United States
Environmental Protection
Agency

National Center for
Environmental Economics
Washington DC 20460

Version 2.0.0
December 23, 2020

SAGE Model Documentation

For Further Information:

Copies of this documentation, source code for the model, and all publicly available data are available at `https://www.epa.gov/environmental-economics/cge-modeling-regulatory-analysis`

Authors and Developers:

Alex Marten, PhD
U.S. Environmental Protection Agency
Office of Policy, National Center for Environmental Economics
Washington, DC

Andrew Schreiber, PhD
U.S. Environmental Protection Agency
Office of Policy, National Center for Environmental Economics
Washington, DC

Ann Wolverton, PhD
U.S. Environmental Protection Agency
Office of Policy, National Center for Environmental Economics
Washington, DC

Contact Information:

Questions related to this document should be addressed to Andrew Schreiber, U.S. Environmental Protection Agency, Office of Policy, Washington, DC 20460 (email: schreiber.andrew@epa.gov).

Suggested Citation:

Marten, A., Schreiber, A., and Wolverton, A. 2020. *SAGE Model Documentation* (2.0.0). U.S. Environmental Protection Agency: https://www.epa.gov/environmental-economics/cge-modeling-regulatory-analysis.

Acknowledgements:

Richard Garbaccio contributed to the initial version of the SAGE model published in Marten and Garbaccio (2018). David A. Evans provided invaluable ongoing review and comments on numerous issues that arose during model construction.

Contents

Executive Summary

In 2017, the EPA's Science Advisory Board (SAB, 2017) recommended that the Agency enhance its regulatory analyses using computable general equilibrium (CGE) models "to offer a more comprehensive assessment of the benefits and costs" of regulatory actions. In response, the EPA has invested in building capacity in this class of economy-wide modeling. A key outcome of this effort is the EPA's CGE model of the U.S. economy, called SAGE. The SAGE model provides an important complement to the analyses typically performed during rule development by evaluating a broader set of economic impacts and offering a more complete estimate of costs.[1]

CGE models, such as SAGE, are aggregate representations of the entire economy. They assume that for some discrete period of time an economy can be characterized by a set of conditions in which supply equals demand in all markets. When the imposition of a regulation alters conditions in one or more markets, the CGE model estimates a new set of relative prices and quantities for all markets that return the economy to a new equilibrium.[2] For example, the model estimates changes in relative prices and quantities for sector outputs and household consumption of goods, services, and leisure that allow the economy to return to equilibrium after the regulatory intervention. In addition, the model estimates a new set of relative prices and demand for factors of production (e.g., labor, capital, and land) consistent with the new equilibrium, which determines changes to household income as a result of the regulation (EPA, 2010). In CGE models, the social cost of the regulation is estimated as the change in economic welfare in the post-regulation simulated equilibrium compared to the pre-regulation "baseline" equilibrium.

Unlike other analytic tools typically used to evaluate the costs of regulations, CGE models account for how effects in directly regulated sectors interact with and affect the behavior of other sectors and consumers. Specifically, they are designed to capture substitution possibilities between production, consumption and trade; interactions between economic sectors; and interactions between a regulation and pre-existing distortions, such as taxes. Figure 1 uses a simplified circular flow diagram to depict how input and output markets are generally connected to each other in CGE models. Following a standard assumption in economics, the model assumes that households maximize their wellbeing, while firms maximize their profits. Households supply factors of production to firms in exchange for income (e.g. wages, profits, and interest payments). Firms use the available factors of production and materials to produce outputs that are then bought and consumed by households.

The SAGE model includes explicit subnational regional representation within the United States. Each region contains multiple representative firms that vary by the commodity they produce and

[1] CGE models may also be able to provide additional information on the benefits of regulatory interventions, though this is a relatively new but active area of research. Note that until the benefits that accrue to society from mitigating environmental externalities can be incorporated in a CGE model, the economic welfare measure is incomplete and needs to be augmented with traditional benefits analysis to develop measures of net benefits.

[2] CGE models are generally focused on analyzing medium- or long-run policy effects since they characterize the new equilibrium (i.e., when supply once again equals demand in all markets). Their ability to capture the transition path of the economy depends on the degree to which they include characteristics of the economy the restrict its ability to adjust instantaneously (e.g., rigidities in capital markets).

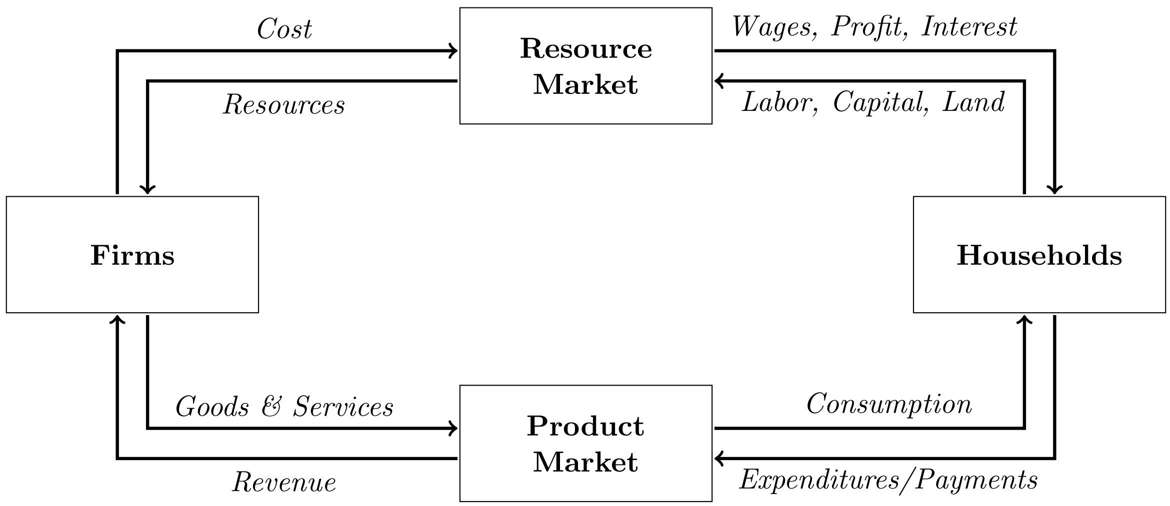


Figure 1: Depiction of the Circular Flow of the Economy

have regional specific production technologies. Each region also has multiple representative households that vary by their income level and have region specific preferences. Within the economy households and firms are assumed to interact in perfectly competitive markets. In addition to households and firms, there is a single government in SAGE that represents all state, local and federal governments within the United States. The government imposes taxes on capital earnings, labor earnings, and production and uses that revenue (in addition to deficit spending) to provide government services, make transfer payments to households, and pay interest on government debt.

Modeling domestic and international trade presents a unique challenge in that the model's structure needs to account for the fact that the United States can be both an importer and an exporter of the same good at both the national and regional level. SAGE handles these cases through the use of the "Armington" approach, which assumes that imported and exported versions of the same good are not perfect substitutes. In SAGE this assumption is applied to both international and cross-regional trade within the United States. In addition, SAGE recognizes that the United States is a relatively large part of the global economy and shifts in our imports and exports have the potential to influence world prices.

SAGE is a forward-looking intertemporal model, which means that households and firms are assumed to make their decisions taking into account what is expected to occur in future years and how current decisions will impact those outcomes. In an intertemporal model care needs to be taken to ensure that in response to a new policy the economy does not instantaneously jump to a new equilibrium in a way that is inconsistent with the rate at which the economy can adjust in practice. SAGE seeks to model a more realistic transition path, in part, by differentiating the flexibility of physical capital by its age. Under this approach the model distinguishes between existing capital constructed in response to previous investments and new capital constructed after the start of the model's simulation. Existing capital is assumed to be relatively inflexible and is used for its original purpose unless a relatively high cost is borne to alter its functionality. New capital is more flexible and easily adjusts to changes in the future. Independent of its vintage, once capital has been constructed in a given region it cannot be moved to another region. While physical capital is not mobile, households can make investments in whatever region of the country they desire.

The dynamics of the baseline economy in SAGE are informed through the calibration of key exogenous parameters in the model. Most importantly are population and productivity growth over time. The model reflects heterogeneity in productivity growth across sectors of the economy consistent with trends that have been historically observed. In addition, the model captures improvements in energy efficiency that are expected for firms and households going forward. Additional baseline characteristics, such as changes to government spending and deficits and changes to international flows of money and investments, are calibrated to key government forecasts or informed by historical trends.

The SAGE model relies on a large number of data sources to calibrate its parameters. The foundation is a state-level dataset produced by IMPLAN that describes the interrelated flows of market goods and factors of production over the course of a year with a high level of sectoral detail. When needed this dataset is augmented by information from other sources, such as the Bureau of Economic Analysis, Energy Information Administration, Federal Reserve, Internal Revenue Service, and the National Bureau of Economic Research. These data are combined with key behavioral parameters for firms and households that are adopted from the published literature or econometrically estimated specifically for the purposes of calibrating SAGE. The result is a static dataset that describes the structure and behavior of the economy in a single year. To develop the forward-looking baseline for the model, additional information on key parameters, such as productivity growth, future government spending, and energy efficiency improvements are incorporated from sources including the Congressional Budget Office and Energy Information Administration.

To ensure that SAGE is consistent with economic theory and reflects the latest science, the EPA initiated a SAB panel to conduct a high-quality technical review, completed in August 2020. The SAB report commended the agency on its development of SAGE, calling it a well-designed open-source model. The report included recommendations for refining and improving the model, including several changes that the SAB advised the EPA to incorporate before using the model in regulatory analysis (denoted as Tier 1 recommendations by the SAB). The SAB's Tier 1 recommendations, including improving the calibration of government expenditures and deficits and the foreign trade deficit; allowing for more flexibility in the consumer demand system; and representing the United States as a large open economy, are incorporated as of version 2.0 of the model.[3] A number of the SAB's medium- and long-run recommendations have also been incorporated into SAGE.

[3]SAGE, version 1.2.2 was peer reviewed by the EPA's Science Advisory Bord (SAB). The final report is available at `https://yosemite.epa.gov/sab/sabproduct.nsf/0/511476D92CEF2AC7852585D6005D373C/$File/EPA-SAB-20-010.pdf`

What's New in Version 2.0?

The current version of the model includes a number of major improvements, many of which implement the top-tier recommendations of the EPA's SAB (2020). Key improvements relative to the previous version of the model (1.2.2) include:

- Improving the calibration of exogenous variables, including government and international accounts.
- Allowing for more flexibility in the consumer demand system so the share of overall spending on different goods varies with changes in income.
- Relaxing the assumption that the United States is a small open economy.
- Allowing for the implementation of variable time steps.
- Allowing for sector-differentiated productivity growth in the baseline.
- Providing additional diagnostic checks and illustrative examples.

1 Introduction

SAGE is a computable general equilibrium (CGE) model of the United States economy developed to aid in the analysis of environmental regulations and policies.[4] It is an intertemporal model with perfect foresight, resolved at the sub-national level. Each of the regions in the model has five households reflective of national income quintiles. Each region has 23 representative firms, with greater disaggregation in the manufacturing and energy sectors that are often impacted by environmental policies. Production technologies are represented with nested CES functions, which may include natural resource inputs. Capital for these firms is represented in a partial putty-clay framework to aid in capturing transition dynamics. A single government agent levies taxes on labor earnings, capital earnings, production, and consumption. The United States is treated as a large open economy that relies on the Armington framework governing both domestic and international trade.

In the following section, technical details on the structure of the model are presented. Section 3 describes the model's calibration. Section 4 discusses the solution algorithm. Section 5 discusses potential options for representing regulations within the model. Section 6 provides a description of how to run the model and describes the verification checks run by the model to test the solution. For a more general description of the model and sensitivity analyses of the model's results we refer the interested reader to Marten et al. (2019).

2 Model Structure

SAGE solves for the set of relative prices that return the economy to equilibrium after the imposition of a policy or other shock, such that all markets clear. This section describes the model's basic structure by first defining the markets in the model, followed by how firms, households, and the government are represented. The section concludes by describing the market clearance conditions that are used to determine equilibrium, where supply equals demand in all markets, along with the closures applied in the model.

2.1 Trade

The United States is modeled as a large open economy. While SAGE does not include the rest of the world explicitly, the model provides a reduced form mechanism to influence world commodity

[4] We use a recursive naming convention, where SAGE stands for SAGE is an Applied General Equilibrium model. Note that CGE and AGE are often used in the economics literature to refer to the same class of models. Shoven and Whalley (1984) describe AGE models as converting the simple two-sector general-equilibrium structure of Arrow and Debreu to a more complex model of the economy that can then be solved computationally to evaluate the welfare and distributional implications of different policies. CGE models have the same aim. Both rely on elasticity estimates to parameterize and simplifying functional forms to solve the models. In addition, both types of models are described as deriving from micro-theoretic foundations (e.g., cost minimization by firms, utility maximization by households, all markets clear in equilibrium), though historically their solution algorithms have differed. AGE models are solved iteratively to approximate the price vector that re-equilibrates the economy. CGE models use macro balancing equations to close the model to solve for a unique solution (Horridge et al., 2013).

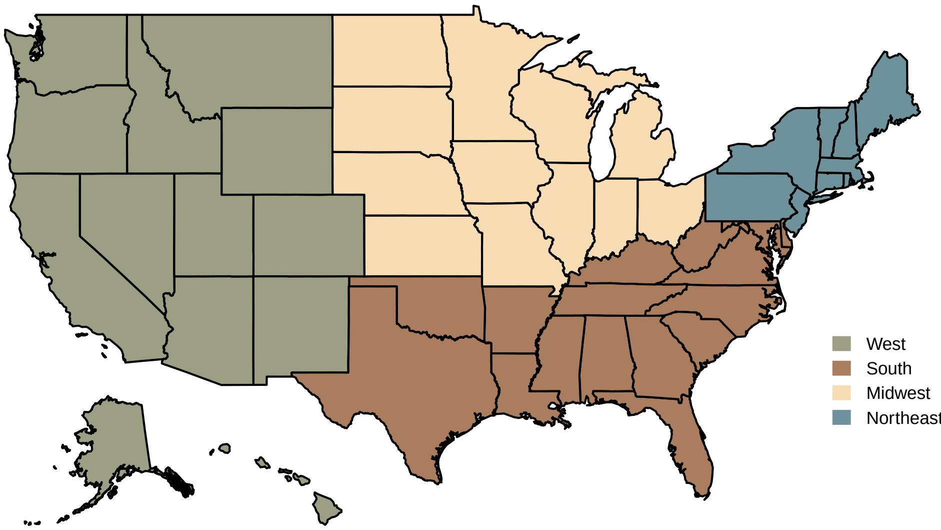


Figure 2: SAGE Regions

prices. There are four subnational regions in the model matching the U.S. Census regions (see Figure 2). Labor and natural resources are not mobile across regions. Capital once installed is not mobile across regions; however, investment is mobile across regions.

Within a region, goods from different origins markets (regional, intra-national imports, international imports) are aggregated using the Armington specification (Armington, 1969). Intra-national trade is pooled at the national level. That is, there exists a single market clearing price for commodities traded across regions, independent of the region of origin or destination.[5] This structure for intra-national trade is similar to other CGE models with subnational detail (e.g., Rausch et al. (2011); Ross (2014)).[6] The Armington aggregate is based on first bundling regional output with intra-national imports and then combining that bundle with international imports. A constant elasticity of transformation (CET) function is used to differentiate regional output between different destination markets (regional, intra-national exports, international exports). This structure is presented in Figure 3.

To implement upward sloping rest of world supply curves for imports into the United States and downward sloping rest of world demand curves for U.S. exports, the model uses a reduced form approach without an explicit representation of the international economy. Specifically, the international demand for U.S. exports and supply of imports into the U.S. market require the use of a fixed factor, of which a rest of world agent is endowed. This specification is not meant to represent a physical process in reality, but is a reduced form assumption that allows the model's

[5] The pooled approach for national trade is due to a lack of well established state-by-state bilateral trade data by commodity.

[6] However, we note that there are examples where estimates of state-by-state bilateral trade matrices have been applied (e.g., Balistreri and Rutherford (2001); Caron and Rausch (2013)).

export demand and import supply curves to be calibrated to exogenous price elasticities. This reduced form approach to modeling a large open economy follows Yuan et al. (2019).

More specifically, the Armington aggregate is defined as

$$a_{t,r,s} = a0_{r,s}\left\{cs\_nf_{r,s}\left(\frac{m_{t,r,s,ftrd}}{m0_{r,s,ftrd}}\right)^{\frac{se\_nf-1}{se\_nf}} + (1-cs\_nf_{r,s})\left[\ cs\_dn_{r,s}\left(\frac{m_{t,r,s,dtrd}}{m0_{r,s,dtrd}}\right)^{\frac{se\_dn-1}{se\_dn}} + (1-cs\_dn_{r,s})\left(\frac{d_{t,r,s}}{d0_{r,s}}\right)^{\frac{se\_dn-1}{se\_dn}}\right]^{\frac{(se\_nf-1)se\_dn}{se\_nf(se\_dn-1)}}\right\}^{\frac{se\_nf}{se\_nf-1}}, \tag{1}$$

where $a_{t,r,s}$ is the Armington composite in period t and region r for commodity s, $m_{t,r,s,trd}$ are imports from market trd, $d_{t,r,s}$ is domestic production consumed locally. Throughout this document a 0 trailing a variable name denotes the value in the benchmark year; benchmark cost shares have the prefix cs; and substitution elasticities have the prefix se. [7] The national market is denoted $dtrd$ and the international market is denoted $ftrd$. The parameter $cs\_nf_{r,s}$ represents the international imports share of the Armington composite, and $cs\_dn_{r,s}$ represents the share of national imports in the domestic-national composite. The substitution elasticity between international imports and the domestic-national composite is se_nf and the substitution elasticity between domestic production and national imports is se_dn. The inputs into the Armington aggregate are determined based on minimizing the price of the composite good, $pa_{t,r,s}$, given the price in the domestic market, $pd_{t,r,s}$, the price in the national market, $pn_{t,s}$, and the price of foreign imports, $pm_{t,s}$.

The CET function to differentiate domestic output across destination markets is defined as

$$y_{t,r,s} + y\_ex_{t,r,s} = y0_{r,s}\left[cs\_dx_{r,s,d}\left(\frac{d_{t,r,s}}{d0_{r,s}}\right)^{\frac{te\_dx+1}{te\_dx}} + cs\_dx_{r,s,dtrd}\left(\frac{x_{t,r,s,dtrd}}{x0_{r,s,dtrd}}\right)^{\frac{te\_dx+1}{te\_dx}} + cs\_dx_{r,s,ftrd}\left(\frac{x_{t,r,s,ftrd}}{x0_{r,s,ftrd}}\right)^{\frac{te\_dx+1}{te\_dx}}\right]^{\frac{te\_dx}{te\_dx+1}}, \tag{2}$$

where $y_{t,r,s}$ is output from production with new capital, $y\_ex_{t,r,s}$ is output from production with extant capital, $x_{t,r,s,trd}$ is exports to market trd, $cs\_dx_{r,s,mkt}$ is the share of output destined for market mkt, and te_dx is the transformation elasticity. Within the production possibilities frontier represented by equation (2), firms select the shares of production destined for each market based on maximizing the price of output, $py_{t,r,s}$, given the price of the commodity in the different destination markets.

[7] In the model, most substitution elasticities vary across sectors as discussed in further detail in Section 3. However, to simplify the exposition, in this document we forgo the sector subscript on substitution elasticities.

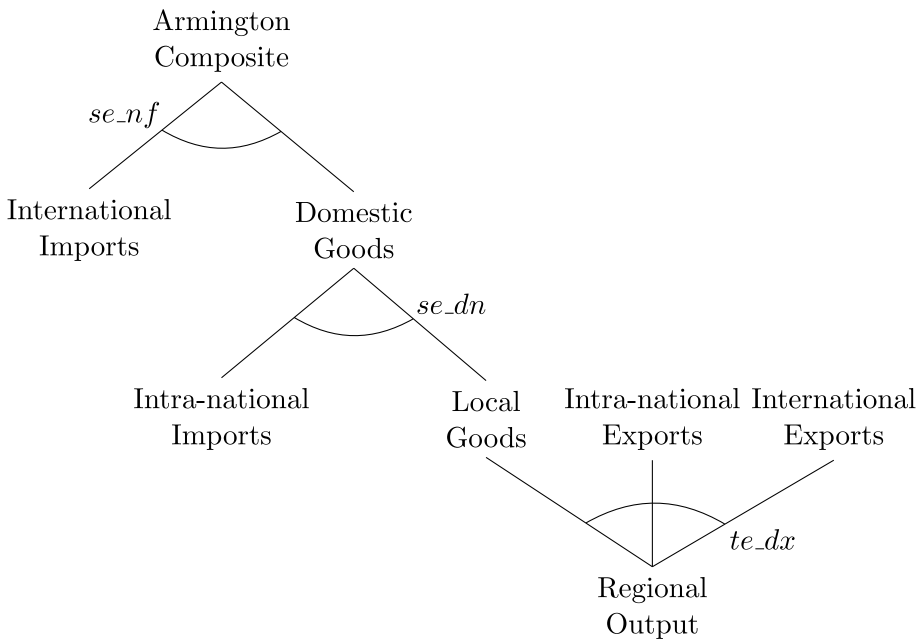


Figure 3: Armington Trade Specification

The supply of imports from the rest of the world is characterized by a Cobb-Douglas production function, defined as

$$ms_{t,s} = \Big(\sum_r m0_{r,s,ftrd}\Big)\left(\frac{fim_{t,s}}{fim0_s}\right)^{cs\_loe\_m_s}\left(\frac{ms\_fx_{t,s}}{(\sum_r m0_{s,ftrd} - fim0_s)}\right)^{(1-cs\_loe\_m_s)}, \tag{3}$$

where $ms_{t,s}$ is the supply of imports into the U.S. market in period t for commodity s, $fim_{t,s}$ is the fixed factor input to import supply and $ms\_fx_{t,s}$ is foreign imports less the fixed factor. cs_m_s denotes the calibrated cost share based on exogenous price elasticities for import supply (as further explained in Section 3). The inputs into the supply of imports at price $pm_{t,s}$ are based on minimizing the costs of the fixed factor, $pfim_{t,s}$, and the price for foreign exchange, pfx_t.

Similarly, the international demand for exports ($xd\_fx_{t,s}$) is characterized by a Cobb-Douglas production function combining exports from the U.S. market ($xd_{t,s}$) and exports from the rest of world agent treated as a fixed factor ($fix_{t,s}$), defined as

$$xd\_fx_{t,s} = \Big(\sum_r x0_{r,s,ftrd} + fix0_s\Big)\left(\frac{fix_{t,s}}{fix0_s}\right)^{cs\_loe\_x_s}\left(\frac{xd_{t,s}}{\sum_r x0_{r,s,ftrd}}\right)^{(1-cs\_loe\_x_s)}, \tag{4}$$

where $cs\_loe\_x_s$ denotes the calibrated cost share based on exogenous price elasticities for export demand. The inputs into the demand of international exports at price pfx_t are based on minimizing the costs of the fixed factor $pfix_{t,s}$ and the U.S. export price, $px_{t,s}$. Notably, when $cs\_loe\_m_s = 0$ and $cs\_loe\_x_s = 0$, the model is translated into a small open economy where $pm_{t,s} = px_{t,s} = pfx_t$.

We use a Cobb-Douglas function without loss of generality to pin down calibrated cost shares and fixed factors based on exogenous price elasticities.

2.2 Production

Production in the model is aggregated to 23 sectors, with greater detail in manufacturing and energy. The sectors in the model and their associated NAICS codes are presented in Table 1. This default disaggregation represents sectors that have historically been the focus of environmental regulations. This set of sectors also maps nicely into the industrial sectors of the U.S. Energy Information Administration's (EIA) National Energy Model System (NEMS) that are used to inform the baseline calibration.

