Wisconsin Open Source Economic Data Consortium

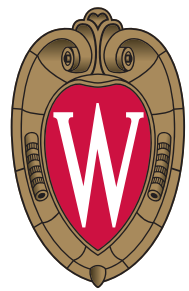
Thomas F. Rutherford
Andrew Schreiber

Department of Agricultural and Applied Economics

Wisconsin Institute for Discovery

University of Wisconsin-Madison

May 18, 2018



Blaise Pascal, Treatise on Vacuum (c.1651)

Je n'ai fait celle-ci plus longue que parce que je n'ai pas eu le loisir de la faire plus courte.

I would have written a shorter letter, but I did not have the time.

Table of Contents

Collaboration with EDF (2017)

An important byproduct of our project will be an open-source dataset suitable for analysis of energy-economy-environment issues in North America. We begin with the national input-output table and downscale to the county level using regional economic statistics from the Bureau of Economic Analysis (sectoral value added and price household expenditure). We also employ data from Census Bureaus (foreign trade statistics) and International Trade Administration for bilateral trade statistics. Input-output tables will further be complemented by physical energy quantities and energy prices from the Department of Energys State Energy Data System (SEDS) of EIA.

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.

- IMPLAN sells both state- and country-level national datasets which are based on public data
- 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.

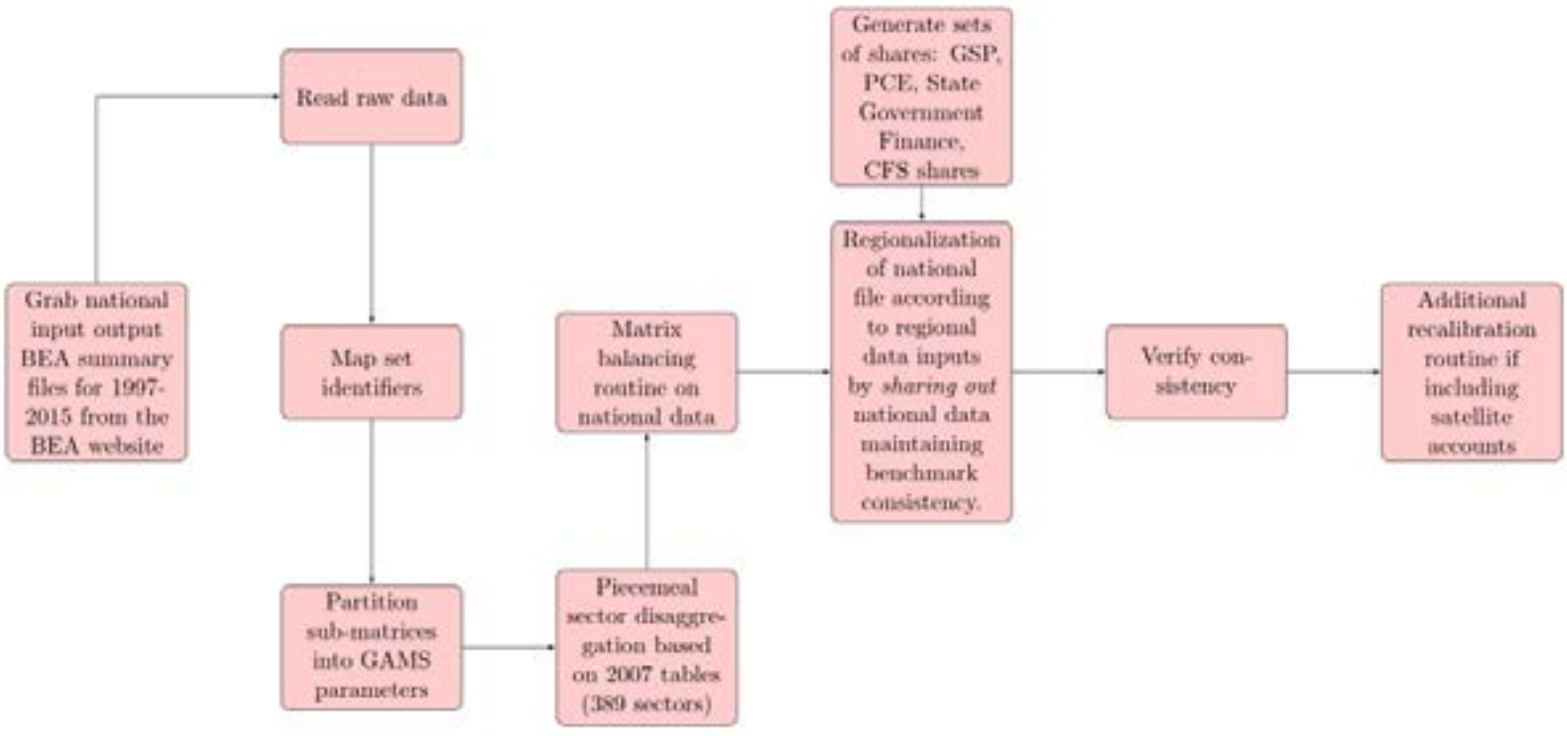
The Energy-Economy Dataset

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.

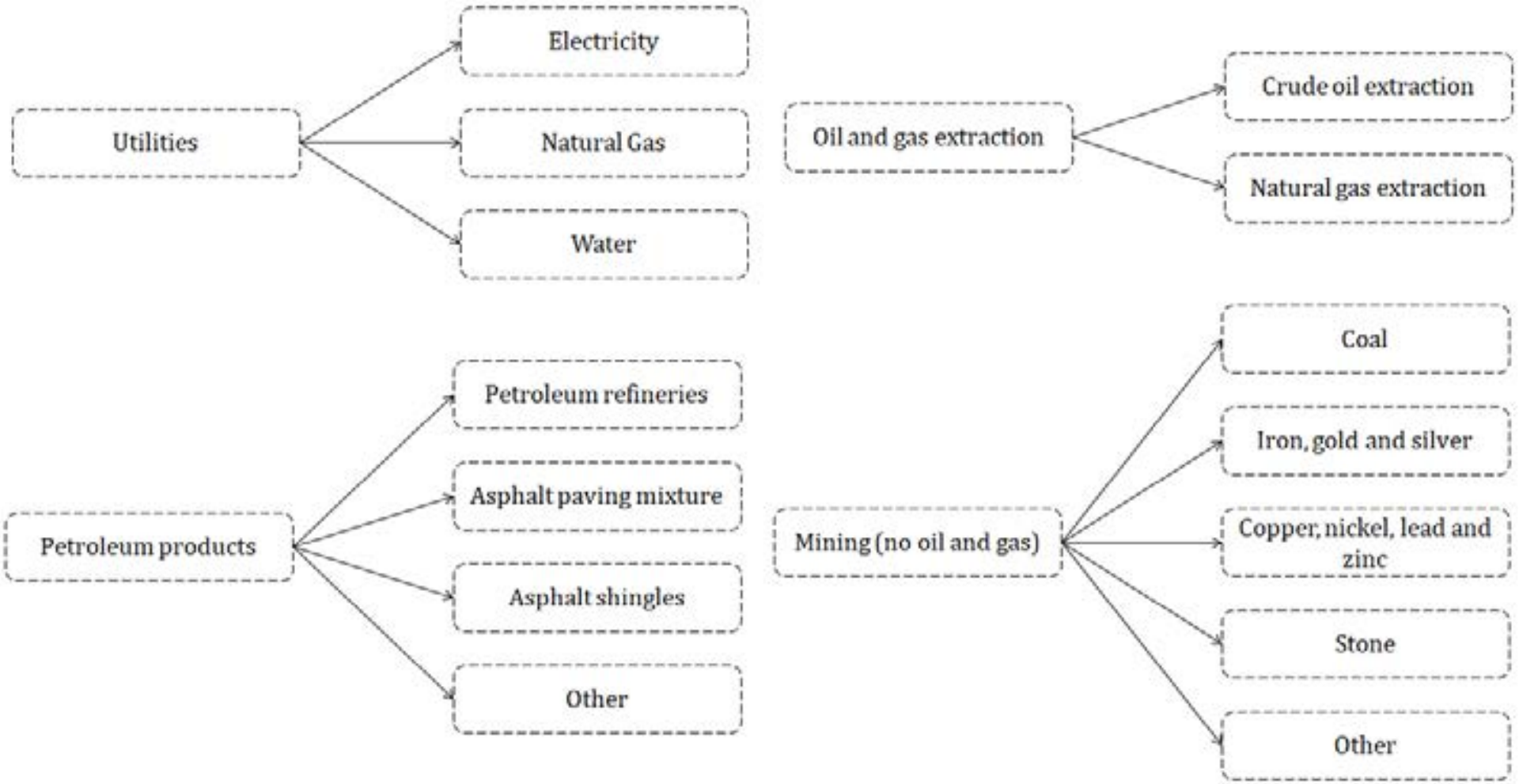
Sector Disaggregation

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



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

- 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.
- Use PCE shares to separate household final consumption.
- Use SGF shares to separate government expenditures.
- GSP shares separate investment demand and exports.
- For a given year then, total domestic absorption must equal:

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

- Generate implicit shares based on absorption totals to enforce identities:

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

- 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

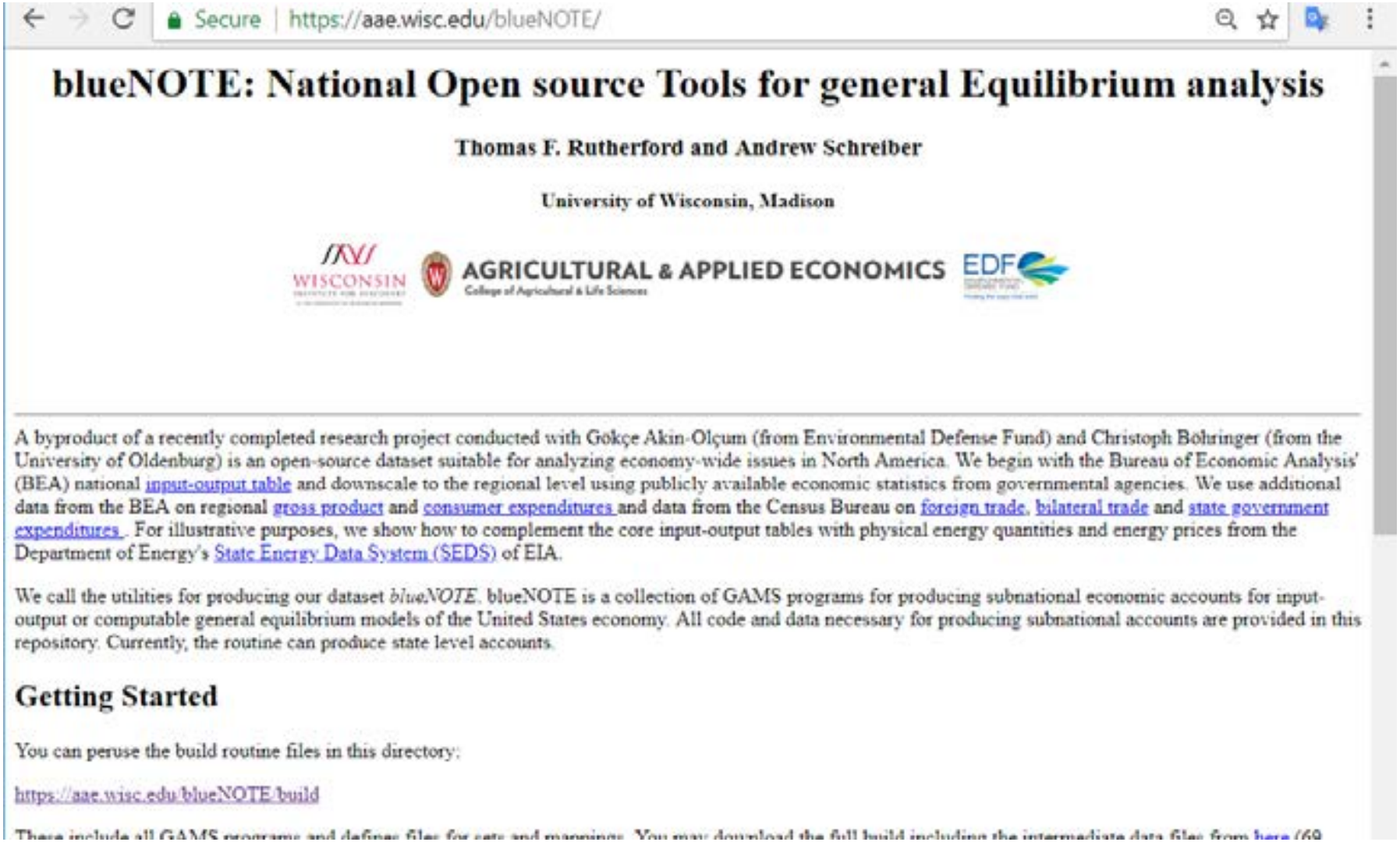
		Production	Exports	Absorption Composite	Margins	Output Market	Regional Market	National Market	Domestic Composite	Factors	Margins Market	Trade	Agents
		$Y_{r,s}$	$X_{r,g}$	$A_{r,g}$	$M_{r,m}$	$p^Y_{r,g}$	$p^D_{r,g}$	p^N_s	$p^A_{r,g}$	$p^L_r, p^K_{r,s}$	$p^M_{r,m}$	p^{FX}	
Production	$Y_{r,s}$					$\bar{ys}_{r,s,g}$							
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.

Web Page Snapshot of the Database



A Cautionary Perspective: Curb Your Enthusiasm*

It is well-known that modelers and policy analysts gain access to policymaking arenas based on what they know. Therefore, critics of models are quick to employ various types of technical standards when evaluating policy models in order to assess validity and reliability of claims to knowledge. This article argues that, in the effort to make models better, overreliance on technical standards misses the important political and policy reasons to model: models call attention to the modelers and to their advice about important policy problems of the day. In this sense, models are used as symbols, as claims to authority, whether or not the underlying knowledge is technically up to snuff. Drawing on the experience of energy policy models, this article explores the problem of models as knowledge versus models as symbols and it examines the muddle that conflicts between them produce. (*Policy Sciences* February 1984, Volume 16, Issue 3, pp 227-243)

