Open Source Tools for Sub-National Applied General Equilibrium Modeling

Andrew Schreiber
Thomas F. Rutherford

Department of Agricultural and Applied Economics
University of Wisconsin-Madison

January 12, 2018

Table of Contents

1. blueNOTE Overview
2. Modeling Framework
3. Comparison: blueNOTE vs. IMPLAN
4. State of the Project

Motivation

Existing subnational models have largely relied on a commercial database (IMPLAN) to characterize base year state and county-level economic activity in the United States.

- $
- Lack of transparency in regionalizing data.
- No mechanisms for understanding how data related assumptions impact model results.

The open-source tools for combining data and building a benchmark equilibrium database will be useful to many research groups across the country. Provide means for making more quantitative evidence based research possible.

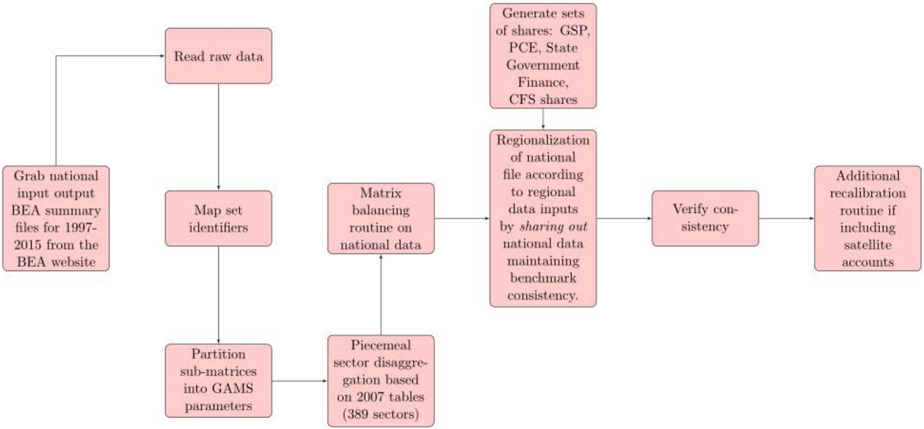
Package Overview

blueNOTE: **N**ational **O**pen source **T**ools for general **E**quilibrium modeling

- Micro-consistent sub-national social accounting matrices.
- All code for the build stream – provides logic and assumptions needed to produced dataset.
- A multi-regional, multi-sectoral computable general equilibrium model.
- Matrix balancing routines for recalibration using additional satellite data.

Build Stream

Figure 1: Build Stream Process



National Tables

National level summary files from 1997-2015:

- Supply tables – byproduct matrices with aggregate imports and trade/transport margins.
- Use tables – includes aggregate intermediate inputs, total taxes, exports, and demand accounts (aggregate household, government purchases and investment).

Use of GAMS to define submatrices and partition into CGE based parameters.

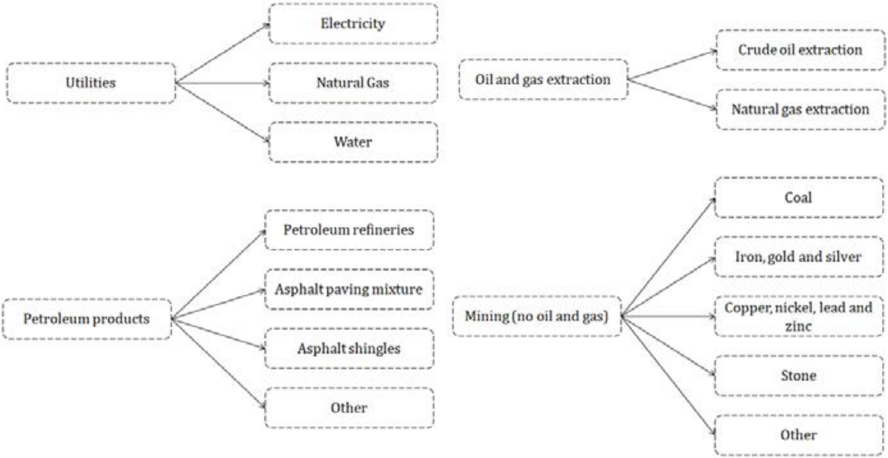
Sector Disaggregation

The routine provides options on the preferred level of sector disaggregation. Sector level detail is leveraged from the 2007 tables with 389 sectors. Level of disaggregation would depend on analysis. Options in the code include:

- *full*: full disaggregation,
- *eng*: energy related sectors,
- *agr*: agricultural sectors

For data in the 2007 tables, disaggregation shares are generated through linking disaggregate sector data with aggregate sector data through particular parameters. Data not in the disaggregate data (margins) are shared according to equal weight. Can use satellite data as well (oil and gas extraction).