2.2.1 Manufacturing and Service Sectors

In SAGE, perfectly competitive firms maximize profits subject to market prices and a given production technology. Due to their parsimony and global regularity, nested constant elasticity of substitution (CES) production functions have become widely used in applied general equilibrium modeling (Brockway et al., 2017), and this is particularly true in the case of CGE models used to analyze energy and environmental policies. Similarly, SAGE makes use of nested CES functions (in calibrated share form) to define the production functions for the sectors represented. The policy response of CGE models based on nested CES production functions may be sensitive to the ordering of the nests, as this choice defines separability of the production functions amongst inputs (Lecca et al., 2011). Thus, there has been much discussion about the hierarchy for nested CES production functions, particularly with regards to capital, K, labor, L, and energy, E, inputs. Much of this discussion has been based on heuristics, although the empirical work of Van der Werf (2008) is a notable exception. Van der Werf (2008) studied the fit of different nesting structures given historical production data for 12 OECD countries between 1978 and 1996. Van der Werf (2008) finds that the nesting structure combining K and L in the lower nest and the KL bundle with E in the top nest, denoted $KL(E)$, provides a significantly better fit to the data compared to the other possible nesting structures. Furthermore, Van der Werf (2008) finds that the structure combining K and E in the lower nest provided the worst fit for the data, a finding that has been corroborated in other single country contexts (e.g., Dissou et al. (2015); Ha et al. (2012); Kemfert (1998)). Other multi- and single-country studies have found that the $KE(L)$ nesting structure may fit the data as well as the $KL(E)$ structure at the aggregate national level (e.g., Markandya and Pedroso-Galinato (2007); Su et al. (2012)). However, Kemfert (1998) finds that in cases where the $KE(L)$ nesting structure finds support at the aggregate national level the specification may actually provide a worse fit than the $KL(E)$ structure when disaggregated sectoral production functions are estimated. We use a structure that combines primary factors K and L in a lower nest, where that value-added bundle is then combined with an energy composite. At the top level of the production function the $KL(E)$ composite is combined with a Leontief composite of material inputs. This structure is similar to

Table 1: Model Sectors

Abbreviation	Description	NAICS Codes	NEMS IDM Code
agf	Agriculture, forestry, fishing and hunting	11	1, 2
gas	Natural gas extraction and distribution	211,* 213111,* 213112,* 2212,	4*
cru	Crude oil extraction	211,* 213111,* 213112*	4*
col	Coal mining	2121, 213113	3
min	Metal ore and nonmetallic mineral mining	2122, 2123, 213114, 213115	5
ele	Electric power	2211	NA
wsu	Water, sewage, and waste	2213	NA
con	Construction	23	6
fbm	Food and beverage manufacturing	311, 312	7
wpm	Wood and paper product manufacturing	321, 322	8, 19
ref	Petroleum refineries	32411	NA
chm	Chemical manufacturing	325	9
prm	Plastics and rubber products manufacturing	326	20
cem	Cement	32731	22
pmm	Primary metal manufacturing	331	12, 13
fmm	Fabricated metal product manufacturing	332	14
cpu	Electronics and technology	334, 335	16, 18
tem	Transportation equipment manufacturing	336	17
bom	Balance of manufacturing	3122, 313, 314, 316, 323, 32412, 3271, 3272, 32732, 32733, 32739, 3274, 3279, 333, 337, 339	10, 15, 21, 23
trn	Non-Truck Transportation	481, 482, 483, 485, 486, 4869, 487, 488, 491, 492, 493	NA
ttn	Truck transportation	484	NA
srv	Services	42, 44, 45, 51, 52, 53, 54, 55, 56, 61, 624, 71, 72, 81	NA
hlt	Healthcare services	621, 622	NA

* Crude oil and natural gas extraction is included as a single sector in the benchmark data. However, we disaggregate this activity into separate sectors for crude oil and natural gas. Details are available in Section 3.1.1.

other CGE models used to analyze energy and environmental policies (e.g., Paltsev et al. (2005); Rausch et al. (2011); Capros et al. (2013); Cai et al. (2015)).

For the energy composite we also use a nested CES function to represent available production technologies. Initial work using energy-explicit CGE models typically combined all energy sources - including primary energy sources and electricity - in a single nest, commonly with a unit substitution elasticity (e.g., Borges and Goulder (1984)). Subsequent efforts separated electricity from other primary energy sources in a two-nest CES structure that defined the energy composite (e.g., Van der Mensbrugghe (1994); Babiker et al. (1997); Paltsev et al. (2005); Rausch et al. (2011) Böhringer et al. (2018)). The assumption of weak seperability between primary energy inputs and electricity is representative of the primary energy choice across fuels for a sector being defined more by the production process or regional fuel supply characteristics than by the price of electricity. Some recent models have even gone a step further using a three-level CES nest to further disaggregate the primary energy composite in order to impose separability between some of the fossil-fuel use decisions in the cost-minimization problem (e.g., Burniaux and Truong (2002); Chateau et al. (2014); Ross (2014)). However, the three-level CES nesting structure has not been applied consistently across models, and evidence of weak separability in the data is lacking empirically (Serletis et al., 2010a).

SAGE applies the two-level energy nesting with the bottom level nest combining refined petroleum products (or by-products), coal, and natural gas. The second level nest combines the primary energy composite with electricity. This nesting structure is presented in Figure 4. More specifically, the production function for manufacturing goods and services produced with new capital is

$$y_{t,r,s} = y0_{r,s}\left[cs\_klem_{r,s}\left(\frac{mat_{t,r,s}}{mat0_{r,s}}\right)^{\frac{se\_klem-1}{se\_klem}} + (1-cs\_klem_{r,s})\left(\frac{kle_{t,r,s}}{kle0_{r,s}}\right)^{\frac{se\_klem-1}{se\_klem}}\right]^{\frac{se\_klem}{se\_klem-1}}, \quad (5)$$

where $mat_{t,r,s}$ is the materials bundle, which is defined as

$$mat_{t,r,s} = mat0_{r,s}\min\left(\frac{id_{t,r,agf,s}}{id0_{r,agf,s}},\ldots,\frac{id_{t,r,srv,s}}{id0_{r,srv,s}}\right). \quad (6)$$

$id_{t,r,ss,s}$ is the demand for intermediate good ss, and $kle_{t,r,s}$ is the energy and value added composite, which is defined as

$$kle_{t,r,s} = kle0_{r,s}\left[cs\_kle_{r,s}\left(\frac{ene_{t,r,s}}{ene0_{r,s}}\right)^{\frac{se\_kle-1}{se\_kle}} + (1-cs\_kle_{r,s})\left(\frac{kl_{t,r,s}}{kl0_{r,s}}\right)^{\frac{se\_kle-1}{se\_kle}}\right]^{\frac{se\_kle}{se\_kle-1}}. \quad (7)$$

$ene_{t,r,s}$ is the electricity and primary energy composite, which is defined as

$$ene_{t,r,s} = ene0_{r,s}\left[cs\_ene_{r,s}\left(\frac{en_{t,r,s}}{en0_{r,s}}\right)^{\frac{se\_ene-1}{se\_ene}} + (1-cs\_ene_{r,s})\left(\frac{id_{t,r,ele,s}}{id0_{r,ele,s}}\right)^{\frac{se\_ene-1}{se\_ene}}\right]^{\frac{se\_ene}{se\_ene-1}}. \quad (8)$$

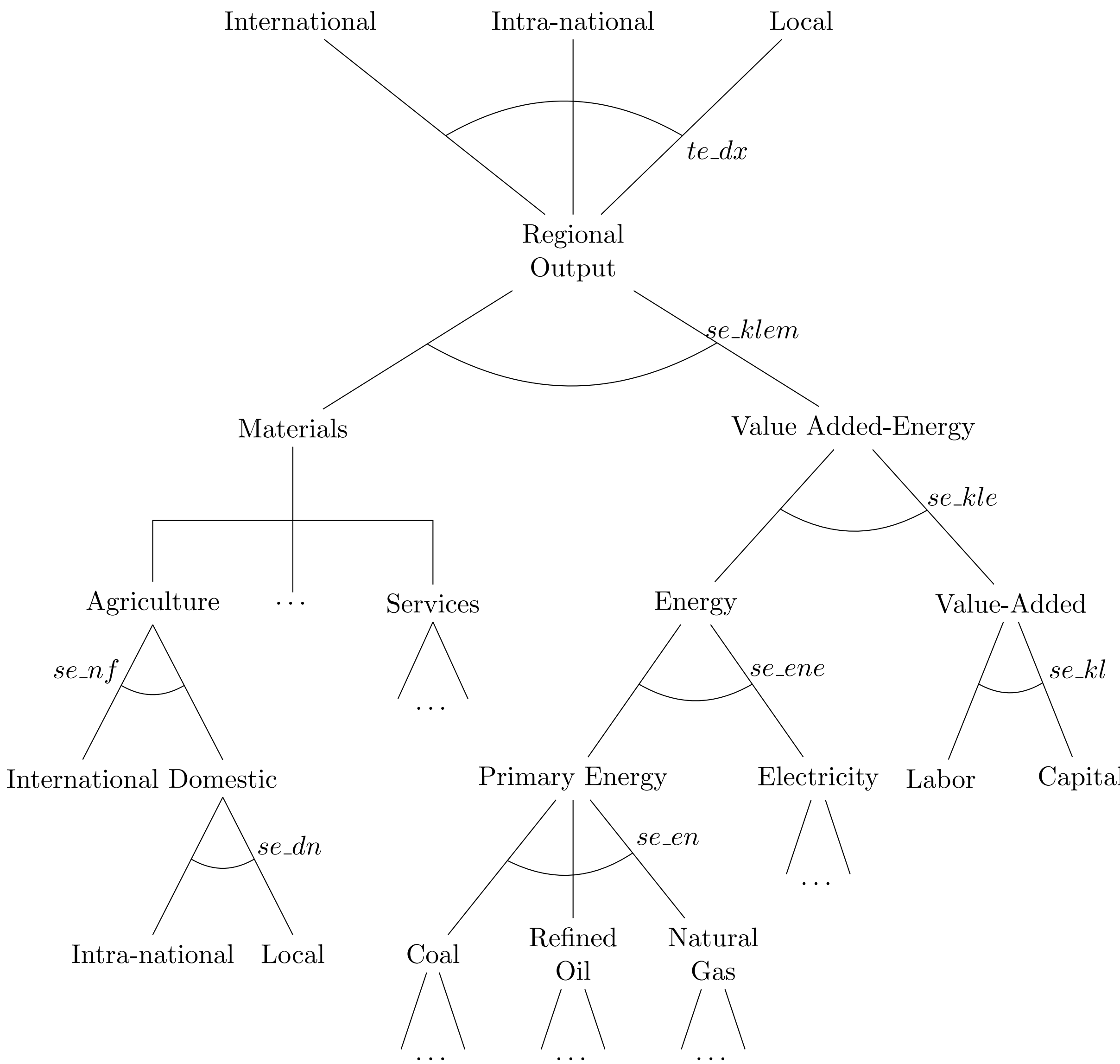


Figure 4: Manufacturing and Services Production Functions

$en_{t,r,s}$ is the primary energy composite, which is defined as

$$en_{t,r,s} = en0_{r,s}\left[cs\_en_{r,col,s}\left(\frac{id_{t,r,col,s}}{id0_{r,col,s}}\right)^{\frac{se\_en-1}{se\_en}} + cs\_en_{r,ref,s}\left(\frac{id_{t,r,ref,s}}{id0_{r,ref,s}}\right)^{\frac{se\_en-1}{se\_en}} + cs\_en_{r,gas,s}\left(\frac{id_{t,r,gas,s}}{id0_{r,gas,s}}\right)^{\frac{se\_en-1}{se\_en}}\right]^{\frac{se\_en}{se\_en-1}}, \tag{9}$$

where $\sum_{ss} cs\_en_{r,ss,s} = 1$. Finally, $kl_{t,r,s}$ is the value added composite, which is defined as

$$kl_{t,r,s} = kl0_{r,s}\left[cs\_kl_{r,s}\left(\frac{kd_{t,r,s}}{kd0_{r,s}}\right)^{\frac{se\_kl-1}{se\_kl}} + (1 - cs\_kl_{r,s})\left(\frac{ld_{t,r,s}}{ld0_{r,s}}\right)^{\frac{se\_kl-1}{se\_kl}}\right]^{\frac{se\_kl}{se\_kl-1}}, \tag{10}$$

where $kd_{t,r,s}$ is demand for new capital, and $ld_{t,r,s}$ is demand for labor. Recall that parameters with the prefix cs are the relevant cost shares in the benchmark year, and parameters with the prefix se are the relevant substitution elasticities.

Markets are assumed to be perfectly competitive, such that firms are price takers. Given market prices, firms seek to maximize profits

$$(1 - ty_{t,r,s})\, py_{t,r,s}y_{t,r,s} - \sum_{ss} pa_{t,r,ss}id_{t,r,ss,s} - (1 + tk_{t,r})\, pr_{t,r}kd_{t,r,s} - pl_{t,r}ld_{t,r,s}, \tag{11}$$

where $py_{t,r,s}$ is the output price based on maximizing returns across destination markets per equation (2), $pa_{t,r,s}$ is the price of the Armington composite, $pr_{t,r}$ is the rental rate for new capital, $pl_{t,r}$ is the wage rate, and $ty_{t,r,s}$, and $tk_{t,r}$ are ad valorem taxes on output and capital income, respectively.[8]

2.2.2 Resource Extraction, Agriculture, and Forestry Sectors

The resource extraction sectors (crude oil, natural gas, coal, and mining) have an additional primary factor input, in this case representing the finite natural resource. In many cases, models have included this resource in a top-level nest with a bundle of non-resource inputs (e.g., Ross (2005); Paltsev et al. (2005); Sue Wing (2006); Rausch et al. (2011); Capros et al. (2013); Ross (2014); Böhringer et al. (2018)). While some models allow for substitution between materials, energy, and value-added in resource extraction sectors (e.g., Sue Wing et al. (2011); Capros et al. (2013)), other models treat energy, labor, and capital as Leontief inputs (e.g., Ross (2014)), although in most cases there is some substitutability allowed between labor and capital (e.g., Ross (2005); Paltsev et al. (2005); Sue Wing (2006); Rausch et al. (2011)). Recent empirical evidence suggests non-zero and statistically significant substitution elasticities between inputs in resource extraction industries (Young (2013); Koesler and Schymura (2015)). Therefore, we maintain the same structure as in

[8] Payroll taxes are included as part of households' tax rate on labor income to capture the limit on Old Age and Survivor's Insurance payments, which causes the marginal ad valorem tax rate to differ across employees based on income.

the standard production nesting albeit with the addition of a fixed resource. The structure of the production functions for the fossil fuel extraction sectors is presented in Figure 5.

We model the agriculture and forestry sectors using a similar production function with land as a fixed factor input. We recognize that there has been an ongoing discussion in the literature related to the degree of flexibility required by a production function to capture the separability, or lack thereof, observed in empirical studies of agricultural sectors (e.g., Higgs and Powell (1990); Zahniser et al. (2012); Simola (2015)). However, the decreasing returns to scale nature of production in the sector, as captured in Figure 5, is common among approaches, independent of the nesting structure applied.

For the resource extraction, agriculture, and forestry sectors the specific form of the production function is

$$y_{t,r,s} = y0_{r,s}\left[cs\_rklem_{r,s}\left(\frac{res_{t,r,s}}{res0_{r,s}}\right)^{\frac{se\_rklem-1}{se\_rklem}} + (1-cs\_rklem_{r,s})\left(\frac{klem_{t,r,s}}{klem0_{r,s}}\right)^{\frac{se\_rklem-1}{se\_rklem}}\right]^{\frac{se\_rklem}{se\_rklem-1}}, \tag{12}$$

where

$$klem_{t,r,s} = klem0_{r,s}\left[cs\_klem_{r,s}\left(\frac{mat_{t,r,s}}{mat0_{r,s}}\right)^{\frac{se\_klem-1}{se\_klem}} + (1-cs\_klem_{r,s})\left(\frac{kle_{t,r,s}}{kle0_{r,s}}\right)^{\frac{se\_klem-1}{se\_klem}}\right]^{\frac{se\_klem}{se\_klem-1}}, \tag{13}$$

and $mat_{t,r,s}$ and $kle_{t,r,s}$ are defined in (6)-(10). The fixed factors, $res_{t,r,s}$, are sector specific and in the baseline fixed at the benchmark level, $res_{t,r,s} = res0_{r,s}\ \forall t$.

The resource extraction, agriculture, and forestry markets are also assumed to be perfectly competitive, such that firms are price takers. Given market prices, firms seek to maximize profits

$$\begin{aligned}(1-ty_{t,r,s})\,py_{t,r,s}y_{t,r,s} - \sum_{ss} pa_{t,r,ss}id_{t,r,ss,s} - (1+tk_{t,r})\,pr_{t,r}kd_{t,r,s} - pl_{t,r}ld_{t,r,s}\\ - (1+tk_{t,r})\,pres_{t,r,s}res_{t,r,s},\end{aligned} \tag{14}$$

where $pres_{t,r,s}$ is the price of the fixed factor resource. It is assumed that returns to the fixed factor face the same ad valorem tax rate as income from physical capital.

2.3 Partial Putty-Clay Capital

To better represent limitations associated with transitioning existing capital stock between sectors or changing its production process, the model considers two capital vintages: existing stock in the benchmark year and new capital formed after the benchmark year. Production with new capital has the flexibility described in Figure 4 and 5. Production with extant capital has a Leontief production structure, as shown in Figure 6.[9] For a profit maximizing firm this means that output

[9]Given the Leontief structure of the production function with extant capital, the nesting pictured in Figure 6 is unnecessary but is retained to make the figure more readable.

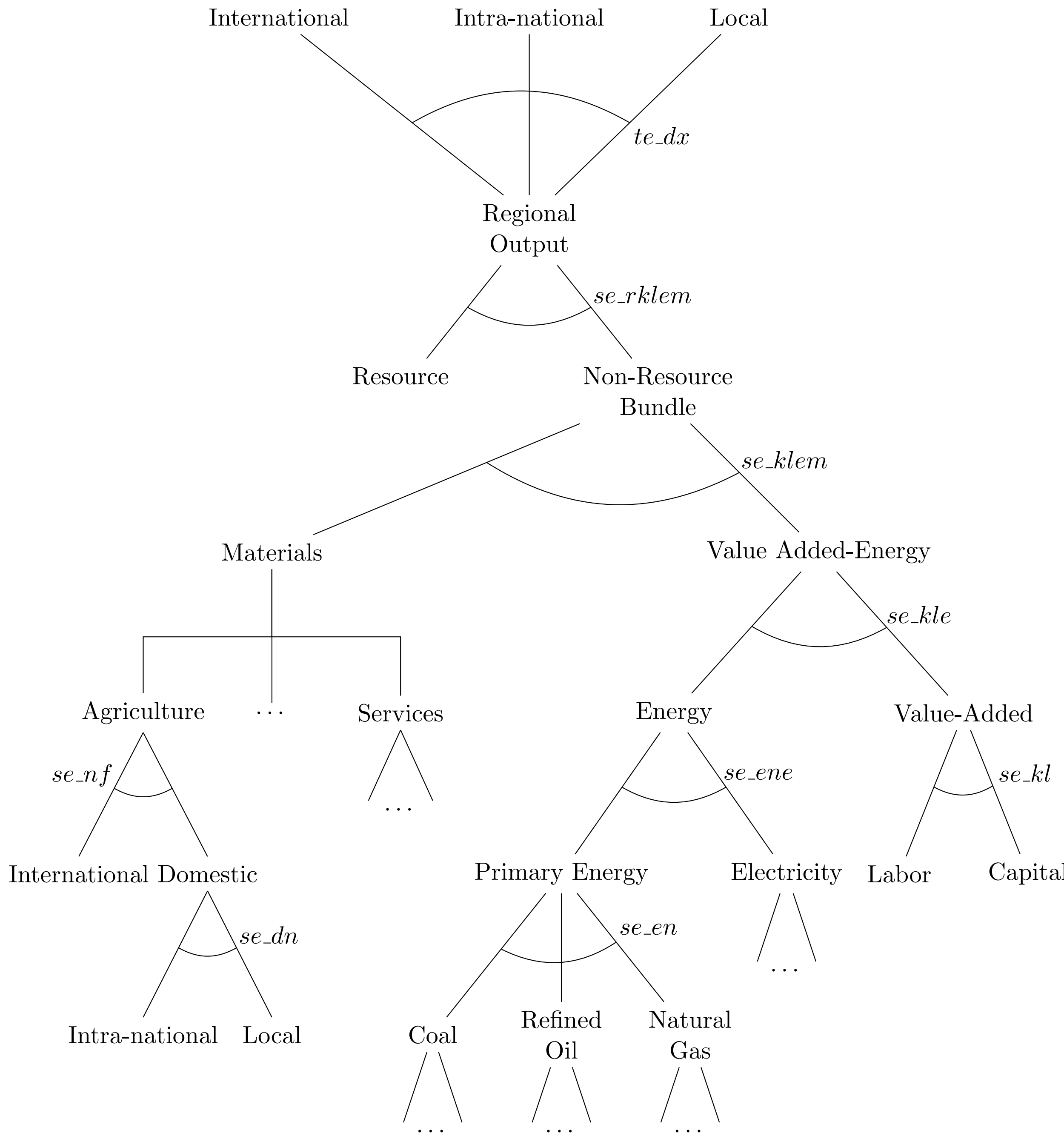


Figure 5: Resource Extraction, Agriculture, and Forestry Production Functions

of commodity s using extant capital is

$$y\_ex_{t,r,s} = y0_{r,s}\frac{kd\_ex_{t,r,s}}{kd0_{r,s}} \tag{15}$$

and demand for intermediate good ss, labor, and fixed factor resources to be used with extant capital will be

$$id\_ex_{t,r,ss,s} = id0_{r,ss,s}\frac{kd\_ex_{t,r,s}}{kd0_{r,s}}, \tag{16}$$

$$ld\_ex_{t,r,s} = ld0_{r,s}\frac{kd\_ex_{t,r,s}}{kd0_{r,s}}, \tag{17}$$

and

$$res\_ex_{t,r,s} = res0_{r,s}\frac{kd\_ex_{t,r,s}}{kd0_{r,s}}. \tag{18}$$

In our partial putty-clay specification, extant capital is primarily sector specific, although it allows a limited potential to shift extant capital across sectors at a cost. This feature is included to match observations that some extant capital (e.g., structures) can be transferred across sectors. To capture this characteristic, sector-specific extant capital, $kd\_ex_{t,r,s}$ is determined by a CET function that transforms a region's extant capital, $k\_ex_{t,r}$, with elasticity te_k_ex. More specifically, given the rental rates for sector-specific extant capital the returns to the stock of extant capital are maximized subject to the production possibilities frontier

$$k\_ex_{t,r} = k0_r\left[\sum_s cs\_kd\_ex_{r,s}\left(\frac{kd\_ex_{t,r,s}}{kd0_{r,s}}\right)^{\frac{te\_k\_ex+1}{te\_k\_ex}}\right]^{\frac{te\_k\_ex}{te\_k\_ex+1}}, \tag{19}$$

where $\sum_s cd\_kd\_ex_{r,s} = 1$.

