimizing agents are assumed to maximize utility subject to their budget constraint. We model demand by assume Cobb-Douglas preferences amongst goods (holding fixed government and investment demands).

$$\begin{aligned} \max_{cd} \quad & U(cd_{r,g}) \\ \text{s.t.} \quad & RA_r = \sum_g p^Y_{r,g} \bar{yh}_{r,g} + p^L_r \sum_s \bar{ld}_{r,s} + \sum_s p^K_{r,s} \bar{kd}_{r,s} + \bar{bop}_r + \bar{adj}^{hh}_r \\ & \quad + \sum_g \left(t^A_{r,g} A_{r,g} + t^M_{r,g} m_{r,g}\right) - \sum_g p^A_{r,g} \left(\bar{g}_{r,g} + \bar{i}_{r,g}\right) \end{aligned}$$

3.2 Equilibrium Conditions

We can alternatively cast the general equilibrium optimization problem as an equilibrium problem in a system of inequalities known as a mixed complementarity problem (MCP). MCP representations of general equilibrium models have been shown to be robust and efficient given the framework's ability to avoid specifying intermediate and definitional variables (Mathiesen, 1985; Rutherford, 1995). A complementarity problem can be defined as given a function $F\colon \mathbb{R}^n \longrightarrow \mathbb{R}^n$, find $z \in \mathbb{R}^n$ such that $F(z) \geq 0$, $z \geq 0$, and $z^T F(z) = 0$. This can equivalently be written using the $\perp$ symbol: $F(z) \geq 0 \perp z \geq 0$.

A model equilibrium requires three sets of conditions previously outlined: zero excess profits, markets must clear, and incomes must balance. The zero profit condition maintains that markets act competitively with free entry and exit. This assumption therefore requires that the price of output (or unit revenue) be less than or equal to the unit cost of inputs, otherwise production is zero. Let $\Pi^Y(p)$ denote a unit profit function for activity Y. The zero profit condition can be written as: $-\Pi^Y(p) \geq 0 \perp Y \geq 0$. In words, if the cost of inputs is greater than the revenue received from outputs, production is zero. Market clearing conditions allow for fluctuating prices to equate supply with demand. Let S denote supply, D denote demand for a good with price p. The associated complementarity condition is therefore: $S(p) - D(p) \geq 0 \perp p \geq 0$, or if supply is greater than demand, arbitrage between producers and consumers force the price to zero.

All profit functions are represented as unit functions. We use Shepard's Lemma to generate netput coefficients by differentiating a given profit function with respect to a given price to characterize net supply. Moreover, reference prices are normalized to unity such that CGE parameters (normally denominated in value) represent quantities. The GAMS code associated with the following conditions can be found in appendix B. We also provide MPSGE code which allows for a concise representation of the equivalent model (Rutherford, 1999).

3.2.1 Zero Profit

Before explicitly writing out equilibrium conditions, we first define intermediate functions representing expressions for equilibrium levels of demands and supplies. The assumed production structure for $Y_{r,s}$ is a nested CES function with Cobb-Douglas technologies for value added. Let the value share of labor relative to total value added be $\alpha^L_{r,s}$ and $c\bar{v}a_{r,s}$ denote the reference cost of value added.

$$\alpha^L_{r,s} = \frac{\bar{ld}_{r,s}}{\bar{ld}_{r,s} + \bar{kd}_{r,s}} = \frac{\bar{ld}_{r,s}}{c\bar{v}a_{r,s}}$$

The unit cost of value added with Cobb-Douglas technologies in *calibrated share form* is characterized as follows.[9]

$$C^{VA}_{r,s} = p^{L\ \alpha^L_{r,s}}_{r} p^{K\ 1-\alpha^L_{r,s}}_{r,s}$$

Given this cost function, we can define the equilibrium labor and capital input demands.

$$LD_{r,s} = c\bar{v}a_{r,s}\frac{\partial C^{VA}_{r,s}}{\partial p^L_r} = \bar{l}d_{r,s}\frac{C^{VA}_{r,s}}{p^L_r}$$
$$KD_{r,s} = c\bar{v}a_{r,s}\frac{\partial C^{VA}_{r,s}}{\partial p^K_{r,s}} = \bar{k}d_{r,s}\frac{C^{VA}_{r,s}}{p^K_{r,s}}$$

The zero profit condition therefore is decomposed into revenue (the value of output) relative to input costs written linearly which embed equilibrium intermediary functions.

$$-\Pi^Y_{r,s} = \sum_g p^A_{r,g}\bar{id}_{r,g,s} + p^L_r LD_{r,s} + p^K_r KD_{r,s} - \sum_g p^Y_{r,g}\bar{ys}_{r,s,g} \geq 0 \quad \perp \quad Y_{r,s} \geq 0 \tag{46}$$

The allocation of output (both sectoral and household) is characterized through a CET function with elasticity of transformation, η_{rg}. Let the value shares of foreign exports, national supply and regional supply, respectively be written as follows.

$$\alpha^X_{r,g} = \frac{\bar{x}_{r,g} - \bar{rx}_{r,g}}{\bar{s}_{r,g}}, \quad \alpha^N_{r,g} = \frac{\bar{xn}_{r,g}}{\bar{s}_{r,s}}, \quad \alpha^D_{r,g} = \frac{\bar{xd}_{r,g}}{\bar{s}_{r,g}}$$

The unit revenue function for output allocation, $X_{r,g}$, is comprised of these three supply markets.

$$R^X_{r,g} = \left(\alpha^X_{r,g} p^{FX\ 1+\eta_{r,g}} + \alpha^N_{r,g} p^{N\ 1+\eta_{r,g}}_{r,g} + \alpha^D_{r,g} p^{D\ 1+\eta_{r,g}}_{r,g}\right)^{1/(1+\eta_{r,g})}$$

Using Shepard's Lemma, we can solve for supply functions for each respective market.

$$S^X_{r,g} = \bar{s}_{r,g}\frac{\partial R^X_{r,g}}{\partial p^{FX}} = (\bar{x}_{r,g} - \bar{rx}_{r,g})\left(\frac{p^{FX}}{R^X_{r,g}}\right)^{\eta_{r,g}}$$
$$S^N_{r,g} = \bar{s}_{r,g}\frac{\partial R^X_{r,g}}{\partial p^N_g} = \bar{xn}_{r,g}\left(\frac{p^N_g}{R^X_{r,g}}\right)^{\eta_{r,g}}$$
$$S^D_{r,g} = \bar{s}_{r,g}\frac{\partial R^X_{r,g}}{\partial p^D_{r,g}} = \bar{xd}_{r,g}\left(\frac{p^D_{r,g}}{R^X_{r,g}}\right)^{\eta_{r,g}}$$

[9] See Rutherford (2002) for more on calibrated share form equations.

The zero profit condition for, $X_{r,g}$, is comprised of revenues associated with each supply market.

$$-\Pi^X_{r,g} = p^Y_{r,g}\bar{s}_{r,g} - p^{FX}S^X_{r,g} - p^N_g S^N_{r,g} - p^D_{r,g}S^D_{r,g} \geq 0 \quad \perp \quad X_{r,g} \geq 0 \tag{47}$$

The total pool of goods forming final demand and intermediate input demand is formed through demands from the foreign (imports), national and regional markets. The assumed nested CES function in figure 6 asserts that margins are applied evenly across demand categories. The value shares for goods nest of the import aggregation function are defined as follows.

$$\theta^N_{r,g} = \frac{\bar{nd}_{r,g}}{\bar{nd}_{r,g} + \bar{dd}_{r,g}}, \quad \theta^M_{r,g} = \frac{(1+\bar{tm}_{r,g})\bar{m}_{r,g}}{\bar{nd}_{r,g} + \bar{dd}_{r,g} + (1+\bar{tm}_{r,g})\bar{m}_{r,g}}$$

The unit cost of the lower nest between regional and national demand for good g is a constant elasticity of substitution function.

$$C^{DN}_{r,g} = \left(\theta^N_{r,g} {p^N_g}^{1-\sigma^D_{r,g}} + (1-\theta^N_{r,g}) {p^d_{r,g}}^{1-\sigma^D_{r,g}}\right)^{1/(1-\sigma^D_{r,g})}$$

The trade off between foreign imports and the domestic good composite is another CES function. Here, we make the distinction between $\bar{tm}_{r,g}$ and $t^M_{r,g}$ which are equivalent in the benchmark, though provides notation for possible tax policy scenarios.

$$C^{DM}_{r,g} = \left(\theta^M_{r,g}\left(\frac{p^{FX}(1+t^M_{r,g})}{(1+\bar{tm}_{r,g})}\right)^{1-\sigma^F_{r,g}} + (1-\theta^M_{r,g}) {C^{DN}_{r,g}}^{1-\sigma^F_{r,g}}\right)^{1/(1-\sigma^F_{r,g})}$$

Input demand functions can be solved for as above. Let $\bar{cdm}_{r,g} = \bar{nd}_{r,g} + \bar{dd}_{r,g} + (1+\bar{tm}_{r,g})\bar{m}_{r,g}$.

$$D^M_{r,g} = \bar{cdm}_{r,g}\frac{\partial C^{DM}_{r,g}}{\partial p^{FX}} = (1+\bar{tm}_{r,g})\bar{m}_{r,g}\left(\frac{C^{DM}_{r,g}(1+\bar{tm}_{r,g})}{p^{FX}(1+t^m_{r,g})}\right)^{\sigma^F_{r,g}}$$

$$D^N_{r,g} = \bar{cdm}_{r,g}\frac{\partial C^{DM}_{r,g}}{\partial p^N_g} = \bar{xn}_{r,g}\left(\frac{C^{DM}_{r,g}}{C^{DN}_{r,g}}\right)^{\sigma^F_{r,g}}\left(\frac{C^{DN}_{r,g}}{p^N_g}\right)^{\sigma^D_{r,g}}$$

$$D^D_{r,g} = \bar{cdm}_{r,g}\frac{\partial C^{DM}_{r,g}}{\partial p^D_{r,g}} = \bar{xd}_{r,g}\left(\frac{C^{DM}_{r,g}}{C^{DN}_{r,g}}\right)^{\sigma^F_{r,g}}\left(\frac{C^{DN}_{r,g}}{p^D_{r,g}}\right)^{\sigma^D_{r,g}}$$

The zero profit condition is then composed of the value of demands and margins relative to the output price.

$$-\Pi^A_{r,g} = p^{FX} D^M_{r,g} + p^N_{r,g} D^N_{r,g} + p^D_{r,g} D^D_{r,g} + \sum_m p^M_{r,m} \bar{md}_{r,m,g}$$
$$- p^A_{r,g}(1 - \bar{ta}_{r,g})\bar{a}_{r,g} - p^{FX} \bar{rx}_{r,g} \geq 0 \quad \perp \quad A_{r,g} \geq 0 \tag{48}$$

The zero profit condition for margin supply can be formulated explicitly because of the assumed Leontief production structure.

$$-\Pi^{MS}_{r,m} = \sum_g \left(p^N_g \bar{nm}_{r,g,m} + p^D_{r,g} \bar{dm}_{r,g,m}\right) - p^M_{r,m} \sum_g \bar{md}_{r,m,g} \geq 0 \quad \perp \quad MS_{r,m} \geq 0 \tag{49}$$

Equations (46), (47), (48) and (49) characterize all necessary zero profit conditions in the WiNDC canonical framework.

3.2.2 Market Clearance

One of the advantages of using intermediate equilibrium expressions for demand and supply in the zero profit conditions is their use in defining market clearance equations. A market clearance equation is necessary for every price in the model (with exception to the numeraire due to Walras' Law). In what follows, each condition is represented as: $S(p) - D(p) \geq 0$. Before itemizing such conditions, let us define one additional activity variable, C_r, which represents aggregate household final demand in region r. Assuming Cobb-Douglas preferences for the representative agent, final demands can be expressed through "zero profit" equations assuming a given level of final consumption. The representative agent's expenditure function is composed of regional goods (where the distinction between imports and regional production is made in $A_{r,g}$). Let the value shares of consumption be defined using reference demands, and $\bar{c}_r$ denote the value of total final demand.

$$\theta^C_{r,g} = \frac{\bar{cd}_{r,g}}{\sum_{g'} \bar{cd}_{r,g'}} = \frac{\bar{cd}_{r,g}}{\bar{c}_r}$$

The unit expenditure function is written as the composite unit price of aggregate final demand.

$$p_r^c = \prod_g {p_{r,g}^A}^{\theta_{r,g}^C}$$

Input demands are found by differentiating p_r^C with respect to a given goods price, $p_{r,g}^A$.

$$D_{r,g}^C = \bar{c}_r \frac{\partial p_r^C}{\partial p_{r,g}^A} = \bar{cd}_{r,g} \left(\frac{p_r^C}{p_{r,g}^A} \right)$$

The zero profit condition associated with the final demand aggregation is written as follows.

$$-\Pi_r^C = \sum_g p_{r,g}^A D_{r,g}^C - p_r^C \bar{c}_r \geq 0 \quad \perp \quad C_r \geq 0 \tag{50}$$

Given this, the market clearance condition for $p_{r,g}^A$ is composed of input demands by agents and sectoral production.

$$A_{r,g}\bar{a}_{r,g} - \bar{g}_{r,g} - \bar{i}_{r,g} - C_r D_{r,g}^C + \sum_s Y_{r,s}\bar{id}_{r,g,s} \geq 0 \quad \perp \quad p_{r,g}^A \geq 0 \tag{51}$$

The output market is supplied by sectoral and household production and demanded by the allocation market, $X_{r,g}$.

$$\sum_s Y_{r,s}\bar{ys}_{r,s,g} + \bar{yh}_{r,g} - X_{r,g}\bar{s}_{r,g} \geq 0 \quad \perp \quad p_{r,g}^Y \geq 0 \tag{52}$$