Sector Disaggregation: Energy Sectors



Matrix Balancing: Huber vs. Least Squares

Optimization techniques used only once in the core build process – to balance the *national* tables to satisfy accounting identities (zero profit, market clearance, and income balance).

- Problem: *how can we minimize changes in the underlying data while enforcing accounting constraints?*

Two options are provided:

- Least squares – minimize the sum of squared percent changes in the data.
- Hybrid approach – minimize percent change in the data subject to a piecewise objective function. The loss function is quadratic in the neighborhood of the data and becomes linear further from the target.

Matrix Balancing: Huber vs. Least Squares

Least squares:

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

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

Hybrid approach based on Huber (1981):

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

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

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

where Φ_{rc} denote the subset of (r, c) with non-zero elements and Φ_{rc}^c are zero elements.

In the hybrid barrier method we retain Huber's loss function for increases from the target value and we add a log term to penalize values which go to zero:

Regionalization Process

The process to go from consistent national tables to state level tables relies on *sharing* data parameters. Shares are based on:

- gross state product (GSP)
- personal consumer expenditures (PCE)
- state government finance tables (SGF)
- commodity flow survey (CFS)

In the first three cases, data are given in aggregate categories. Categories are mapped to sectors in national data. Shares are generated such that:

$$\sum_{r} \delta_{yr,r,s} = 1 \quad \forall \quad (yr, s)$$

Regionalization Process

1. Use GSP shares to separate production data: sectoral supply with byproducts, intermediate demand and value added. Split aggregate value added based on labor and capital accounts in GSP data.
2. Use PCE shares to separate household final consumption.
3. Use SGF shares to separate government expenditures.
4. GSP shares separate investment demand and exports.
5. For a given year then, total domestic absorption must equal:

$$= HHDem_{r,g} + GovDem_{r,g} + Inv_{r,g} + \sum_{s} IDem_{r,g,s}$$

6. Generate implicit shares based on absorption totals to enforce identities:

$$= Abs_{r,g} / \sum_{rr} Abs_{rr,g}$$

7. Use implicit shares to separate imports and margin demand.

Regionalization Process

In order to maintain zero profit and market clearance in the data, we determine demand/supply from/to the state vs. national markets by imposing *regional purchase coefficients* based on commodity flow survey data.

- Regional purchase coefficients (RPC) are found by assigning aggregate categories in CFS data to blueNOTE sectors. The dataset provides a metric on how much of a given good is retained in a given state or shipped to other states.
- $RPC_{r,g} \in [0, 1]$. I.e. an $RPC_{r,g} = 0.4$ would indicate 40% of a given good's domestic demand was sourced in the state. The rest came from the national market.

State level or national level domestic demand is defined by either the supply or demand side of the market to maintain zero profit in either the export or absorption markets.

Margins are supplied by both the state and national markets.

Trade

The dataset is currently structured for a pooled national market. Explicit bilateral trade flows cannot be determined using CFS data:

- Wittwer (2017) shows that CFS data provide information on the value of goods between transport nodes, which may or may not be in line with production origins or consumption destinations.
- Points to need of gravity based estimates.

Satellite Information

Matrix balancing routines are provided (similar to those in the national case) which can enforce certain totals in the dataset if needed. For energy applications we use the State Energy Data System (SEDS) data.

- It's been pointed out that BEA data tends to under-report energy related demands. Use SEDS to impose both energy demands (which match emission levels) and supplies.
- Electricity supplied by alternative technologies for bottom up representation. Separate electricity production accounts by energy technologies.
- Adjust trade margins to be in tune with electricity mark ups.