Capital, regardless of vintage, is assumed to depreciate at the annual rate δ. The law of motion for extant capital reflects this ongoing depreciation, such that

$$k\_ex_{t+1,r} = (1-\delta)k\_ex_{t,r}, \tag{20}$$

where $k\_ex_{0,r} = k0_r$. The law of motion for the new capital stock reflects both depreciation and new regional investment, such that

$$k_{t+1,r} = (1-\delta)k_{t,r} + inv_{t,r}, \tag{21}$$

where $inv_{t,r}$ is investment in region r in year t and $k_{0,r} = 0$. The aggregate investment good is composed of commodity output not otherwise used in final demand, government expenditures or exported away. The formation of new physical capital via the aggregate investment good is assumed

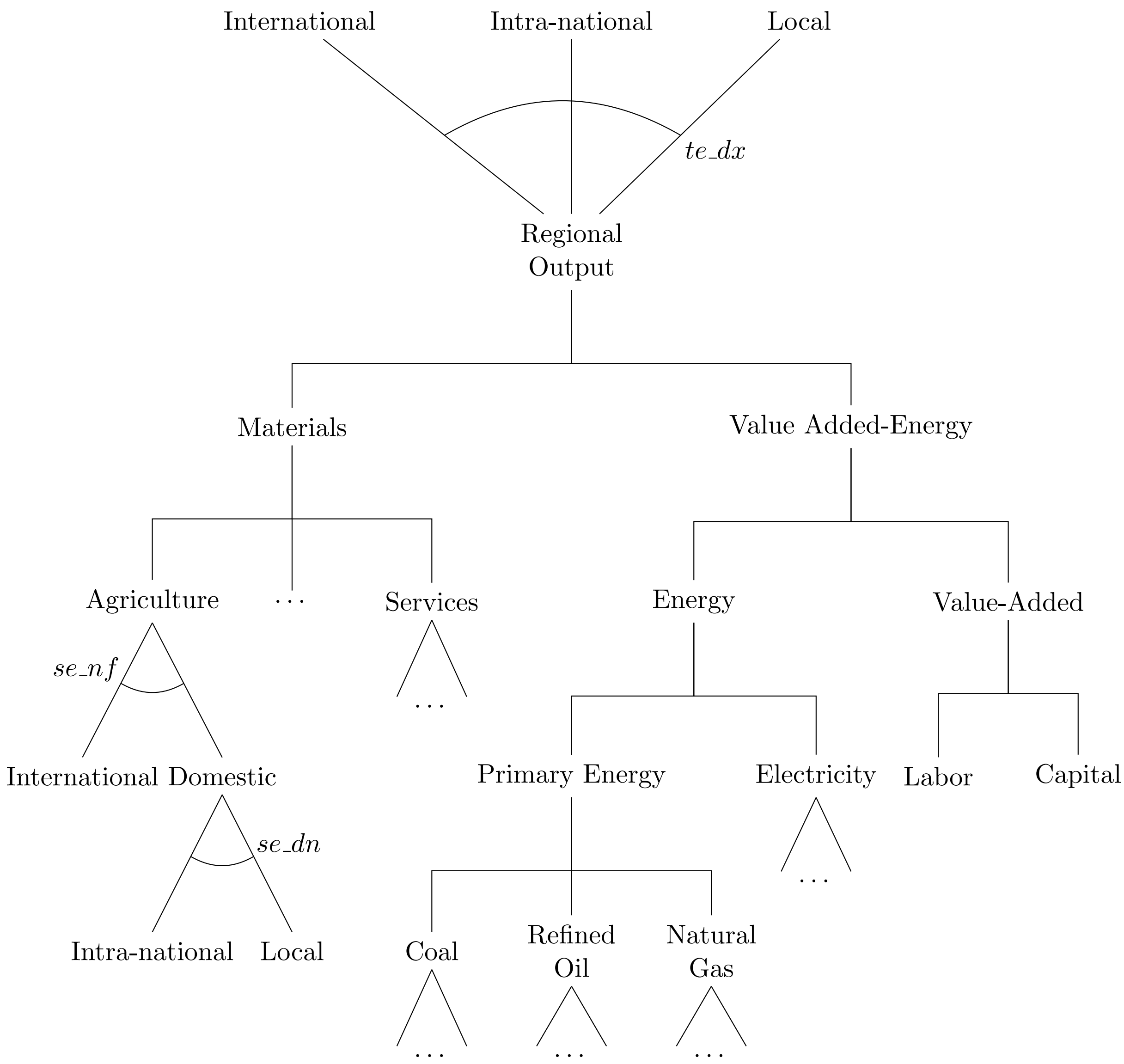


Figure 6: Manufacturing and Services Production Functions with Extant Capital

Table 3: SAGE Representative Households

Household	Benchmark Income
hh1	$\leq$ \$25,000
hh2	\$25,000-\$50,000
hh3	\$50,000-\$75,000
hh4	\$75,000-\$150,000
hh5	$\geq$ \$150,000

to operate through a CES function with an elasticity of substitution, *se_inv*, such that

$$inv_{t,r} = inv0_r \left[\sum_s cs\_inv_{r,s} \left(\frac{i_{t,r,s}}{i0_{r,s}} \right)^{\frac{se\_inv-1}{se\_inv}} \right]^{\frac{se\_inv}{se\_inv-1}}, \tag{22}$$

where $i_{t,r,s}$ is investment demand for commodity s, and $\sum_s cs\_inv_{r,s} = 1$ for each region.

2.4 Households

Each region has five representative households differentiated by benchmark income. Benchmark incomes for the representative households are presented in Table 3. Based on the underlying economic data in our social accounting matrix, these represent the closest approximation to national income quintiles possible.

Each representative household seeks to maximize intertemporal welfare, which is defined for household h in region r as

$$W_{r,h} = \sum_{t=0}^{\infty} \beta^t n_{t,r,h} u \left(\frac{cl_{t,r,h}}{n_{t,r,h}} \right), \tag{23}$$

where β is the discount factor, $n_{t,r,h}$ are the number of households represented by this agent, $cl_{t,r,h}$ is the consumption-leisure composite, and $u(\cdot)$ is the intra-temporal utility function. The discount factor is defined as

$$\beta = \frac{1}{1+\rho}, \tag{24}$$

where ρ is the pure rate of time preference.

The intra-temporal utility function is isoelastic, such that

$$u(cl_{t,r,h}) = \frac{cl_{t,r,h}^{1-\eta}}{1-\eta}, \tag{25}$$

where η represents the inverse of the intertemporal substitution elasticity of full consumption. Intra-temporal household preferences are defined by a nested CES-LES (linear expenditure system) utility function as presented in Figure 7. The top level CES representation of the leisure-consumption trade-off allows us to calibrate the elasticity of substitution to empirical labor supply elasticities. This nested CES-LES approach also assumes weak separability between leisure and consumption of other commodities (see Caron et al. (2017) for another example this type of utility function). Non-leisure consumption is assumed to follow a linear expenditure system (LES) demand structure. This LES structure provides for non-homothetic preferences that can be calibrated to empirical income elasticities while maintaining global regularity conditions that more flexible functional forms are often missing (Ho et al., 2020; Sands et al., 2017; Lofgren et al., 2002).[10] A linear expenditure demand system allows for non-homothetic preferences with linear Engel curves that do not pass through the origin when subsistence demands are positive (Stone, 1954). Engel curves describe the proportion of income spent on a particular commodity as income changes.

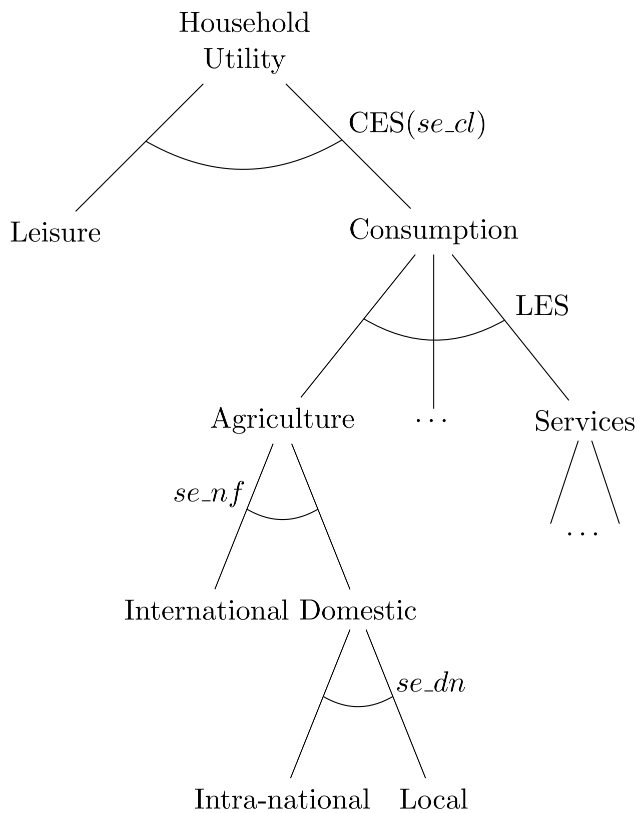


Figure 7: Household Consumption

More specifically, this framework disaggregates non-leisure commodity demands into discretionary ($dcd_{t,r,s,h}$) and subsistence demands ($scd_{t,r,s,h}$) such that total non-leisure commodity demand ($cd_{t,r,s,h}$) is defined as

$$cd_{t,r,s,h} = dcd_{t,r,s,h} + scd_{t,r,s,h}. \tag{26}$$

Each household in the model must achieve an exogenously specified minimum level of consumption, denoted here as subsistence spending, for each commodity demand to contribute to positive utility levels. To incorporate this demand sub-system into the model, calibrated subsistence demands

[10]Note that CES functional forms imply unitary income elasticities. This violates Engel's Law for food demand, which states that consumers increase their expenditures for food products as their income grows at a decreasing rate. Since household income is assumed to grow over time in the baseline, this assumption "imposes a baseline growth path at odds with historical experience" (SAB, 2020). The LES functional form relaxes this assumption.

($scd_{t,r,s,h}$) are held fixed and appear in the budget constraint but do not contribute to positive utility levels (and are hence not included in Figure 7). Once agents have achieved the minimum level of consumption, additional demand is endogenous in the model, denoted here as discretionary spending, as captured in Figure 7. Aggregate discretionary consumption ($dc_{t,r,h}$) is combined with leisure in the top-level nest of the utility function using a CES function. More information about the calibration of our demand system is presented in Section 3.3.6.

Intra-temporal household preferences over full discretionary consumption are defined as

$$cl_{t,r,h} = cl0_{r,h} \left[cs\_cl_{r,h} \left(\frac{dc_{t,r,h}}{dc0_{r,h}} \right)^{\frac{se\_cl-1}{se\_cl}} + (1 - cs\_cl_{r,h}) \left(\frac{leis_{t,r,h}}{leis0_{r,h}} \right)^{\frac{se\_cl-1}{se\_cl}} \right]^{\frac{se\_cl}{se\_cl-1}}, \tag{27}$$

where $leis_{t,r,h}$ is leisure and $dc_{t,r,h}$ is the discretionary final goods consumption composite, defined as

$$dc_{t,r,h} = dc0_{r,h} \prod_{s} \left(\frac{cd_{t,r,s,h} - scd_{t,r,s,h}}{dcd0_{r,s,h}} \right)^{cs\_les_{t,r,h,s}}, \tag{28}$$

where $cs\_les_{t,r,h,s}$ is the value share for discretionary spending, adjusted for subsistence demands. The subsistence portion of commodity demand, $scd_{t,r,s,h}$, is assumed to grow at an exogenous rate consistent with income growth so that the calibrated income elasticities are relatively stable over the model's time horizon (see Section 3.3.6 for more details).

Households seek to maximize welfare in (23) subject to a budget constraint where sources of income are represented on the right-hand side and expenditures on the left-hand side of the equation, net of taxes. Disaggregation of exogenous government and international accounts implies that the household budget constraint must also track these same categories of net transfers from the government, net transfers from the rest of the world, and payments to service government debt.

$$\begin{aligned} pkh_{t+1} invh_{t,r,h} + pcl_{t,r,h} cl_{t,r,h} + \sum_s ((1 + tc_{t,r}) pa_{t,r,s} sd_{t,r,s,h}) = \\ prh_t kh_{t,r,h} + prh\_ex_t kh\_ex_{t,r,h} + \sum_s pres_{t,r,s} rese_{t,r,s,h} \\ + (1 - tl_{t,r,h} - tfica_{t,r,h}) \, pl_{t,r} te_{t,r,h} + (tl_{t,r,h} - tl\_avg_{t,r,h}) \, pl_{t,r} l_{t,r,h} \\ + pfx_t \, (inc\_row_{t,r,h} + tran\_row_{t,r,h} + curactbal_{t,r,h}) \, int\_share_{r,h} \\ + cpi_t transfers_{t,r,h} + cpi_t \, (gint_{t,r,h} - deficit_{t,r,h} + icnadj_{t,r,h}) \, g\_share_{r,h} \end{aligned}, \tag{29}$$

In the household budget constraint, the left-hand side variables are defined as follows: pkh_t is the price of new capital stock for households, $invh_{t,r,h}$ is the level of investment in new capital stock in period t, $pcl_{t,r,h}$ is the unit cost of full consumption (i.e., discretionary consumption and leisure) inclusive of any consumption taxes, and $sd_{t,r,s,h}$ is subsistence demand or the minimum amount of demand required for positive utility levels.

The right-hand side of the budget constraint consists of several components. Earnings from new capital are calculated as the after tax rate of return on new capital, prh_t, multiplied by household holdings of new capital stock, $kh_{t,r,h}$. Households also earn income from investments in extant capital, which consists of the average national return on households' extant capital stock, prh_ex_t, multiplied by their stock of extant capital, $kh\_ex_{t,r,h}$.[11] prh_ex_t is further defined as

$$prh\_ex_t = \frac{\sum_r pr\_ex\_agg_{t,r} k\_ex_{t,r}}{\sum_r k\_ex_{t,r}}, \tag{30}$$

where $pr\_ex\_agg_{t,r}$ are the returns to extant capital in a region.

The model explicitly tracks investment returns by household and region in each time step. To do this, we characterize a second "intermediary" household specific capital stock that allows households to invest in capital formation in regions other than where they live. This capital stock behaves consistently with the law of motion for capital and evolves according to

$$kh_{t+1,r,h} = (1-\delta)kh_{t,r,h} + invh_{t,r,h}. \tag{31}$$

This capital stock reflects capital ownership by region, household and time period. It supplies a national market for capital that is cleared by firm capital demands. Essentially, households invest into a national investment fund that optimally (and without transaction costs) allocates those investments across regions. Since new investment is mobile across regions there is a single price for new capital investments by households and they earn a single rate of return that is inclusive of regionally depreciating assets. These will be equal to the average price and rate of return to new capital across regions.

Earnings from the ownership of fixed resources are captured in the household budget constraint as the price of the fixed resource used by sector s, $pres_{t,r,s}$, multiplied by the household's endowment of that fixed resource, $rese_{t,r,s,h}$.

To capture income from labor supply, $te_{t,r,h}$ is defined as the household's effective time endowment, $pl_{t,r}$ (i.e., the wage rate) is the price of labor, and $l_{t,r,h}$ is the household labor supply. Since households are assumed to "purchase" leisure at its opportunity cost (i.e., the wage rate), the household labor supply, $l_{t,r,h}$, will be determined according to the time endowment constraint

$$te_{t,r,h} = leis_{t,r,h} + l_{t,r,h}. \tag{32}$$

The population and the time endowment are assumed to grow at exogenous rates, discussed in further detail in Section 3.4.

Labor income must also be adjusted for taxes to be adequately represented in the budget constraint. Therefore, the value of a household's time endowment is $(1 - tl_{t,r,h} - tfica_{t,r,h})\, pl_{t,r} te_{t,r,h}$, which accounts for two taxes that are collected from households on labor: the personal income tax on labor earnings and Federal Insurance Contribution Act taxes, with ad valorem rates $tl_{t,r,h}$ and $tfica_{t,r,h}$, respectively. The FICA tax is collected in this manner, as opposed to a payroll tax on

[11] It is assumed that past investment was also mobile across space, such that households have historically invested nationally similar to the treatment of new capital. So the return households earn on their extant capital stock is a weighted average of regional rental rates for extant capital stock.

the firm side, to allow the effective FICA tax rate to incorporate the limit on the Old-Age, Survivors, and Disability Insurance tax. The additional component in the budget constraint related to labor taxes, $(tl_{t,r,h} - tl\_avg_{t,r,h})\, pl_{t,r} l_{t,r,h}$, adjusts the household's after tax income so that it only reflects the average tax rate collected on labor supply, while continuing to allow the use of effective marginal tax rates when modeling behavioral incentives. The adjustment represents the difference in the tax payment that would be collected by the government on labor income using the marginal tax rate for all labor income versus the average income tax rate. In the absence of this adjustment, the excess tax collected at the marginal rates would be returned through $incadj_{t,r,h}$ and indexed based on the consumer price index, which can affect welfare estimates and incidence in particular.

Households also receive net transfers from the rest of the world. These are calculated as the price of foreign exchange, pfx_t, multiplied by the sum of their share of net income from the rest of the world, $inc\_row_{t,r,h}$, their share of net taxes and transfers from the rest of the world, $tran\_row_{t,r,h}$, and their share of the current account balance, $curactbal_{t,r,h}$. This sum is then adjusted by the modeled incidence of those foreign transactions on households, $int\_share_{r,h}$.

Net government transfers to households are captured by $transfers_{t,r,h}$, which is assumed to be indexed to the consumer price index, cpi_t, as is the case with most federal government transfer programs. Payments to service government debt are captured in the budget constraint through the sum of the interest received on government debt, $gint_{t,r,h}$, the purchase of government debt, $deficit_{t,r,h}$, and the balance of other inter-institutional transfers, $incadj_{t,r,h}$. This sum is indexed by the cpi_t and adjusted by the modeled incidence of those government transactions on households, $g_share_{r,h}$.

The consumer price index, cpi_t, is defined as

$$cpi_t = \frac{\sum_{r,s,h} (1 + tc_{t,r})\, pa_{t,r,s} cd0_{r,s,h}}{\sum_{r,s,h} (1 + tc0r)\, cd0_{r,s,h}}, \tag{33}$$

and $cd_{t,r,s,h}$ is demand for commodity s. $int\_share_{r,h}$ and $g_share_{r,h}$ are defined as

$$int\_share = \frac{\sum_s m0_{r,s,ftrd} - x0_{r,s,ftrd}}{\sum_{r,s} m0_{r,s,ftrd} - x0_{r,s,ftrd}} \frac{c0_{r,h}}{\sum_h c0_{r,h}} \tag{34}$$

and

$$g\_share = \frac{c0_{r,h}}{\sum_{r,h} c0_{r,h}}, \tag{35}$$

such that the incidence is predominately assumed to follow the incidence of consumption.

2.5 Government, Taxes and the Rest of World

There is a single national government in the model that imposes ad valorem taxes on capital income, production, wage income, and consumption, $tk_{t,r}$, $ty_{t,r,s}$, $tl_{t,r,h}$, and $tc_{t,r}$, respectively. The taxes are region specific, the production tax is also sector specific, and the labor income tax rates are household specific. While they remain constant over time in the baseline, we allow for the possibility of future changes in tax rates in the policy simulations.

Government purchases in region r are assumed to be Leontief, such that

$$gov_{t,r} = gov0_r \min_s \left(\frac{g_{t,r,s}}{g0_{r,s}} \right), \tag{36}$$

where $g_{t,r,s}$ is public demand for commodity s in region r, and $gov_{t,r}$ is the composite public consumption good.

The government's budget constraint is

$$\begin{aligned}\sum_r pgov_{t,r} gov_{t,r} + \quad & \sum_h cpi_t transferst, r, h + cpi_t gint_t + pfx_t gint\_row_t + cpi_t incadj_t \\ = \sum_r \sum_s \Bigg\{ \quad & ty_{t,r,s} py_{t,r,s} \left(y_{t,r,s} + y\_ex_{t,r,s} \right) \\ & + tk_{t,r} \left[pr_{t,r} kd_{t,r,s} + pr\_ex_{t,r,s} kd\_ex_{t,r,s} + pres_{t,r,s} \left(res_{t,r,s} + res\_ex_{t,r,s} \right) \right] \Bigg\} \\ + \sum_r \sum_h & \left[\left(tl_{t,r,h} + tfica_{t,r,h} \right) pl_{t,r} l_{t,r,h} + tc_{t,r} pa_{t,r,s} cd_{t,r,s,h} \right] + cpi_t deficit_{t,r,h},\end{aligned} \tag{37}$$

where $pgov_{t,r}$ is the unit cost of government consumption based on (36).

The government's budget is balanced through lump sum transfers, $incadj_t$, which are shared out to households based on their share of national consumption in the benchmark dataset, as discussed in Section 2.4. Exogenous government and international accounts have been disaggregated to allow for differentiated assumptions about their growth and provide greater transparency over the assumptions within the model. The specific level of disaggregation selected for these accounts also allows the exogenous variables in SAGE to line up well with other government forecasts to which these variables are calibrated (see Section 3.4 for more details).

A rest-of-the-world agent is characterized as being endowed with fixed factors for export demands and import supplies that demands foreign exchange. The agent's budget constraint is specified as

$$pfx_t row_t = \sum_s \left(pfix_{t,s} fix0_s + pfim_{t,s} fim0_s \right) \tag{38}$$

where $pfx_t row_t$ denotes the income level of the rest-of-the-world agent, such that row_t is the rest-of-the-world demand for foreign exchange in each period.

2.6 Market Clearance

Given firm, household, and government behavior, along with the capital dynamics described in the preceding sections, prices in equilibrium are assumed to clear all markets.

The price of the Armington aggregate, $pa_{t,r,s}$, is a composite price index derived from three prices: the price for domestic output consumed regionally, $pd_{t,r,s}$, the price of commodities imported from the national market, $pn_{t,r,s}$, and the price for imported commodities from the foreign market, $pm_{t,r,s}$. $pa_{t,r,s}$ clears the intermediary market for the aggregate demand of commodities, such that

$$a_{t,r,s} = \sum_{ss}(id_{t,r,s,ss} + id\_ex_{t,r,s,ss}) + \sum_{h}(cd_{t,r,s,h}) + i_{t,r,s} + g_{t,r,s}. \tag{39}$$

The price of domestic output consumed domestically, $pd_{t,r,s}$, clears the domestic market, such that

$$\frac{y\_ex_{t,r,s} + y_{t,r,s}}{y0_{r,s}} \left(\frac{pd_{t,r,s}}{py_{t,r,s}}\right)^{te\_dx} = \frac{d_{t,r,s}}{d0_{r,s}}, \tag{40}$$

where the left hand side defines the optimal share of output supplied to the domestic market based on the output transformation function in (2). The price of labor, $pl_{t,r}$, clears the labor market, such that

$$\sum_{h} l_{t,r,h} = \sum_{s} ld_{t,r,s} + ld\_ex_{t,r,s}. \tag{41}$$

The rental rate for sector specific extant capital, $pr\_ex_{t,r,s}$, clears the market for extant capital, such that

$$\frac{k\_ex_{t,r}}{k0_r} \left(\frac{pr\_ex_{t,r,s}}{pr\_ex\_agg_{t,r}}\right)^{te\_k\_ex} = \frac{kd\_ex_{t,r,s}}{kd0_{r,s}}, \tag{42}$$

where the left hand side defines the optimal share of extant capital supplied to sector s based on the extant transformation function in (19). The rental rate for new capital, $pr_{t,r}$, clears the regional market for new capital, such that

$$k_{t,r} = \sum_{s} kd_{t,r,s}. \tag{43}$$

The price of new capital, $pk_{t,r}$, clears the regional investment market, such that

$$k_{t-1,r}\,(1-\delta) + inv_{t-1,r} = k_{t,r}. \tag{44}$$

The rental rate households' earn on their holdings, prh_t, clears the national investment market, such that

$$\sum_{r,h} kh_{t,r,h} = \sum_{r} k_{t,r}. \tag{45}$$

The price of commodities on the national market, $pn_{t,s}$, clears the market for national trade, such that

$$\sum_{r} x_{t,r,s,dtrd} = \sum_{r} m_{t,r,s,dtrd}. \tag{46}$$

The price of imports, $pm_{t,s}$, clears the market for foreign import market, such that

$$ms_{t,s} = \sum_{r} m_{t,r,s}. \tag{47}$$

The price of the import fixed factor, $pfim_{t,s}$, clears the market for foreign fixed import factors, such that[12]

$$fim0_s = fim_{t,s}. \tag{48}$$

The price of exports, $px_{t,s}$, clears the market for foreign export market, such that

$$\sum_r x_{t,r,s} = xd_{t,s}. \tag{49}$$

The price of the export fixed factor, $pfix_{t,s}$, clears the market for foreign fixed export factors, such that

$$fix0_s = fix_{t,s}. \tag{50}$$

The price of foreign exchange, pfx_t, clears the foreign exchange market, such that

$$\sum_s xd\_fx_{t,s} + inc\_row_t + tran\_row_t - gint\_row_t - curactbal_t = \sum_s ms\_fx_{t,s} + row_t. \tag{51}$$

Finally, the rental rate for sector-specific fixed factors, $pres_{t,r,s}$, clears the market for sector-specific fixed factors, such that

$$\sum_h rese_{t,r,s,h} = res_{t,r,s} + res\_ex_{t,r,s}. \tag{52}$$

Given that the CES and CET functions defining much of the model's structure are homothetic, the prices for composite goods (e.g., $py_{t,r,s}$ and $pcl_{t,r,h}$) are defined by their unit cost.

2.7 Closures

This section summarizes the main model closures, which are needed to ensure the model is well specified and that there are enough equations to solve for the endogenous variables in the model. These include the government account, trade accounts, intertemporal no-arbitrage condition, and the terminal condition for the finite time horizon model. While some of these are presented above, they are repeated here to provide a complete accounting in one section.

The government budget constraint in (37) is balanced through lump-sum transfers with households, where the endogenous transfers are distributed according to shares of benchmark consumption per equation (35). The government budget is balanced via lump-sum transfers to avoid altering the marginal incentives in the model through the speculative choice of which tax(es) to adjust.