The regional level of supply must be greater than or equal to the amount regionally demanded plus margin formulation in region r for good g.

$$X_{r,g}S_{r,g}^D - A_{r,g}D_{r,g}^D - \sum_m MS_{r,m}\bar{dm}_{r,g,m} \geq 0 \quad \perp \quad p_{r,g}^D \geq 0 \tag{53}$$

Similarly, though summing across regions to create a *pooled* national market, the level of national supply across regions must satisfy the total domestic demand for both goods and margins.

$$\sum_r \left(X_{r,g}S_{r,g}^N - A_{r,g}D_{r,g}^N - \sum_m MS_{r,m}\bar{nm}_{r,g,m} \right) \geq 0 \quad \perp \quad p_g^N \geq 0 \tag{54}$$

We assume labor to be freely mobile across sectors, but fixed to a region, r.

$$\sum_{s} \left(\bar{ld}_{r,s} - Y_{r,s} LD_{r,s}\right) \geq 0 \quad \perp \quad p^L_r \geq 0 \tag{55}$$

Capital is assumed to be sector specific, and therefore the market clearing condition for capital must hold for each region-sector pairing.

$$\bar{kd}_{r,s} - Y_{r,s} KD_{r,s} \geq 0 \quad \perp \quad p^K_{r,s} \geq 0 \tag{56}$$

The market for margins must clear for each region and margin type. The market clearance condition requires that margins supplied through the regional and national market must greater than or equal to the total level of margin demand in the absorption goods market.

$$MS_{r,m} \sum_{g} \bar{md}_{r,m,g} - \sum_{g} A_{r,g} \bar{md}_{r,m,g} \geq 0 \quad \perp \quad p^M_{r,m} \geq 0 \tag{57}$$

The intermediary zero profit condition on C_r allows us to concisely represent the market for aggregate final demand. Because p^C_r denotes the composite price of a unit of aggregate demand, the total quantity demanded is simply the composite price divided into total household income RA_r.

$$C_r \bar{c}_r - \frac{RA_r}{p^C_r} \geq 0 \quad \perp \quad p^c_r \geq 0 \tag{58}$$

The price of foreign exchange is determined through difference between total imports and exports net of re-exports and reference balance of payments.

$$\sum_{r} \bar{bop}_r + \sum_{r,g} \left(X_{r,g} S^X_{r,g} + A_{r,g} \left(\bar{rx}_{r,g} - D^M_{r,g}\right)\right) \geq 0 \quad \perp \quad p^{FX} \geq 0 \tag{59}$$

Equations (51) - (59) describe the full set of market clearing conditions needed for specifying the WiNDC model.

3.2.3 Income Balance

The final equilibrium condition requires that expenditures do not exceed income levels. This is characterized, in conjunction with the determination of price p_r^C, by defining the income level for the representative agent. Here, we hold fixed government and investment demands, and allow tax revenue to accrue to the representative household.

$$RA_r = \sum_g p^Y_{r,g} \bar{yh}_{r,g} + p^{FX}\left(\bar{bop}_r + \bar{adj}_r^{hh}\right) + p^L_r \sum_s \bar{ld}_{r,s} + \sum_s p^K_{r,s} \bar{kd}_{r,s}$$
$$+ \sum_g A_{r,g}\left(p^{FX} t^M_{r,g} D^M_{r,g} + p^A_{r,g} t^A_{r,g} \bar{a}_{r,g}\right) - \sum_g p^A_{r,g}\left(\bar{g}_{r,g} + \bar{i}_{r,g}\right) \quad (60)$$

With fixed government and investment demanding agents, equation (60) represents the only income balance constraint needed for specifying an equilibrium.

4 The BlueNOTE Module

Depending on the intention of the analysis, it may be appropriate to incorporate additional data if aspects of the regional economic structure are known. In this paper, we consider the economic implications of state level climate policy to illustrate the blueNOTE module and modeling framework. The build routine generates the module by recalibrating the core database using known aspects of the energy system from the State Energy Data System (SEDS). This is done in three steps. First, `readseds.gms` converts raw SEDS data into a GAMS readable format. The dataset is then aggregated to match the sectoral and region schemes in the climate policy analysis. Finally, `bluenote.gms` reads in the generated SEDS data, separates the natural gas and crude oil extraction sector[10], enforces state level energy supplies, demands and trade and adjusts the remaining accounts using optimization based matrix balancing techniques to accommodate the new economic data. The blueNOTE database is generated to the database directory pending a micro-consistency check for a single year (as chosen in `run.bat`).

[10]This is done so by exploiting the relative production value of crude oil to natural gas by year and region to generate shares.

SEDS provides price and quantity data by state for energy producing sectors. The data distinguishes between the supply and demand prices for energy and electricity. The supply price of energy is assumed to be the minimum price across regions and demanding sectors and given a lack of data for crude oil prices, its supply price is assumed to be half of the corresponding price for refined oil. The difference in the demand and supply prices for energy characterizes markups between wholesale and final demand across aggregated categories (e.g. industrial, residential). New margin totals are shared out to input output sectors using existing margins in the input output data. Similarly, because SEDS reports energy input demands (and calculated emissions totals) by aggregate sectors, we share out such demands given core input output intermediate input demands, adjusting the scale of input demands, not the regional and sectoral composition of use. Energy supplies are added explicitly given price and quantity data. The data also allow us to keep track of net interstate flows of electricity to require that national demand of electricity relative to national supply in the economic accounts follow similar magnitudes.

Notably, converting demand and supply of energy into value terms requires converting energy quantities from BTUs (British Thermal Unit) to KWHs (Kilowatt Hours). To do so, we impose heat rates which provide a conversion scaler by technology and year.[11] We also translate energy quantities to emissions through carbon dioxide content coefficients and verify their consistency with reports from the U.S. Environmental Protection Agency.

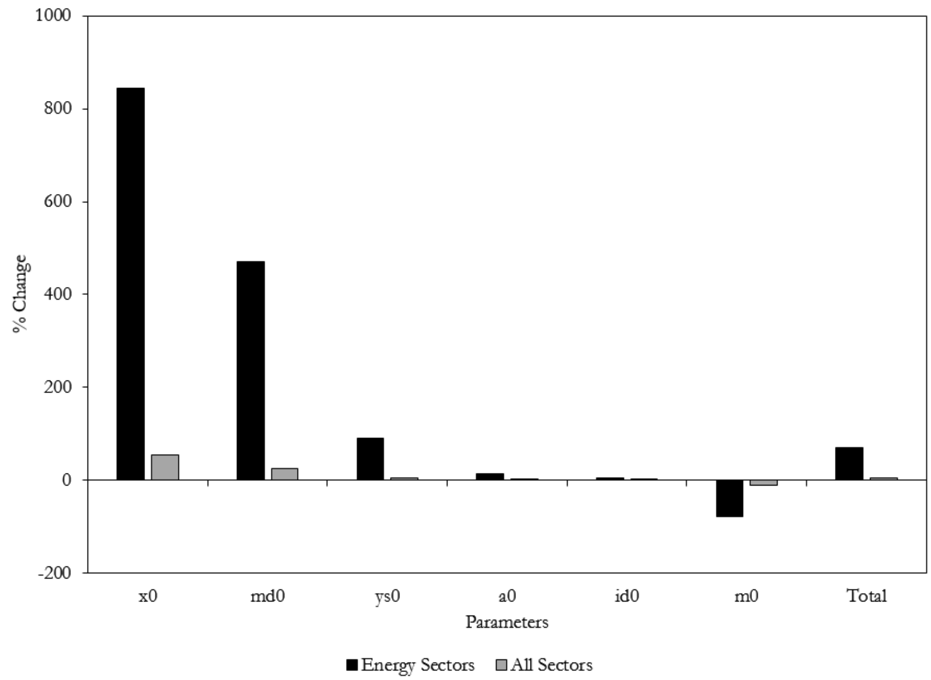
The final output of this sub-routine is a single year's worth of re-calibrated data. Figure 7 describes the change in a subset of CGE parameters following the recalibration routine. The percent change in energy sector data (those explicitly treated in `readseds.gms`) are largest for exports and margin adjustments. However, when compared with the percent change across all sectors, magnitudes are smaller.

4.1 Application: State Level Climate Policy

In the United States, disinterest at the federal level for comprehensive climate policy has prompted some states and cities to tackle climate efforts individually or through collective ef-

[11] Data is taken from the U.S. Energy Information Administration at: https://www.eia.gov/electricity/annual/html/epa_08_01.html.

Figure 7: Percent Change in Data Due to Re-calibration



forts. Even pre-Paris Agreement, instances of state action aimed at curbing carbon emissions existed. The Regional Greenhouse Gas Initiative (RGGI) encompasses a collection of northeastern states with the intent of reducing carbon emissions from the electricity sector. It is the first instance of a market based cap and trade system in the United States for reducing carbon emissions for fossil fuel generating plants. California's AB-32 is a more comprehensive program seeking to reduce economy wide emissions through a combination of policy frameworks (e.g. cap and trade, fuel standards, renewable energy standards). There are also instances of proposed legislation (not yet implemented) for a comprehensive state level carbon tax in Massachusetts, New York, Oregon, Rhode Island, Vermont, and Washington.

Aside from existing frameworks, some state governors have committed to The US Climate Alliance after the United States exited from the Paris Agreement.[12] The US Climate Alliance is composed of a collection of state governors committed to the goals of the Paris Agreement of reducing greenhouse gas emissions by 26-28% below 2005 levels by 2025. In states without state level elected officials interested in climate policy, many city mayors (402) have joined Climate Mayors whom preside over roughly 20% of the total population of the United States.[13]

[12]See: https://www.usclimatealliance.org/
[13]See: http://climatemayors.org/

New York, Hawaii, California, Illinois, Texas, Arizona, North Carolina, Colorado, District of Columbia Tennessee, Oregon, New Mexico, Alaska all have over 20% of their total state population living in cities with mayors committed to the Paris Agreement goals.[14] A report providing a state by state analysis on the potential for state level carbon tax found that 14 states have the potential for a carbon tax, 11 states have challenging legal or idealogical commitments for action, and the rest have very challenging environments for social, legal and/or economic reasons (Bauman and Komanoff, 2017).

Whether or not these coalitions or individual state level interest in climate action amount to regulation is uncertain, however, this prompts interesting questions on the effectiveness of subnational climate policy. The inability for states to impose border sanctions between one another reduces the scope for limiting carbon leakage to states that have no interest in imposing restrictions on carbon for its constituents. We define carbon leakage as the increase in emissions in states not covered by climate policy relative to emissions decreases in covered states. The level of carbon leakage will determine the effectiveness of reducing emissions country wide. Caron et al. (2015) studies the leakage implications of California's cap and trade program using IMPLAN and GTAP (Global Trade and Analysis Project) data. Bilateral trade flows between states are taken from IMPLAN's gravity model of trade, replacing state level electricity trade with data from the National Renewable Energy Laboratory's ReEDS model (Short et al., 2009). Without any measure of border adjustments, the authors calculate a 45% leakage rate with an imposed carbon price of $15 per ton of CO2.

We note that there are additional complicating factors like overlapping regulations. (e.g. Böhringer and Behrens (2015)). While California and RGGI states may have the potential to adopt some type additional climate policy, overlapping emissions regulations will impact the performance of the existing program.[15] Here, we abstract away from existing regulation and seek to understand how a binding cap on state level economy wide carbon emissions translates into country level emissions reductions. We assess this question using a variety of data and model related assumptions to understand their relative impact on model results.

[14] Own calculations using population data from the Census Bureau.

[15] For instance, Akin-Olcum et al. (2018) show that if New York imposes an additional price adder on top of the RGGI permit price for its electricity generators, the permit price is driven downward which creates a wedge in welfare between New York households and other RGGI states.