Table of Contents

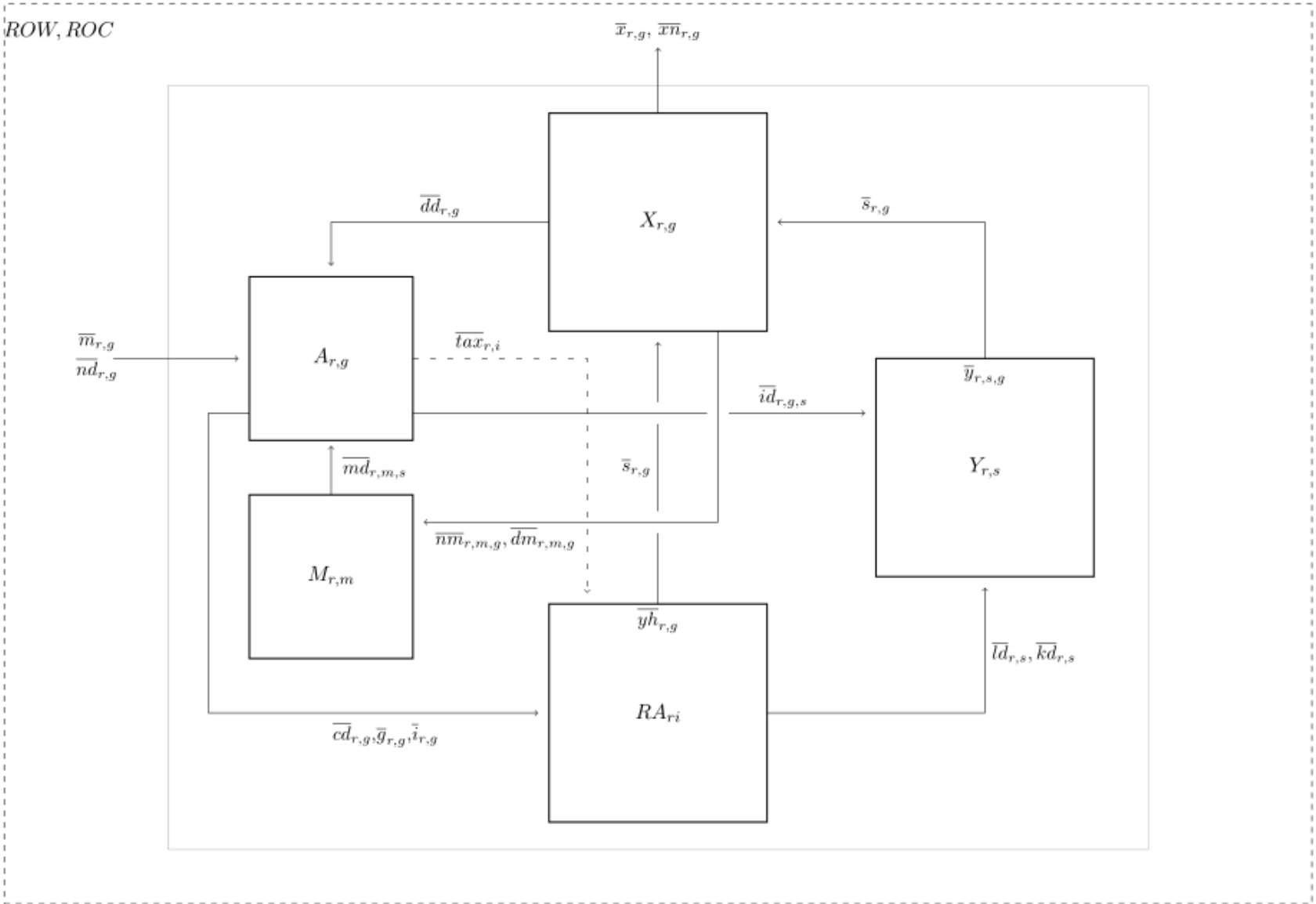
Model/Data Syntax: Variables

Table 2: Nomenclature in the Regional CGE Model

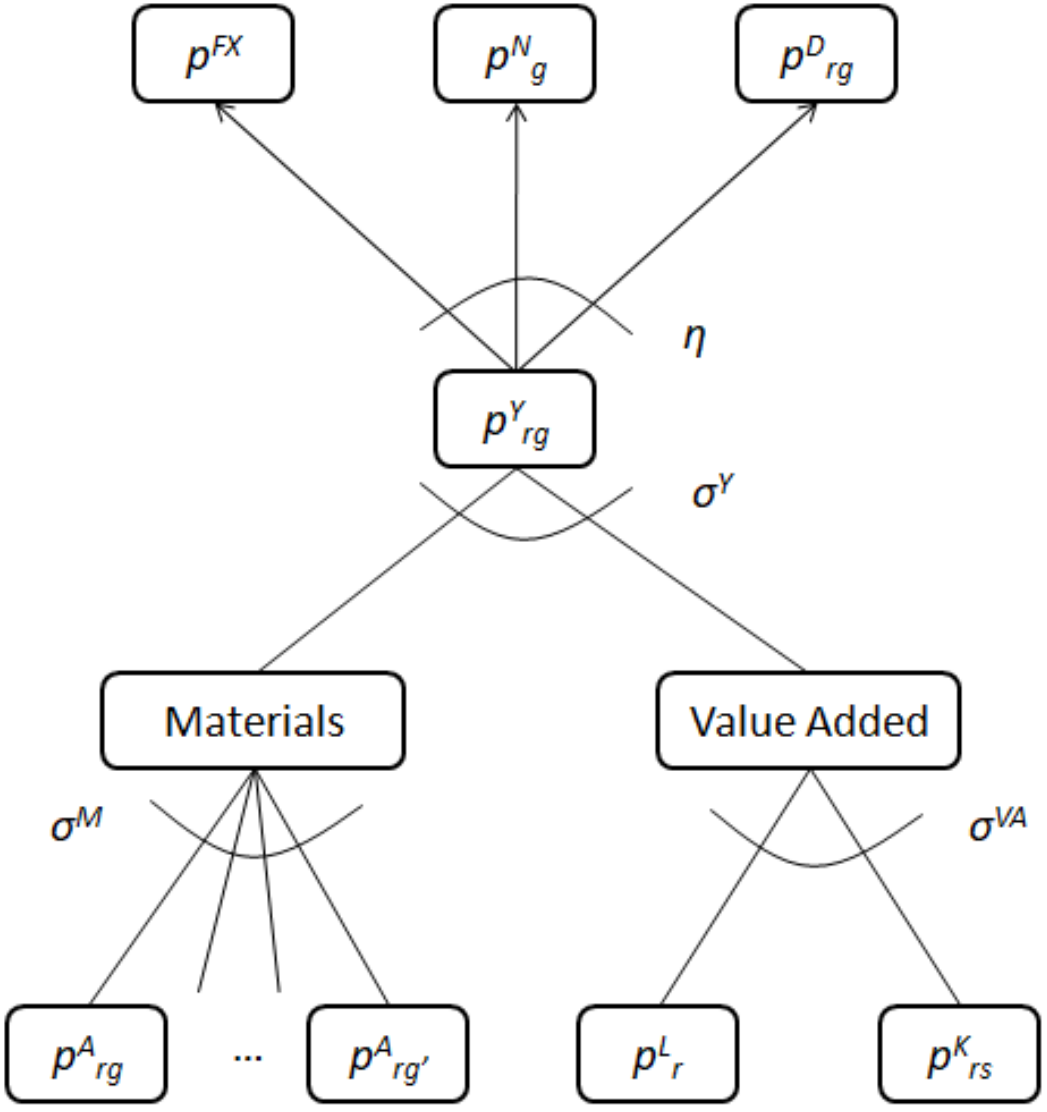
Type	Item	Description
Activity Levels:	$Y_{r,s}$	Sectoral output
	$A_{r,g}$	Armington composite
	$X_{r,g}$	Supply allocation
	$MS_{r,m}$	Margin supply
Prices:	$p^{Y}_{r,g}$	Output market price
	$p^{A}_{r,g}$	Armington composite price index
	$p^{D}_{r,g}$	State level market price for goods
	p^{N}_{g}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