The domestic trade account is closed each period with a single national price per commodity, per the market clearance condition in (46). Each region's overall balance of payments (across all commodities) in the domestic trade market is not required to be zero in a given period. Deviations from zero are the result of net investment flows in or out of the region in combination with differences between the region's tax payments and receipt of public funds.

The foreign trade account, across all commodities, is closed each period by the price of foreign

[12] Note that in the model, we let the fixed factors for imports and exports grow with the implicit growth rate. $fim0_s$ and $fix0_s$ do not include a t subscript here for exposition purposes.

exchange, per the market clearance condition in (51). The balance of payments is exogenously specified.

To ensure that the model does not allow for intertemporal arbitrage opportunities, the following constraint is placed on the price of capital

$$pk_{t,r} = pr_{t,r} + pk_{t+1,r}(1 - \delta). \tag{53}$$

In other words, the price of capital in period t must equal the return it receives in period t plus the present value of the depreciated asset in period $t + 1$. This is equivalent to an equilibrium price of capital that is equal to the present value of returns it will earn over its lifetime. A similar constraint is placed on the price of new capital stock held by households

$$pkh_t = prh_t + pkh_{t+1}(1 - \delta). \tag{54}$$

To close the finite approximation to the infinite time problem we generally follow Lau et al. (2002). The capital stock in the post-terminal period, kt_r, is introduced as an endogenous variable with associated price, pkt_r. The post-terminal capital stock is determined by requiring that regional investment is growing at the rate of aggregate regional output growth, such that

$$\frac{inv_{T,r}}{inv_{T-1,r}} = \frac{\sum_s y_{T,r,s} + y\_ex_{T,r,s}}{\sum_s y_{T-1,r,s} + y\_ex_{T-1,r,s}}, \tag{55}$$

where T is the terminal period. The price of terminal capital stock is determined by requiring the law of motion for capital to hold, such that

$$k_{T,r}\,(1 - \delta) + inv_{T,r} = kt_r, \tag{56}$$

where households' share of the post-terminal capital stock is assumed to be equivalent to their share of the capital stock in the last period of the model, such that

$$kht_{r,h} = \frac{kh_{T,r,h}}{\sum_{r,h} kh_{T,r,h}}, \tag{57}$$

where the price of terminal capital for the households is equal to the weighted average price of terminal capital

$$pkht = \frac{\sum_r pkt_r kt_r}{kt_r}. \tag{58}$$

3 Calibration and Data

There are multiple sets of data and parameters that define the calibration of the model. The benchmark social accounting matrix; the substitution and transformation elasticities in the model's production and utility functions; parameters defining the transformation and depreciation of capital stocks; tax rates; and the parameters defining the baseline projection. This section describes the sources of each of these in turn.

3.1 Benchmark Data

The benchmark data is based on IMPLAN's 2016 database of the U.S. economy aggregated up to the sectors in Table 1 for each of the regions in Figure 2, representative households in Table 3, and a single government.[13] The data are used to define the benchmark year values and cost shares. In addition, the benchmark dataset uses information from BEA's National Income and Product Account tables for the government deficit, $deficit0$, government interest payments to domestic agents, $gint0$, government interest payments to the rest of the world, $gint\_row0$, the current account balance, $curactbal0$, net income from the rest of the world, inc_row0, and net taxes and transfers from the rest of the world, $tran_r ow0$.

The remainder of this section describes additional transformations and modifications made to the database to conform to the structure of the model. Smaller transformations include:

- Household exports, which are primarily purchases by foreign tourists, are shared out across commodities based on final good consumption shares and transferred from households to sector-specific foreign exports.
- Government production (make and use) is integrated with private sector production.
- Investment demand, $i0_{r,s}$, is determined as the residual that would lead the goods market clearance condition in (39) to hold.

3.1.1 Crude Oil and Natural Gas Extraction Disaggregation

The underlying IMPLAN data does not distinguish between crude oil and natural gas extraction. Therefore, we disaggregate the single IMPLAN oil and gas extraction sector into separate natural gas extraction and crude oil extraction sectors. To determine the natural gas share of consumption/use we assume that crude oil serves as an intermediate input only to the petroleum refining sector and that natural gas is the only intermediate input (between the two) to all other sectors. We make the same assumptions for household and government consumption and investment demand. In the IMPLAN data, some of the intermediate inputs to the petroleum refining sector are natural gas. To determine that share and the natural gas share of production and trade we minimize the sum of squared deviations for those shares from observed values or assumed shares conditional on market clearance conditions and the assumption of weakly positive domestic use of production. The observed or assumed shares we try to match are derived as follows:

1. The observed share of natural gas production by region is defined using EIA data on crude oil and natural gas production by state aggregated up to the regional level. To arrive at a

[13]IMPLAN Group, LLC, 16740 Birkdale Commons Parkway, Suite 206, Huntersville, NC 28078, www.IMPLAN.com

value share we multiply state-level production quantities by EIA data on state-level wellhead prices for crude oil and city gate natural gas prices as a proxy for natural gas wellhead prices (which are not available).

2. The shares of natural gas international imports and exports by region are defined using census data on state-level international imports and exports of crude oil and natural gas aggregated to the regional level.

3. A region's intra-national import share of natural gas is assumed to be similar to the region's share of natural gas use relative to the region's total crude oil and natural gas use. A region's intra-national export share of natural gas is assumed to be similar to the share of natural gas production in the region.

4. The observed share of natural gas used as an intermediate input in the refining sector is estimated based on national annual averages of crude oil and natural gas inputs to the sector collected by EIA and converted to values using the Brent and Henry Hub average annual prices as reported by EIA.

3.1.2 Filtering and Balancing Benchmark

To improve the computational performance of the model we, follow what is standard practice for many CGE models and, filter out small values and rebalance the SAM. We remove any value less than 5×10^{-5} and any intermediate input whose cost share is less than 5×10^{-5}. This translates into filtering values of less than \$50,000 or less than .005% of the total costs of production, respectively. The relatively small size of such values can impeed computational performance, while there removal does not meaningfully affect the results of the model.

After filtering small values the SAM is rebalanced by minimizing the squared percent deviation from the original values weighted by the original values. Specifically we solve for new values of intermediate input demand, $id0_{r,ss,s}$, labor demand, $ld0_{r,s}$, capital demand, $kd0_{r,s}$, imports, $m0_{r,s,trd}$, exports, $x0_{r,s,trd}$, household consumption, $cd0_{r,s,h}$, government spending, $g0_{r,s}$, investment, $i0_{r,s}$, capital endowment, labor endowment, household savings, and lump sum government transfers, $tran0_{r,h}$. This optimization is subject to the market clearance conditions in (39), (41), (43), and (46), the budget constraints in (29) and (37), the balance of payment sharing in (34), the zero profit condition

$$(1 - ty_{r,s})y0_{r,s} = \sum_{s} id0_{r,ss,s} + ld0_{r,s} + (1 + tk_r)\, kd0_{r,s}, \tag{59}$$

the requirement that regional investment equals household savings

$$\sum_{s} i0_{r,s} = \sum_{h} kh0_{t+1,r,h} - kh0_{t,r,h}, \tag{60}$$

consistent with the original data set, weakly positive domestic own use

$$y0_{r,s} > \sum_{trd} x0_{r,s,trd}, \tag{61}$$

and where household savings is consistent with steady-state growth. The balancing occurs prior to distinguishing between types of capital - new, extant - and fixed factor resources, as covered in the next section. Therefore, the notation is somewhat simpler.

3.1.3 Natural Resources

Capital returns in the benchmark SAM are disaggregated into returns on man-made capital and natural resources. The disaggregation is based on estimates of the returns to natural resources as a share of gross surplus for those sectors. As described in greater detail below, these shares are assumed to be approximately 25% for the oil and natural gas extraction sectors, 40% for the coal mining sector, 40% for the agricultural and forestry sectors, and 40% for other mining sectors.

Through 2009, the U.S. Energy Information Administration (EIA) collected information on the performance of major U.S. energy-producing companies. Based on the most recent survey, they estimated that the total upstream costs (lifting costs plus finding costs) between 2007 and 2009 for crude oil and natural gas companies included in the survey was $33.76 per barrel of oil equivalent (EIA, 2011). EIA reports that the U.S. produced 1.95 billion barrels of crude oil[14] and 3.67 billion barrels of oil equivalent of natural gas[15] in 2009.[16] The U.S. Bureau of Economic Analysis (BEA) estimates that in 2009 the output value for the oil and natural gas extraction sectors was $220 billion with gross operating expenditures of $123 billion.[17] Combined, these estimates suggest that the output value of the sector exceeded the upstream costs by $30 billion, which is 25% of the gross operating surplus.

An alternative approach to defining the share of gross operating surplus due to rents paid to natural resource ownership is to consider royalty payments. The United States has widespread private ownership of minerals, including crude oil and natural gas. In 2012 an estimated 77% of onshore crude oil and natural gas production revenue was associated with privately owned minerals for which $22 billion in private royalties were paid (Fitzgerald and Rucker, 2016). In 2012, $8.5 billion in federal royalty payments were collected from onshore and offshore oil and gas production, according to the U.S. Department of Interior's Natural Resources Revenue Data.[18] The BEA estimates that in 2012 value added for the crude oil and natural gas extraction sectors, less employee compensation and production taxes, was $157 billion.[19] Private and federal royalties represented approximately 19% of this remaining value added. Brown et al. (2016) found evidence that private

[14] `https://www.eia.gov/dnav/pet/PET_CRD_CRPDN_ADC_MBBLPD_A.htm`

[15] `https://www.eia.gov/dnav/ng/hist/n9070us2A.htm`

[16] Natural gas was converted to equivalent barrels of oil at 0.178 barrels per thousand cubic feet following EIA (2011).

[17] `https://www.bea.gov/industry/input-output-accounts-data`

[18] `https://revenuedata.doi.gov/explore/#federal-revenue`

[19] `https://www.bea.gov/industry/input-output-accounts-data`

royalty rates may not represent full rent on the natural resource, potentially due to monopsony power and long-term contracts. Similarly, government royalty rates may not represent the full rent associated with the nonrenewable resource. Therefore, 19% likely represents a lower bound on the rents associated with crude oil and natural gas resources.

Sue Wing (2001), based on BEA rent estimates from 1994[20] and before the growth in shale production, estimated resource rents to be approximately 45% for crude oil and natural gas production. Technological progress such as horizontal drilling and hydraulic fracturing likely placed downward pressure on the resource rents (e.g., Farzin (1992); Lin and Wagner (2007)). Given the breadth of technical progress in these markets, 45% therefore may be a reasonable upper bound.

The share of gross surplus associated with coal resources is approximated using average extraction cost estimates from Jordan et al. (2018) along with additional information on operation costs for coal companies. Jordan et al. (2018) estimate average per-ton extraction costs for coal by region based on 10-K filings from large publically traded coal companies (Figure 1 from their paper). Based on there estimates of extraction costs and regional coal production levels, the extraction costs for the industry in 2012 were approximately \$37 billion. This value does not include consumption of fixed capital, sales, or general administrative costs. Using the 10-K fillings for the same publicly traded coal companies evaluated in Jordan et al. (2018), these additional costs were on average 20% of extraction operating costs in 2012. BEA estimates that in 2012 the total output value for the coal mining sector was \$52 billion with \$19 billion in gross operating surplus.[21] Estimating the payments to the resource as the difference between the total output value and extraction costs scaled to include other costs yields \$8 billion, which is approximately 40% of gross surplus for the sector. Notably, extraction operating costs may include some royalty payments, which may represent returns to the resource, leading to an underestimated share of gross surplus associated with resource payments. Conversely, the estimates of variable input costs do not include expenditures associated with mine closures, which can be large, leading to an overestimate of the share of gross surplus associated with resource payments. We note that, while based on data from the early 1990s, Sue Wing (2001) similarly estimated resource payments to be 40% of gross surplus in the coal sector.

Remaining mineral and metal mining activity is aggregated into another mining sector (*min*). Approximately two-thirds of the output value from the sector is attributable to stone mining and quarrying (NAICS 21231) or sand and gravel mining (NAICS 21232). Of the remaining third of the sector's output value, copper ore mining (NAICS 212234) accounts for approximately half. Due to a lack of recent data that would facilitate an exercise similar to those conducted for the other mining sectors we assume the share of gross surplus attributable to the resource is 40%, following the coal sector.

Disaggregating the returns to agricultural and forestry land as 40% of gross operating surplus is consistent with rental data from the U.S. Department of Agriculture (USDA). The USDA estimates

[20] `https://apps.bea.gov/scb/account_articles/national/0494od2/maintext.htm`
[21] `https://www.bea.gov/industry/input-output-accounts-data`

that the rental value in 2016 for cropland and pastureland is $136 and $13 per acre, respectively,[22] and that there is approximately 245 million acres of cropland[23] and 528 million acres of pastureland.[24] The BEA estimates gross surplus in 2016 for the agricultural sectors to be $103 billion. Using the USDA estimates to compute the total rent paid to agricultural land and dividing by the BEA gross surplus estimate, suggests that land rental values are up to 40% of gross surplus for the sector. It is worth noting that the USDA rent per acre estimates may include the returns to some structures, potentially making them an overestimate of the returns to land. Since the agriculture and forestry sectors are combined in the default aggregation of SAGE this assumption is also implicitly applied to the returns to land for the forestry sector, which accounts for less than 9% of the gross surplus for the combined sector.

3.2 Taxes

As previously noted, the model explicitly includes business taxes/subsidies, $ty_{r,s}$, personal labor income taxes, $tl_{r,h}$, and capital income taxes, tk_r. The taxes are introduced into the dataset prior to aggregation to the model's regions. When aggregating the dataset, taxes are set to keep the tax revenue constant between the disaggregated and aggregated datasets. Production taxes net of any subsidies, $ty_{r,s}$, are based on the average rate observed in the IMPLAN database. The production tax rates are presented in Table 4. Based on the design of the IMPLAN database these values also include sales and excise taxes. A placeholder exists for consumption taxes, tc_r, in the model's code to allow for future development work that may move the sales and excise taxes out of production taxes. In the current version of the model, explicit consumption taxes are set to zero. Therefore, as it currently stands, sales and excise taxes are applied on the supply side of the market as opposed to the demand side and are associated with the sector that submits the tax payment and not necessarily the sector that produces the taxed commodity.

Personal income taxes on labor are differentiated across regions and households. Effective marginal Federal Insurance Contribution Act (FICA) taxes are also differentiated across regions and households. This allows the payroll tax rates to capture the annual limit on Old Age and Survivor's Insurance (OASI) taxes, which would not be possible if the payroll taxes were collected on the firm side due to the model's structure. Data from the U.S. Census Bureau's Current Population Survey (CPS) Annual Social and Economic Supplement (ASEC) is used to create a representative sample of tax returns. These sample returns are then run through NBER's Taxsim model version 27 to estimate marginal tax rates for wage income and FICA for each sample return (Feenberg and Coutts, 1993).[25] For each region and household we compute the weighted average effective marginal tax rate from the sample returns by weighting the Taxsim results by the CPS

[22] https://quickstats.nass.usda.gov/results/ABF12C63-5DDA-3745-A0B3-C91279A860D1

[23] https://www.fsa.usda.gov/Assets/USDA-FSA-Public/usdafiles/NewsRoom/eFOIA/crop-acre-data/zips/2016-crop-acre-data/2016_fsa_acres_data_aug2016_dr6.zip

[24] https://www.nrcs.usda.gov/wps/portal/nrcs/detail/national/landuse/rangepasture/?cid=nrcsdev11_001074

[25] http://users.nber.org/~taxsim/taxsim27/

Table 4: Tax/Subsidy Rates on Production

	nor	sou	mid	wes
agf	0.02	0.01	0.01	0.00
cru	0.05	0.12	0.13	0.17
col	0.03	0.07	0.09	0.10
min	0.02	0.03	0.03	0.03
ele	0.09	0.08	0.08	0.08
gas	0.05	0.09	0.06	0.12
wsu	-0.09	-0.01	-0.04	-0.03
con	0.01	0.01	0.01	0.01
fbm	0.04	0.04	0.01	0.03
wpm	0.01	0.01	0.01	0.01
ref	0.01	0.01	0.01	0.01
chm	0.02	0.01	0.01	0.02
prm	0.01	0.01	0.01	0.01
cem	0.01	0.01	0.01	0.01
pmm	0.01	0.01	0.01	0.01
fmm	0.01	0.01	0.01	0.01
cpu	0.01	0.01	0.01	0.01
tem	0.01	0.00	0.00	0.00
bom	0.01	0.02	0.01	0.02
trn	0.03	0.03	0.04	0.04
ttn	0.01	0.01	0.01	0.01
srv	0.05	0.05	0.05	0.05
hlt	0.01	0.01	0.01	0.01

Table 5: CPS to NBER Taxsim Income Mapping

Taxsim Variable	Description	CPS Variable(s)*
pwage	Wage and salary income of primary taxpayer	ws_val, semp_val, frse_val, -hiemp×phip_val
swage	Wage and salary income of spouse	ws_val, semp_val, frse_val, -hiemp×phip_val
dividends	Qualified dividend income	qual_frac×div_val
stcg	Short term capital gains or losses	NA**
ltcg	Long term capital gains or losses	NA**
otherprop	Other property income	rnt_val, (1 − qual_frac) ×div_val
nonprop	Other non-property income	oi_val, ed_val
pensions	Taxable pensions and IRA distributions	rtm_val
gssi	Gross social security benefits	ss_val, ssi_val, srvs_val, dsab_val
ui	Unemployment compensation	uc_val
transfers	Other non-taxable transfer income	paw_val, wc_val, vet_val, csp_val, fin_val

* Except for the primary and spouse wage and salary income, for married taxpayers each Taxsim variable is the sum of the CPS variables for both the primary taxpayer and their spouse.
** The CPS ASEC does not include information on imputed capital gains after 2010.

ASEC earned income and applying the supplement weights.

From the CPS, the filing status variable (filestat) and the dependent status variable (dep_stat) are used to distinguish between single/head of household taxpayers and dependent taxpayers. All married taxpayers are assumed to file jointly, and the person records for each couple are identified using the a_spouse variable. The dep_row variable in the CPS is used to assign non-filing dependents to taxpayers, along with the ages of the dependents. This information is used to populate the Taxsim variables used to assess personal exemptions, the Dependent Care Credit, the Child Credit, and the Earned Income Tax Credit.

The income variables in the CPS ASEC are mapped to the Taxsim variables as described in Table 5. For married couples, all income values entered into Taxsim are joint earnings, except in the case of wage and salary income, which are kept separate. In both cases the employee-paid portion of employer-provided health insurance plans are subtracted from wages and saleries. Dividend income reported in the CPS is split between qualified and ordinary dividends based on the aggregate share of dividends that are qualified, $qual\_frac$, from the IRS individual income tax returns line item totals.[26] The CPS no longer includes imputed capital gains; therefore they are omitted from the submission to Taxsim. This limitation may bias the weighted average effective marginal tax rates downwards for the household representing the top income quintile (where nearly all capital gains accrue) if the inclusion would cause some households to be in a higher tax bracket.

The implicit deductions for each filer are computed as the difference between adjusted gross

[26] `https://www.irs.gov/pub/irs-soi/16inlinecount.pdf`

income (agi) and taxable income (tax_inc) as reported in the CPS minus personal exemption deductions accounting for the phase out. From this value, we subtract property and state taxes. We submit either this value or zero, whichever is higher, to Taxsim as potential sources of itemized deductions. Property taxes in the CPS ASEC (prop_tax) are associated with household records so we divide those taxes equally amongst all tax filing units in a household.

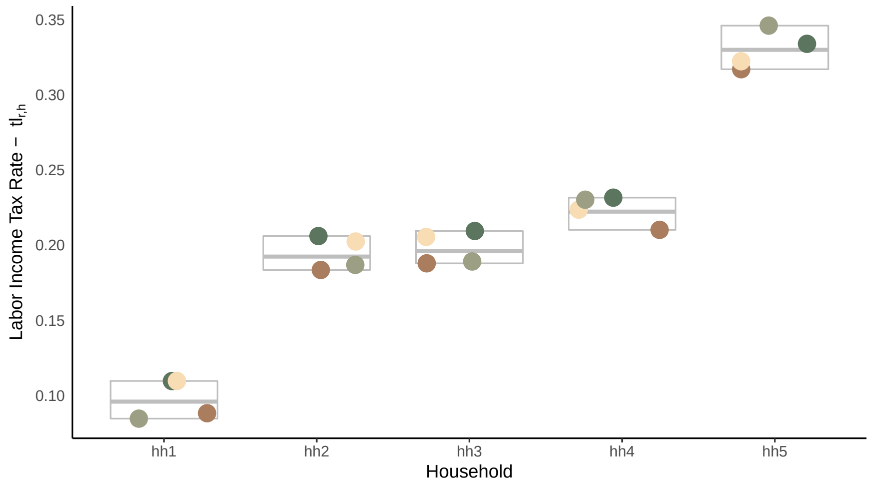
For each representative filer, Taxsim returns the effective marginal tax rate for primary earner wage income. Using the CPS ASEC person weights and primary earner wages, a weighted average of the effective marginal tax rates for wage income are computed for each region and representative household in the model. Primary earner wages are used as the weight because each married couple has two returns in our sample that are the same except for switching the primary and secondary earner. Since the entire FICA tax is collected on the household side in the model the labor income and FICA tax rates are adjusted to account for the fact that employees do not pay income or FICA taxes on employer paid shares of the FICA tax.

The personal labor income tax rates by region and household are presented in Figure 8a, and the FICA tax rates are presented in 8b.[27] The crossbars represent the national income-weighted, average effective marginal tax rate for the representative household.

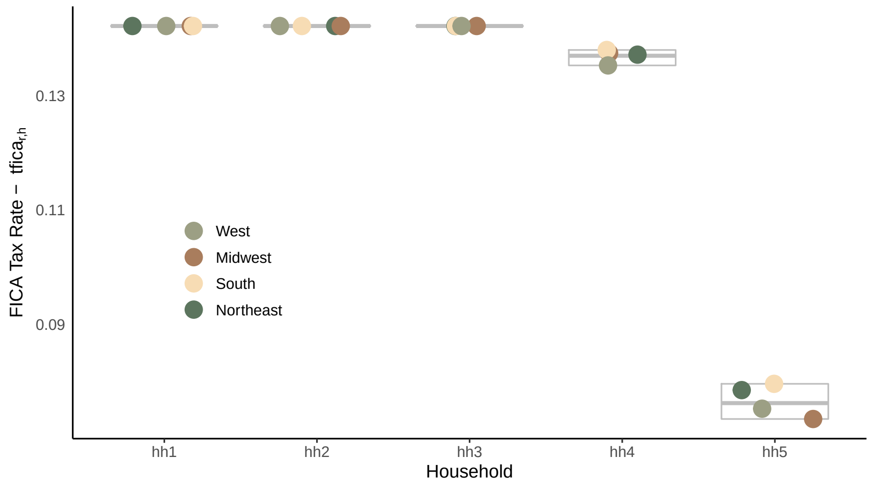
The average individual income tax payment for household h in region r, $tl\_avg0_{r,h}$, is computed following the same procedure outlined above for the effective marginal individual income tax rate on labor income. However, in this case the rate used is the average individual income tax rate calculated by Taxsim. The average individual income tax rates by region and household are presented in Figure 8c. These estimates are consistent with recent estimates by the U.S. Congressional Budget Office (CBO), noting that the estimates for SAGE are slightly higher due to the inclusion of state income taxes (CBO, 2018).