4.2 Embodied Carbon

To reduce the dimensionality of the dataset, we aggregate sectors according to embodied carbon. Embodied carbon is both the amount of carbon emitted directly through production or indirectly through the supply chain of using carbon intensive goods as inputs or in final demand. We calculate the level of embodied carbon in the output of each sector-region pairing in blueNOTE by solving a system of linear equations. Let $e^{Y}_{r,s}$ denote the embodied carbon in output of sector s in region r, $e^{P}_{r,g}$ be the embodied carbon in domestic goods, $e^{PM}_{r,g}$ the embodied carbon in imports of good g into region r, $e^{PA}_{r,g}$ denote the embodied carbon in the Armington supply of goods, e^{PN}_{g} be the embodied carbon in nationally produced goods and $e^{PMRG}_{r,g}$ be the embodied carbon in margin demand. Given our equilibrium accounting identities, we can formulate the following system of linear equations to solve for embodied carbon coefficient. Let $\bar{co}2_{r,g,s}$ be the level of CO2 emissions from sector s in region r for using input g. The total carbon content in sectoral output is characterized by both direct emissions and carbon intensive inputs.

$$e^{Y}_{r,s} \sum_{g} \bar{ys}_{r,s,g} = \sum_{g} \bar{co}2_{r,g,s} + \sum_{g} e^{PA}_{r,g} \bar{id}_{r,g,s}$$

Because production includes byproducts in the WiNDC build stream (a given sector can have multiple types of output), we translate the carbon content of sectoral output to the carbon content of a given goods state level supply.

$$e^{P}_{r,g} \sum_{s} \bar{ys}_{r,s,g} = \sum_{s} e^{Y}_{r,s} \bar{ys}_{r,s,g}$$

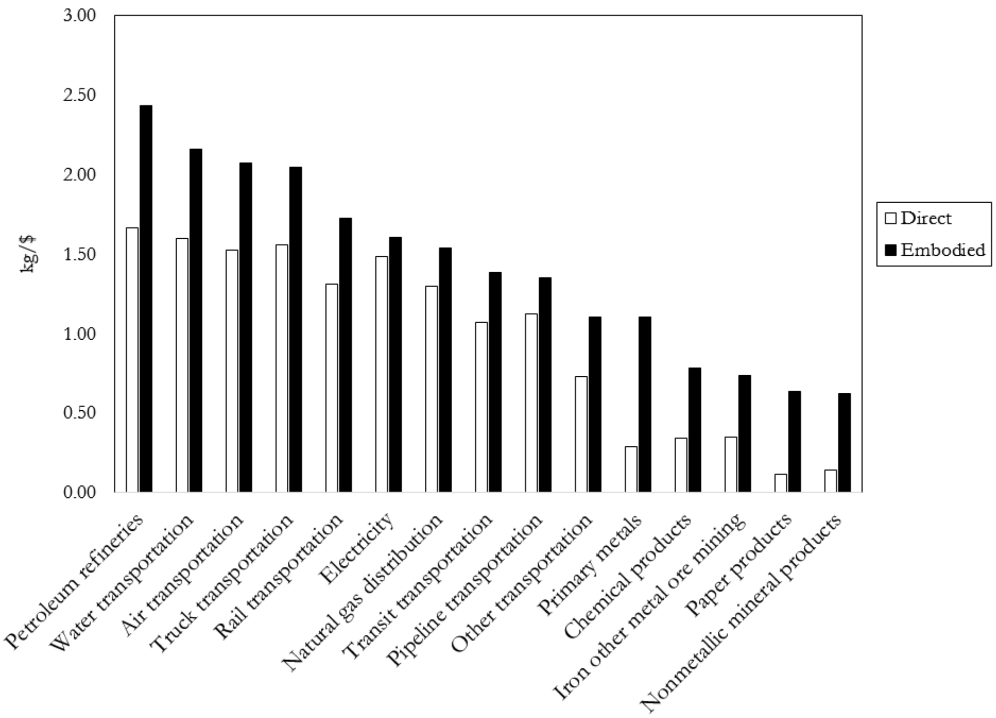
We assume that imports have twice the carbon content of locally produced goods.

$$e^{PM}_{r,g} = 2 e^{P}_{r,g}$$

In a pooled national market, the carbon content of a given goods is composed of a weighted average of all regions supplying the pooled market.

$$e^{PN}_{g} \sum_{r} \bar{xn}_{r,g} = \sum_{r} e^{P}_{r,g} \bar{xn}_{r,g}$$

Figure 8: Top 15 most carbon intensive sectors in the United States



The carbon content of margins are supplied through the national and state level markets.

$$e_{r,m}^{PMRG} \sum_g \bar{md}_{r,m,g} = \sum_g \left(e_g^{PN} \bar{nm}_{r,g,m} + e_{r,g}^{P} \bar{dm}_{r,g,m} \right)$$

Finally, the embodied carbon in the Armington supply of a good g into region r to be used as input or final demand is composed of the carbon content from regional and national demand, imports and margins.

$$e_{r,g}^{PA} \left(\bar{a}_{r,g} + \bar{rx}_{r,g} \right) = e_{r,g}^{P} \bar{dd}_{r,g} + e_g^{PN} \bar{nd}_{r,g} + e_{r,g}^{PM} \bar{m}_{r,g} + \sum_m e_{r,m}^{PMRG} \bar{md}_{r,m,g}$$

We solve this problem for embodied carbon coefficients using linear programming. Averaged across the United States, the top 15 most carbon intensive sectors are included in figure 8 reported in kilograms per dollar of CO2. The most carbon intensive sectors include oil refining, transportation, electricity, natural gas distribution, metal processing, paper products and chemical product manufacturing. These estimates are in line with those reported by Böhringer et al. (2018).

Each state has a different profile of carbon intensity by sector depending on technologies (direct emissions) and the regional configuration of where inputs are sourced (indirect emissions). For comparison, figure 9 reports the two most carbon intensive goods (excluding transport sectors) in the United States and ranks the top/bottom 10 states according to carbon intensity. The figure reports embodied carbon at the good level, or rather the composition of the Armington good. Carbon is embodied in goods coming from the state market, national market or foreign markets and through margins (primarily through transport margins).

States with the largest carbon content in refined oil import most of their oil from foreign countries whereas states with the least carbon content produce oil locally. The ranking of states in electricity generation follows the composition of generating technologies. States like California or Vermont that have larger shares of renewable technologies relative to Wyoming or West Virginia have a smaller carbon content. The sectoral aggregation used for the policy analysis reflects these calculations. The energy/emission intensive sectors are defined as those with embodied carbon greater than .45 kilograms per dollar.

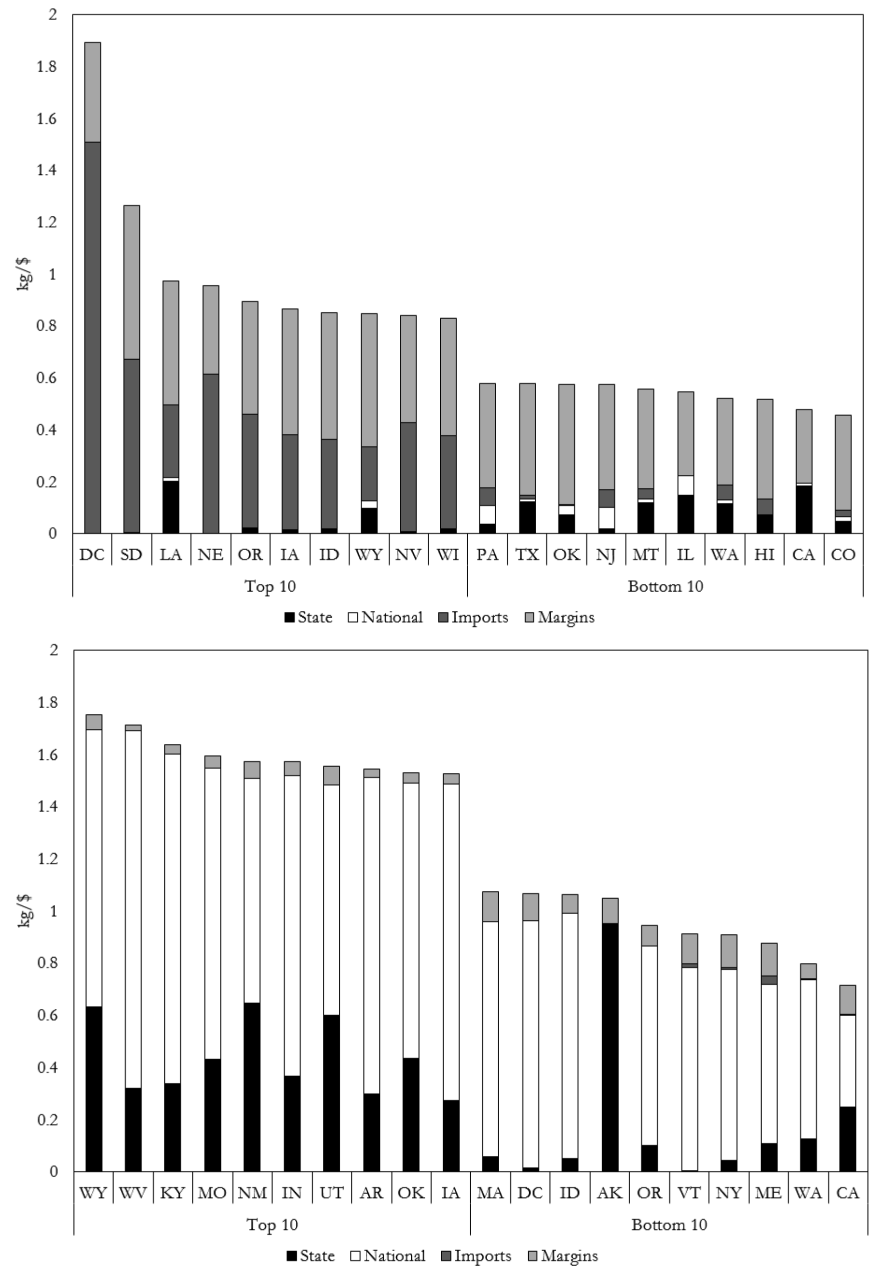
4.3 Gravity Model of Trade

In an attempt to make the leakage calculation more accurate, we estimate bilateral trade flows between all WiNDC sectors and states in the United States using a gravity model (subsequently aggregating to match blueNOTE schema). The gravity model, a framework traditionally used to predict foreign trade flows (e.g. Sapir (1981), Abrams et al. (1980)), asserts that the trade in value terms from region i to region j depends on economic forces in both origin and destination nodes, and on those which may aid or restrict the flow of goods from origin to destination (Bergstrand, 1985).[16]

The specification used here relates bilateral trade flows (Y_{ij}) between regions i and j to total regional gross product (both GDP_i and GDP_j), the distance between regions ($Dist_{ij}$), and additional impedance factors between regions i and j (generally written as A_{ij}^f for factor f). We specify the model with a multiplicative error term. Note that the following specification

[16]Notably, this is not a paper asserting the correctness of econometric approaches (e.g. Mátyás (1997), Egger (2000)) or the theoretical underpinnings of the gravity model (e.g. Balistreri and Hillberry (2006), Anderson and Van Wincoop (2003), Anderson and Yotov (2010)). Rather, we seek to provide *some* type of empirically estimated measure of trade with available data which can be reconciled within blueNOTE, due to the lack of subnational trade metrics reported in the United States.

Figure 9: Embodied carbon in refined oil (top) and electricity (bottom) at the good level



provides a cross sectional approach for estimating trade elasticities for a given sector.

$$Y_{ij} = a_0 \left(GDP_i\right)^{\beta_1} \left(GDP_j\right)^{\beta_2} \left(Dist_{ij}\right)^{\beta_3} \left(\prod_f A_{ij}^{f\ \beta_f} \right) u_{ij}$$

Log-linearizing the model yields an additive expression. We use OLS to estimate the model where $\beta_0 = \ln a_0$, $\epsilon_{ij} = \ln u_{ij}$ and X_{ij}^f denotes additional control variables. X_{ij}^f contains variables for regional contiguity and same dominant language in origin and destination regions.

$$lnY_{ij} = \beta_0 + \beta_1 \ln(GDP_i) + \beta_2 \ln(GDP_j) + \beta_3 \ln(Dist_{ij}) + \sum_f \beta_f X_{ij}^f + \epsilon_{ij} \tag{61}$$

We identify the elasticities using D-Level input output data from Statistics Canada for 2014. Canadian elasticities are then used to generate a set of predicted trade levels for the United States. The D-Level input output dataset provides bilateral trade between all Canadian provinces for 230 sectors (not including fictive industries). Lacking a panel dataset, we obtain more explanatory power by running the model for WiNDC sectors. For instance, the D-Level tables have four separate agricultural sectors. Trade flows in each of these disaggregated accounts are used as separate observations for agricultural elasticities. Distance in kilometers (as the crow flies) between provinces is taken from Anderson and Van Wincoop (2003). Table 7 provides an example of estimation results. Remaining sectoral results can be found in appendix C. In each case, the model does a good job of explaining the variation in bilateral trade in Canada. Elasticities on GDP in both origin and destination regions are statistically significant and positive across sectors. The distance elasticity is unambiguously negative, and larger in absolute value for service sectors and goods not easily traded (the most negative elasticity is for the housing sector).

We use the elasticities to generate a set of bilateral trade flows between states in the United States. In this illustrative calculation, we don't make use of uncertainty in parameter estimates. This can, however, represent an added measure of sensitivity to final results. We fit a bilateral trade matrix by using aggregate regional GDP directly from WiNDC for 2014 and distance between states as calculated with GIS shapefiles of population weighted centroids. Core accounts are reassigned given these fitted values. For instance, this includes state level demands and supplies ($\bar{d}d_{r,g}$, $\bar{x}d_{r,g}$, and $\bar{d}m_{r,g,m}$) and national market demands and supplies ($\bar{n}d_{r,g}$, $\bar{x}n_{r,g}$, $\bar{n}m_{r,g,m}$) which were initially calculated using regional purchase coefficients based on the Commodity Flow Survey. Bilateral trade flows are therefore estimated such that totals are maintained from the pooled national market framework. Let $MRT_{r,rr,g}$ characterize a variable for multi-regional trade. We estimate trade flows which satisfy the

Table 7: Gravity Estimates: All Sectors - Accomodations

	(1) *all*	(2) *adm*	(3) *agr*	(4) *air*	(5) *alt*	(6) *amb*	(7) *amd*
lnFromGDP	0.713***	1.014***	0.437***	1.132***	0.428***	0.597***	0.703***
	(0.0638)	(0.0782)	(0.0687)	(0.0617)	(0.0637)	(0.0852)	(0.0540)
lnToGDP	0.549***	0.810***	0.551***	0.739***	0.369***	0.553***	0.768***
	(0.0603)	(0.0816)	(0.0735)	(0.0813)	(0.0604)	(0.0795)	(0.0637)
lnDist	-1.339***	-1.370***	-1.322***	-1.089***	-0.759***	-1.469***	-1.220***
	(0.121)	(0.145)	(0.148)	(0.141)	(0.128)	(0.167)	(0.0990)
Contiguity	-0.302	0.284	0.188	-0.575	-0.185	-0.532	0.0805
	(0.272)	(0.393)	(0.302)	(0.349)	(0.286)	(0.412)	(0.254)
Language	0.483***	0.393*	0.798***	0.719***	0.199	0.625***	0.674***
	(0.159)	(0.224)	(0.196)	(0.214)	(0.147)	(0.225)	(0.161)
Constant	-6.618***	-8.854***	-5.622***	-11.26***	-8.374***	-5.381***	-7.816***
	(0.909)	(1.423)	(0.909)	(1.436)	(0.825)	(1.268)	(0.781)
Observations	49100	700	2000	300	600	400	200
R2	0.439	0.594	0.510	0.522	0.415	0.552	0.585

Notes: Standard errors, clustered by origin destination pairs, are in parentheses with * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Sectors are described as follows: all - all sectors, adm - administrative and support services, agr - farms, air - air transportation, alt - apparel and leather and allied products, amb - ambulatory health care services, amd - accommodation.

following constraints by penalizing deviations from fitted values.

$$\sum_{rr} MRT_{r,rr,g} = \bar{xn}_{r,g}$$

$$\sum_{rr} MRT_{rr,r,g} = \bar{nd}_{r,g} + \sum_{m} \bar{nm}_{r,g,m}$$

See figure 10 for a visual on electricity trade from Wisconsin (excluding in state trades). Trade is prominent for nearby states due to the large distance elasticity (in absolute terms) for electricity, as well as states that are bigger (e.g. California, Texas, New York).