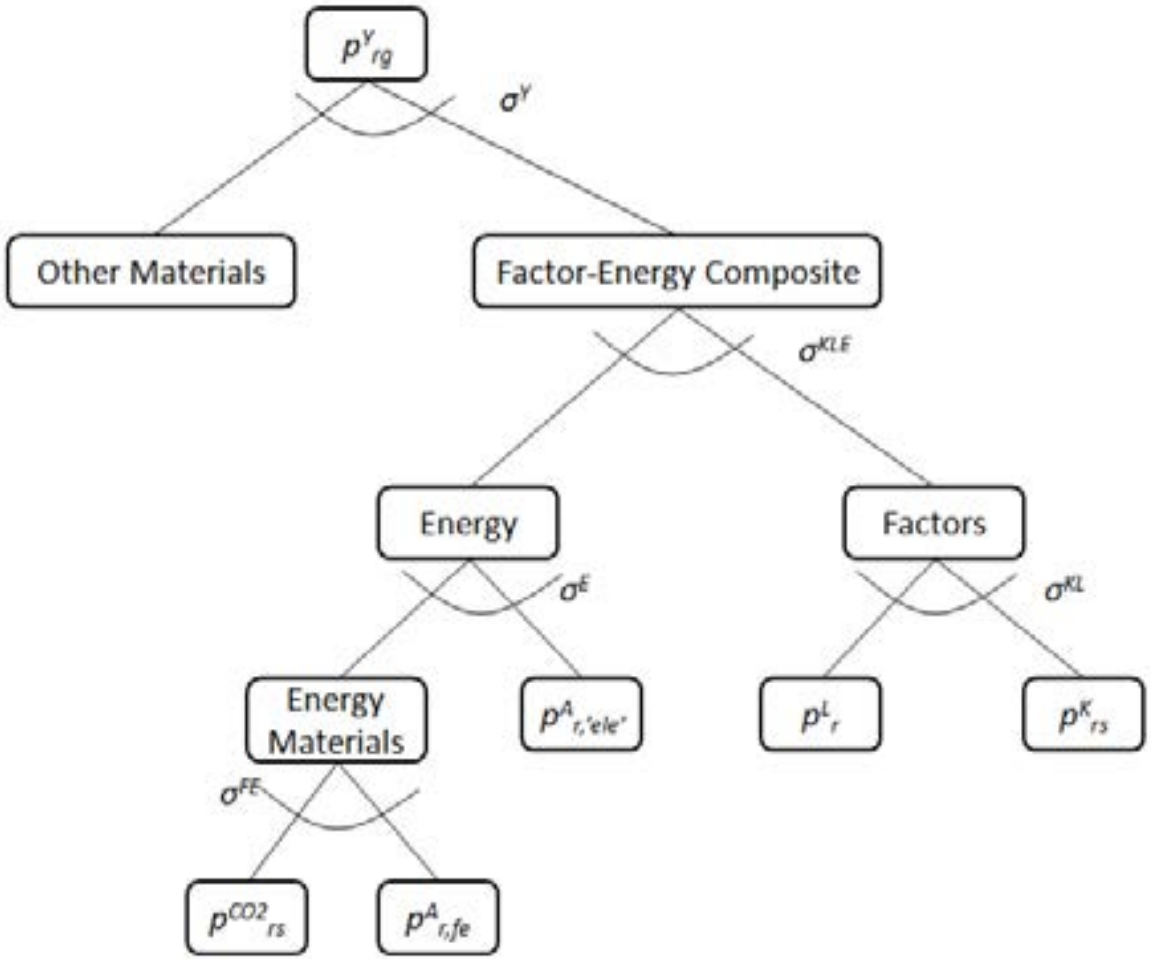


Accounting Model Overview



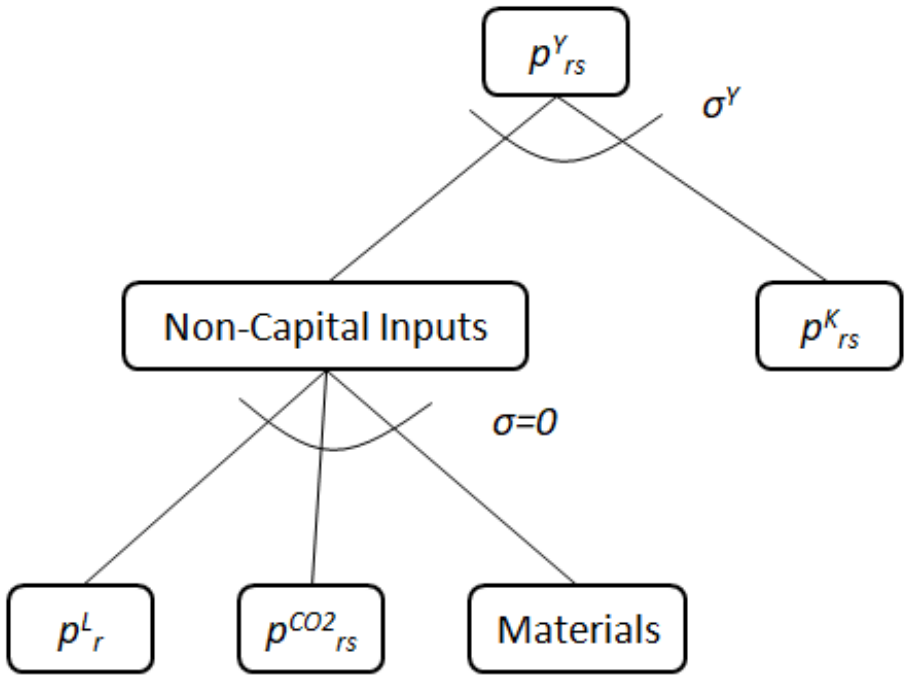
Flexible Functional Form

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



Extension: Energy-Economic Model

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



Extension: Electricity

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.

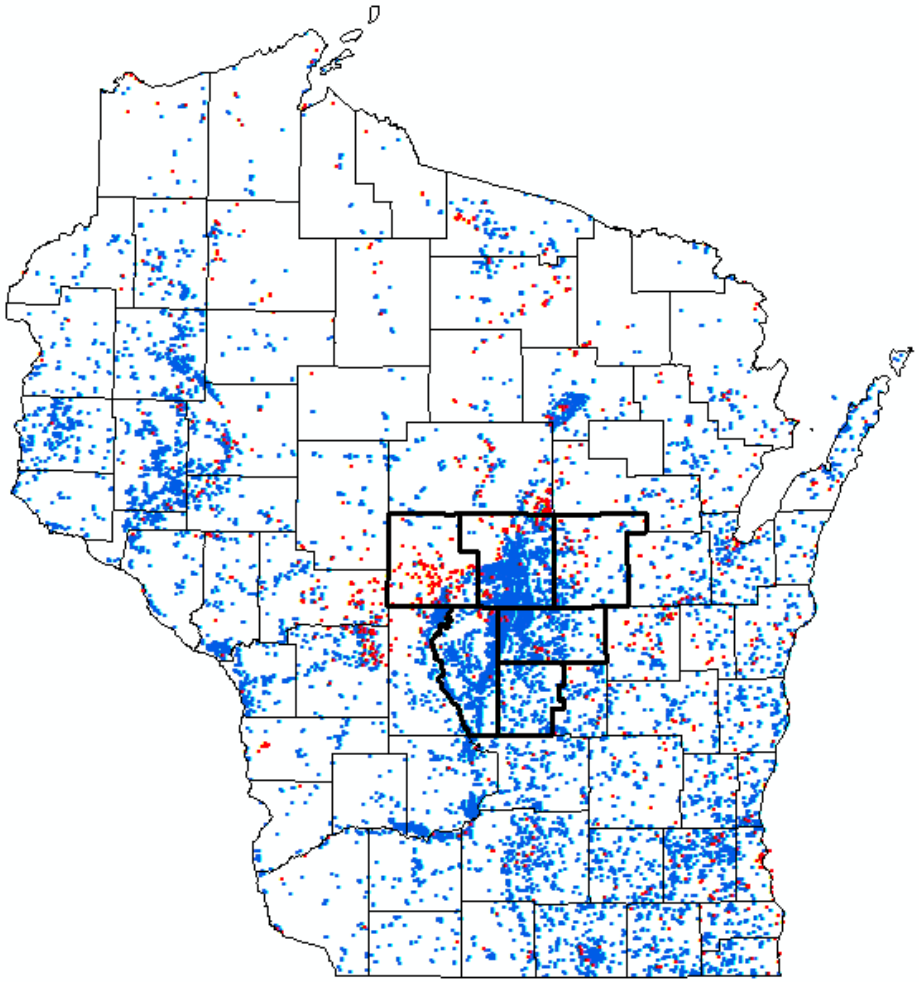
- Each generation technology produces electricity at the same output price.
- Must separate each input component for different technologies. I.e. Coal mining inputs are used in coal electricity generation, and natural gas is used in gas related electricity production.