The effective marginal tax rate on capital income is calculated as a weighted average of corporate and personal income tax rates. The exercise described above for determining personal labor income tax rates is replicated for qualified dividends, interest income, and other business income, such as ordinary dividends and income from sole proprietorships and partnerships. In these cases, a national weighted average for the effective marginal tax rate is calculated with weights based on the income category being considered. The corporate income tax is based on an assessment of the average effective marginal corporate income tax rate by the U.S. CBO (CBO, 2017). Specifically the average effective marginal corporate income tax rate is set to 0.186. It is assumed that capital income passed on to households in the form of interest payments or dividends are subject to both the effective corporate income tax rate and the effective personal income tax rate for those types of income. In contrast, capital returns associated with sole proprietorships and partnerships is assumed to be only subject to the effective personal income tax rate on those types of income. Based on those assumptions, the effective marginal tax rate on capital income, tk_r, is calculated

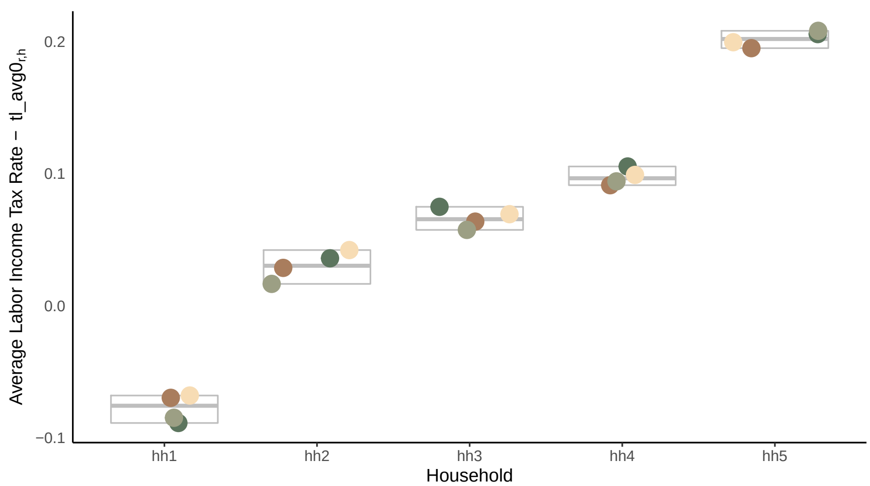
[27] Recent CBO estimates of the effective marginal federal income tax rate on wages without the Tax Cut and Jobs Act (TCJA) is approximately 20% in 2016, which is consistent with the income weighted average effective marginal federal income tax rate of approximately 20% used in the SAGE calibration (`https://www.cbo.gov/system/files/2019-01/54911-MTRchartbook.pdf`).



(a) Labor Income Effective Marginal Tax Rate by Household and Region



(b) FICA Effective Marginal Tax Rate by Household and Region



(c) Average Individual Income Tax Rate by Household and Region

Figure 8: Household Effective Marginal Labor Tax Rates

Table 6: Tax Rates on Capital Income

Region	tk
nor	0.33
sou	0.33
mid	0.33
wes	0.33

as a weighted average of the effective marginal tax rates on capital income distributed as interest, qualified dividends, ordinary dividends, and other business income, where the weights are the IRS individual income tax returns line-item totals for the types of capital income.[28] The values of the capital income tax rate based on these calculations are presented in Table 6.

3.3 Substitution Elasticities

In the calibrated CES and CET functions, the input-output data are used to define the benchmark value shares, and the free parameters are defined by the substitution elasticity parameters. The list of substitution elasticities included in the model is presented in Table 7.

3.3.1 Armington Elasticities

The sector-specific Armington elasticities between national and foreign goods, se_nf, are based on the estimates included in the GTAP database (Hertel et al., 2008). The GTAP elasticities are based on econometrically estimated substitution elasticities between imports across foreign sources, se_m, by Hertel et al. (2007) and using the "rule of two." The rule, first proposed by Jomini et al. (1991) and applied widely in CGE modeling, suggests that the elasticity of substitution across foreign sources is twice as large as the elasticity of substitution between domestic and imported commodities[29], such that

$$se\_nf = \frac{se\_m}{2}. \tag{62}$$

In cases where more than one of the 57 GTAP sectors map into a single SAGE sector, we use value-weighted averages based on GTAP v9 imports by the United States at world prices (Narayanan et al., 2016).

To define the elasticity of substitution between domestic goods and intra-national imports we follow the work of Caron and Rausch (2013). They provide a framework for estimating U.S. intra-national trade elasticities of substitution based on empirical estimates of international and domestic border effects. Specifically, they note that the relative strength of the intra-national and international border effects, α, is defined by the ratio of one minus the substitution elasticities

[28] https://www.irs.gov/pub/irs-soi/16inlinecount.pdf

[29] Using a back-casting experiment, Liu et al. (2004) found no evidence to reject the rule of two, providing additional support for its continued use.

Table 7: Elasticity Parameters

Parameter	Description
Trade	
se_nf	Elasticity of substitution between national and foreign goods
se_dn	Elasticity of substitution between domestic goods and national imports
te_dx	Transformation elasticity between domestically consumed and exported goods
Standard Production	
se_klem	Substitution elasticity between material inputs and energy-value-added
se_kle	Substitution elasticity between energy and value added
se_kl	Substitution elasticity between capital and labor
se_ene	Substitution elasticity between electricity and primary energy
se_en	Substitution elasticity among primary energy sources
Resource Extraction, Agriculture, and Forestry Specific	
se_rklem	Substitution elasticity between resource and materials-energy-value-added
Putty-Clay Capital	
te_k_ex	Transformation elasticity of sector differentiated extant capital
se_inv	Substitution elasticity in aggregate investment bundle
Household	
se_cl	Substitution elasticity between consumption bundle and leisure
eta	Inverse intertemporal substitution elasticity of consumption

Table 8: SAGE Elasticities

Sector	se_kl	se_kle	se_klem	se_ene	se_en	se_nf	se_dn
agf	1.07	0.40	0.98	0.68	0.33	2.45	4.13
bom	0.36	0.19	0.56	0.68	0.33	4.01	7.06
cem	0.20	0.25	0.81	0.68	0.33	2.90	4.98
chm	0.24	0.72	0.94	0.68	0.33	3.30	5.73
col	0.79	0.42	0.22	0.68	0.33	3.05	5.26
con	0.17	0.15	0.61	0.68	0.33	1.90	3.12
cpu	0.10	1.06	0.64	0.68	0.33	4.40	7.79
cru	0.79	0.42	0.22	0.68	0.33	7.30	13.20
ele	1.00	0.46	0.68	0.01	0.23	2.80	4.80
fbm	0.22	0.19	0.63	0.68	0.33	2.66	4.53
fmm	0.18	1.01	0.11	0.68	0.33	3.75	6.57
gas	0.79	0.42	0.22	0.68	0.33	2.80	4.80
hlt	0.58	0.16	0.80	0.77	0.10	1.90	3.12
min	0.79	0.42	0.22	0.68	0.33	0.90	1.25
pmm	0.18	1.01	0.11	0.68	0.33	3.74	6.56
prm	0.12	0.18	0.68	0.68	0.33	3.30	5.73
ref	0.73	0.38	0.42	0.68	0.33	2.10	3.49
srv	0.31	0.27	0.66	0.77	0.10	1.90	3.12
tem	0.18	0.16	0.38	0.68	0.33	3.46	6.02
trn	0.54	0.46	0.73	0.25	0.25	1.90	3.12
ttn	0.14	0.42	0.22	0.25	0.25	1.90	3.12
wpm	0.12	0.24	0.67	0.68	0.33	3.06	5.28
wsu	1.00	0.46	0.68	0.68	0.33	2.80	4.80

between intra-national sources, se_d, and international sources, se_m, such that

$$\alpha = \frac{1 - se\_d}{1 - se\_m}. \tag{63}$$

Given an estimate for α and se_m, this relationship may be used to solve for the substitution elasticity across domestic sources, se_d. We follow Caron and Rausch (2013) and apply the rule of two to calibrate the substitution elasticity between locally produced goods in the region and intra-national imports, such that $se\_dn = se\_d/2$. Given this relationship, along with (62) and (63), we can solve for the substitution elasticity between locally produced goods

$$se\_dn = \frac{1}{2} - \alpha \left(\frac{1}{2} - se\_nf \right). \tag{64}$$

Coughlin and Novy (2013) estimate both intra-national and international border effects for the U.S. Based on their results we assume that α is 1.868. The SAGE values for se_nf and se_dn are presented in Table 8.

We also follow Caron and Rausch (2013) in setting the transformation elasticity of output between domestic use, national exports, and international exports, te_dx, to 2.

3.3.2 Production Elasticities of Substitution

Koesler and Schymura (2015) provide empirical estimates of the capital-labor substitution elasticities (se_kl), (capital-labor)-energy substitution elasticities (se_kle), and (capital-labor-energy)-materials substitution elasticities (se_klem) at the industry level using a CES nesting structure that is consistent with our standard production structure in Figure 4 and the resource dependent sectors' production structure in Figure 5. The estimates are calculated with a panel dataset, covering 1995 to 2007, allowing for the estimation of long-run elasticities, which have been previously applied to CGE modeling (e.g., Böhringer et al. (2016)). The 34 sectors estimated by Koesler and Schymura (2015) are roughly consistent with our default aggregation, though notably they have more detail in the service sectors and less detail in the resource extraction sectors. For cases where a one-to-one mapping between their sectors and SAGE's sectors is not possible we use a weighted average of the Koesler and Schymura (2015) elasticities, where the weighting is by the U.S. sectoral output value in the last year of their dataset. For some sectors, the estimation routine of Koesler and Schymura (2015) returned non-finite values for se_kl. Therefore, for the electricity and refining sectors we use values from the recent study by Young (2013), which estimates sector-specific value-added substitution elasticities for the United States[30] Koesler and Schymura (2015) also reported a non-finite value for se_kle in the refining sector, in which case we apply the total industry value. The SAGE values for se_kl, se_kle, and se_klem are presented in Table 8. In general, a larger value for the substitution elasticity suggests a greater degree of substitutability between the inputs.

The interfuel substitution elasticities are based on estimates from Serletis et al. (2010a), which provide the most recent estimates for the United States based on contemporary data disaggregated across the industrial, commercial, electricity, and residential sectors. For the primary energy substitution elasticity, se_en, in the industrial sectors we use the Allen elasticity across refined petroleum and natural gas, as coal expenditures represent a small share of overall energy expenditures in those sectors. For the electricity sector (ele), the primary energy substitution elasticity is set equal to the estimate of the Allen substitution elasticity between coal and natural gas, as refined petroleum inputs represent a very small share. The results of Serletis et al. (2010a) suggest there are few substitution possibilities between refined petroleum and natural gas in the commercial sectors, so the substitution elasticity in the services and healthcare sectors (srv and hlt) is set to be commensurate with that finding. The substitution elasticity between the primary energy composite and electricity, se_ene, is a weighted average of the Allen substitution elasticity estimates for electricity and primary fuels from Serletis et al. (2010a). The weights represent the sector's national primary fuel expenditures in the model's benchmark year based on EIA's State Energy Data System.[31] We assign values from the industrial sector to the manufacturing and resource extraction sectors in the model.[32] We assign values from the commercial sector to the services and healthcare sectors

[30]We use the non-normalized generalized method of moments estimates from Young (2013).

[31]`https://www.eia.gov/state/seds/`

[32]Following this same procedure but using the meta-analysis results of Stern (2012) for the industrial sector produces

(srv and hlt). For the electricity sector we assume that the nest combining electricity and primary energy inputs is essentially Leontief. For the transportation sectors we base the substitution elasticities on the estimates of Serletis et al. (2010b) for high-income countries. The SAGE values for *se_en* and *se_ene* are presented in Table 8.

3.3.3 Resource Extraction, Agriculture, and Forestry

In sectors with a fixed factor input, including the resource extraction sectors and the agriculture and forestry sectors, the elasticity of substitution between the fixed factor resource and other inputs, *se_rklem*, is calibrated to match a long-run supply elasticity based on the benchmark conditions, similar to Balistreri and Rutherford (2001). In partial equilibrium with fixed prices for all non-resource inputs and a fixed quantity for the resource, the elasticity of supply for a given sector is given by

$$\eta = -\sigma_{res}, \tag{65}$$

where σ_{res} is the Allen own-price elasticity of substitution (Hertel and Tsigas, 2002). In the nesting structure for sectors with a fixed factor, as depicted in Figure 5, the Allen own-price elasticity for sector s in region r is

$$\sigma_{res} = -se\_rklem_{r,s}\left(\theta_{r,s,res}^{-1} - 1\right), \tag{66}$$

where $\theta_{r,s,res}$ is the benchmark resource cost share of total costs (Keller, 1976). Combining (65) and (66) provides the calibrated substitution elasticity for a given elasticity of supply

$$se\_rklem_{r,s} = \frac{\eta}{\theta_{r,s,res}^{-1} - 1}. \tag{67}$$

The endogenous supply elasticity in the model is a function of the share of production from new capital in the sector and the endogenously determined value shares, which differ from $\theta_{r,s,res}$. As production with extant capital becomes a smaller share of total production over time, the endogenous supply elasticity increases towards the long-run value to which the function is calibrated. However, as demand for the sector's commodity expands over time the value share of production from variable inputs increases (akin to a stock effect on marginal extraction costs), which in the case of the CES production function places downward pressure on the endogenous supply elasticity.

Arora (2014) examines the natural gas supply elasticity in the United States before and after the expansion of shale gas production through hydraulic fracturing, finding evidence of more elastic supply in recent years. Based on these estimates, Arora and Cai (2014) suggest a long-run supply elasticity of 0.5 for natural gas production as a reference case in CGE modeling. We apply a long-run supply elasticity of 0.5 for the natural gas extraction sector (*gas*).

U.S. oil supply is also considered to be inelastic. Huntington (1992) reviewed expectations of U.S. crude oil supply elasticities through the elasticities implicitly used in energy modeling systems of the time and found an average long-run elasticity of 0.40. There is evidence that in recent

similar values for *se_en* and *se_ene*.

decades, the oil supply has been more inelastic than those implied expectations (Greene and Liu, 2015). Krichene (2002) estimates the long-run world crude oil supply elasticity to be 0.25 over the period 1918-1999, with a lower elasticity estimates of 0.10 when the sample was restricted to the later years. This is relatively consistent with recent estimates of short-run world crude oil supply of 0.10 by Caldara et al. (2018) and 0.15 by Baumeister and Hamilton (2019). Caldara et al. (2018) provides evidence that short-run supply elasticities may be lower in non-OPEC nations relative to the world value. However, Bjørnland et al. (2017) finds that the supply elasticity for shale wells in the U.S. (which are responsible for around 60% of U.S. oil production[33]) may be notably larger, in the range of 0.3 to 0.9 depending on well characteristics. Finally, using a long-run supply elasticity of 0.25, Beckman et al. (2011) find that the GTAP-E model was able to adequately capture the variance of oil price responses to supply and demand shocks based on historical observations. Based on this evidence, we apply a long-run supply elasticity of 0.15 for the crude oil extraction sector (*cru*).

The supply of coal in the United States is generally thought to be elastic. For example, Balistreri and Rutherford (2001) use a long-run supply elasticity of 1.9 to calibrate an energy-detailed CGE model. This value is consistent with the long-run supply elasticity in other previous modeling exercises (Golombek et al. (1995); Brown and Huntington (2003)). Empirical elasticities of coal supply elasticities are limited. Dahl and Duggan (1996) survey the literature and find a range of estimates between 0.05 and 7.9 for the United States, with a median value of 0.79. However, data used in the included studies all end in the early 1970s. In a study of coal supply in Australia, Beck et al. (1991) find a long-run supply elasticity of 1.9. Econometric analyses conducted by EIA staff (EIA (2001)) find coal supply elasticities in the range of 1.5 to 3.0. Haggerty et al. (2015) calculate an average supply elasticity of 2.4 from the results of econometric analyses underlying recent versions of EIA's National Energy Modeling System. Based on this evidence, we apply a long-run supply elasticity of 2.4 for the coal mining sector (*col*).

The long-run supply elasticity for the aggregate other mineral and metal mining sector (*min*) is also likely to be elastic.[34] Empirical estimates of supply elasticities for stone, sand, and gravel mining are extremely limited. However, past investigations by the U.S. International Trade Commission (ITC) found the short-run supply elasticity for cement and clinker to be between 2 and 4, suggesting the supply of stone inputs is likely to be fairly elastic (ITC, 2014a). There appear to be no recent estimates of the supply elasticity for copper, though older estimates suggest that the supply is elastic. For example, Foley and Clark (1981) estimate the long-run supply elasticity of copper in the United States to be 6. While refractory minerals represent a smaller share of the sector, a recent ITC investigation concluded the supply elasticity to be in the range of 5 to 7 (ITC, 2014b). Similarly, the ITC found that pure magnesium and alloy magnesium have a short-run supply elasticity of 1.5 to 3 and 3 to 5, respectively (ITC, 2011). Based on this evidence, we apply

[33] `https://www.eia.gov/tools/faqs/faq.php?id=847&t=6`

[34] As reported previously, approximately two-thirds of the output value from the sector is attributable to stone mining and quarrying (NAICS 21231) or sand and gravel mining (NAICS 21232). Of the remaining third of the sector's output value, copper ore mining (NAICS 212234) accounts for approximately half.

a long-run supply elasticity of 5 for the other mineral and metal mining sector (*min*).

The agriculture and forestry sector is dominated by crop and livestock production and therefore, we focus on empirical estimates of long-run supply elasticities in those areas. The majority of U.S. cropland is associated with the production of grains, with corn and soybeans as the dominant crops. Kim and Moschini (2018) estimate the long-run supply elasticity of corn and soybeans in the United States to be 0.4 and 0.3, respectively. These results are consistent with those of Hendricks et al. (2014), who find a long-run supply elasticity for both corn and soybeans in the United States of 0.3. While older studies also find a long-run supply elasticity for corn of 0.3, the estimate for soybeans is higher at 1.6 (Shideed and White, 1989).[35]

For elasticities in livestock production, Kaiser (2012) estimates a long-run elasticity of hog supply of 0.3. Boetel et al. (2007) estimate a long-run supply elasticity of breeding stock with respect to the hog price of 0.6. Marsh (2003) and Sarmiento and Allen (2000) estimate long-run cattle supply elasticities of 0.6 to 2.8 and 0.3 to 2.9, respectively. These ranges are roughly consistent with previous estimates of cattle supply elasticities (e.g., Rucker et al. (1984) and Buhr and Kim (1997)). Little empirical evidence exists for the long-run supply elasticity of poultry (e.g., broilers) in the United States.[36] Based on this evidence, we apply a long-run supply elasticity of 0.5 for the agriculture and forestry sector (*agf*).

3.3.4 Large Open Economy Elasticities

Similar to the natural resource sectors, the large open economy assumption is operationalized by calibrating the fixed factors in equations 3 and 4 so that model behavior is consistent with exogenous price elasticities. The price elasticities used for international export demand and import supply are produced by tracing out export demand and import supply functions with the GTAPinGAMS package for 2011 (Lanz and Rutherford, 2016) and fitting an isoelastic function to the simulated data. Table 9 reports the generated price elasticities for export demand and import supply by sector (in columns labeled as GTAP). Import supply elasticities are noticeably higher indicating a flatter supply curve (and in many cases well approximated by the small open economy assumption).

In equation 67, the elasticity of substitution is solved in terms of an exogenously set supply elasticity and value shares derived from the underlying social accounting matrix. In this case, however, the value shares of the export demand and import supply fixed factors are not known. Therefore, the elasticity of substitution is set equal to one and the procedure solves for the value shares. Rearranging terms, costs shares are solved for as,

$$cs\_loe\_x_s = \frac{1}{|\epsilon_s^x|} \tag{68}$$

[35] Iqbal and Babcock (2018) find global long-run supply elasticity estimates of 0.2 and 0.6, respectively. However, Roberts and Schlenker (2013) find a slightly lower global supply elasticity for corn of around 0.1. For non-grain U.S. agricultural production, a significant portion of production value in attributed to California. Russo et al. (2008) study long-run supply elasticities of Californian horticulture and generally find estimates of less than 1, with values of 0.7 for almonds, 0.2 for walnuts, and 0.4 for tomatoes.

[36] Kapombe and Colyer (1998), Holt and Aradhyula (1998), and Holt and McKenzie (2003) all find evidence of a short-run supply elasticity of 0.1.

Table 9: Large Open Economy Price Elasticities

Sector	Export Demand		Import Supply	
	GTAP	SAGE	GTAP	SAGE
agf	-4.02	-4.01	108.10	108.13
bom	-7.32	-7.32	192.47	192.47
cem	-5.27	-5.27	153.56	153.67
chm	-5.79	-5.79	194.73	194.73
col	-5.05	-5.05	725.14	727.77
cpu	-7.98	-7.98	188.47	188.47
cru	-10.15	-10.15	10.00	10.00
ele	-5.60	-5.60	210.53	210.53
fbm	-4.67	-4.67	140.24	140.23
fmm	-6.99	-6.99	193.31	193.32
gas	-27.04	-27.04	167.77	167.77
hlt	-3.74	-3.74	37.10	37.10
min	-1.47	-1.47	231.66	231.93
pmm	-6.75	-6.75	215.72	215.72
prm	-5.79	-5.79	194.73	194.73
ref	-3.63	-3.63	65.03	65.03
srv	-3.77	-3.77	75.17	75.19
tem	-5.74	-5.74	156.73	156.73
trn	-2.12	-2.12	313.90	313.90
ttn	-3.21	-3.21	92.22	92.22
wpm	-5.61	-5.61	132.07	132.07

and

$$cs\_loe\_m_s = \frac{1}{1 + \epsilon_s^m} \tag{69}$$

where ϵ_s^x is the price elasticity of export demand for sector s and ϵ_s^m is the price elasticity of import supply for sector s. The reference levels of the fixed factors for exports ($fix0_s$) and imports ($fim0_s$) are then defined as,

$$fix0_s = \frac{cs\_loe\_x_s}{(1 - cs\_loe\_x_s)} \sum_r x0_{r,s,ftrd} \tag{70}$$

and

$$fim0_s = cs\_loe\_m_s \sum_r m0_{r,s,ftrd}. \tag{71}$$

Given the dynamic baseline assumptions in the model, the model's endogenous price elasticities for export demand and import supply are slightly different from the calibration points. Table 9 also reports the average implicit price elasticities across regions in SAGE in the base year of the model.

3.3.5 Partial Putty-Clay Elasticities

The elasticity of transformation for extant capital introduces a small amount of flexibility to shift extant capital across sectors within a model region. We set this parameter to 1.5 to capture the observation that some extant capital can be re-purposed in other sectors, though not excessively so. Lacking good empirical work on this topic, we chose the elasticity to restrict the shift of a given sector's extant capital stock for use in another sector to be roughly smaller than 5% (for non fixed factor sectors) for reasonable sized policy shocks. The substitution elasticity in the aggregate investment good bundle is set to be 0.05. Our formulation roughly follows Yuan et al. (2019), who assume that aggregate investment is composed of fixed proportions of commodity output, while allowing for a CES representation to explore this assumption in sensitivity.[37]

3.3.6 Consumption Elasticities

For non-leisure consumption, the model is calibrated to exogenous income elasticities as estimated in the literature.[38] The linear expenditure system requires that total commodity demand be divided between discretionary and subsistence level spending. Because the underlying social accounting matrix only reports total commodity demands, we determine subsistence levels that are consistent with exogenously specified income elasticities. We illustrate this procedure in a simplified context. Consider the LES utility function,

$$U(cd_{r,1,h}, \ldots, cd_{r,n,h}) = \prod_s (cd_{r,s,h} - sd0_{r,s,h})^{cs\_les_{r,s,h}}. \tag{72}$$

Maximizing this utility function subject to a budget constraint yields the following demand function,

$$cd_{r,s,h} = sd0_{r,s,h} + \frac{cs\_les_{r,s,h}(I_{r,h} - \sum_s ps_{r,s} sd0_{r,s,h})}{pa_{r,s}}, \tag{73}$$

where $pa_{r,s}$ is the regional price for commodity s and $I_{r,h}$ is the total income level for each agent. This demand function for total non-leisure commodity expenditures is broken into the subsistence demand component and the discretionary demand component. In this equation, there are two unknowns: $sd0_{r,s,h}$ and $cs\_les_{r,s,h}$. The benchmark data for SAGE only contains total commodity demands, $cd0_{r,s,h}$. Assuming reference prices are unity, algebra reveals the following LES budget shares by rearranging equation 73,

$$cs\_les_{r,s,h} = \frac{cd_{r,s,h} - sd0_{r,s,h}}{I_{r,h} - \sum_{ss} sd0_{r,ss,h}}. \tag{74}$$

[37]Data produced by the Bureau of Labor Statistics suggests that the composition of investment is to some extent responsive to changes in prices over time (see: https://www.bls.gov/emp/data/input-output-matrix.htm). However, we are unaware of studies characterizing the magnitude of this price responsiveness and as such, provide the option to perform sensitivity to the assumption made above.

[38]Technically, the model's implicit elasticity of consumption to aggregate consumption expenditures is calibrated to match empirical estimates of the same behavioral response. But for simplicity of exposition, this elasticity is referred to as the income elasticity in section 3.3.6.