Notably, the model design must change to accommodate bilateral trade flows as a pooled national market is no longer needed to fully represent the economic system. The total supply of goods from sectoral and household production is allocated to either the foreign market of the local market. The new level of local supply is a gross value composed of both allocated goods previously directed to the national market and those destined for the local market.[17] The demand of goods from other regions, $p^D_{r',g}$ replaces the p^N_g arm of figure 6.

[17] Letting $\hat{xd}_{rg}$ be the new level of local supply, then: $\hat{xd}_{rg} = \bar{xd}_{rg} + \bar{xn}_{rg}$.

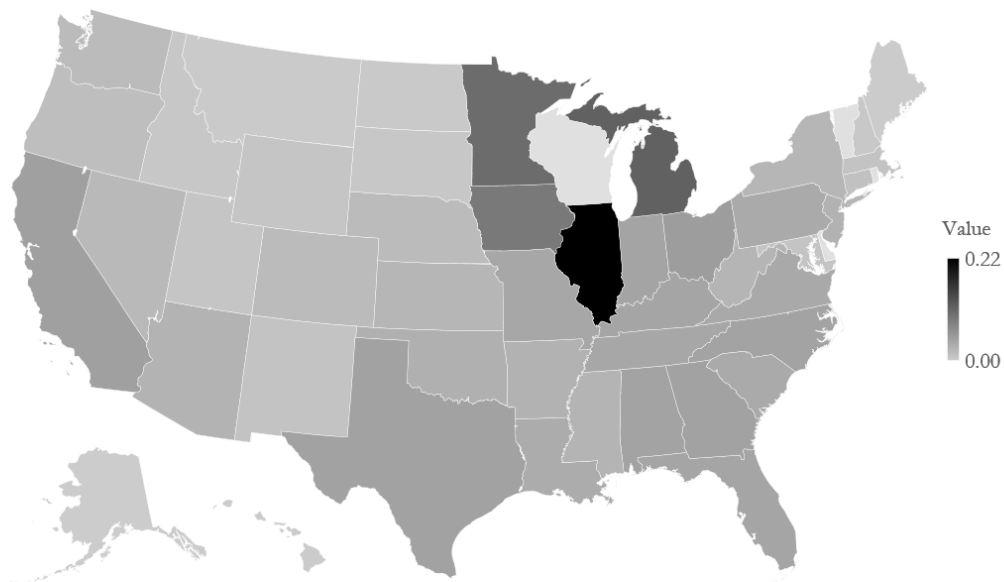


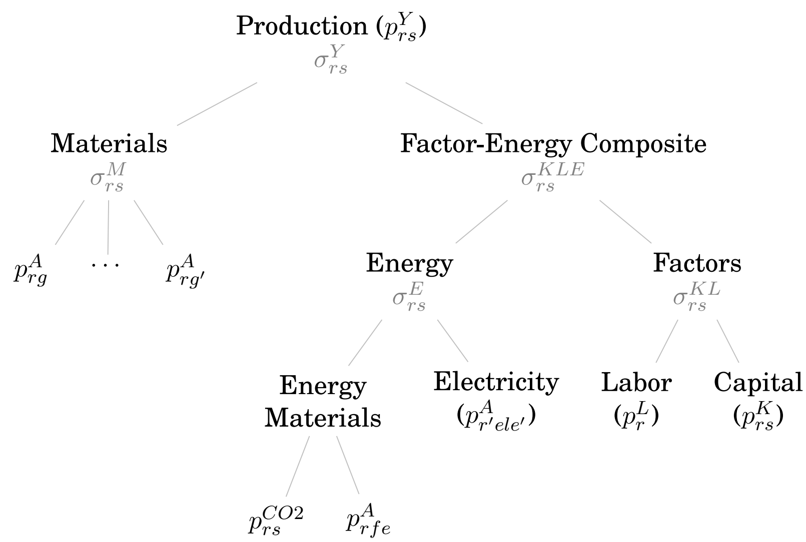
Figure 10: Illustrative trade visualization: Wisconsin electricity exports

4.4 KLEM Production Structure and Government

There are two final alterations to the canonical WiNDC structure used in the blueNOTE module. The first embeds energy based substitution in the production of goods using a "KLEM" (capital, labor, energy and materials) production structure. Figure 11 describes this production structure. Non-energy material inputs are distinguished as being non-substitutable with the energy, value added composite ($\sigma^Y = 0$). We assume that energy is substitutable with labor and capital according to the elasticity of substitution, σ^{KLE}. σ^{KL} characterizes the trade-off between labor and capital as before. Energy materials (coal, natural gas, refined oil and crude oil) are all substitutable with electricity according to σ^E and CO2 (with shadow price p_{rs}^{CO2}) must be used in fixed proportions with energy material input demands ($\sigma^{FE} = 0$).

The second change separates the representative household agent into a household and government sector (still holding investment fixed). As it stands, the core WiNDC build lacks the "fourth quadrant" of the social accounting matrix, or data on transfers between government and households. We use aggregate totals discerned through income balance constraints and impose an equal yield constraint on the government for modeling lump sum payments due to additional permit revenues.

Figure 11: KLEM based production structure



4.5 Policy Analysis

The effectiveness of subnational state level climate policies will depend on how much of the reductions in pollution in the affected state(s) are offset by increases elsewhere due to relatively cheaper environments. In order to reliably perform these calculations, we reduce the dimensionality of the model by aggregating sectors according to table 8. Most carbon emissions are associated with electricity generation which is reflected in our aggregation. Disaggregate sectors are either electricity generating or sectors with embodied carbon carbon greater than 1 kg/$. The remaining emission intensive sectors with higher levels of embodied carbon are aggregated into *eint*. Other sectors are characterized by other manufacturing, services or government and import adjustment sectors (here labeled rest of economy).[18]

We study the leakage rates due to a 20% decrease in carbon for affected states. Table 9 describes our simulation scenarios which depend on the configuration of states seeking climate action. The configuration of states are additive with exception to California or RGGI states and are included depending on the "likelihood" of some type of state level climate policy. The final simulation scenario (*Climate Mayors*) seeks to understand the impacts if city level policies were implemented (proxied by state level reductions).

[18] The blueNOTE aggregation scheme required separating coal mining, electric utilities, oil refineries, crude oil extraction and natural gas extraction from their respective aggregated WiNDC sectoral scheme.

Table 8: Sectoral Aggregation

Symbol	Description
oil	Petroleum refineries
cru	Crude oil extraction
gas	Natural gas extraction
col	Coal mining
ele	Electric power generation, transmission, and distribution
trn	Transportation
pmt	Primary metals
con	Construction
eint	Energy/Emission intensive sectors (embodied carbon > .45 kg per $)
omnf	Other manufacturing sectors
osrv	Other services
roe	Rest of the economy

Table 9: Policy Configurations

Name	Description	Included States
CA	California	CA
RGGI	Regional Greenhouse Gas Initiative States	CT, DE, MA, MD, ME, NH, NY, RI, VT
CA-RGGI	California and RGGI States	CA, CT, DE, MA, MD, ME, NH, NY, RI, VT
History	States with a history of attempted climate action	CA, CT, DE, MA, MD, ME, NH, NY, RI, VT, WA, OR
State Alliance	States with attempted past action and those in the State Alliance	CA, CT, DE, MA, MD, ME, NH, NY, RI, VT, WA, OR, CO, HI, MN, NJ, NC, VA
Carbon Center	States with attempted past action those in the State Alliance, or those deemed with some potential or challenging per the Carbon Tax Center's report	CA, CT, DE, MA, MD, ME, NH, NY, RI, VT, WA, OR, CO, HI, MN, NJ, NC, VA DC, FL, NM, MI, NV, WI, AR, SC
Climate Mayor	States with attempted past action, in the State Alliance, in the Carbon Tax Center report, or have at least 20% of their population in cities with mayors joining Climate Mayors	CA, CT, DE, MA, MD, ME, NH, NY, RI, VT, WA, OR, CO, HI, MN, NJ, NC, VA, DC, FL, NM, MI, NV, WI, AR, SC TX, AZ, TN, AK

There are three main sensitivity parameters in our model: trade flows (bilateral or pooled national market), carbon permit trade, and production structure. Trade flows are either characterized by explicit bilateral trade flows of each sector and between each region per the gravity estimation routine or in a pooled national market described in the canonical framework. Explicit representation of bilateral trade allows us to characterize inter-regional impacts more precisely as opposed to averaged impacts in a pooled market. In reducing carbon emissions in affected states, we present results assuming that states are implementing policies individually (roughly corresponding to the Paris Agreement goals) or forming carbon markets across affected states through tradable permits. Notably, in both cases, emissions are tradable across sectors. Permit trade will equalize permit prices across affected regions or sectors. Finally, we consider result sensitivity to the assumed production structure, with the canonical framework (factor substitution) relative to a more realistic energy based production structure (factor and

Table 10: Carbon Emissions and Leakage Rates: Gravity Trade and KLEM Production Structure (%)