Extension: Hydrology-Economy Interactions

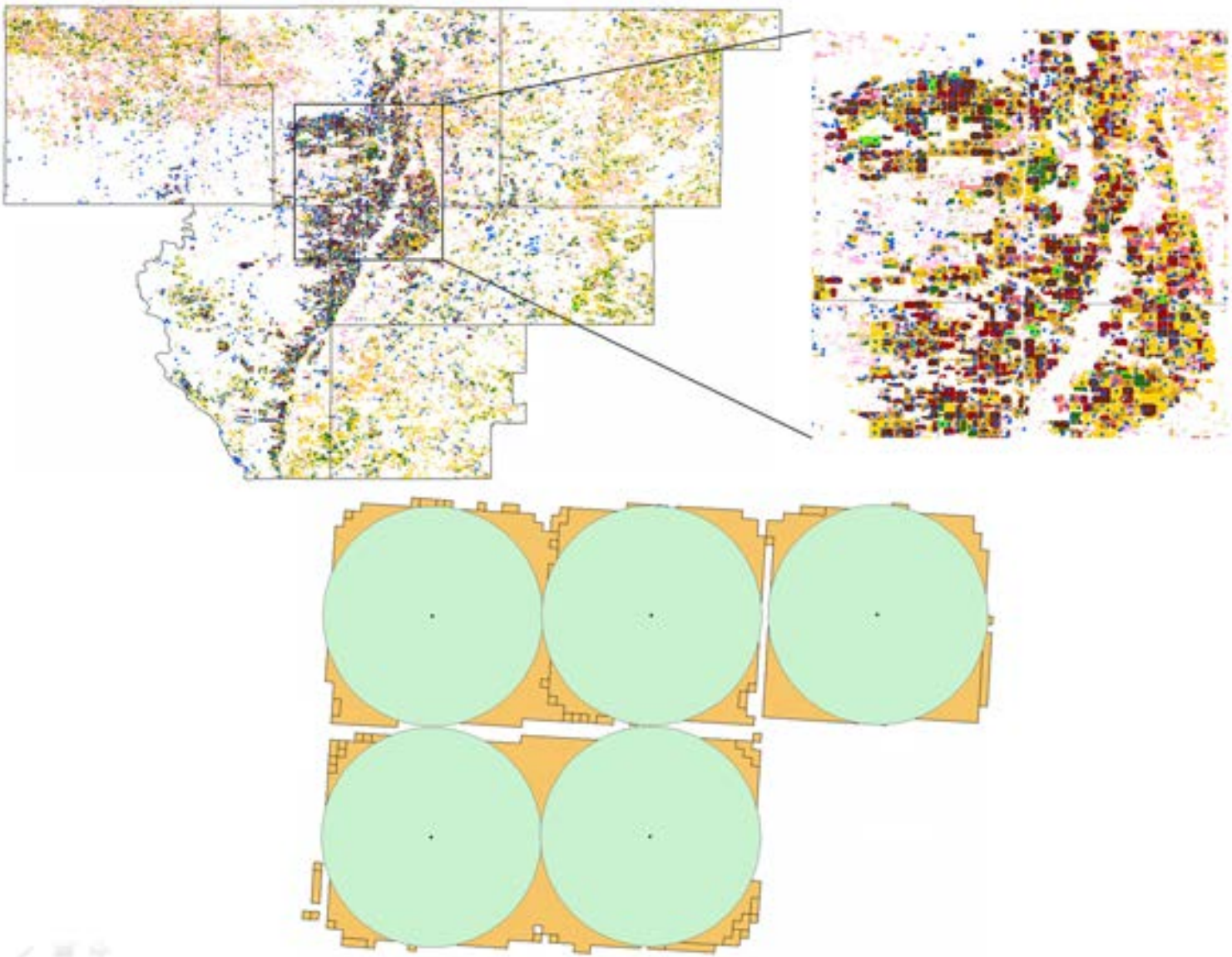
Water withdrawal satellite accounts are derived from Wisconsin Department of Natural Resources GIS data on registered water withdrawal wells in Wisconsin.

- Groundwater, surface water from the Great Lakes and non-Great Lakes.
- Aggregate use categories are provided: Irrigation, Public Supply, etc.
- Irrigation water withdrawals are mapped to crop types using the Cropland Data Layer (CDL) from the National Agricultural Statistical Service.

Irrigation in the Wisconsin Central Sands



Data: Water Mapping



Groundwater Statistics

Table 5: Groundwater Withdrawals

		Adams	Marquette	Portage	Waupaca	Waushara	Wood
Withdrawals (bill. gallons)	*Grains*	6.7	0.5	6.0	0.5	5.0	0.2
	Vegetables, fruit, and nuts	6.8	0.1	7.0	0.2	5.9	0.9
	Oil seeds	1.1	0.1	0.7	0.1	1.1	
	Other crops	0.4	0.1	0.7	0.1	0.6	0.1
	Raw milk			0.1	0.1		
	Public administration	0.3	0.1	2.7	1.4	0.2	1.6
Gallons per dollar output	*Grains*	850.2	53.5	389	29.9	421.5	17.4
	Vegetables, fruit, and nuts	127.4	6	57.9	17.7	102.6	21.9
	Oil seeds	313.1	16.3	149.6	18.3	244.8	8.1
	Other crops	97.7	19.1	68.3	12	97.9	3.4
	Raw milk		1.1	2.4	0.6	0.3	
	Public administration	2.4	2.1	4.4	3.4	1.6	0.8

Notes: Groundwater withdrawals are provided in billions of gallons for the year 2011. Empty elements in tier 1 are associated with non-empty elements of tier 2 due to rounding. Some sectors were omitted due to negligible levels of withdrawals. Surface water withdrawals are excluded from the table. See table 22 in appendix E for more information.
Sources: Wisconsin DNR, USDA Cropland Data Layer.

Table of Contents

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.

Table of Contents

Integration for specific policy analysis:

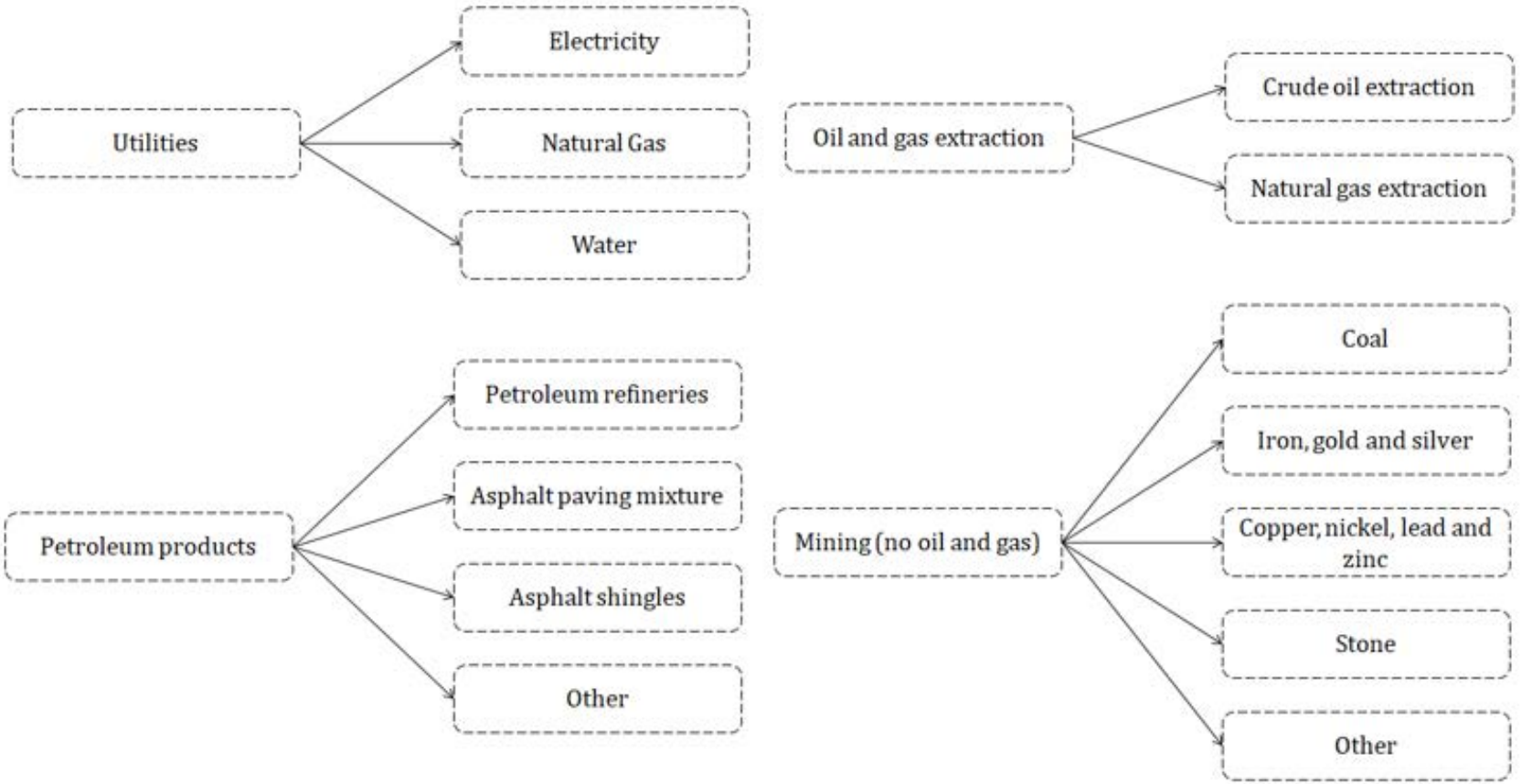
- Bilateral trade flows: developed a method for generating inter-state trade flows for all US sectors. We provide options for a pooled national market and gravity based trade estimates. Trade elasticities are *estimated* based on Canadian trade statistics.
- SEDS (State Energy Data System): blueNOTE recalibrated to match physical energy quantity demands and prices. Matched with state and sector level carbon emissions. Example policy analysis: subnational climate policy.
- NASS (National Agricultural Statistical Survey): blueNOTE recalibrated to match agricultural census data on the market value of sales. Potential policy applications: drought.
- GTAP (Global Trade and Analysis Project): blueNOTE sectors aggregated to the GTAP aggregation and recalibrated to match import and export totals. State disaggregation of bilateral trade to other countries is based on USA Trade Online data. Potential policy applications: international trade issues (e.g. trade tariffs). *Provides state level distributional effects of national trade policy.*