Income elasticities of demand are then defined as,

$$\epsilon_s^{les} = cs\_les_{r,s,h} \frac{I_{r,h}}{cd_{r,s,h}}. \tag{75}$$

Equations 74 and 75 are used to calibrate the nested non-leisure component of the demand system. Note that we can substitute in benchmark levels of $cd_{r,s,h}$ ($cd0_{r,s,h}$) and $I_{r,h}$ directly from the underlying social accounting matrix. We pin down the remaining degree of freedom using estimated expenditure elasticities. To do so, we use the approach developed in Aguiar and Bils (2015) to estimate expenditure elasticities using data from the Consumer Expenditure Survey (CEX) provided by the Bureau of Labor Statistics. Note that in our simple calibration above, income equals aggregated non-leisure commodity expenditures. In SAGE, that is not the case. Therefore, we use the approach in Aguiar and Bils (2015) to estimate expenditure elasticities that do not account for leisure demand. The authors estimate log-linear Engel curves controlling for household demographics (etc. age, number of earners, family size) for different commodity groups and instrument for total expenditures with total income to correct for reporting measurement error. The approach identified by Aguiar and Bils (2015) has the advantage of not relying on price data jointly with expenditure data to estimate a complete set of income elasticities (relative to estimating an LES consumer demand system directly). The original estimates in the paper are reported using data from 1994-1996. We update these estimates using CEX data from 2013-2017.[39] We use this range of data to capture income elasticities that correspond with the reference year (2016) in SAGE. In our adapted estimation routine, we expand the set of commodity groupings relative to the original paper to aggregate to SAGE sectors. Estimated elasticities are mapped to SAGE sectors as weighted averages based on the Personal Consumer Expenditure (PCE) bridge file to input output accounts provided by the BEA.[40] Table 10 reports the reference elasticities used in this calibration procedure and Table 11 reports the mapping shares between the CEX categories and SAGE sectors as proxied by the PCE bridge file.

To ensure that the aggregate national subsistence demands are independent of the regional aggregation used in the model, the calibration routine is first run at the national level with a single household to generate total calibrated subsistence demands by commodity group. The calibration for the default aggregation of the model (that is distinguished by both regional and household heterogeneity) is then conducted but constrained to match the subsistence demand totals as estimated in the single region-household case. Because the system becomes overidentified when constraining the aggregate subsistence demand levels to equal the values in the national model calibration, we use a least squares calibration routine. This procedure will naturally allow calibrated income elasticities to vary by region and income group. However, the least squares routine penalizes deviations away from the empirically estimated elasticities while still being subject to the calibration constraints (equations 74 and 75).

[39] For replication code from Aguiar and Bils (2015), see `https://www.aeaweb.org/articles?id=10.1257/aer.20120599`.

[40] See: `https://www.bea.gov/industry/industry-underlying-estimates`.

Table 10: Expenditure Elasticities

Estimated Expenditure Elasticities (methodology from Aguiar and Bils (2015))		Mapped Elasticities in SAGE	
CEX Category	Elasticity	Sector	Elasticity
Alcoholic Beverages	1.1	agf	1.02
Food and Other Beverages	0.85	min	1.27
Tobacco Products	0.34	ele	0.31
Clothing and Apparel	1.13	gas	0.32
Personal Care	1.07	wsu	0.26
Reading	0.74	fbm	0.86
Education	1.5	wpm	1.4
Medical Treatment	1.16	ref	0.65
Entertainment	1.39	chm	1.17
Electricity Utilities	0.31	prm	1.17
Natural Gas Utilities	0.32	cem	1.58
Heating Fuels	0.31	pmm	1.55
Telephone	0.65	fmm	1.41
Water Utilities	0.26	cpu	1.31
Housing	0.45	tem	1.07
Housing Supplies and Furnishings	1.58	bom	1.27
Vacation Home Rentals	0.6	trn	0.97
Transportation Fuels	0.66	ttn	1.58
Vehicle Maintenance	0.78	srv	0.79
Vehicle Financing	0.27	hlt	1.17
Vehicle Services	1.1		

Given the dynamic nature of the model and the baseline calibration that deviates from the assumptions in the simplified static household problem above, the model's endogenous income elasticities differ slightly from the calibration points. Figure 9 presents the income elasticities implicit in the model's baseline along with the empirically estimated targets from the calibration procedure. The largest, though still relatively small, deviations from the calibration targets occur in sectors where the presence of fixed factors causes higher relative price growth in the outer simulation years.

The consumption-leisure substitution elasticity is determined jointly with the time endowment in the model to match observed estimates of the compensated and uncompensated labor supply elasticities in a static setting. Consider the demand system in (27) and the simplified budget constraint

$$pcl_{t,r,h}cl_{t,r,h} = (1 - tl_{t,r,h} - tfica_{t,r,h})\, pl_{t,r}te_{t,r,h} + pl_{t,r}tl\_refund_{t,r,h} + \pi_{t,r,h}, \qquad (76)$$

where $\pi_{t,r,h}$ represents non-labor income net of savings and $tl\_refund_{t,r,h}$ represents a hypothetical "refund" from the government to the household for the difference between the marginal and average

Table 11: CEX to SAGE Mapping Shares

SAGE Sector	CEX Category	Share
agf	Food and Other Beverages	0.69
agf	Entertainment	0.31
cem	Housing Supplies and Furnishings	0.99
cem	Vehicle Maintenance	0.01
chm	Medical Treatment	0.77
chm	Personal Care	0.15
chm	Housing Supplies and Furnishings	0.06
chm	Entertainment	0.01
chm	Reading	0.00
chm	Food and Other Beverages	0.00
chm	Heating Fuels	0.00
chm	Vehicle Maintenance	0.00
chm	Transportation Fuels	0.00
cpu	Entertainment	0.47
cpu	Housing Supplies and Furnishings	0.27
cpu	Telephone	0.11
cpu	Clothing and Apparel	0.05
cpu	Personal Care	0.05
cpu	Medical Treatment	0.03
cpu	Vehicle Maintenance	0.02
cpu	Vehicle Services	0.00
fmm	Entertainment	0.49
fmm	Housing Supplies and Furnishings	0.37
fmm	Personal Care	0.11
fmm	Vehicle Maintenance	0.03
fmm	Vehicle Services	0.00
fbm	Food and Other Beverages	0.73
fbm	Alcoholic Beverages	0.13
fbm	Tobacco Products	0.09
fbm	Entertainment	0.04
hlt	Medical Treatment	0.98
hlt	Education	0.02
min	Housing Supplies and Furnishings	0.63
min	Food and Other Beverages	0.18
min	Heating Fuels	0.13
min	Entertainment	0.06
bom	Clothing and Apparel	0.52
bom	Housing Supplies and Furnishings	0.26
bom	Entertainment	0.13
bom	Medical Treatment	0.06
bom	Reading	0.01
bom	Personal Care	0.01
bom	Vehicle Maintenance	0.00
bom	Vehicle Services	0.00
ref	Transportation Fuels	0.94
ref	Heating Fuels	0.05
ref	Medical Treatment	0.00
prm	Vehicle Maintenance	0.45
prm	Housing Supplies and Furnishings	0.44
prm	Vehicle Services	0.04
prm	Personal Care	0.02
prm	Entertainment	0.02
prm	Medical Treatment	0.02
prm	Reading	0.01
pmm	Housing Supplies and Furnishings	0.93
pmm	Vehicle Maintenance	0.04
pmm	Entertainment	0.03
srv	Housing	0.34
srv	Food and Other Beverages	0.12
srv	Vehicle Financing	0.12
srv	Entertainment	0.10
srv	Education	0.07
srv	Medical Treatment	0.05
srv	Vehicle Maintenance	0.04
srv	Housing Supplies and Furnishings	0.04
srv	Telephone	0.04
srv	Vehicle Services	0.03
srv	Personal Care	0.02
srv	Reading	0.01
srv	Water Utilities	0.01
srv	Vacation Home Rentals	0.01
srv	Clothing and Apparel	0.00
trn	Vehicle Services	0.71
trn	Vacation Home Rentals	0.22
trn	Vehicle Maintenance	0.06
trn	Food and Other Beverages	0.00
trn	Housing Supplies and Furnishings	0.00
trn	Entertainment	0.00
tem	Vehicle Services	0.89
tem	Vehicle Maintenance	0.09
tem	Entertainment	0.02
ttn	Housing Supplies and Furnishings	1.00
gas	Natural Gas Utilities	1.00
wsu	Water Utilities	1.00
ele	Electricity Utilities	1.00
wpm	Housing Supplies and Furnishings	0.66
wpm	Clothing and Apparel	0.16
wpm	Personal Care	0.15
wpm	Reading	0.02
wpm	Heating Fuels	0.01
wpm	Entertainment	0.01

tax rate. We use the notation of $tl\_refund_{t,r,h}$ in this section to emphasize that the tax rate relevant for behavioral choices is the marginal rate. Assuming labor income taxes are constant over time, $tl_{t,r,h} + tfica_{t,r,h} = tl_{t+1,r,h} + tfica_{t+1,r,h}\ \forall\ t$, the Marshallian demand for leisure is

$$\begin{aligned} leis_{t,r,h} =& leis0_{r,h} \left(\frac{\pi_{t,r,h} + (1 - tl_{t,r,h} - tfica_{t,r,h})\, pl_{t,r} te_{t,r,h} + pl_{t,r} tl\_refund_{t,r,h}}{\pi 0_{r,h} + (1 - tl_{t,r,h} - tfica_{t,r,h} + pl0_r tl\_refund0_{r,h})\, pl0_r te0_{r,h}} \right) \left(\frac{pl_{t,r}}{pl0_r} \right)^{-se\_cl} \\ & \times \left[cs\_cl_{r,h} \left(\frac{pc_{t,r,h}}{pc0_{r,h}} \right)^{1-se\_cl} + (1 - cs\_cl_{r,h}) \left(\frac{pl_{t,r}}{pl0_r} \right)^{1-se\_cl} \right]^{-1}. \end{aligned} \tag{77}$$

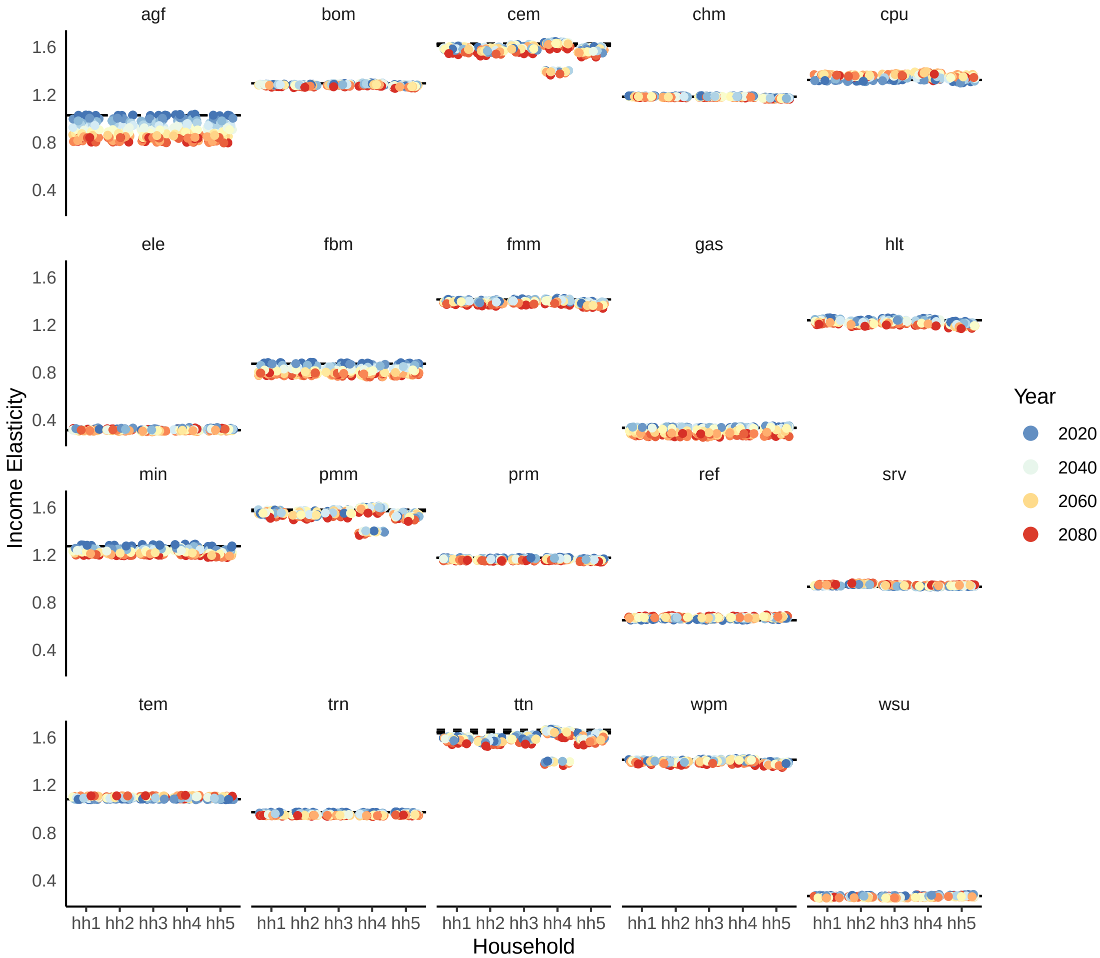


Figure 9: Calibrated Income Elasticities

The uncompensated price elasticity of leisure demand, μ_l, may be obtained from (77), such that

$$\begin{aligned}
\mu_{t,r,h}^{leis} \equiv& \frac{\partial leis_{t,r,h}}{\partial pl_{t,r}} \frac{pl_{t,r}}{leis_{t,r,h}} \\
=& \frac{\left(1 - tl_{t,r,h} - tfica_{t,r,h}\right) pl_{t,r} te_{t,r,h} + pl_{t,r} tl_refund_{t,r,h}}{\pi_{t,r,h} + \left(1 - tl_{t,r,h} - tfica_{t,r,h}\right) pl_{t,r} te_{t,r,h} + pl_{t,r} tl_refund_{t,r,h}} \\
& - \left(1 - cs_cl_{r,h}\right) \left(\frac{pl0_r}{pl_{t,r}}\right)^{se_cl-1} e\left(pc_{t,r,h}, pl_{t,r}\right)^{1-se_cl} \\
& + se_cl \left[\left(1 - cs_cl_{r,h}\right) \left(\frac{pl0_r}{pl_{t,r,h}}\right)^{se_cl-1} e\left(pc_{t,r,h}, pl_{t,r}\right)^{1-se_cl} - 1\right],
\end{aligned} \tag{78}$$

where

$$e\left(pc_{t,r,h}, pl_{t,r}\right) = \left[cs\_cl_{r,h} \left(\frac{pc_{t,r,h}}{pc0_{r,h}}\right)^{1-se\_cl} + \left(1 - cs\_cl_{r,h}\right) \left(\frac{pl_{t,r}}{pl0_r}\right)^{1-se\_cl}\right]^{\frac{1}{se\_cl-1}}. \tag{79}$$

The first two components of (78) define the income elasticity of leisure,

$$\begin{aligned}\mu^{I}_{t,r,h} =& \frac{\left(1 - tl_{t,r,h} - tfica_{t,r,h}\right) pl_{t,r} te_{t,r,h} + pl_{t,r} tl\_refund_{t,r,h}}{\pi_{t,r,h} + \left(1 - tl_{t,r,h} - tfica_{t,r,h}\right) pl_{t,r} te_{t,r,h} + pl_{t,r} tl\_refund_{t,r,h}} \\ & - \left(1 - cs\_cl_{r,h}\right) \left(\frac{pl0_r}{pl_{t,r}}\right)^{se\_cl-1} e\left(pc_{t,r,h}, pl_{t,r}\right)^{1-se\_cl},\end{aligned} \tag{80}$$

and the third component represents the substitution effect, or the compensated price elasticity of leisure demand,

$$\mu^{leis|\bar{cl}}_{t,r,h} = se\_cl \left[\left(1 - cs\_cl_{r,h}\right) \left(\frac{pl0_r}{pl_{t,r,h}}\right)^{se\_cl-1} e\left(pc_{t,r,h}, pl_{t,r}\right)^{1-se\_cl} - 1\right] \tag{81}$$

This may be verified through the Hicksian demand function via the Slutsky equation. Given the definition of labor supply, $te_{t,r,h} - leis_{t,r,h}$, the compensated labor supply elasticity, or substitution effect, is

$$\epsilon^{l|\bar{cl}}_{t,r,h} = -\mu^{leis|\bar{cl}}_{t,r,h} \frac{leis_{t,r,h}}{te_{t,r,h} - leis_{t,r,h}}. \tag{82}$$

And the uncompensated labor supply elasticity is

$$\epsilon^{l}_{t,r,h} = -\mu^{leis}_{t,r,h} \frac{leis_{t,r,h}}{te_{t,r,h} - leis_{t,r,h}}, \tag{83}$$

which, may be written as

$$\epsilon^{l}_{t,r,h} = -\left(\mu^{I}_{t,r,h} + \mu^{leis|\bar{cl}}_{t,r,h}\right) \frac{leis_{t,r,h}}{te_{t,r,h} - leis_{t,r,h}}. \tag{84}$$

We define the share of the time endowment spent on leisure as $\phi_{t,r,h} = leis_{t,r,h}/te_{t,r,h}$ and rewrite (82) and (84) as

$$\epsilon^{l|\bar{cl}}_{t,r,h} = \frac{-\phi_{t,r,h}}{1 - \phi_{t,r,h}} \mu^{leis|\bar{cl}}_{t,r,h} \tag{85}$$

and

$$\epsilon^{l}_{t,r,h} = \frac{-\phi_{t,r,h}}{1 - \phi_{t,r,h}} \left(\mu^{I}_{t,r,h} + \mu^{leis|\bar{cl}}_{t,r,h}\right). \tag{86}$$

Substituting (85) into (86) yields

$$\epsilon^{l}_{t,r,h} = \frac{-\phi_{t,r,h}}{1 - \phi_{t,r,h}} \mu^{I}_{t,r,h} + \epsilon^{l|\bar{cl}}_{t,r,h}. \tag{87}$$

From (80), the benchmark year income elasticity of leisure is

$$\mu^{I}_{0,r,h} = \frac{\left(1 - tl_{0,r,h} - tfica_{0,r,h}\right) pl0_r te0_{r,h} + pl0_r tl\_refund0_{r,h}}{\pi0_{r,h} + \left(1 - tl_{0,r,h} - tfica_{0,r,h}\right) pl0_r te0_{r,h} + pl0_r tl\_refund0_{r,h}} - \left(1 - cs\_cl_{r,h}\right). \tag{88}$$

Assuming that in the benchmark prices are normalized to unity such that the effective labor price is

$(1 - tl_{0,r,h} - tfica_{0,r,h})$ and given the definition of $cs\_cl_{r,h}$ and an estimate of the income elasticity of labor, $\hat{\epsilon}^I$, (76) and (88) may be substituted into (87) to yield the calibrated benchmark value of leisure

$$leis0_{r,h} = -\frac{dc0_{r,h}\hat{\epsilon}^I}{(1 - tl_{0,r,h} - tfica_{0,r,h})\,(1 + \hat{\epsilon}^I) + \frac{tl\_refund0_{r,h}}{l0_{r,h}}}. \tag{89}$$

From (81), the benchmark uncompensated leisure demand elasticity is

$$\mu_{0,r,h}^{\bar{leis}|\bar{cl}} = -se\_cl \cdot cs\_cl_{r,h}. \tag{90}$$

Substituting (90) into (85) yields the calibrated version of the elasticity of substitution between consumption and leisure,

$$se\_cl = \frac{\hat{\epsilon}^{l|\bar{cl}} cl0_{r,h} l0_{r,h}}{leis0_{r,h} dc0_{r,h}}, \tag{91}$$

where $\hat{\epsilon}^{l|\bar{cl}}$ is the empirical estimate of the substitution elasticity. The observed labor earnings are combined with the calibrated benchmark value of leisure in (89) to determine the time endowment $te0_{r,h} = l0_{r,h} + leis0_{r,h}$.

To calibrate the time endowment and the substitution elasticity between consumption and leisure, we use the conclusions from the literature review by McClelland and Mok (2012) on estimates of the income and substitution effects for the United States. Specifically, they conclude that estimates on the order of $\hat{\epsilon}^I = -0.05$ and $\hat{\epsilon}^{l|\bar{cl}} = 0.20$ are representative of the most recent empirical evidence. Given the dynamic nature of the model and the baseline calibration that deviates from the assumptions in the simplified static household problem above, the model's endogenous labor supply elasticities differ slightly from the calibration points. Figure 10 presents the substitution and income effects implicit in the model's baseline.

In the households' welfare maximization problem the additively separable nature of the intertemporal welfare function in (23) and the isoelastic form of the intra-temporal utility function (25) mean the elasticity of intertemporal substitution will be $1/\eta$. In a recent review of over 1,400 estimates of the elasticity of intertemporal substitution for the United States, Havranek et al. (2015) find a mean value of 0.6. Based on this evidence, we set the value of η to 1.66.

3.4 Dynamic Baseline

The model's baseline is a result of the economic structures and parameters previously defined, along with exogenous growth assumptions regarding: productivity, population, government accounts, foreign accounts, and energy use. Each of these components is discussed in turn, followed by a presentation of baseline indicators from the default version of the model.

As discussed in Section 2.7, the model is closed using a terminal condition that assumes the economy converges to a steady state in the very long-run. Because the model includes fixed factors of production, achieving convergence to a steady state requires that growth trends to zero in the very long-run and all quantities and prices remain constant. The assumptions that allow the model

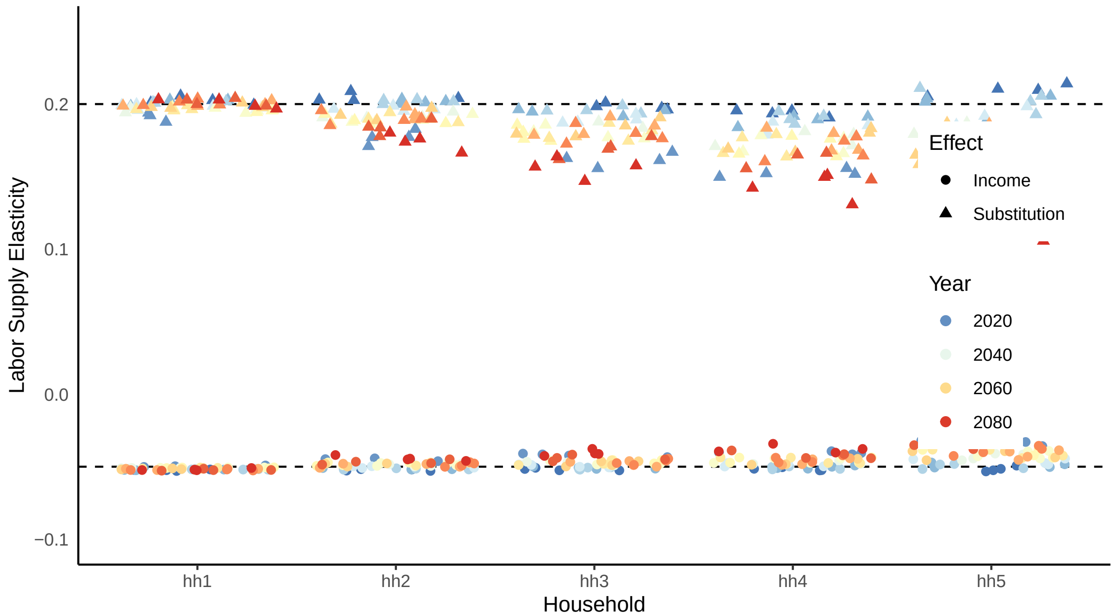


Figure 10: Calibrated Labor Supply Elasticities

to meet these terminal conditions are described in each of the Subsections below. Care is taken to ensure that the assumptions required for the model's terminal conditions do not meaningfully impact the results of the model in the time frame typically considered in regulatory analysis. However, in instances where a particularly long time-horizon is relevant for the analysis these specification can be revisited if necessary.

Also relevant to the baseline are assumptions regarding the initial private return to capital, $rbar$, which is set to 0.045. The interest rate reflects the average after-tax rate of return on private capital. Given the capital tax in Section 3.2 the social return on private capital in the model is approximately 0.07, which is consistent with the average pre-tax rate of return on capital observed between 1960 and 2014 (CEA, 2017). The depreciation rate, δ, is set to 0.05, which is the average U.S. capital depreciation rate from 1950 to 2014 as estimated by Feenstra et al. (2015). This rate is applied to both new and extant capital.