		CA		**RGGI**		**CA-RGGI**		**History**		**State Alliance**		**Carbon Center**		**Climate Mayors**	
		No Trade	*Trade*	*No Trade*	*Trade*	*No Trade*	*Trade*	*No Trade*	*Trade*	*No Trade*	*Trade*	*No Trade*	*Trade*	*No Trade*	*Trade*
Emissions	*Alabama*	0.0	0.0	0.1	0.1	0.1	0.1	0.2	0.2	0.4	0.5	0.9	1.0	1.2	1.2
(% Change)	*Alaska*	1.3	1.3	1.4	1.4	1.6	1.6	1.3	1.3	1.5	1.6	1.7	1.8	-20.0	-15.6
	Arizona	0.0	0.0	0.1	0.1	0.1	0.2	0.2	0.2	0.4	0.5	-20.0	-27.5	-20.0	-25.5
	Arkansas	-0.1	-0.1	0.0	0.0	0.2	0.2	0.3	0.3	0.7	0.6	2.1	1.9	-20.0	-25.2
	California	-20.0	-20.0	0.5	0.5	-20.0	-19.6	-20.0	-19.7	-20.0	-18.0	-20.0	-16.7	-20.0	-15.0
	Colorado	0.4	0.4	0.3	0.3	0.5	0.5	0.5	0.5	-20.0	-28.1	-20.0	-26.5	-20.0	-24.5
	Connecticut	-2.2	-2.2	-20.0	-19.8	-20.0	-20.2	-20.0	-20.6	-20.0	-18.2	-20.0	-16.7	-20.0	-15.0
	D.C.	-1.6	-1.6	-0.4	-0.6	-0.3	-0.5	-0.3	-0.5	0.1	-0.2	-20.0	-15.1	-20.0	-13.6
	Delaware	-3.0	-3.0	-20.0	-25.7	-20.0	-26.1	-20.0	-26.5	-20.0	-24.0	-20.0	-22.3	-20.0	-20.5
	Florida	0.1	0.1	0.2	0.2	0.3	0.3	0.4	0.4	0.7	0.7	-20.0	-21.2	-20.0	-19.2
	Georgia	-2.4	-2.4	-1.8	-1.8	-1.8	-1.8	-1.7	-1.7	-1.2	-1.2	2.6	2.7	6.7	5.4
	Hawaii	0.1	0.1	0.4	0.4	0.4	0.4	0.5	0.5	-20.0	-15.3	-20.0	-14.1	-20.0	-12.7
	Iowa	-3.5	-3.5	-2.4	-2.5	-2.4	-2.5	-2.2	-2.3	-0.4	-0.4	3.2	3.1	3.4	2.5
	Idaho	0.1	0.1	3.1	2.8	3.5	3.3	4.5	4.2	5.1	4.4	7.2	6.6	7.6	6.4
	Illinois	-2.6	-2.6	-2.6	-2.6	-2.5	-2.5	-2.8	-2.8	-2.0	-2.0	-1.1	-1.2	-2.0	-2.1
	Indiana	0.2	0.2	0.1	0.1	0.1	0.1	0.1	0.1	0.3	0.3	0.6	0.6	0.7	0.7
	Kansas	2.3	2.3	1.2	1.2	2.0	1.9	0.8	0.7	0.6	0.4	-0.5	-1.2	-2.1	-2.0
	Kentucky	0.0	0.0	-0.1	-0.1	0.0	0.0	-0.1	-0.1	0.0	0.0	0.0	0.0	-0.2	-0.3
	Louisiana	0.4	0.4	0.0	0.0	0.1	0.1	-0.4	-0.4	-0.7	-0.6	-1.4	-1.6	-0.6	-1.0
	Massachusetts	-0.7	-0.7	-20.0	-17.9	-20.0	-18.3	-20.0	-18.7	-20.0	-16.5	-20.0	-15.0	-20.0	-13.2
	Maryland	-0.9	-0.9	-20.0	-26.2	-20.0	-26.7	-20.0	-27.1	-20.0	-24.6	-20.0	-22.4	-20.0	-20.4
	Maine	-0.8	-0.8	-20.0	-23.0	-20.0	-23.4	-20.0	-23.8	-20.0	-21.7	-20.0	-20.2	-20.0	-18.2
	Michigan	-5.8	-5.8	-5.6	-5.7	-5.5	-5.5	-5.4	-5.5	-4.2	-4.4	-20.0	-22.1	-20.0	-19.4
	Minnesota	0.5	0.5	0.2	0.2	0.7	0.7	0.8	0.8	-20.0	-24.9	-20.0	-23.2	-20.0	-21.3
	Missouri	-0.1	-0.1	0.0	0.0	0.0	0.1	0.1	0.1	0.3	0.4	0.9	1.1	1.2	1.3
	Mississippi	1.0	1.0	0.9	0.8	1.0	1.0	-0.4	-0.5	-0.7	-0.9	-1.7	-2.2	-3.0	-3.0
	Montana	0.2	0.2	0.2	0.2	0.0	0.0	-0.9	-0.9	-1.0	-1.0	-1.4	-1.4	-2.8	-2.8
	North Carolina	0.0	0.0	0.0	0.0	0.1	0.1	0.1	0.1	-20.0	-24.7	-20.0	-22.7	-20.0	-20.8
	North Dakota	0.1	0.1	0.2	0.2	0.3	0.3	0.4	0.4	0.5	0.5	0.8	0.8	1.0	0.9
	Nebraska	0.0	0.0	0.3	0.3	0.4	0.3	0.4	0.4	0.8	0.8	1.6	1.7	1.9	1.8
	New Hampshire	-0.5	-0.5	-20.0	-19.9	-20.0	-20.2	-20.0	-20.5	-20.0	-19.0	-20.0	-18.0	-20.0	-16.8
	New Jersey	-1.0	-1.0	0.4	0.4	0.5	0.5	0.6	0.5	-20.0	-17.0	-20.0	-15.3	-20.0	-13.5
	New Mexico	6.6	6.6	6.4	6.3	7.4	7.3	7.0	6.9	8.2	7.9	-20.0	-27.9	-20.0	-25.4
	Nevada	0.1	0.1	0.1	0.1	0.3	0.3	0.3	0.3	0.5	0.5	-20.0	-22.9	-20.0	-21.0
	New York	-4.6	-4.6	-20.0	-17.9	-20.0	-18.3	-20.0	-18.6	-20.0	-16.0	-20.0	-13.7	-20.0	-11.8
	Ohio	-0.2	-0.2	-0.2	-0.2	-0.2	-0.2	-0.2	-0.2	0.2	0.2	0.7	0.7	0.9	0.9
	Oklahoma	8.4	8.4	7.8	7.7	8.5	8.4	8.1	8.0	9.8	9.6	8.7	8.2	10.7	9.0
	Oregon	0.1	0.1	0.2	0.2	0.4	0.4	-20.0	-19.6	-20.0	-18.0	-20.0	-16.7	-20.0	-15.1
	Pennsylvania	-1.9	-1.9	-2.0	-2.1	-2.0	-2.0	-2.0	-2.0	-1.4	-1.5	-1.1	-1.2	-1.0	-1.1
	Rhode Island	-0.2	-0.2	-20.0	-18.9	-20.0	-19.4	-20.0	-19.8	-20.0	-17.4	-20.0	-15.8	-20.0	-13.7
	South Carolina	-0.2	-0.2	0.0	0.0	0.1	0.1	0.1	0.1	0.9	0.9	-20.0	-24.5	-20.0	-22.5
	South Dakota	0.1	0.1	0.2	0.2	0.3	0.3	0.4	0.4	0.6	0.7	1.1	1.1	1.4	1.4
	Tennessee	-0.1	-0.1	-0.6	-0.6	-0.8	-0.8	-1.2	-1.2	-1.0	-1.0	-0.5	-0.5	-20.0	-22.5
	Texas	0.7	0.7	0.2	0.2	0.4	0.4	0.1	0.1	0.3	0.2	0.4	0.3	-20.0	-24.5
	Utah	1.5	1.5	1.7	1.6	1.6	1.6	0.4	0.4	0.4	0.2	0.0	-0.3	-1.1	-1.2
	Virginia	-0.3	-0.3	0.1	0.1	0.1	0.1	0.2	0.1	-20.0	-23.3	-20.0	-21.3	-20.0	-19.3
	Vermont	0.1	0.1	-20.0	-19.0	-20.0	-19.4	-20.0	-19.7	-20.0	-17.5	-20.0	-16.0	-20.0	-14.0
	Washington	0.4	0.4	0.1	0.1	0.9	0.9	-20.0	-17.5	-20.0	-15.8	-20.0	-14.2	-20.0	-12.6
	Wisconsin	-0.1	-0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.3	0.4	-20.0	-24.2	-20.0	-22.3
	West Virginia	-0.2	-0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.4	0.4	1.0	1.1	1.2	1.2
	Wyoming	3.3	3.3	3.1	3.1	3.3	3.3	2.5	2.5	1.5	0.4	1.2	1.5	1.1	1.1
	Total	-1.1	-1.1	-1.1	-1.1	-2.3	-2.3	-2.7	-2.7	-4.5	-4.5	-7.1	-7.1	-10.8	-10.8
Leakage (%)		12.4	12.4	22.1	21.3	15.9	15.6	13.5	13.2	10.8	10.2	7.4	7.1	4.4	4.2

energy substitution). The main results here assume the KLEM cost function.

Table 10 characterizes impacts to emissions levels and leakage rates for a model with gravity based bilateral trade. Policy scenarios differ by permit trade across affected states (*No Trade* or *Trade*). Emissions impacts are presented in percentage change from reference emissions levels. In both the California and RGGI scenarios, a 20% reduction in emissions leads to a roughly 1% reduction in country wide emissions. As more states are included in emissions reductions policies, the total level of US emissions increases in absolute value leading to roughly an 11% reduction in the case of *Climate Mayors*. Emissions increases or decreases in states due to reductions elsewhere in the country tend to increase with more

Table 11: Carbon Emissions and Leakage Rates: Pooled National Market and KLEM Production Structure (%)

		CA		RGGI		CA-RGGI		History		State Alliance		Carbon Center		Climate Mayors	
		No Trade	*Trade*	*No Trade*	*Trade*	*No Trade*	*Trade*	*No Trade*	*Trade*	*No Trade*	*Trade*	*No Trade*	*Trade*	*No Trade*	*Trade*
Emissions	*Alabama*	0.3	0.3	0.8	0.8	1.2	1.2	1.4	1.4	2.3	2.3	3.6	3.5	4.6	4.4
(% Change)	*Alaska*	0.1	0.1	0.2	0.2	0.3	0.3	0.3	0.3	0.5	0.5	0.8	0.7	-20.0	-14.3
	Arizona	*	*	0.8	0.7	*	*	*	*	*	*	*	*	-20.0	-24.1
	Arkansas	0.2	0.2	0.7	0.6	0.9	0.9	1.1	1.1	1.8	1.8	-20.0	-27.7	-20.0	-25.8
	California	-20.0	-20.0	0.3	0.3	-20.0	-17.3	-20.0	-17.6	-20.0	-16.3	-20.0	-15.2	-20.0	-14.3
	Colorado	0.1	0.1	0.3	0.3	0.4	0.5	0.5	0.5	-20.0	-25.1	-20.0	-23.7	-20.0	-22.6
	Connecticut	0.1	0.1	-20.0	-20.6	-20.0	-24.0	-20.0	-24.3	-20.0	-21.0	-20.0	-18.2	-20.0	-16.1
	D.C.	0.2	0.2	0.8	0.7	1.0	1.1	1.2	1.2	1.9	1.9	-20.0	-13.1	-20.0	-12.3
	Delaware	0.0	0.0	-20.0	-40.4	-20.0	-43.1	-20.0	-43.4	-20.0	-40.6	-20.0	-37.0	-20.0	-32.5
	Florida	0.2	0.2	0.7	0.7	0.9	1.0	1.1	1.1	1.8	1.8	-20.0	-20.7	-20.0	-19.0
	Georgia	0.2	0.2	5.4	4.7	6.1	6.2	7.7	7.7	12.8	12.0	22.4	21.5	30.9	27.4
	Hawaii	0.2	0.2	0.7	0.7	1.0	1.0	1.2	1.2	-20.0	-14.6	-20.0	-12.7	-20.0	-11.3
	Iowa	0.3	0.3	5.6	4.9	6.3	6.4	7.8	7.9	12.9	12.3	22.0	21.3	30.3	27.4
	Idaho	0.8	0.8	3.5	2.7	4.2	3.9	4.9	4.5	5.8	5.1	7.8	7.4	9.1	8.2
	Illinois	1.5	1.5	0.6	0.6	1.9	1.9	1.6	1.6	2.1	2.0	2.9	2.8	2.1	2.0
	Indiana	0.2	0.2	0.4	0.4	0.6	0.6	0.7	0.7	1.2	1.2	1.9	1.9	2.5	2.4
	Kansas	0.2	0.2	0.1	0.1	0.3	0.3	0.3	0.2	0.4	0.4	0.7	0.7	1.0	1.0
	Kentucky	0.1	0.1	0.3	0.3	0.4	0.4	0.4	0.4	0.7	0.7	1.4	1.4	1.9	1.8
	Louisiana	0.0	0.0	-0.4	-0.4	-0.4	-0.4	-0.6	-0.6	-1.1	-1.1	-1.7	-1.7	-2.1	-2.0
	Massachusetts	0.3	0.3	-20.0	-20.1	-20.0	-23.3	-20.0	-23.6	-20.0	-20.7	-20.0	-17.7	-20.0	-15.5
	Maryland	0.5	0.5	-20.0	-25.3	-20.0	-27.2	-20.0	-27.4	-20.0	-26.4	-20.0	-25.4	-20.0	-24.5
	Maine	0.4	0.4	-20.0	-33.2	-20.0	-35.6	-20.0	-35.9	-20.0	-34.1	-20.0	-32.1	-20.0	-30.3
	Michigan	0.7	0.7	2.8	2.3	3.8	3.5	4.6	4.2	6.3	5.4	-20.0	-22.1	-20.0	-19.8
	Minnesota	0.3	0.3	0.6	0.6	1.0	1.0	1.1	1.1	-20.0	-23.0	-20.0	-21.3	-20.0	-19.9
	Missouri	0.2	0.2	0.6	0.6	0.8	0.8	0.9	1.0	1.6	1.6	2.6	2.6	3.3	3.2
	Mississippi	0.8	0.8	1.9	1.8	2.9	2.9	3.2	3.2	4.9	4.7	7.1	6.8	8.7	8.1
	Montana	0.1	0.1	0.1	0.1	0.2	0.2	0.2	0.2	0.4	0.3	0.5	0.5	0.7	0.6
	North Carolina	0.2	0.2	0.6	0.6	0.8	0.9	1.0	1.0	-20.0	-25.1	-20.0	-23.0	-20.0	-21.3
	North Dakota	0.1	0.1	0.2	0.2	0.3	0.3	0.4	0.4	0.6	0.6	1.0	0.9	1.2	1.2
	Nebraska	*	*	*	*	*	*	*	*	*	*	*	*	*	*
	New Hampshire	1.1	1.1	-20.0	-18.0	-20.0	-19.5	-20.0	-19.7	-20.0	-18.8	-20.0	-18.1	-20.0	-17.5
	New Jersey	0.4	0.4	0.9	0.8	1.4	1.4	1.6	1.6	-20.0	-17.5	-20.0	-15.0	-20.0	-13.1
	New Mexico	0.3	0.3	0.6	0.6	0.9	0.9	1.0	1.0	1.5	1.5	-20.0	-27.0	-20.0	-25.4
	Nevada	0.2	0.2	0.6	0.6	0.8	0.9	1.0	1.0	1.6	1.6	-20.0	-21.6	-20.0	-19.9
	New York	0.4	0.4	-20.0	-14.9	-20.0	-17.3	-20.0	-17.4	-20.0	-15.2	-20.0	-12.7	-20.0	-11.0
	Ohio	0.3	0.3	0.7	0.6	1.0	1.0	1.2	1.2	1.9	1.9	3.0	2.9	3.7	3.6
	Oklahoma	0.1	0.1	0.1	0.1	0.2	0.2	0.2	0.1	0.2	0.1	0.4	0.4	0.5	0.5
	Oregon	0.2	0.2	0.9	0.8	1.2	1.2	-20.0	-21.1	-20.0	-19.0	-20.0	-17.0	-20.0	-15.4
	Pennsylvania	0.0	0.0	-0.1	-0.1	0.0	-0.1	-0.1	-0.1	0.1	0.1	0.7	0.6	1.2	1.1
	Rhode Island	0.4	0.4	-20.0	-22.9	-20.0	-24.3	-20.0	-24.5	-20.0	-23.8	-20.0	-23.1	-20.0	-21.5
	South Carolina	0.4	0.4	1.8	1.7	2.4	2.5	2.8	2.9	4.7	4.6	-20.0	-31.3	-20.0	-29.5
	South Dakota	0.2	0.2	0.7	0.7	1.0	1.0	1.2	1.2	2.0	1.9	3.0	3.0	3.9	3.6
	Tennessee	-0.5	-0.5	-0.8	-0.8	-0.5	-0.5	-0.4	-0.4	0.2	0.2	1.0	1.0	-20.0	-21.4
	Texas	0.2	0.2	0.4	0.4	0.6	0.6	0.6	0.7	1.0	1.0	1.4	1.4	-20.0	-23.9
	Utah	0.2	0.2	0.3	0.3	0.5	0.5	0.5	0.5	0.8	0.8	1.3	1.3	1.6	1.5
	Virginia	0.2	0.2	0.8	0.8	1.1	1.1	1.2	1.3	-20.0	-24.4	-20.0	-21.8	-20.0	-19.7
	Vermont	0.3	0.3	-20.0	-23.7	-20.0	-26.8	-20.0	-27.1	-20.0	-24.6	-20.0	-22.0	-20.0	-19.9
	Washington	*	*	1.0	0.9	*	*	-20.0	-17.1	-20.0	-15.4	-20.0	-13.6	-20.0	-12.3
	Wisconsin	0.1	0.1	0.5	0.5	0.7	0.7	0.9	0.9	1.4	1.4	-20.0	-23.5	-20.0	-22.0
	West Virginia	0.1	0.1	0.4	0.4	0.6	0.6	0.7	0.7	1.2	1.3	2.2	2.2	2.9	2.8
	Wyoming	0.5	0.5	1.4	1.3	2.2	2.2	2.6	2.7	4.8	4.6	8.8	8.2	13.3	11.6
	Total	-1.1	-1.1	-0.9	-0.9	-2.0	-1.9	-2.3	-2.3	-3.9	-3.9	-6.4	-6.4	-10.2	-10.2
Leakage (%)		18.2	18.2	39.0	36.7	27.5	27.6	26.6	26.7	22.1	21.5	15.8	15.5	9.8	9.4

Notes: A * indicates states which lacked numerical precision in calculating final results.

states included in the climate policies (e.g. Idaho, Georgia, Nebraska) and depend on regional trading patterns. This amounts to varying levels of leakage. The highest leakage rates are calculated under the RGGI policy of 21-22%, through including more states leads to lower levels of leakage. The leakage rates calculated here are smaller than those reported by Caron et al. (2015) and likely due to differences in the treatment of electricity trade.