Bilateral Trade

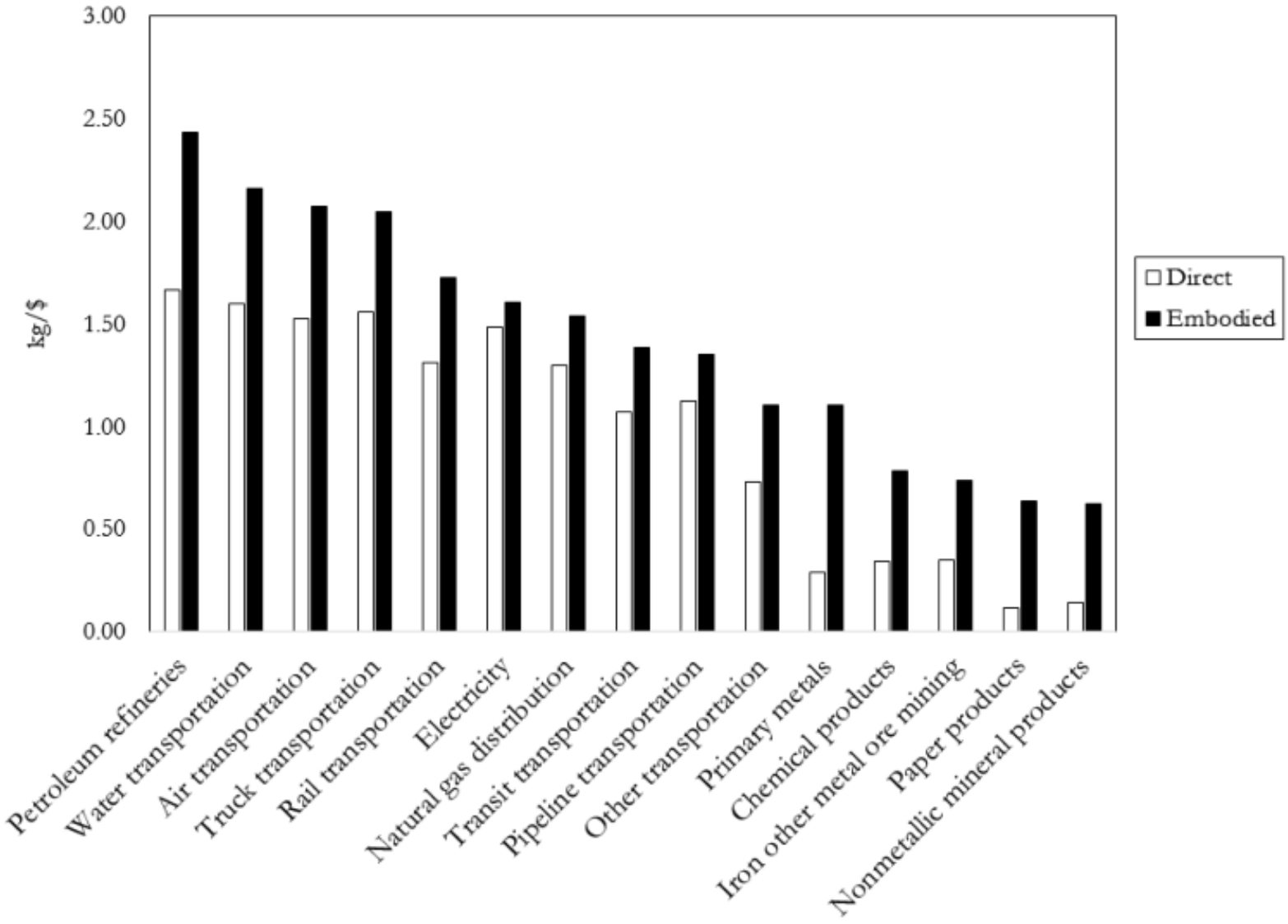
Estimated with Canadian D-Level input output data for 2014 for each blueNOTE sector. Trade from region i to j depends on economic forces in both origin and destination nodes, and forces that aid or restrict the flow of goods from origin to destination.

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

SEDS

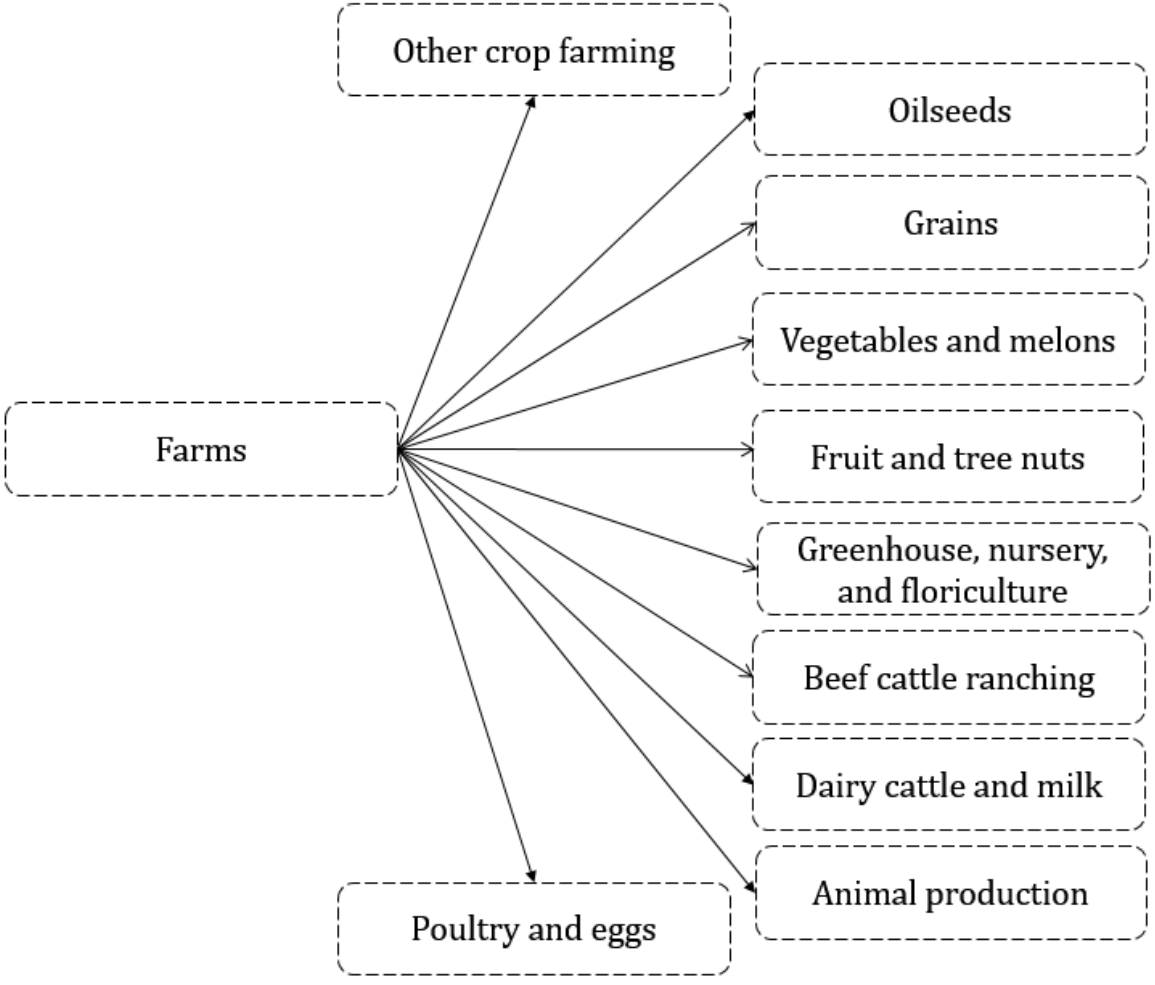

Utilities
Electricity
Natural Gas
Water
Oil and gas extraction
Crude oil extraction
Natural gas extraction
Petroleum products
Petroleum refineries
Asphalt paving mixture
Asphalt shingles
Other
Mining (no oil and gas)
Coal
Iron, gold and silver
Copper, nickel, lead and zinc
Stone
Other