The pure rate of time preference, ρ, for households is a determinant of their savings rate and therefore, important for defining the baseline. Based on the isoelastic form of the intra-temporal utility function, the pure rate of time preference, ρ, in (24) is calibrated via the Ramsey formula given the specification of η, $rbar$, and the expected labor productivity and population growth rates over the first four model periods. Based on these assumed parameter values, ρ is about 0.018. Opinions on the most appropriate value for ρ vary. Some argue on ethical grounds that it should be equal to or very near zero (e.g., Ramsey (1928), Stern (2007)). Others rely on descriptive approaches to backout the implied pure rate of time preference, which typically imply somewhat higher values between 0.02 and 0.03 (Nordhaus, 2007). See Arrow et al. (2013) for a summary of these different approaches.

3.4.1 Productivity Growth

The foundation for productivity growth in the model is Harrod neutral (i.e., labor augmenting) technological progress. Aggregate economy-wide labor productivity growth is calibrated to match the estimates used in the CBO Long Term Budget Projection, which extends to 2050.[41] As previously noted, to meet the terminal conditions for the solution algorithm, growth in the very long-run needs to converge to zero. Therefore, the aggregate labor productivity growth rate in the last period of the CBO projection is extrapolated to 2060 at which point it linearly declines to zero by 2080.

Future productivity growth is not expected to be uniform across the sectors of the economy. To calibrate exogenous labor productivity growth, $l\_prod_{t,s}$, in SAGE historic differences in labor productivity growth across sectors are extrapolated in a manner consistent with the assumptions regarding economy-wide productivity growth based on the CBO projection.

The Integrated Industry-level Production Account data from BEA and the U.S. Bureau of Labor Statistics (BLS), covering 1998 to 2017, provide estimates of historic integrated labor productivity growth by sector in addition to data on gross output. This information is used to developed output-weighted average labor productivity growth estimates for each of the SAGE sectors. Prior to 2050, the baseline assumes that these historic differences in productivity will persist. After that point, the variance across sectors is calibrated to linearly decline to zero by 2070 (i.e., the productivity growth rate in each sectors converges to the mean growth rate).

In each model year, the sector-specific distribution of growth rates is then scaled by a constant proportion so that the aggregate economy-wide labor productivity growth matches the calibrated aggregate projections previously discussed. This scaling to match the aggregate productivity growth assumption is conducted conditional on relative labor demand across sectors in the benchmark year and persisting throughout the modeling horizon. This condition will not hold in the baseline as heterogeneity in sectoral growth rates alone will cause sectors' shares of aggregate labor demand to shift over time. However, simultaneously solving for both the model's baseline and these calibrated scaling parameters is computationally infeasible. While restricting sectors' share of labor demand to the benchmark levels in this scaling procedure will lead the baseline growth rate of aggregate labor productivity to be higher than the intended calibration point from the CBO projections, in practice the difference is relatively small. Figure 11 presents the growth rate of economy-wide labor productivity in the SAGE baseline and the CBO calibration points. The differences due to use of the benchmark labor shares will be most prevalent in the later years. The larger differences between the SAGE baseline and the CBO calibration points in the first two simulation periods are due to the representation of large but short-term COVID related impacts on labor markets in 2020 in the CBO projection. The SAGE baseline smooths out these impacts due to the larger time steps and forward looking nature of the model.

Figure 12 presents the exogenous sector specific labor productivity growth rates for SAGE in addition to the historical growth rates used for the calibration. While there is temporal heterogene-

[41] https://www.cbo.gov/data/budget-economic-data

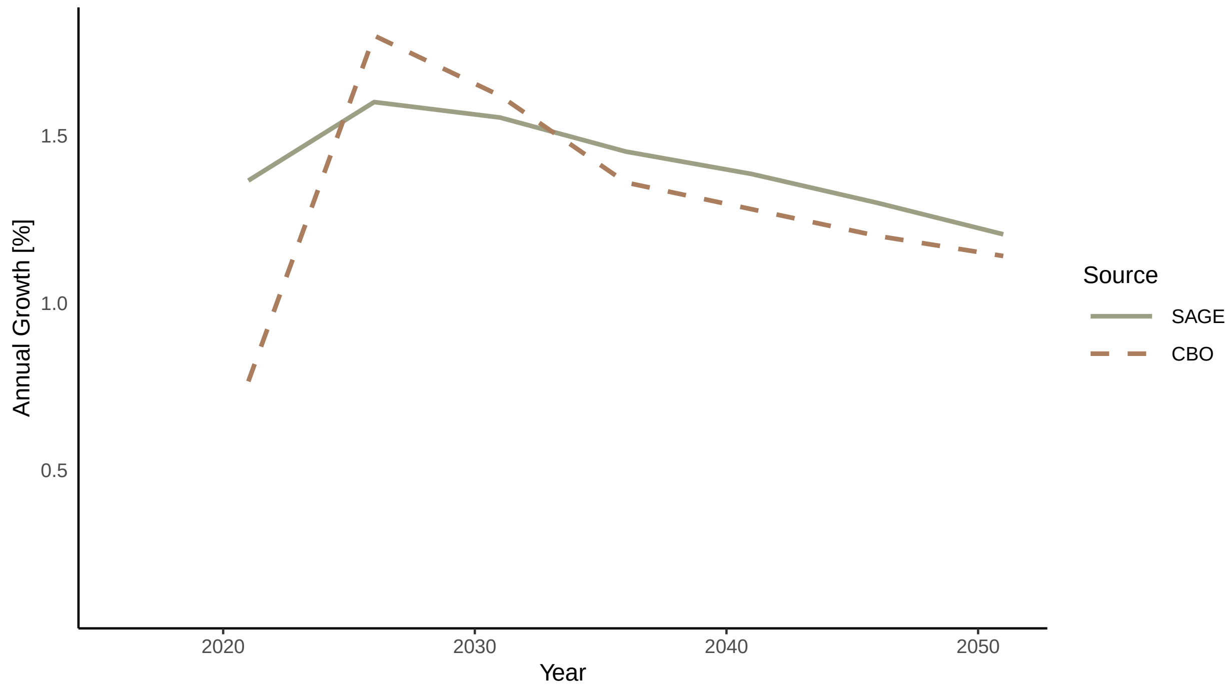


Figure 11: Aggregate Labor Productivity Growth in SAGE Baseline and CBO Target

ity in the exogenous growth rates the figure presents average annual growth rates over two periods to more easily facilitate a comparison.

3.4.2 Population Growth

Population in the model is assumed to grow at an exogenous rate equal to the growth rate of the labor force in the CBO Long-Term Budget Outlook, which extends to 2050. As previously noted, to meet the terminal conditions for the solution algorithm growth in the very long-run needs to converge to zero. Therefore, the population growth rate in the last period of the CBO projection is extrapolated to 2060 at which point it linearly declines to zero by 2080.

Assumed population growth is presented in Figure 16.

3.4.3 Government Accounts

The government agent in SAGE represents all federal, state, and local governments in the United States. Real government expenditures are exogenously specified as is the level of deficit financing. The relevant expenditure variables are government consumption, $gov_{t,r}$, interest payments, $gint_t$ and $gint\_row_t$, and transfer payments, $transfers_{t,r,h}$. Where possible we calibrate the variables to CBO's budget projections.[42] However, CBO's budget projections only cover the federal portion of the government expenditure variables and in many cases are only presented in the 10-year budget outlook, requiring extrapolation for the longer time horizon in SAGE.

[42] https://www.cbo.gov/data/budget-economic-data

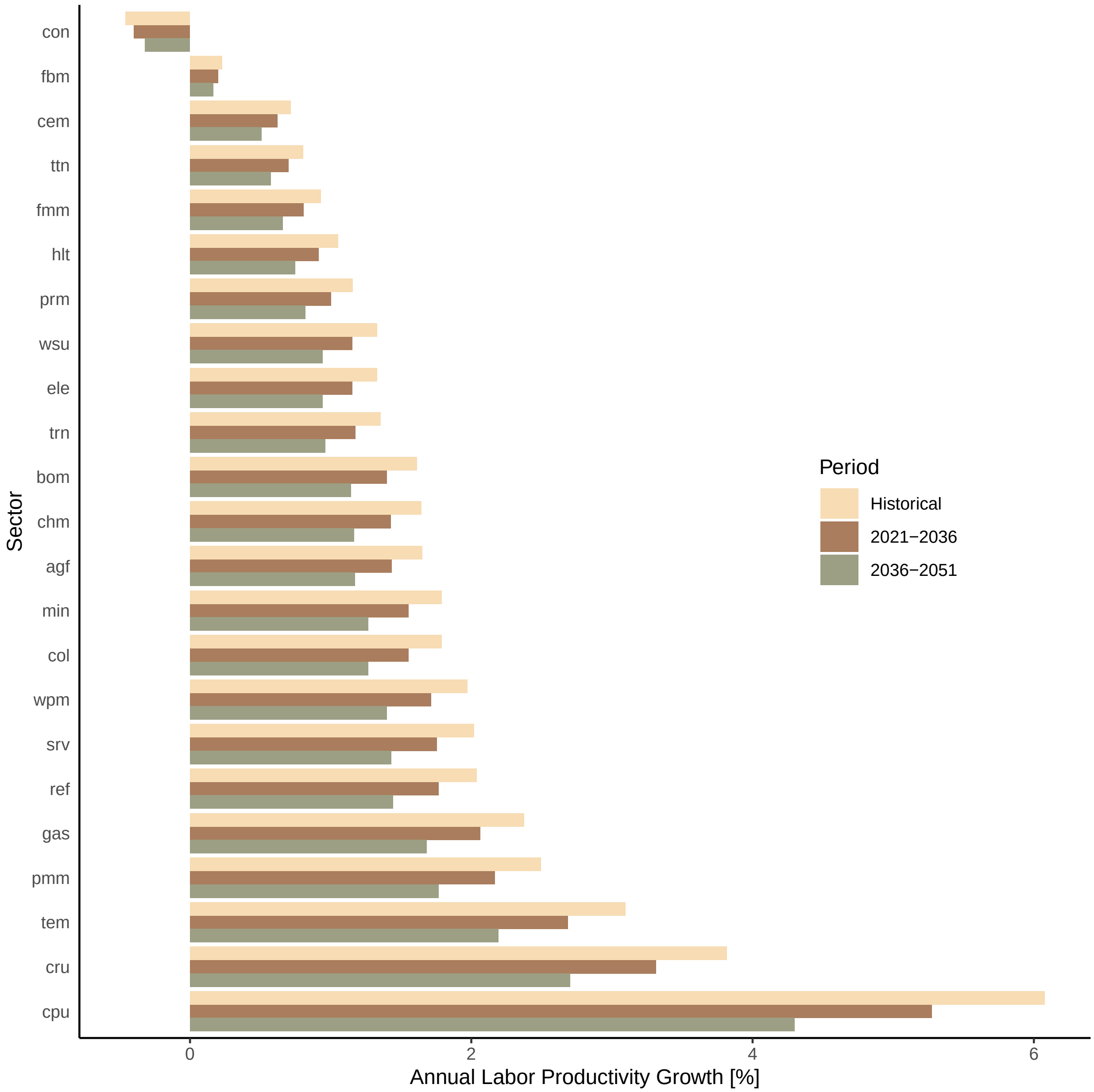


Figure 12: Historical and Calibrated Sectoral Labor Productivity Growth Rates

To set the government deficit we follow an approach similar to Jorgenson et al. (2013). Initially, the federal deficit is calibrated to grow at the rate projected in the CBO 10-year budget projection. In the benchmark year, state and local governments have a small deficit which is assumed to follow real GDP growth over the 10-year budget projection period. At the end of that projection period all government deficits are assumed to linearly decline to zero over the next 40 years, after which point they remain at zero.

The growth of interest payments both domestically, $gint_t$, and to the rest of the world, $gint\_row_t$, are based on the growth of government debt. Since the current level of all government debt is not part of the SAM, the U.S. Federal Reserve estimates of federal debt[43] and state and local government debt[44] are used to set the starting point for total public debt. The growth of debt is based on the deficit projection described above. Total interest payments are then assumed to follow the growth of debt. The baseline does not attempt to project changes in the share of debt held domestically compared to the rest of the world. Therefore, domestic interest payments and foreign interest payments are assumed to grow at the same rate.

Both real federal government consumption and transfer payments follow a similar calibration process. In both cases they are assumed to initially grow at a rate consistent with the CBO 10-year budget projection.[45] Afterwards, federal government consumption and transfer payments are assumed to grow at the rate of GDP. There are no comprehensive authoritative projections for state and local government accounts equivalent to CBO's projections for the federal government. Therefore, we assume the state and local government consumption and transfers will grow at the rate of real GDP throughout the time horizon.

The real GDP growth rate projection used to calibrate these exogenous payment projections is based on the extended projection in the CBO Long-Term Budget Projection out to 2050 and the implied growth rate based on aggregate labor productivity and population growth rate assumptions beyond 2050. The latter portion ensures that the projection of government expenditures and transfer payments is consistent with the exogenous assumptions regarding productivity and population, which is particularly important for ensuring that the terminal conditions in the model will hold.

An alternative approach to projecting government accounts would be to exogenously specify the government accounts as a share of GDP and allow the level to be endogenously determined. Since government consumption does not directly enter the utility functions for agents in the model, the baseline for SAGE uses exogenously projected levels to facilitate consistent changes in welfare when modeling policy changes. For a similar reason, the government budget constraint is endogenously balanced through lump sum transfers with households.

[43] `https://fred.stlouisfed.org/series/GFDEBTN#0`

[44] `https://fred.stlouisfed.org/series/SLGSDODNS#0`

[45] For historic years that may be included in the baseline due the benchmark year, CBO estimates of realized values are used.

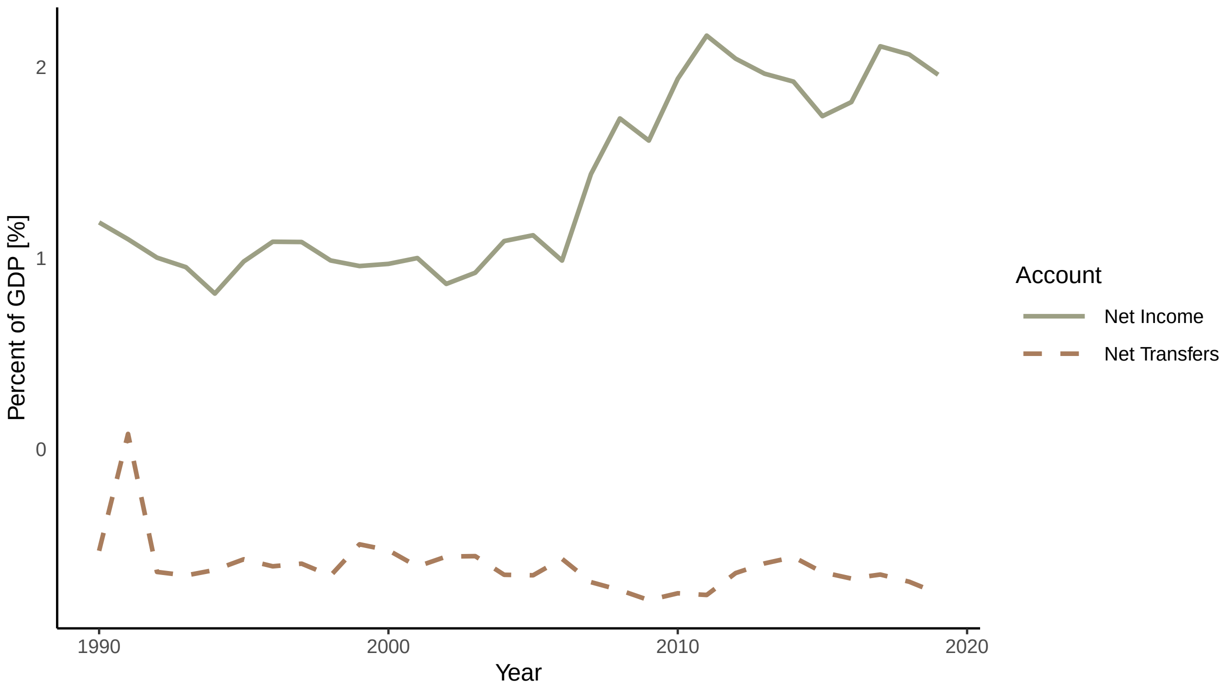


Figure 13: Historical Net Income and Net Transfers from Rest of the World as Percent of GDP

3.4.4 Foreign Accounts

There are four variables associated with the rest of the world that are exogenously specified: the current account balance, $curactbal_t$, net taxes and transfers from the rest of the world, $tran\_row_t$, net income for the rest of the world, inc_row_t, and government interest paid to the rest of the world, $gint\_row_t$.

To set the current account balance, $curactbal_t$, we follow an approach similar to Jorgenson et al. (2013) and the one used for the government deficit. Initially, the current account balance is calibrated to grow at the rate projected by the CBO in their economic projections. At the end of that projection period the current account balance is assumed to linearly decline to zero over the next 40 years, after which point it remains at zero.

Both net taxes and transfers from the rest of the world, $tran\_row_t$, and net income for the rest of the world, inc_row_t, are calibrated in a similar fashion. Historically, net taxes and transfers from the rest of the world have been a roughly constant relative to GDP at around 0.5%, see Figure 13.[46] Net income from the rest of the world relative to GDP was consistently around 1% from 1990 until 2008, after which it increased to around 2% where it has remained for the last decade. Given the relatively stable relationship over the last decade for both variables, they are calibrated to follow real GDP growth similar to the method described in Section 3.4.3.

The exogenous projection of government interest paid to the rest of the world is described in Section 3.4.3.

[46]Figure 13 is based on data from the BEA National Income and Product Account tables.

3.4.5 Baseline Energy Use

The cost shares in the production functions are adjusted to capture expected technological change in the energy intensity of production based on EIA's Annual Energy Outlook (AEO) forecast.[47] To get the unit energy consumption (UEC) we divide the total energy consumption in the AEO by the real value of shipments for each sector. The National Energy Modeling System (NEMS) used for the AEO only allows for limited fuel switching within the industrial sectors, so changes in the UEC over time predominately represent exogenous forecasts regarding technological change in energy efficiency. We use the average growth rate of the UEC in the AEO over SAGE model periods to calibrate the cost shares in the production function. The UEC relative to its value in the benchmark year is denoted as $ene\_growth_{t,s}$.

The change in energy efficiency is assumed to be capital embodied. Therefore, the change is represented as a shift from energy use to capital such that the "benchmark" values for intermediate and capital inputs as well as the cost shares are time dependent. The partial putty-clay framework needs to be accounted for to ensure that the overall UEC trend in SAGE is consistent with AEO, since only production with new capital is associated with the improvements and the goal is to match the overall UEC trend in AEO. The energy-related intermediate inputs and capital benchmark values for production with new capital are calibrated, such that

$$id0_{t,r,ss,s} = ene\_factor_{t,s} id0_{0,r,ss,s} \quad ss \in sene \tag{92}$$

and

$$kd0_{t,r,s} = kd0_{0,r,s} + \frac{(1 - ene\_factor_{t,s}) \sum_{ss \in sene} id0_{0,r,ss,s}}{1 + tk0_r}, \tag{93}$$

where

$$ene\_factor_{t,s} = \frac{ene\_growth_{t,s} q\_base_t - (1-\delta)^t}{q\_base_t - (1-\delta)^t} \tag{94}$$

$sene \in (col, gas, ref, ele)$ is the set of primary energy commodities plus electricity, and q_base_t reflects an approximation of general growth of the economy by capturing the cumulative growth of the effective labor force (i.e., aggregate economy-wide labor productivity growth plus population growth). The relevant cost shares, cs_kle and cs_kl, become time dependent and are adjusted to be consistent with (92) and (93).[48]

The mapping from the AEO sectors to the SAGE sectors is presented in Table 12. For the non-truck transportation sector, trn, the UEC growth rate is based on the average growth rate of air transportation fuel efficiency as forecast by the AEO, since this represents a large share of the energy consumption for the sector. For the truck transportation sector, ttn, the UEC growth rate is based on the average growth rate of truck freight transportation fuel efficiency as forecast by the AEO. No changes in the energy intensity of the electricity sector are assumed.

[47] The calibration is conducted with the most recent AEO forecast that includes a representation of the benchmark year for SAGE.

[48] Outside of this section we exclude the time subscript on the benchmark values and cost shares to simplify the exposition.

Table 12: Unit Energy Consumption SAGE and AEO Mapping

SAGE Sector	AEO Sectors
agf	agg
col	ming
min	ming
ele	
gas	ming
cru	ming
wsu	bmf
con	cns
fbm	fdp
wpm	ppm, wdp
ref	ref
chm	bch
prm	pli
cem	cem
pmm	ism, aap
fmm	fbp
cpu	cmpr, eei
tem	teq
bom	bmf, ggr, mchi
trn	
ttn	
srv	comm
hlt	comm

Figure 14 presents the exogenous sector-specific UEC growth rates for SAGE. While there is temporal heterogeneity in the exogenous growth rates the figure presents average annual growth rates over two periods to more easily facilitate a comparison.

Household and government energy consumption shares are assumed to change over time to match the energy intensity forecasts in AEO (the same assumption is used for subsistence demands). Consumption shares of electricity and natural gas are assumed to grow at the same average rate as in the AEO forecast, and are reflected by the indices $cd\_ene\_growth_{t,ele}$ and $cd\_ene\_growth_{t,gas}$, respectively. The consumption share of refined petroleum is assumed to grow based on the average consumption share growth rate of light-duty vehicle fuel expenditures, and is reflected by the index $cd\_ene\_growth_{t,ref}$. This is assumed to represent a shift towards other consumption goods in proportion to their benchmark consumption shares, such that

$$cd0_{t,r,h,s} = cd\_ene\_growth_{t,s} cd0_{0,r,h,s} \; s \in sene \tag{95}$$

and

$$cd0_{t,r,h,s} = cd0_{0,r,h,s} + \left\{ \sum_{ss \in sene} \left[1 - cd\_ene\_growth_{t,ss}\right] cd0_{0,r,h,ss} \right\} \frac{cd0_{0,r,h,s}}{\sum_{ss \notin sene} cd0_{0,r,h,ss}} \; s \notin sene. \tag{96}$$

Figure 15 presents the exogenous household UEC growth rates for SAGE. While there is temporal heterogeneity in the exogenous growth rates, the figure presents average annual growth rates over two periods to more easily facilitate a comparison. The same assumption is used for both discretionary and subsistence demands. Government consumption is subject to the same treatment and growth rates.