We contrast the gravity based results with those from a model with a pooled national market in table 11. Country wide emissions implications are smaller relative to the KLEM production case because leakage rates are higher in the pooled market. Leakage rates range from 39% in the RGGI case to 10% in the Climate Mayors scenario. Notably, the same general trend holds across production structures that leakage rates tend to be lower when permits are tradable across regions. Importantly, however, this data and modeling related difference leads to notably different results from the same policy simulation of to 50%.

These trends are contrasted with results using a canonical production structure that does not allow for energy substitution. Therefore, any output impacts due to emissions policies cannot be mitigated through substituting to less carbon intensive energy leading to higher prices for linked sectors. Higher state wide prices results in much higher leakage rates. Gravity based rates (ranging up to 54%) are lower than pooled rates (ranging up to 75%), through leakage decreases with the inclusion of more states. We conclude this section with welfare implications. Permit revenues are assumed to accrue to the representative government agent. As a result of an equal yield constraint, any additional revenue received by the government is transfered back to the household agent via a lump sum. Unsurprisingly, the distribution of the percent change in household welfare in states that do impose climate policies ($-12\% - 1\%$) is shifted leftward from the range of welfare impacts in cases without climate policies ($-9\% - 5\%$).[19] However, there are significant inter-regional impacts. For instance, Wyoming has a wide range of potential welfare implications ($\pm 20\%$) depending on which configuration of states attempt to implement policy.

Our policy analysis on state level climate policies serves to highlight the capabilities of the WiNDC build stream. This calculation shows the importance of flexibility when preparing a data set for policy analysis and the potential importance on final results. We've treated this question in a relatively simple modeling framework. More work should be done to achieve a comprehensive analysis of state level climate action. For instance, we refrain from discussing two step revenue recycling (through the tax system) which could have large implications for potential double dividends throughout the country. Moreover, our treatment of the electricity sector deserves more attention. Restricting emissions on electricity generation would require a more detailed representation of the types of technologies used to produce electricity across each state, trade, and ramping costs and constraints for satisfying peak and non-peak electricity demands. The results also hinge on assumed trade elasticities in the Armington framework. For instance, an improvement could be made to look at the importance of this assumed

[19] These ranges exclude D.C. and Wyoming as being outliers.

trade structure relative to others (e.g. Balistreri et al. (2011)) on the leakage rate. Aside from revenue recycling, however, these issues are methodological ones. Answering these questions is well within the capacity of possible extensions to blueNOTE module.

References

Abrams, R. K. et al. (1980). International trade flows under flexible exchange rates. *Economic Review 65*(3), 3–10.

Akin-Olcum, G., C. Boehringer, T. Rutherford, and A. Schreiber (2018). Economic and Environmental Interactions between Alternative Carbon Pricing Policies in New York. *Working Paper*, 1–28.

Anderson, J. E. and E. Van Wincoop (2003). Gravity with gravitas: a solution to the border puzzle. *American economic review 93*(1), 170–192.

Anderson, J. E. and Y. V. Yotov (2010). The changing incidence of geography. *American Economic Review 100*(5), 2157–86.

Armington, P. S. (1969). A theory of demand for products distinguished by place of production. *Staff Papers 16*(1), 159–178.

Balistreri, E. J. and R. H. Hillberry (2006). Trade frictions and welfare in the gravity model: how much of the iceberg melts? *Canadian Journal of Economics/Revue canadienne d'économique 39*(1), 247–265.

Balistreri, E. J., R. H. Hillberry, and T. F. Rutherford (2011). Structural estimation and solution of international trade models with heterogeneous firms. *Journal of International Economics 83*(2), 95–108.

Bauman, Y. and C. Komanoff (2017). Opportunities for Carbon Taxes at the State Level. *A Carbon Tax Center Report*, 1–120.

Bergstrand, J. H. (1985). The gravity equation in international trade: some microeconomic foundations and empirical evidence. *The review of economics and statistics*, 474–481.

Böhringer, C. and M. Behrens (2015). Interactions of emission caps and renewable electricity support schemes. *Journal of Regulatory Economics 48*(1), 74–96.

Böhringer, C., J. C. Carbone, and T. F. Rutherford (2016). The strategic value of carbon tariffs. *American Economic Journal: Economic Policy 8*(1), 28–51.

Böhringer, C., J. C. Carbone, and T. F. Rutherford (2018). Embodied carbon tariffs. *The Scandinavian Journal of Economics 120*(1), 183–210.

Böhringer, C. and A. Löschel (2006). Promoting renewable energy in europe: A hybrid computable general equilibrium approach. *The Energy Journal*, 135–150.

Böhringer, C. and T. F. Rutherford (2008). Combining bottom-up and top-down. *Energy Economics 30*(2), 574–596.

Calzadilla, A., K. Rehdanz, and R. S. Tol (2010). The economic impact of more sustainable water use in agriculture: A computable general equilibrium analysis. *Journal of Hydrology 384*(3-4), 292–305.

Caron, J., S. Rausch, N. Winchester, et al. (2015). Leakage from sub-national climate policy: The case of california's cap-and-trade program. *Energy Journal 36*(2), 167–190.

Egger, P. (2000). A note on the proper econometric specification of the gravity equation. *Economics Letters 66*(1), 25–31.

Harrison, G. W., T. F. Rutherford, and D. G. Tarr (1997). Quantifying the uruguay round. *The Economic Journal 107*(444), 1405–1430.

Huber, P. J. (1964). Robust estimation of a location parameter. *The annals of mathematical statistics*, 73–101.

Kitwiwattanachai, A., D. Nelson, and G. Reed (2010). Quantitative impacts of alternative east asia free trade areas: a computable general equilibrium (cge) assessment. *Journal of Policy Modeling 32*(2), 286–301.

Lanz, B. and T. F. Rutherford (2016). Gtapingams: Multiregional and small open economy models. *Journal of Global Economic Analysis 1*(2), 1–77.

Mathiesen, L. (1985). Computation of Economic Equilibria by a Sequence of Linear Complementarity Problems. In *Economic Equilibrium: Model Formulation and Solution*, pp. 144–162. Springer.

Mátyás, L. (1997). Proper econometric specification of the gravity model. *The world economy 20*(3), 363–368.

Rutherford, T. (2002). Lecture notes on constant elasticity functions. pp. 1–32.

Rutherford, T. F. (1995). Extension of GAMS for Complementarity Problems Arising in Applied Economics. *Journal of Economic Dynamics and Control 19*(8), 1299–1324.

Rutherford, T. F. (1999). Applied general equilibrium modeling with mpsge as a gams subsystem: An overview of the modeling framework and syntax. *Computational economics 14*(1-2), 1–46.

Rutherford, T. F. and A. Schreiber (2016). Using IMPLAN Social Accounts for Applied General Equilibrium Modeling. *Technical Manual*, 1–31.

Sapir, A. (1981). Trade benefits under the eec generalized system of preferences. *European Economic Review 15*(3), 339–355.

Short, W., N. Blair, P. Sullivan, and T. Mai (2009). Reeds model documentation: Base case data and model description. *Golden, CO: National Renewable Energy Laboratory*.

A Sectoring Schemes

Table 12: Core Sectoring Scheme

NAICS 2 Code	blueNOTE Index	Description
Agriculture	*agr*	Farms
	fof	Forestry, fishing, and related activities
Mining	*oil*	Oil and gas extraction
	min	Mining, except oil and gas
	smn	Support activities for mining
Utilities	*uti*	Utilities
Construction	*con*	Construction
Manufacturing	*fbp*	Food and beverage and tobacco products
	tex	Textile mills and textile product mills
	alt	Apparel and leather and allied products
	wpd	Wood products
	ppd	Paper products
	pri	Printing and related support activities
	pet	Petroleum and coal products
	che	Chemical products
	pla	Plastics and rubber products
	nmp	Nonmetallic mineral products
	pmt	Primary metals
	fmt	Fabricated metal products
	mch	Machinery
	cep	Computer and electronic products
	eec	Electrical equipment, appliances, and components
	mot	Motor vehicles, bodies and trailers, and parts manufacturing
	ote	Other transportation equipment
	fpd	Furniture and related products
	mmf	Miscellaneous manufacturing
Wholesale Trade	*wht*	Wholesale trade
Retail Trade	*mvt*	Motor vehicle and parts dealers (in the supply table)
	fbt	Food and beverage stores (in the supply table)
	gmt	General merchandise stores (in the supply table)
	ott	Other retail
Transportation	*air*	Air transportation
	trn	Rail transportation
	wtt	Water transportation
	trk	Truck transportation
	grd	Transit and ground passenger transportation
	pip	Pipeline transportation
	otr	Other transportation and support activities
	wrh	Warehousing and storage
Information	*pub*	Publishing industries, except internet (includes software)
	mov	Motion picture and sound recording industries
	brd	Broadcasting and telecommunications
	dat	Data processing, internet publishing, and other information
Finance and Insurance	*bnk*	Federal Reserve banks, credit intermediation
	sec	Securities, commodity contracts, and investments
	ins	Insurance carriers and related activities
	fin	Funds, trusts, and other financial vehicles
Real Estate	*hou*	Housing
	ore	Other real estate
	rnt	Rental and leasing services and lessors of intangible assets
Professional Services	*leg*	Legal services
	com	Computer systems design and related services
	tsv	Misc. professional, scientific, and technical services
Management	*man*	Management of companies and enterprises
Administrative	*adm*	Administrative and support services
	wst	Waste management and remediation services
Education	*edu*	Educational services
Health Care	*amb*	Ambulatory health care services
	hos	Hospitals
	nrs	Nursing and residential care facilities
	soc	Social assistance
Arts and Recreation	*art*	Performing arts, spectator sports, museums
	rec	Amusements, gambling, and recreation industries
Accomodation	*amd*	Accommodation
	res	Food services and drinking places
Other Services	*osv*	Other services, except government
Public	*fdd*	Federal general government (defense)
	fnd	Federal general government (nondefense)
	fen	Federal government enterprises
	slg	State and local general government
	sle	State and local government enterprises
	use	Scrap, used and secondhand goods (only in the use table)
	oth	Rest-of-the-world adjustment (only in the use table)

B State Level Model GAMS Code: MPSGE and MCP

MPSGE and MCP Code for State Level Model (mcp.gms)

```
...

$model:mge

$sectors:
      Y(r,s)$y_(r,s)          !     Production
      X(r,g)$x_(r,g)          !     Disposition
      A(r,g)$a_(r,g)          !     Absorption
      C(r)                    !     Aggregate final demand
      MS(r,m)                 !     Margin supply

$commodities:
      PA(r,g)$pa_(r,g)  !     Regional market (input)
      PY(r,g)$py_(r,g)  !     Regional market (output)
      PD(r,g)$pd_(r,g)  !     Local market price
      PN(g)             !     National market
      PL(r)             !     Wage rate
      PK(r,s)$pk_(r,s)  !     Rental rate of capital
      PM(r,m)           !     Margin price
      PC(r)             !     Consumer price index
      PFX               !     Foreign exchange

$consumer:
      RA(r)             !     Representative agent

$prod:Y(r,s)$y_(r,s)  s:0 va:1
      o:PY(r,g)   q:ys0(r,s,g)
      i:PA(r,g)   q:id0(r,g,s)
      i:PL(r)     q:ld0(r,s)  va:
      i:PK(r,s)   q:kd0(r,s)  va:

$prod:X(r,g)$x_(r,g)  t:4
      o:PFX       q:(x0(r,g)-rx0(r,g))
      o:PN(g)     q:xn0(r,g)
      o:PD(r,g)   q:xd0(r,g)
      i:PY(r,g)   q:s0(r,g)

$prod:A(r,g)$a_(r,g)  s:0 dm:4  d(dm):2
      o:PA(r,g)   q:a0(r,g)         a:RA(r)     t:ta(r,g)   p:(1-ta0(r,g))
      o:PFX       q:rx0(r,g)
      i:PN(g)     q:nd0(r,g)  d:
      i:PD(r,g)   q:dd0(r,g)  d:
      i:PFX       q:m0(r,g)   dm:   a:RA(r)     t:tm(r,g)   p:(1+tm0(r,g))
      i:PM(r,m)   q:md0(r,m,g)

$prod:MS(r,m)
      o:PM(r,m)   q:(sum(gm, md0(r,m,gm)))
      i:PN(gm)    q:nm0(r,gm,m)
      i:PD(r,gm)  q:dm0(r,gm,m)

$prod:C(r)  s:1
      o:PC(r)     q:c0(r)
      i:PA(r,g)   q:cd0(r,g)

$demand:RA(r)
      d:PC(r)     q:c0(r)
      e:PY(r,g)   q:yh0(r,g)
      e:PFX       q:(bopdef0(r) + hhadj(r))
      e:PA(r,g)   q:(-g0(r,g) - i0(r,g))
```

```
      e:PL(r)     q:(sum(s,ld0(r,s)))
      e:PK(r,s)   q:kd0(r,s)

$offtext
$sysinclude mpsgeset mge
```