Data syntax: Sets & Parameters

Table 1: Set Notation in the Regional CGE Model

Type	Item	Description
Sets:	s,g	Sectors/Goods
	r	Regions
	m	Margin type

Table 3: Parameters in the Regional CGE Model

Parameter	Description
$\bar{ys}_{r,g,s}$	Sectoral supply
$\bar{id}_{r,s,g}$	Intermediate demand
$\bar{ld}_{r,s}$	Labor demand
$\bar{kd}_{r,s}$	Capital demand
$\bar{s}_{r,g}$	Aggregate supply
$\bar{xn}_{r,g}$	National supply
$\bar{xd}_{r,g}$	State level supply
$\bar{x}_{r,g}$	Foreign exports
$\bar{nm}_{r,m,g}$	National margin supply
$\bar{dm}_{r,m,g}$	State level margin supply
$\bar{m}_{r,g}$	Imports
$\bar{nd}_{r,g}$	National demand
$\bar{dd}_{r,g}$	State level demand
$\bar{md}_{r,m,g}$	Margin demand
$\bar{a}_{r,g}$	Armington supply
$\bar{ta}_{r,g}$	Tax net subsidy rate on intermediate demand
$\bar{tm}_{r,g}$	Import tariff
$\bar{cd}_{r,g}$	Final demand
$\bar{yh}_{r,g}$	Household production
$\bar{bop}_{r}$	Balance of payments
$\bar{g}_{r,g}$	Government demand
$\bar{i}_{r,g}$	Investment demand

Social Accounting Matrix

Table 4: 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	$p^D_{r,g}$			$\bar{dd}_{r,g}$	$\bar{dm}_{r,m,g}$								
National Market	p^N_g			$\bar{nd}_{r,g}$	$\bar{nm}_{r,m,g}$								
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						$\bar{yh}_{r,g}$				$\bar{ld}_{r,s}, \bar{kd}_{r,s}$		$\bar{bop}_r$	

Data Overview

The build routine provides:

- Social accounting matrices for all 50 states from 1997-2014.
- Based on summary files of 57 sectors.
- Option for disaggregation using the 2007 389 sectoring scheme and additional satellite accounts.
- Regionalization achieved mainly regional level gross state product and expenditure accounts.
- Trade is imposed in national pooled market using regional purchase coefficients generated by commodity flow survey data.
- Option for recalibrating dataset to match totals from satellite accounts.

Table of Contents

1. blueNOTE Overview
2. Modeling Framework
3. Comparison: blueNOTE vs. IMPLAN
4. State of the Project

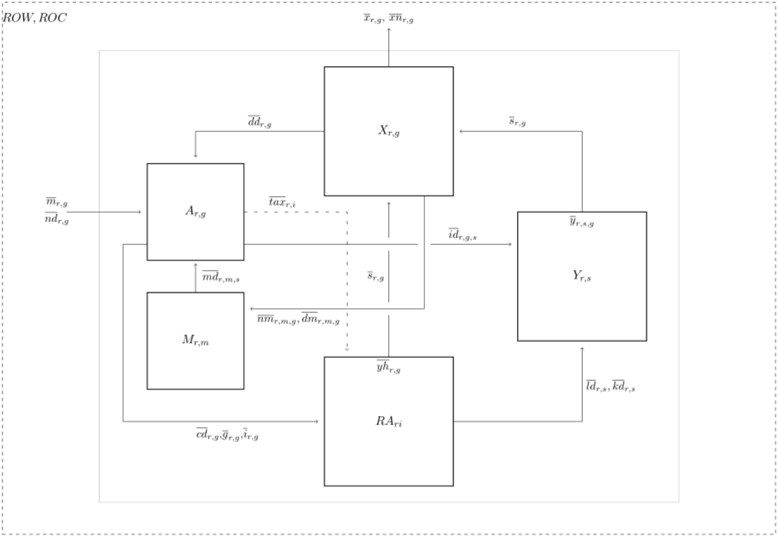
Model/Data Syntax: Variables

Table 2: Nomenclature in the Regional CGE Model

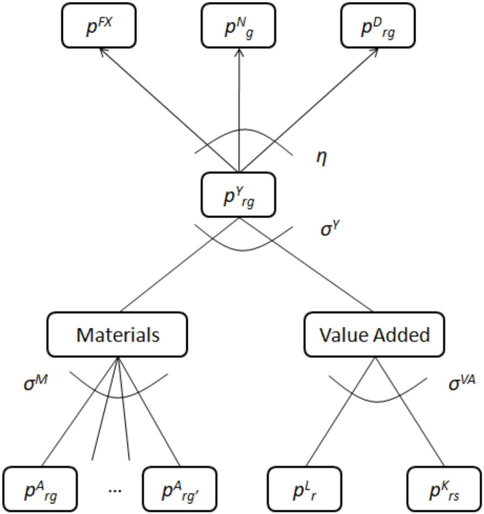
Type	Item	Description
Activity Levels:	$Y_{r,s}$	Sectoral output
	$A_{r,g}$	Armington composite
	$X_{r,g}$	Supply allocation
	$MS_{r,m}$	Margin supply
Prices:	$p^{Y}_{r,g}$	Output market price
	$p^{A}_{r,g}$	Armington composite price index
	$p^{D}_{r,g}$	State level market price for goods
	p^{N}_{g}	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
	GOV_{r}	Representative government