SEDS: Embodied Carbon

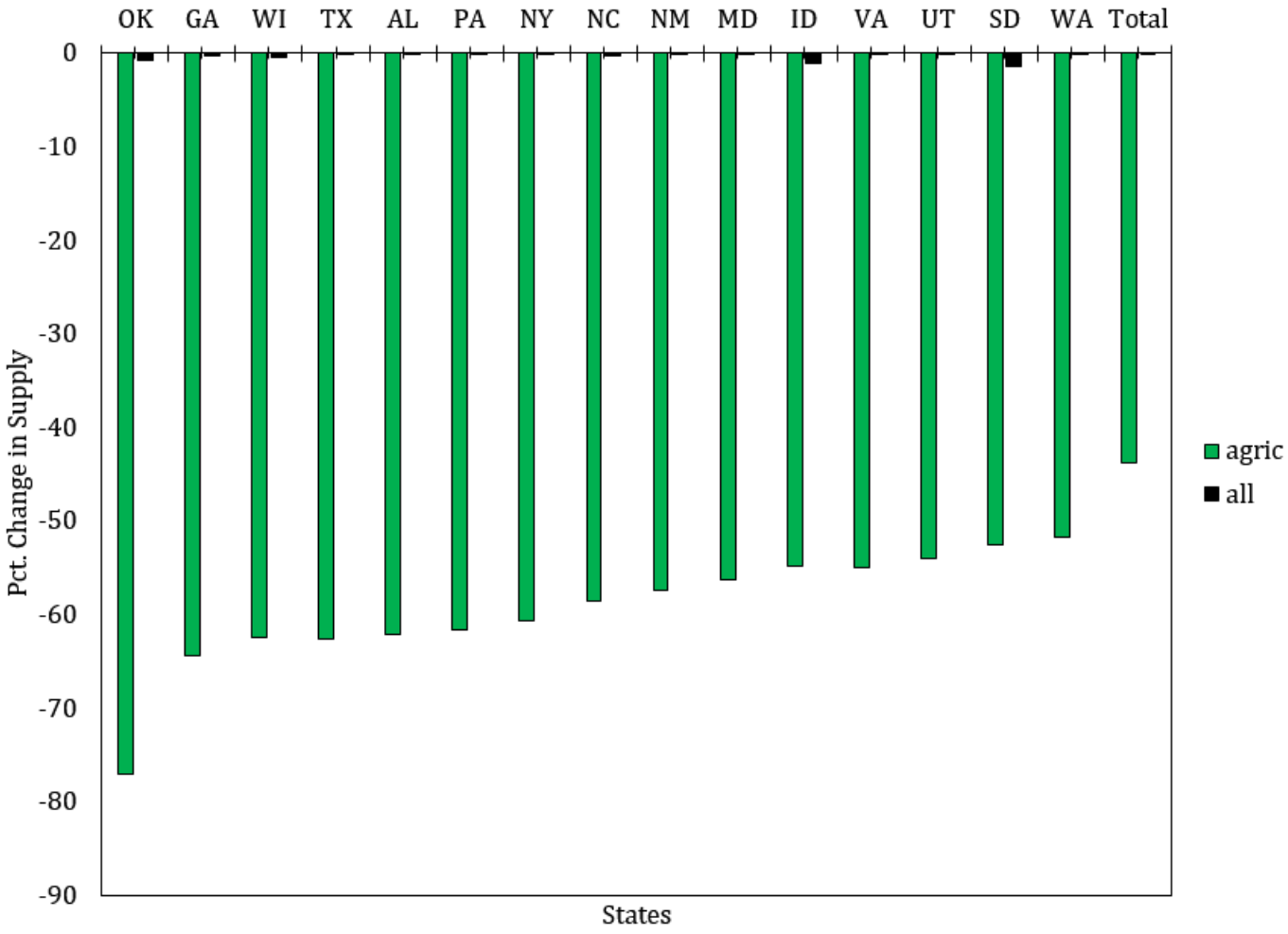


NASS

2007 BEA tables provides the following (can further disaggregate if needed):



Percent Change in Supply (Bottom 15 States)



GTAP Integration

- Mapping from BEA sectoring scheme to GTAP sectors.
- Comparing trade totals from 2011, core blueNOTE data have a 0.9 correlation coefficient across imports and exports with GTAP data. E.g. raw imports and exports are already *close* to GTAP totals.
- To enforce consistency between datasets, blueNOTE is recalibrated to match national import and export totals for the United States in GTAP. State level disaggregation of these imports and exports on origin/destination country in the GTAP database achieved using data from USA Trade Online.

Table of Contents

Future work

- Disaggregate single regional household representative agent account. *Plan:* American Community Survey provides income groups for households at the state level. Household elasticities can be estimated using access to individual level census data at Wisconsin (Wang Jin).
- Further regional disaggregation. *Plan:* BEA reports metro gross product. County business patterns and NASS.
- Further development of trade. *Plan:* Team up with Ed Balistreri (Iowa State University) for expertise on trade modeling.
- Documentation and training materials. *Plan:* Martha Loewe.
- Get students and other users without GAMS licenses access to the data and build stream. *Plan:* Julia/Jump (alternative free optimization software). NEOS – free optimization server housed at Wisconsin (Adam Christiansen).

Current Work Agenda

- Representation of a dataset with bilateral trade flows between Canadian provinces and US states
- Integration of household data from the American Community Survey (PUMS)
- Direct access to US Census records could improve elasticity estimates for both trade and households
- Data set construction and reconciliation tools based on commercial modeling language (GAMS), yet this should not restrict access for non-commercial users (NEOS).

Commodity Flow Survey

GET EMAIL UPDATES

Site Resources
- Press Release

Are You In a Survey?
- Is this form legitimate?
- Get help with your form

Related Sites
- American FactFinder
- Business Help Site
- Economic Census
- NAICS
- Bureau of Transportation Statistics
- North American Transportation Statistics Database

Contact Us
Need help with finding or

Commodity Flow Survey

The Commodity Flow Survey (CFS), undertaken through a partnership between the Census Bureau and the Bureau of Transportation Statistics the Economic Census. The CFS produces data on the movement of goods in the United States. It provides information on commodities shipped destination of shipments of commodities from manufacturing, mining, wholesale, and selected retail and services establishments.

2017 Commodity Flow Survey Respondents

Collection for the 2017 Commodity Flow Survey is ongoing. For inquiries on the Commodity Flow Survey, please visit the following:

- New To complete your 2017 Survey Questionnaire through our online reporting system, please click here.
- New To contact the Commodity Flow Survey Staff about the 2017 Survey Questionnaire, please click here.

2012 Data

New The 2012 CFS Public Use Microdata file is now available through the Public Use Microdata tab on this page.

- 2012 CFS Final estimates are now available on the American FactFinder.
- For more information on metropolitan areas in the CFS, please visit the Economic Census CFS Areas.
- For instructions on how to navigate through American FactFinder for the CFS: Navigate Through AFF for CFS [PDF, 2.7 MB]
- The 2012 CFS data are also available in a print report series:

  U.S. Summary Report [PDF, 8.3 MB]

American Community Survey (ACS)

About the Survey
Respond to the Survey
News & Updates
Data
Data Tables & Tools
Data via FTP
Summary File Data
PUMS Data
Variance Replicate Tables
Race/Ethnicity and American Indian & Alaska Native Data
Custom Tables
Guidance for Data Users
Geography & ACS
Technical Documentation
Methodology
Library
Operations and Administration

PUMS Data

Supporting documentation for the data below is available on the PUMS Documentation page.