...

```
equations
      profit_Y(r,s)           Zero profit: production
      profit_X(r,g)           Zero profit: disposition
      profit_A(r,g)           Zero profit: absorption
      profit_C(r)             Zero profit: final demand
      profit_MS(r,m)          Zero profit: margin supply

      market_PA(r,g)          Market clearance: absorption
      market_PY(r,g)          Market clearance: output
      market_PD(r,g)          Market clearance: local market
      market_PN(g)            Market clearance: national market
      market_PL(r)            Market clearance: labor
      market_PK(r,s)          Market clearance: capital
      market_PM(r,m)          Market clearance: margin
      market_PC(r)            Market clearance: consumption
      market_PFX              Market clearance: foreign exchange

      income_RA(r)            Income balance: representative agent;


parameter   alpha(r,s)  Labor value share;

alpha(r,s)$ld0(r,s) = ld0(r,s)/(ld0(r,s)+kd0(r,s));

$macro CVA(r,s)         (PL(r)**alpha(r,s)*PK(r,s)**(1-alpha(r,s)))
$macro AL(r,s)          (ld0(r,s)*cva(r,s)/PL(r))
$macro AK(r,s)          (kd0(r,s)*cva(r,s)/PK(r,s))

parameter   alphax(r,g) Export value share
            alphad(r,g) Local supply share
            alphan(r,g) National supply share;

alphax(r,g)$(x0(r,g)-rx0(r,g)) = (x0(r,g)-rx0(r,g))/s0(r,g);
alphad(r,g)$xd0(r,g) = xd0(r,g)/s0(r,g);
alphan(r,g)$xn0(r,g) = xn0(r,g)/s0(r,g);

$macro RX(r,g)          ((alphax(r,g)*PFX**5+alphan(r,g)*PN(g)**5+alphad(r,g)*PD(r,g)**5)**(1/5))
$macro AX(r,g)          ((x0(r,g)-rx0(r,g))*(PFX/RX(r,g))**4)
$macro AN(r,g)          (xn0(r,g)*(PN(g)/RX(r,g))**4)
$macro AD(r,g)          (xd0(r,g)*(PD(r,g)/RX(r,g))**4)

parameter   thetan(r,g) National share of domestic absorption
            thetam(r,g) Domestic share of absorption;

thetan(r,g)$nd0(r,g) = nd0(r,g)/(nd0(r,g)+dd0(r,g));
thetam(r,g)$m0(r,g) = (1+tm0(r,g))*m0(r,g)/(nd0(r,g)+dd0(r,g)+m0(r,g)*(1+tm0(r,g)));

$macro CDN(r,g)         ((thetan(r,g)*PN(g)**(1-2)+(1-thetan(r,g))*PD(r,g)**(1-2))**(1/(1-2)))
$macro CDM(r,g)         (((1-thetam(r,g))*CDN(r,g)**(1-4)+thetam(r,g)*(PFX*(1+tm(r,g))/(1+tm0(r,g
    )))**(1-4))**(1/(1-4)))

$macro DN(r,g)          (nd0(r,g)*(CDN(r,g)/PN(g))**2*(CDM(r,g)/CDN(r,g))**4)
$macro DD(r,g)          (dd0(r,g)*(CDN(r,g)/PD(r,g))**2*(CDM(r,g)/CDN(r,g))**4)
$macro MD(r,g)          (m0(r,g)*(CDM(r,g)*(1+tm0(r,g))/(PFX*(1+tm(r,g))))**4)

$macro CD(r,g)          (cd0(r,g)*PC(r)/PA(r,g))

profit_Y(y_(r,s))..           sum(g,PA(r,g)*id0(r,g,s)) + PL(r)*AL(r,s) + PK(r,s)*AK(r,s)
                              =e= sum(g, PY(r,g)*ys0(r,s,g));
```

```
profit_X(x_(r,g))..           PY(r,g)*s0(r,g) =e= PFX*AX(r,g) + PN(g)*AN(r,g) + PD(r,g)*AD(r,g);

profit_A(a_(r,g))..           PN(g)*DN(r,g) + PD(r,g)*DD(r,g) + PFX*(1+tm(r,g))*MD(r,g) +
                              sum(m,PM(r,m)*md0(r,m,g))
                              =e= PA(r,g)*(1-ta(r,g))*a0(r,g) + PFX*rx0(r,g);

profit_C(r)..                 sum(g, PA(r,g)*CD(r,g)) =e= PC(r)*c0(r);

profit_MS(r,m)..              sum(gm, PN(gm)*nm0(r,gm,m) + PD(r,gm)*dm0(r,gm,m))
                              =e= PM(r,m)*sum(gm,md0(r,m,gm));

market_PA(pa_(r,g))..         A(r,g)*a0(r,g) =e= g0(r,g) + i0(r,g) + C(r)*CD(r,g) +
                              sum(y_(r,s), Y(r,s)*id0(r,g,s));

market_PY(py_(r,g))..         sum(y_(r,s), Y(r,s)*ys0(r,s,g)) + yh0(r,g) =e= X(r,g)*s0(r,g);

market_PD(pd_(r,g))..         X(r,g)*AD(r,g) =e= A(r,g)*DD(r,g) +
                              sum(m, MS(r,m)*dm0(r,g,m))$gm(g);

market_PN(g)..                sum(r,X(r,g)*AN(r,g)) =e= sum(r, A(r,g)*DN(r,g)) +
                              sum((r,m), MS(r,m)*nm0(r,g,m))$gm(g);

market_PL(r)..                sum(s,ld0(r,s)) =e= sum(s, Y(r,s)*AL(r,s));

market_PK(pk_(r,s))..         kd0(r,s) =e= Y(r,s)*AK(r,s);

market_PM(r,m)..              MS(r,m)*sum(gm, md0(r,m,gm)) =e= sum(g, A(r,g)*md0(r,m,g));

market_PC(r)..                C(r)*c0(r) =e= RA(r)/PC(r);

market_PFX..                  sum(r, bopdef0(r)) + sum((r,g),X(r,g)*AX(r,g)) +
                              sum(a_(r,g), A(r,g)*rx0(r,g))
                              =e= sum((r,g),A(r,g)*MD(r,g));

income_RA(r)..                RA(r) =e= sum(g,PY(r,g)*yh0(r,g)) + PFX*(bopdef0(r)+hhadj(r)) -
                                sum(g, PA(r,g)*(g0(r,g)+i0(r,g))) +
                                PL(r)*sum(s,ld0(r,s)) + sum(pk_(r,s), PK(r,s)*kd0(r,s)) +
                                sum(a_(r,g), A(r,g)*( MD(r,g)*PFX*tm(r,g) +
                                a0(r,g)*PA(r,g)*ta(r,g) ));

model mcp /
      profit_Y.Y, profit_X.X, profit_A.A, profit_C.C, profit_MS.MS,

      market_PA.PA, market_PY.PY, market_PD.PD, market_PN.PN,
      market_PL.PL, market_PK.PK, market_PM.PM, market_PC.PC,
      market_PFX.PFX,

      income_RA.RA /;

                                         ...
```

C Gravity Results

Table 13: Gravity Estimates: Arts and Recreation - Computer Systems Design

	(8) *art*	(9) *bnk*	(10) *brd*	(11) *cep*	(12) *che*	(13) *col_min*	(14) *com*
lnFromGDP	0.514***	0.862***	1.075***	0.548***	0.994***	0.568***	1.091***
	(0.0669)	(0.0811)	(0.0762)	(0.0749)	(0.0790)	(0.136)	(0.103)
lnToGDP	0.498***	0.695***	0.841***	0.497***	0.613***	0.639***	0.827***
	(0.0651)	(0.0747)	(0.0748)	(0.0764)	(0.0665)	(0.153)	(0.0953)
lnDist	-1.194***	-1.476***	-1.349***	-0.963***	-1.190***	-1.090***	-1.496***
	(0.124)	(0.147)	(0.160)	(0.154)	(0.129)	(0.318)	(0.165)
Contiguity	-0.189	-0.607	-0.724*	-0.530*	-0.211	-1.077*	-0.566
	(0.277)	(0.380)	(0.369)	(0.290)	(0.305)	(0.588)	(0.438)
Language	0.639***	0.311	0.229	0.425**	0.348*	1.224***	0.102
	(0.167)	(0.207)	(0.221)	(0.163)	(0.176)	(0.415)	(0.286)
Constant	-6.574***	-6.852***	-9.462***	-8.605***	-9.930***	-9.287***	-8.003***
	(0.802)	(1.339)	(1.686)	(0.741)	(1.036)	(1.869)	(1.603)
Observations	800	900	700	800	1400	100	300
R2	0.524	0.463	0.593	0.465	0.545	0.397	0.567

Notes: Standard errors, clustered by origin destination pairs, are in parentheses with * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Sectors are described as follows: art - arts and recreation, bnk - federal reserve banks, brd - broadcasting and telecommunications, cep - computer and electronic products, che - chemical products, col_min - coal mining, com - computer systems design and related services.

Table 14: Gravity Estimates: Construction - Food and Beverage Stores

	(15) *con*	(16) *dat*	(17) *edu*	(18) *eec*	(19) *ele_uti*	(20) *fbp*	(21) *fbt*
lnFromGDP	0.404***	0.634***	0.427***	0.711***	0.318***	0.889***	0.822***
	(0.124)	(0.0675)	(0.114)	(0.0736)	(0.107)	(0.0658)	(0.111)
lnToGDP	0.404***	0.511***	0.386***	0.464***	0.314***	0.616***	0.674***
	(0.124)	(0.0637)	(0.113)	(0.0633)	(0.105)	(0.0586)	(0.136)
lnDist	-1.885***	-1.114***	-1.724***	-0.953***	-1.233***	-1.219***	-1.916***
	(0.253)	(0.123)	(0.237)	(0.126)	(0.195)	(0.124)	(0.214)
Contiguity	-1.218**	-0.761**	-1.262**	-0.171	1.340***	-0.226	-0.276
	(0.530)	(0.303)	(0.480)	(0.305)	(0.427)	(0.318)	(0.613)
Language	1.144***	0.346*	0.935***	0.245	0.391	0.0788	0.392
	(0.345)	(0.183)	(0.296)	(0.161)	(0.261)	(0.158)	(0.353)
Constant	-0.745	-7.694***	-1.857	-9.224***	-4.130***	-8.300***	-2.101
	(1.677)	(0.921)	(1.467)	(0.888)	(1.441)	(1.180)	(1.922)
Observations	2000	400	400	1200	200	3300	100
R2	0.582	0.465	0.589	0.473	0.516	0.484	0.648

Notes: Standard errors, clustered by origin destination pairs, are in parentheses with * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Sectors are described as follows: con - construction, dat - data processing, edu - education, eec - electrical equipment and components, ele_uti - electric power generation and transmission, fbp - food and beverage and tobacco products, fbt - food and beverage stores.

Table 15: Gravity Estimates: Federal Government (Defense) - Furniture and Related Products

	(22) *fdd*	(23) *fen*	(24) *fin*	(25) *fmt*	(26) *fnd*	(27) *fof*	(28) *fpd*
lnFromGDP	0.411***	0.540***	0.760***	0.881***	0.506***	0.289***	0.872***
	(0.142)	(0.116)	(0.0991)	(0.0707)	(0.106)	(0.0696)	(0.0733)
lnToGDP	0.411***	0.446***	0.609***	0.532***	0.464***	0.412***	0.490***
	(0.142)	(0.116)	(0.0913)	(0.0600)	(0.105)	(0.0730)	(0.0646)
lnDist	-2.200***	-2.116***	-1.828***	-1.097***	-1.880***	-1.181***	-1.141***
	(0.304)	(0.234)	(0.164)	(0.116)	(0.214)	(0.152)	(0.123)
Contiguity	-1.463**	-1.010**	-0.592	0.163	-1.080**	0.0566	0.187
	(0.617)	(0.508)	(0.450)	(0.292)	(0.453)	(0.317)	(0.340)
Language	1.195***	1.253***	0.558**	0.419***	0.991***	0.492**	0.473***
	(0.387)	(0.314)	(0.252)	(0.157)	(0.283)	(0.202)	(0.173)
Constant	1.558	0.262	-3.711***	-9.407***	-1.228	-4.633***	-8.958***
	(2.224)	(1.770)	(1.395)	(0.942)	(1.461)	(1.089)	(1.025)
Observations	100	300	500	1700	1100	700	700
R2	0.619	0.584	0.541	0.545	0.526	0.386	0.618

Notes: Standard errors, clustered by origin destination pairs, are in parentheses with * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Sectors are described as follows: fdd - federal government (defense), fen - federal government enterprises, fin - funds and other financial vehicles, fmt - fabricated metal products, fnd - federal non-defense, fof - forestry and fishing, fpd - furniture and related products.