Model Flows

Figure 1: The Regional Economic Structure

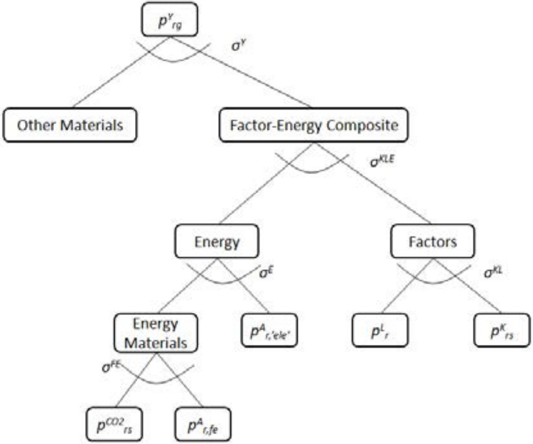


Accounting Model Overview



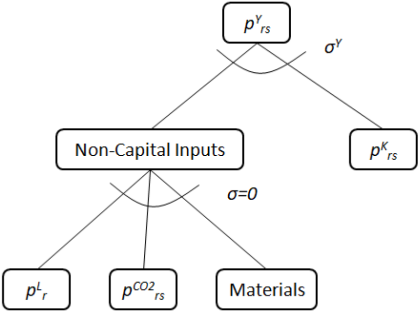
Extensions: EDF Model Energy-Economic Model Structure

Re-arranging energy based inputs, we can tailor the production function for non energy sectors to match KLEM based technologies.



Extensions: EDF Energy-Economic Model Structure

Fossil fuel and electricity production activities are calibrated to capital value shares and exogenous supply elasticities.



Extensions: Electricity Sector in the EDF Model

It may be of interest to include a more detailed representation of the electricity sector. Using SEDS we can decompose the electricity sector by generating technology: coal, natural gas, oil, nuclear, hydro, wind, solar, geothermal.

- Each generation technology produces electricity at the same output price.
- Must separate each input component for different technologies. I.e. Coal mining inputs are used in coal electricity generation, and natural gas is used in gas related electricity production.

Table of Contents

1. blueNOTE Overview
2. Modeling Framework
3. Comparison: blueNOTE vs. IMPLAN
4. State of the Project

blueNOTE vs. IMPLAN

Strengths:

- Transparency: build stream provides all code and data sources to generate regionalized dataset.
- Margin detail: markups are explicitly captured in blueNOTE which is particularly important for electricity related modeling.
- Flexibility: routine provides tools for calibrating model to satellite data tables.

Current version of the build lacks detailed household and government accounts.

- No household groupings by income or government accounts depending on local, state or federal levels. Distinction is given by region.

Given these differences, how would model results compare to equivalent policy simulations?

Basic IMPLAN CGE Model

For basic simulation exercises not reliant on detailed revenue recycling mechanisms, results should be similar if IMPLAN uses comparable procedures for producing regional social accounting matrices.

- Production structure similar to blueNOTE model. Same elasticities and sectoring schemes are employed.
- Slight differences in material goods composition.
- Differences in household and government accounts.
- No explicit representation of margins.

Simulation Exercises

We've merged 2010 state level IMPLAN data for the entire United States and enforced a pooled national market closure in the data. We compare this to the blueNOTE version. Both are represented simply without changes to the core structure.

- No bilateral trade flows in either model explicitly represented.

This is where the current research exists. We are trying to get a sense on how blueNOTE lines up with IMPLAN. Find the sectors that it does a good job at and find those that it doesn't. Are there noticeable differences in the results of policy simulations relying on both datasets? If so, which types?

Table of Contents

1. blueNOTE Overview
2. Modeling Framework
3. Comparison: blueNOTE vs. IMPLAN
4. State of the Project

Dataset Limitations

Aside from those already discussed:

- We stop at 2014 due to underlying data from the BEA.
- We can project to new years but this is contingent on a variety of things. Degree of uncertainty is high.

A Partial Work Agenda

- Representation of trade flows and integration of gravity model based estimates.
- Integration of household data from the American Community Survey (PUMS) – Public Use Microdata Sample. Direct access to US Census records could improve both trade and household data collection. Developing better household accounts will provide more explicit mechanisms for revenue recycling.
- Data set construction and reconciliation tools based on commercial modeling language (GAMS), yet this should not restrict access for non-commercial users (NEOS).

Thank you! Please email me with any questions or comments.
Email: schreiber@aae.wisc.edu