PUMS Data 2005 - Current

Available through the American FactFinder website

2011-2015 ACS 5-year PUMS

2015 ACS 1-year PUMS

2010-2014 ACS 5-year PUMS

2014 ACS 1-year PUMS

2009-2013 ACS 5-year PUMS

2011-2013 ACS 3-year PUMS

2013 ACS 1-year PUMS

2008-2012 ACS 5-year PUMS

2010-2012 ACS 3-year PUMS

2012 ACS 1-year PUMS

2007-2011 ACS 5-year PUMS

2009-2011 ACS 3-year PUMS

Available through the FTP site

2004 ACS PUMS

2003 ACS PUMS

2002 ACS PUMS

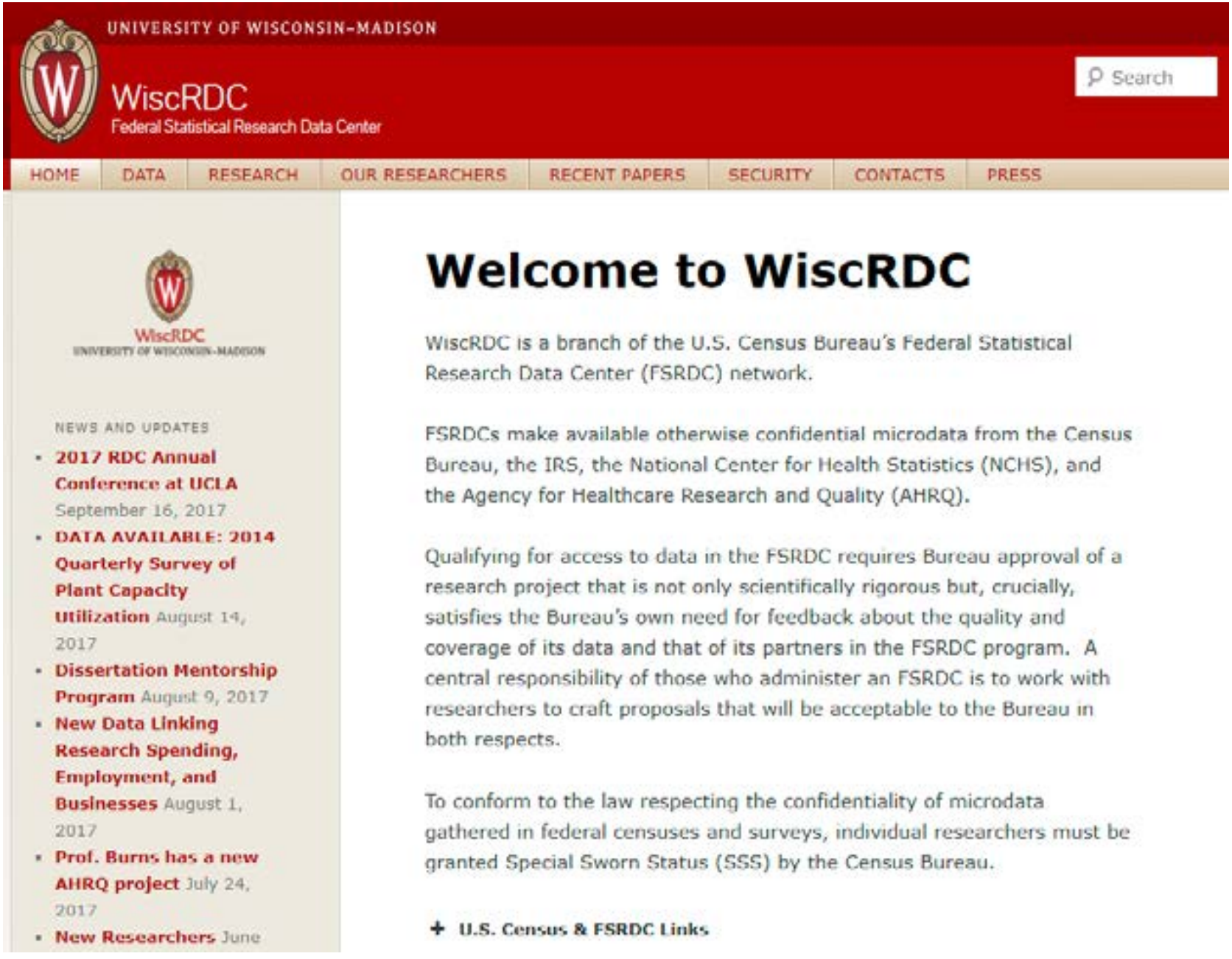
2001 ACS PUMS

2000 ACS PUMS

1996-98 PUMS Files Available on DVD

Microdata for select test areas is available on DVD.

- To request a DVD, email your name and mailing address and specify dataset (PUMS 1996-1998) to acso.dvd.order@census.gov. You may be contacted for additional information before your request is filled.


UNIVERSITY OF WISCONSIN-MADISON
WiscRDC
Federal Statistical Research Data Center
Search
HOME
DATA
RESEARCH
OUR RESEARCHERS
RECENT PAPERS
SECURITY
CONTACTS
PRESS
WiscRDC
UNIVERSITY OF WISCONSIN-MADISON
NEWS AND UPDATES
2017 RDC Annual Conference at UCLA September 16, 2017
DATA AVAILABLE: 2014 Quarterly Survey of Plant Capacity Utilization August 14, 2017
Dissertation Mentorship Program August 9, 2017
New Data Linking Research Spending, Employment, and Businesses August 1, 2017
Prof. Burns has a new AHRQ project July 24, 2017
New Researchers June
Welcome to WiscRDC
WiscRDC is a branch of the U.S. Census Bureau's Federal Statistical Research Data Center (FSRDC) network.
FSRDCs make available otherwise confidential microdata from the Census Bureau, the IRS, the National Center for Health Statistics (NCHS), and the Agency for Healthcare Research and Quality (AHRQ).
Qualifying for access to data in the FSRDC requires Bureau approval of a research project that is not only scientifically rigorous but, crucially, satisfies the Bureau's own need for feedback about the quality and coverage of its data and that of its partners in the FSRDC program. A central responsibility of those who administer an FSRDC is to work with researchers to craft proposals that will be acceptable to the Bureau in both respects.
To conform to the law respecting the confidentiality of microdata gathered in federal censuses and surveys, individual researchers must be granted Special Sworn Status (SSS) by the Census Bureau.
+ U.S. Census & FSRDC Links

NEOS Server: State-of-the-Art Solvers for Numerical Optimization

The **NEOS Server** is a free internet-based service for solving numerical optimization problems. Hosted by the Wisconsin Institute for Discovery at the University of Wisconsin i Madison, the NEOS Server provides access to more than 60 state-of-the-art solvers in more than a dozen optimization categories. Solvers hosted by the University of Wiscons Madison run on distributed high-performance machines enabled by the HTCondor software; remote solvers run on machines at Arizona State University, the University of Klag Austria, and the University of Minho in Portugal.

The **NEOS Guide** website complements the NEOS Server, showcasing optimization case studies, presenting optimization information and resources, and providing backgroun information on the NEOS Server.

Latest NEOS News

June 27, 2017	**The NEOS Server will be down for maintenance Wednesday, July 5th from 8am - 10am CDT.**
June 10, 2017	You can use your NEOS user account to submit authenticated jobs through the web or XML-RPC interfaces. More information is available in our RPC API.
May 28, 2017	We have re-enabled user accounts with improved functionality and security. You will need to reset your password to access an existing account.
Jan 1, 2017	***Sign up for a NEOS user account to receive better service!*** When you Sign Up for a user account and Sign In before you submit your jobs, yo

NEOS Guide

- NEOS Case Studies
- NEOS Optimization Guide
- NEOS Server Information
- Optimization Resources, LP FAQ and NLP FAQ

NEOS Server

- Submit a job to NEOS
- View Job Queue and Job Results
- User's Guide to the NEOS Server
- NEOS Server FAQ
- NEOS Support

Advanced Tools

- Statistics: solvers, web sites, cluster
- Job Archives (password required)
- Downloads: Client Tools (GitHub) and Kestr