Table 16: Gravity Estimates: Natural Gas Distribution - Legal Services

	(29) *gas_uti*	(30) *grd*	(31) *hos*	(32) *hou*	(33) *ins*	(34) *led_min*	(35) *leg*
lnFromGDP	0.467***	0.718***	0.695***	0.885***	1.060***	0.231**	1.294***
	(0.133)	(0.0580)	(0.117)	(0.110)	(0.0960)	(0.0970)	(0.107)
lnToGDP	0.467***	0.585***	0.527***	0.554***	0.855***	0.313***	0.596***
	(0.133)	(0.0568)	(0.109)	(0.105)	(0.0876)	(0.0978)	(0.102)
lnDist	-1.657***	-1.173***	-2.035***	-2.378***	-1.833***	-0.617**	-1.655***
	(0.286)	(0.104)	(0.232)	(0.206)	(0.173)	(0.256)	(0.164)
Contiguity	-1.257**	-0.238	-0.665	-1.055**	-0.852*	0.196	0.322
	(0.486)	(0.269)	(0.537)	(0.476)	(0.440)	(0.413)	(0.478)
Language	0.962***	0.584***	1.043***	1.024***	-0.455*	0.0333	0.759**
	(0.345)	(0.154)	(0.311)	(0.301)	(0.245)	(0.274)	(0.314)
Constant	-3.277*	-8.103***	-1.838	0.0180	-5.892***	-8.031***	-7.526***
	(1.913)	(0.777)	(1.794)	(1.648)	(1.646)	(1.489)	(1.595)
Observations	100	600	200	300	600	300	100
R2	0.557	0.563	0.630	0.621	0.718	0.199	0.771

Notes: Standard errors, clustered by origin destination pairs, are in parentheses with * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Sectors are described as follows: gas_uti - natural gas distribution, grd - transit and ground transportation, hos - hospitals, ins - insurance carriers, led_min - lead and other mining, leg - legal services.

Table 17: Gravity Estimates: Management Services - Nonmetallic Mineral Products

	(36) *man*	(37) *mch*	(38) *mmf*	(39) *mot*	(40) *mov*	(41) *mvt*	(42) *nmp*
lnFromGDP	1.011***	0.913***	0.729***	0.610***	0.808***	0.608***	0.728***
	(0.100)	(0.0694)	(0.0660)	(0.0832)	(0.0835)	(0.138)	(0.0645)
lnToGDP	0.739***	0.644***	0.475***	0.470***	0.467***	1.001***	0.427***
	(0.105)	(0.0604)	(0.0658)	(0.0781)	(0.0717)	(0.158)	(0.0576)
lnDist	-1.710***	-1.133***	-1.085***	-1.073***	-1.156***	-1.942***	-1.137***
	(0.150)	(0.125)	(0.128)	(0.158)	(0.143)	(0.229)	(0.115)
Contiguity	-0.509	0.0827	-0.126	-0.324	-0.575	-0.444	0.153
	(0.424)	(0.304)	(0.330)	(0.317)	(0.370)	(0.719)	(0.301)
Language	0.0783	0.601***	0.282*	0.466***	0.121	0.633*	0.520***
	(0.258)	(0.168)	(0.167)	(0.173)	(0.194)	(0.366)	(0.158)
Constant	-5.530***	-10.11***	-8.226***	-7.972***	-8.175***	-3.536*	-7.783***
	(1.459)	(0.977)	(0.965)	(0.887)	(1.153)	(2.111)	(0.923)
Observations	300	1200	700	1600	700	100	800
R2	0.592	0.587	0.488	0.455	0.576	0.626	0.493

Notes: Standard errors, clustered by origin destination pairs, are in parentheses with * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Sectors are described as follows: man - management services, mch - machinery, mmf - miscellaneous manufacturing, mov - motion picture and sound industries, mvt - motor vehicle and parts dealers, nmp - nonmetallic mineral products.

Table 18: Gravity Estimates: Nursing - Other Mining

	(43) *nrs*	(44) *oil*	(45) *ore*	(46) *ore_min*	(37) *osv*	(48) *ote*	(49) *oth_min*
lnFromGDP	0.392***	0.763***	0.915***	0.208**	0.578***	0.616***	0.399***
	(0.135)	(0.0974)	(0.140)	(0.104)	(0.0758)	(0.0796)	(0.0708)
lnToGDP	0.392***	0.614***	0.751***	0.348***	0.475***	0.574***	0.451***
	(0.135)	(0.105)	(0.125)	(0.110)	(0.0813)	(0.0840)	(0.0732)
lnDist	-2.040***	-1.130***	-1.997***	-0.466**	-1.562***	-1.236***	-1.064***
	(0.276)	(0.174)	(0.246)	(0.233)	(0.160)	(0.163)	(0.145)
Contiguity	-1.347**	-0.0795	-0.460	0.115	-0.633*	-0.709*	0.210
	(0.573)	(0.389)	(0.793)	(0.410)	(0.342)	(0.388)	(0.328)
Language	1.151***	1.083***	0.514	-0.0324	0.749***	0.276	0.641***
	(0.357)	(0.296)	(0.389)	(0.172)	(0.207)	(0.207)	(0.183)
Constant	0.607	-9.494***	-4.244*	-9.304***	-3.923***	-7.113***	-6.632***
	(1.942)	(1.189)	(2.332)	(1.176)	(1.169)	(1.165)	(0.927)
Observations	100	800	100	400	1600	900	400
R2	0.621	0.300	0.660	0.193	0.490	0.403	0.392

Notes: Standard errors, clustered by origin destination pairs, are in parentheses with * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Sectors are described as follows: nrs - nursing and residential care facilities, oil - oil and natural gas extraction, ore - other real estate, ore_min - ore mining, osv - other serves (except government), ote - other transportation equipment, oth_min - other mining.

Table 19: Gravity Estimates: Other Petroleum Products - Primary Metals

	(50) *oth_pet*	(51) *otr*	(52) *ott*	(53) *pav_pet*	(54) *pip*	(55) *pla*	(56) *pmt*
lnFromGDP	1.102***	0.757***	0.815***	1.102***	0.720***	0.724***	0.732***
	(0.130)	(0.0630)	(0.0617)	(0.130)	(0.118)	(0.0674)	(0.0829)
lnToGDP	0.897***	0.579***	0.474***	0.897***	0.863***	0.493***	0.543***
	(0.136)	(0.0587)	(0.0743)	(0.136)	(0.152)	(0.0623)	(0.0802)
lnDist	-1.417***	-1.364***	-1.711***	-1.417***	-1.169***	-1.131***	-0.938***
	(0.259)	(0.106)	(0.133)	(0.259)	(0.257)	(0.124)	(0.166)
Contiguity	0.689	-0.184	-0.288	0.689	0.312	-0.243	0.195
	(0.766)	(0.300)	(0.346)	(0.766)	(0.789)	(0.297)	(0.359)
Language	0.914**	0.370**	0.631***	0.914**	0.893**	0.208	0.365**
	(0.381)	(0.175)	(0.199)	(0.381)	(0.384)	(0.152)	(0.172)
Constant	-11.20***	-6.354***	-3.344***	-11.20***	-10.83***	-7.783***	-10.25***
	(2.211)	(1.024)	(1.177)	(2.211)	(2.053)	(0.945)	(1.040)
Observations	100	600	1000	100	200	1300	1400
R2	0.662	0.498	0.629	0.662	0.505	0.473	0.435

Notes: Standard errors, clustered by origin destination pairs, are in parentheses with * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Sectors are described as follows: oth_pet - other petroleum products, otr - other transportation and support activities, ott - other retail, pav_pet - asphalt paving mixture manufacturing, pip - pipeline transportation, pla - plastics and rubber products, pmt - primary metals.

Table 20: Gravity Estimates: Paper Products - Rental and Leasing Services

	(57) *ppd*	(58) *pri*	(59) *pub*	(60) *rec*	(61) *ref_pet*	(62) *res*	(63) *rnt*
lnFromGDP	0.569***	0.848***	0.804***	0.607***	0.741***	0.749***	0.858***
	(0.0691)	(0.0839)	(0.0692)	(0.0684)	(0.122)	(0.0513)	(0.0515)
lnToGDP	0.410***	0.485***	0.524***	0.572***	0.607***	0.813***	0.572***
	(0.0671)	(0.0737)	(0.0690)	(0.0805)	(0.111)	(0.0598)	(0.0483)
lnDist	-1.175***	-1.226***	-1.284***	-1.694***	-1.701***	-1.344***	-1.247***
	(0.116)	(0.136)	(0.137)	(0.114)	(0.210)	(0.110)	(0.102)
Contiguity	-0.0719	-0.348	-0.243	-0.0169	0.130	0.0521	-0.276
	(0.327)	(0.383)	(0.378)	(0.346)	(0.649)	(0.261)	(0.271)
Language	-0.234	0.422**	0.361*	0.297	0.354	0.519***	0.304**
	(0.176)	(0.187)	(0.187)	(0.229)	(0.358)	(0.170)	(0.144)
Constant	-5.818***	-7.975***	-7.179***	-2.595**	-3.946**	-5.639***	-7.252***
	(0.981)	(1.243)	(1.170)	(1.111)	(1.890)	(0.930)	(0.957)
Observations	600	300	700	200	600	200	600
R2	0.493	0.482	0.595	0.706	0.500	0.775	0.548

Notes: Standard errors, clustered by origin destination pairs, are in parentheses with * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Sectors are described as follows: ppd - paper products, pri - printing and related support activities, pub - publishing industries, rec - recreation industries, ref_pet - oil refineries, res - food services and drinking places, rnt - rental and leasing services.

Table 21: Gravity Estimates: Securities and Investments - Stone Mining

	(64) *sec*	(65) *shn_pet*	(66) *sle*	(67) *slg*	(68) *smn*	(69) *soc*	(70) *stn_min*
lnFromGDP	0.760***	1.102***	0.425***	0.425***	0.799***	0.387***	0.580***
	(0.0991)	(0.130)	(0.137)	(0.137)	(0.113)	(0.124)	(0.130)
lnToGDP	0.609***	0.897***	0.425***	0.425***	0.596***	0.387***	0.454***
	(0.0913)	(0.136)	(0.137)	(0.137)	(0.115)	(0.124)	(0.112)
lnDist	-1.828***	-1.417***	-2.124***	-2.124***	-1.021***	-1.893***	-1.342***
	(0.164)	(0.259)	(0.282)	(0.282)	(0.245)	(0.257)	(0.223)
Contiguity	-0.592	0.689	-1.384**	-1.384**	0.948	-1.254**	-0.0650
	(0.450)	(0.766)	(0.588)	(0.588)	(0.682)	(0.525)	(0.696)
Language	0.558**	0.914**	1.227***	1.227***	1.490***	1.078***	0.0863
	(0.252)	(0.381)	(0.377)	(0.377)	(0.373)	(0.330)	(0.350)
Constant	-3.711***	-11.20***	0.762	0.762	-10.80***	-0.416	-5.199**
	(1.395)	(2.211)	(1.950)	(1.950)	(2.088)	(1.715)	(1.987)
Observations	500	100	500	500	200	300	100
R2	0.541	0.662	0.600	0.600	0.417	0.617	0.522

Notes: Standard errors, clustered by origin destination pairs, are in parentheses with * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Sectors are described as follows: sec - securities and investments, shn_pet - shingle manufacturing, sle - state and local government enterprises, slg - state and local government, smn - support activities for mining, soc - social assistance, stn_min - stone mining.

Table 22: Gravity Estimates: Textile Mills - Water Transportation

	(71) *tex*	(72) *trk*	(73) *trn*	(74) *tsv*	(75) *wht*	(76) *wpd*	(77) *wrh*	(78) *wst*	(79) *wtt*
lnFromGDP	0.513***	0.777***	1.052***	0.941***	1.163***	0.663***	0.847***	0.755***	0.305***
	(0.0727)	(0.0814)	(0.0670)	(0.0770)	(0.0631)	(0.0649)	(0.0568)	(0.151)	(0.103)
lnToGDP	0.432***	0.737***	0.553***	0.610***	0.658***	0.557***	0.752***	0.639***	0.485***
	(0.0648)	(0.0753)	(0.0811)	(0.0759)	(0.0581)	(0.0623)	(0.0603)	(0.163)	(0.0981)
lnDist	-0.992***	-1.662***	-1.145***	-1.526***	-1.348***	-1.292***	-1.367***	-1.883***	-1.410***
	(0.128)	(0.156)	(0.150)	(0.155)	(0.127)	(0.116)	(0.113)	(0.279)	(0.180)
Contiguity	-0.499	-0.183	-0.305	-0.296	0.0477	-0.0994	0.0685	-0.912	-0.590
	(0.323)	(0.459)	(0.374)	(0.394)	(0.287)	(0.280)	(0.296)	(0.809)	(0.595)
Language	0.100	0.104	0.327	0.753***	0.166	0.339**	0.358**	0.329	0.233
	(0.167)	(0.239)	(0.197)	(0.207)	(0.165)	(0.167)	(0.169)	(0.419)	(0.322)
Constant	-7.386***	-4.599***	-9.639***	-6.751***	-7.775***	-6.827***	-7.837***	-3.719	-2.602
	(0.957)	(1.377)	(1.397)	(1.311)	(1.214)	(1.013)	(1.026)	(2.501)	(1.706)
Observations	600	300	200	800	900	1700	400	100	200
R2	0.454	0.627	0.531	0.505	0.745	0.514	0.654	0.544	0.452

Notes: Standard errors, clustered by origin destination pairs, are in parentheses with * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Sectors are described as follows: tex - textile mills, trk - truck transportation, trn - rail transportation, tsv - miscellaneous technical services, wht - wholesale trade, wpd - wood products, wrh - warehousing and storage, wst - waste management and remediation services, wtt - water transportation.