The growth of natural gas consumed per unit of electricity produced in the baseline is roughly consistent with forecasts from AEO. However, the growth of coal consumed per unit of electricity produced, absent any adjustment, would be higher than AEO forecasts due to regulatory and market changes. Therefore, we adjust the cost share of coal in electricity production to be consistent with the share of electricity generated from coal in the AEO forecast. The parameter col_ele_growth represents the share of fossil fuel inputs from coal in the electricity sector based on the AEO forecast. The reduction in the cost share is offset by an increase in the cost share of capital and labor, which would be associated with the alternative non-fossil fuel sources of generation growing in the AEO forecasts. Specifically, the intermediate, capital, and labor inputs are adjusted over time, such that

$$id0_{t,r,col,ele} = col\_ele\_factor_t id0_{0,r,col,ele}, \tag{97}$$

$$kd0_{t,r,ele} = kd0_{0,r,ele} + \left[1 - col\_ele\_factor_t id0_{0,r,col,ele}\right] \frac{kd0_{0,r,ele}}{kl0_{r,ele}}, \tag{98}$$

and

$$ld0_{t,r,ele} = ld0_{0,r,ele} + \left[1 - col\_ele\_factor_t id0_{0,r,col,ele}\right] \frac{ld0_{0,r,ele}}{kl0_{r,ele}}, \tag{99}$$

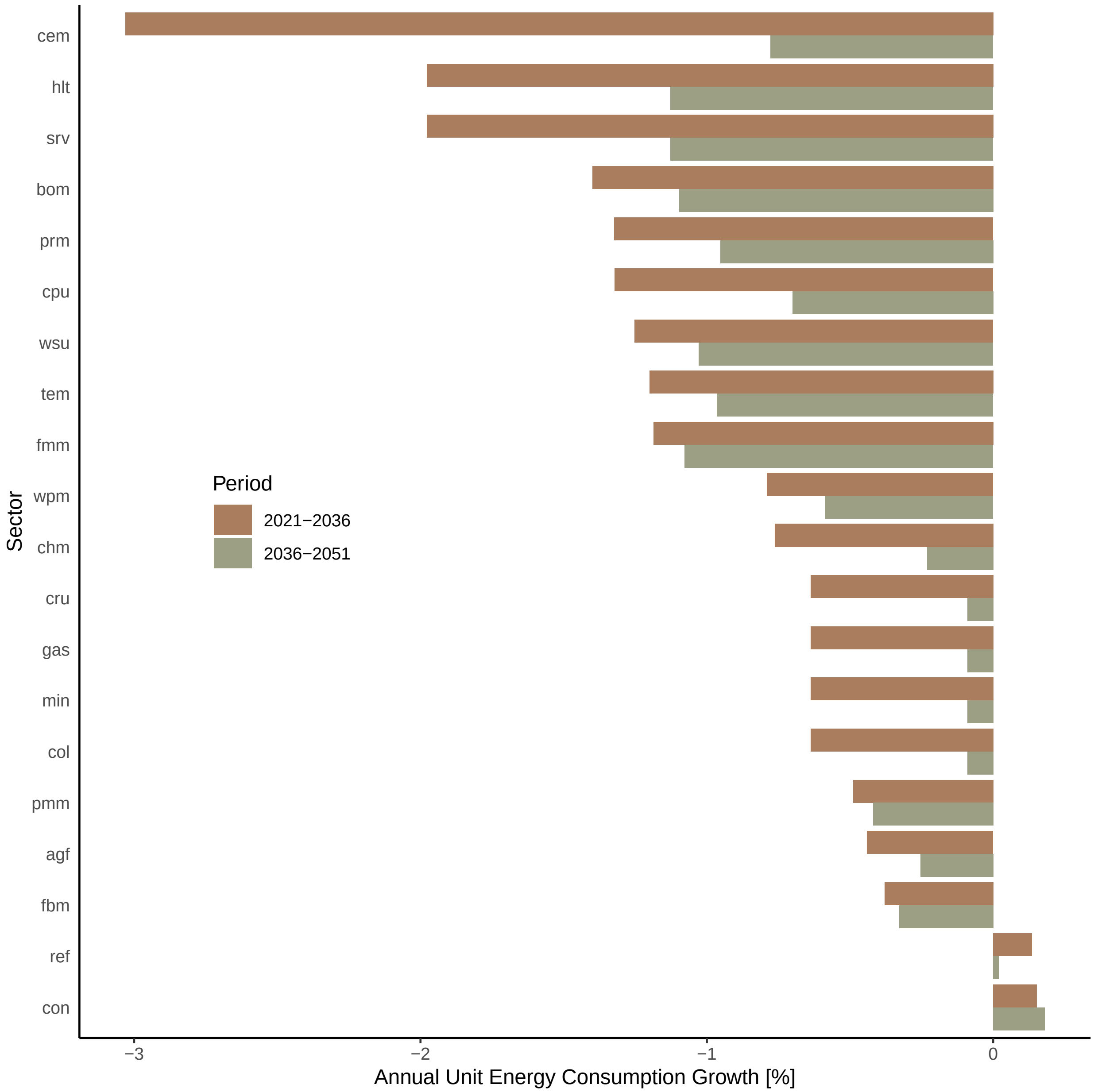


Figure 14: Calibrated Sectoral Unit Energy Consumption Growth Rates

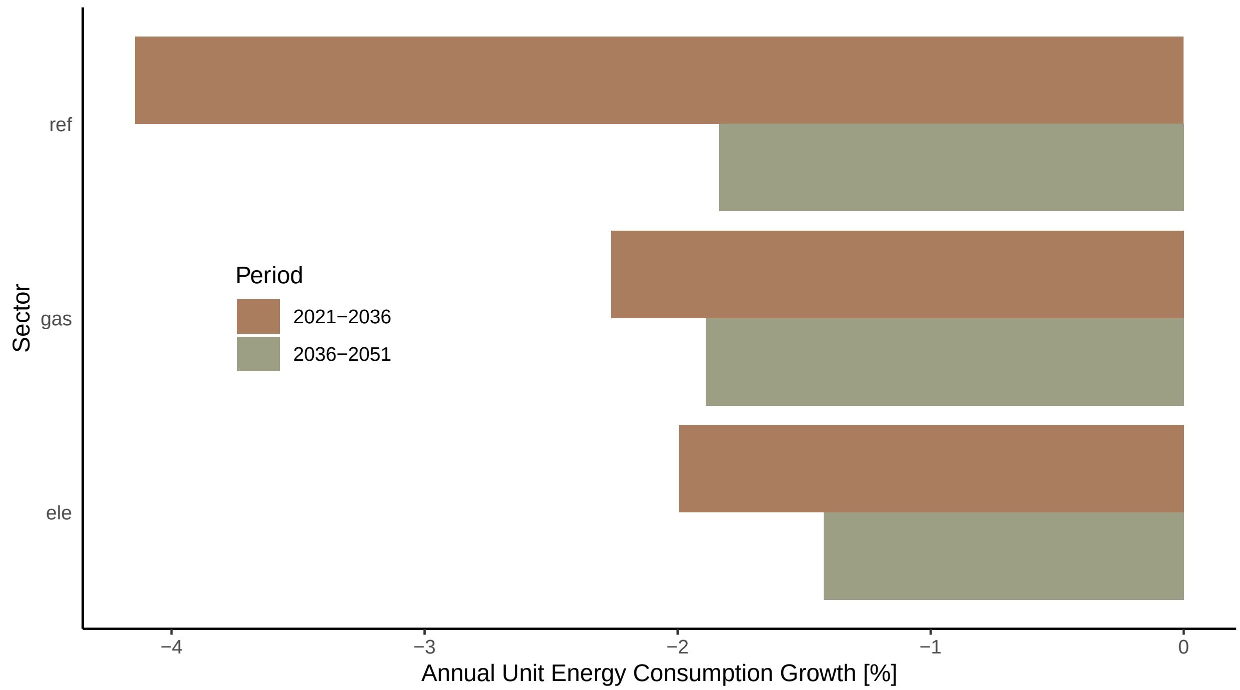


Figure 15: Calibrated Household Unit Energy Consumption Growth Rates

where

$$col\_ele\_factor_{t,s} = \frac{col\_ele\_growth_t q\_base_t - (1-\delta)^t}{q\_base_t - (1-\delta)^t}, \tag{100}$$

and q_base_t reflects an approximation of general growth of the economy by capturing the cumulative growth of the effective labor force (i.e., aggregate economy-wide labor productivity growth plus population growth). The cost shares cs_en, cs_ene, and cs_kle are also adjusted accordingly.

3.4.6 Baseline Visualization

This section presents plots for how a number of key variables and indicators evolve over the model's baseline. Figure 16 presents the growth rates of key economic variables including real GDP, the capital stock, realized economy-wide labor productivity (i.e., real GDP per unit of labor input), labor supply (i.e., total hours worked), and population.

Figure 17 presents the expenditure side GDP accounts as a percent of GDP in the baseline. Figure 18 presents the government accounts as a percent of GDP in the baseline. Figure 19 presents the foreign transaction accounts as a percent of GDP in the baseline.

Figure 20a presents the average annual growth rates of output by sector for the near- and medium-term in the baseline. While there is temporal heterogeneity in the exogenous growth rates, the figure presents average annual growth rates over two periods to more easily facilitate a comparison. Figure 20b similarly presents the average annual growth rate of real prices by commodity in the national market.

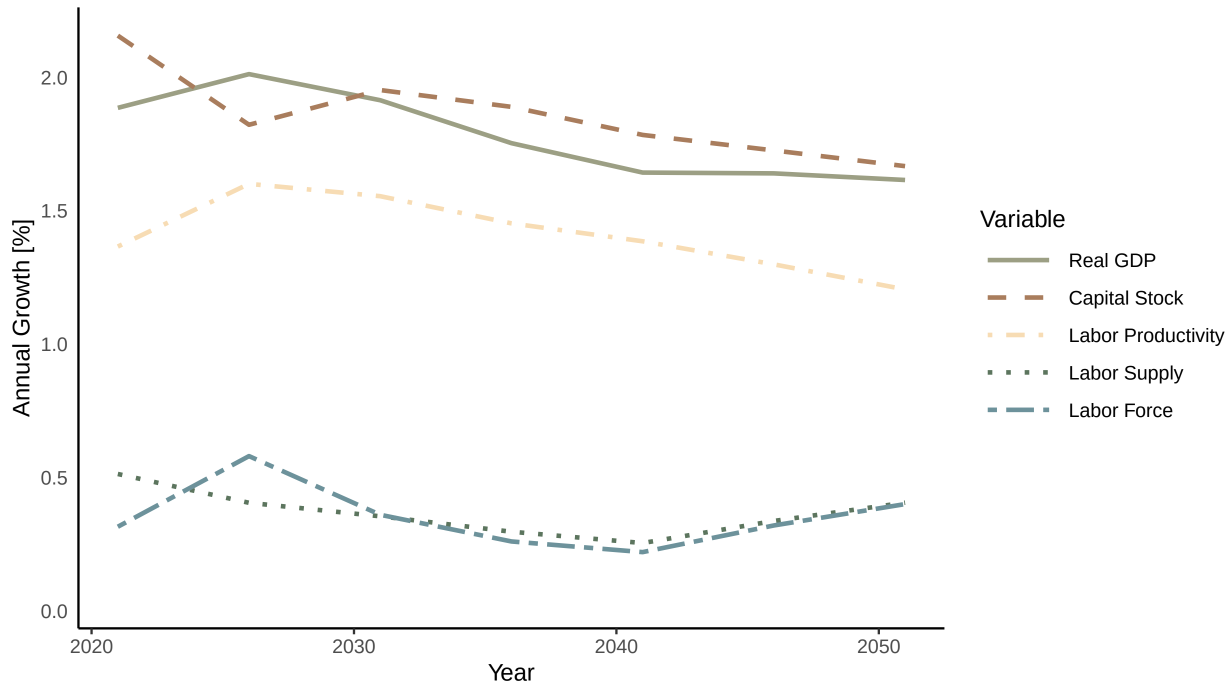


Figure 16: Growth Rate of Key Baseline Variables

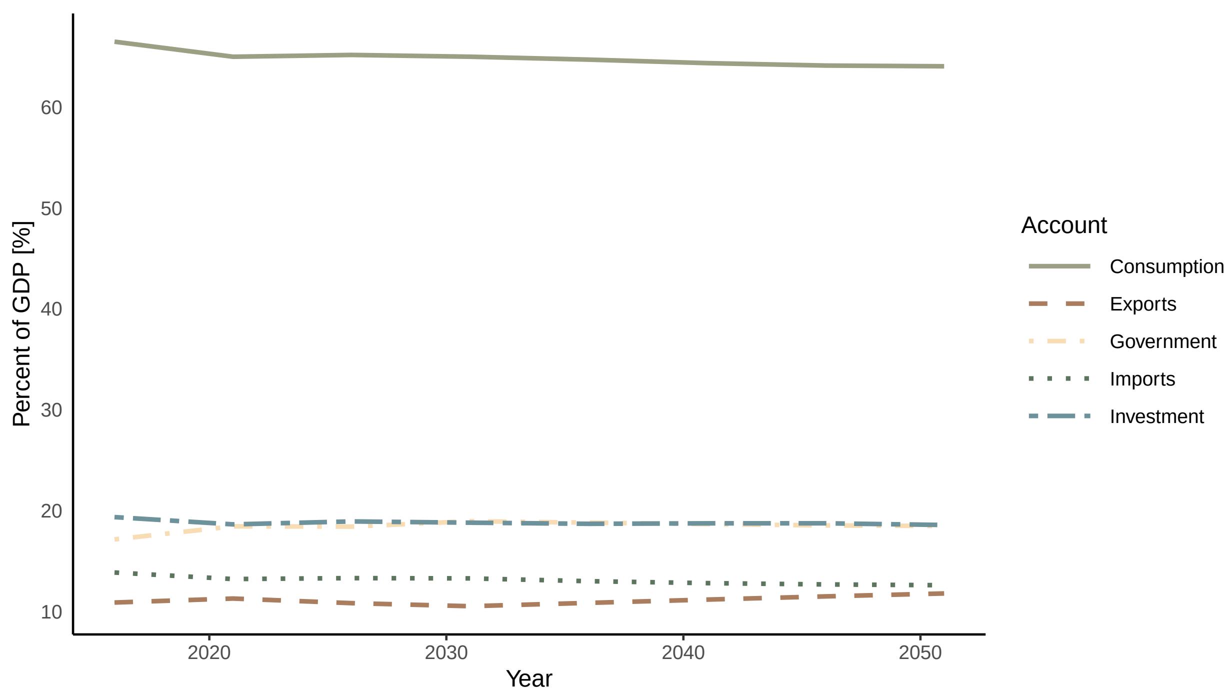


Figure 17: Expenditure-Side GDP Accounts in Baseline

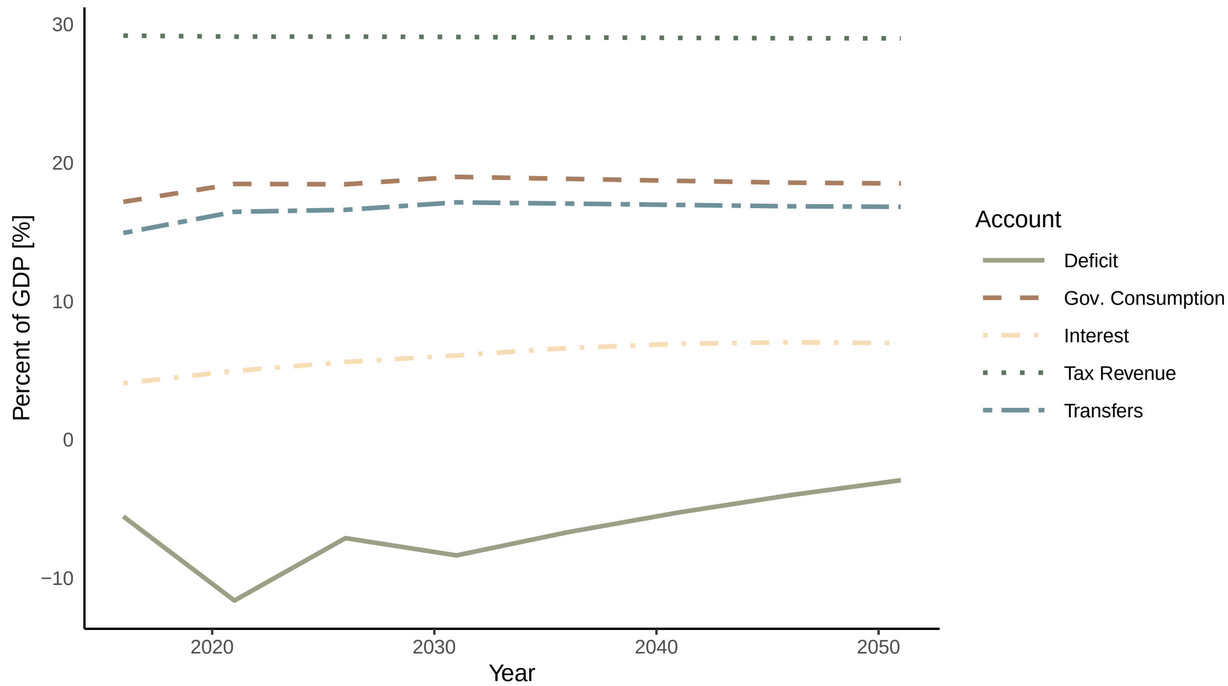


Figure 18: Government Accounts in Baseline

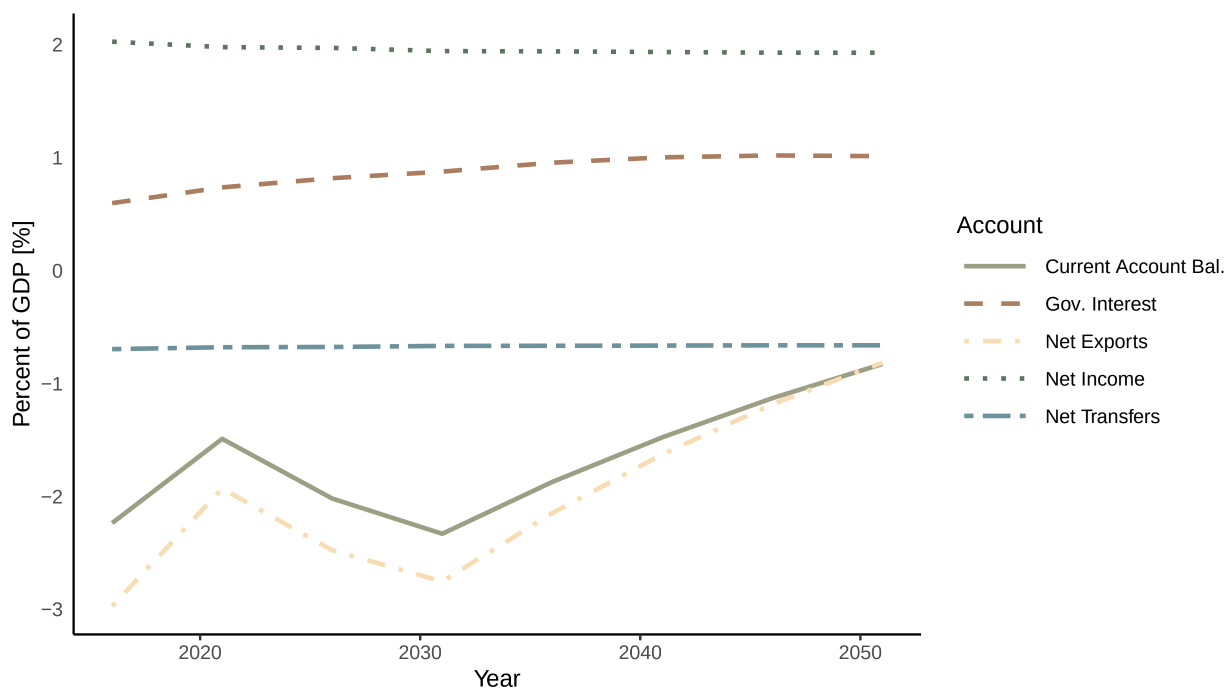
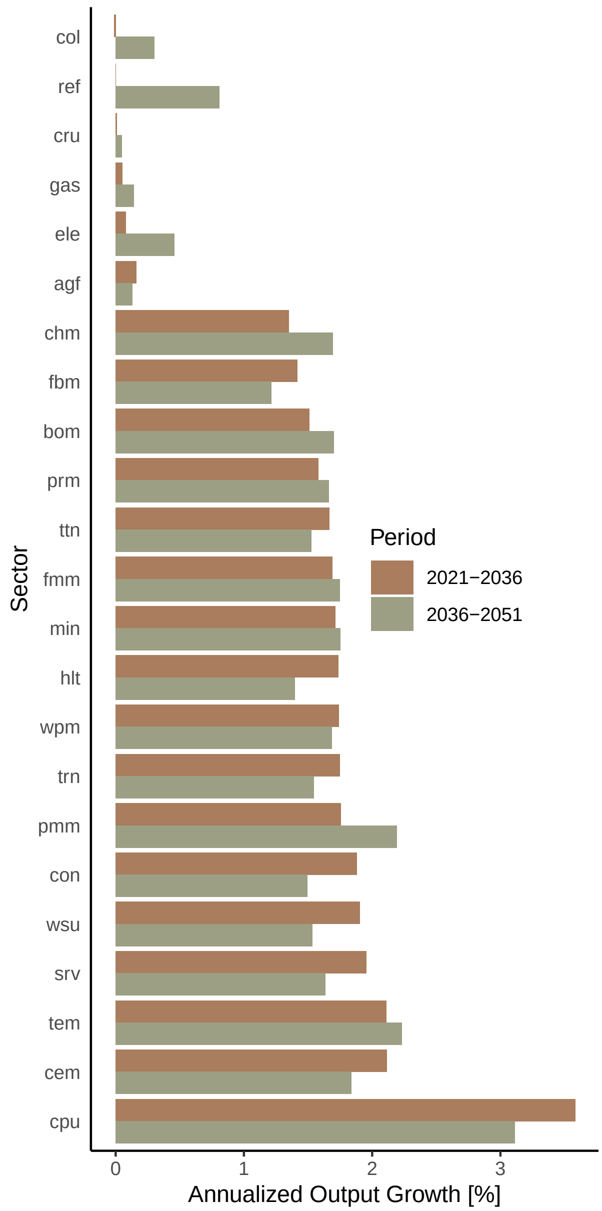
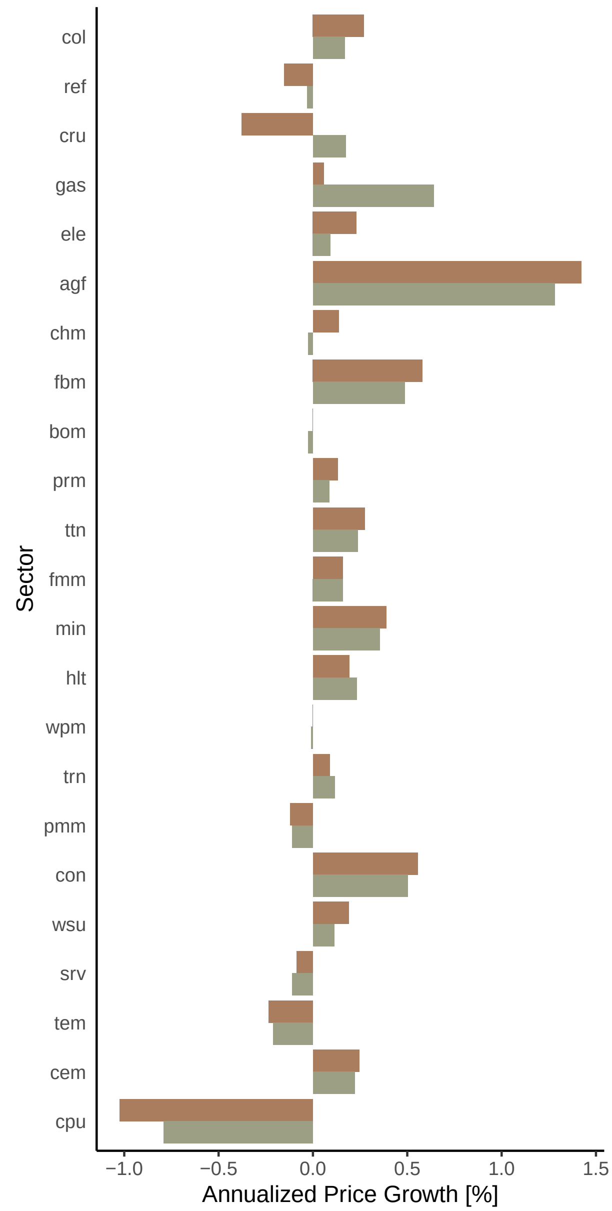


Figure 19: Foreign Accounts in Baseline



(a) Change in Sectoral Output



(b) Change in National Market Prices

Figure 20: Sectoral Change in Baseline

4 Solution

To solve the model, the primal version of the problem in Section 2 is converted to a series of non-linear equations that define profit maximizing firm behavior, welfare maximizing household behavior, market clearance, balanced budgets, and perfect competition following Mathiesen (1985) and Rutherford (1999).

Given the assumption of constant returns to scale, one can solve for the constant unit cost function of producing good z denoted as $C^z_{t,r,z}$. Perfect competition may then be represented along with profit maximization by zero-profit conditions that assume the unit cost function under optimal behavior is at least as great as the price for the good. If it is the case that the unit cost function is greater than the price such that profits are negative, it must be the case that the quantity produced is zero, providing the complementarity condition. This will hold for production with both new and extant capital and provision of the Armington aggregate, government goods, and investment. The zero-profit conditions associated with these activities are

$$C^{y}_{t,r,s}\left(pa_{t,r,agf},\ldots,pa_{t,r,srv},pr_{t,r},pres_{t,r,s},pl_{t,r},tk_{t,r},ty_{t,r,s}\right)\geq py_{t,r,i}\quad\perp\quad y_{t,r,i}\geq 0,\tag{101}$$

$$C^{y\_ex}_{t,r,i}\left(pa_{t,r,agf},\ldots,pa_{t,r,srv},pr\_ex_{t,r,s},pres_{t,r,s},pl_{t,r},tk_{t,r},ty_{t,r,s}\right)\geq py_{t,r,i}\quad\perp\quad y\_ex_{t,r,i}\geq 0,\tag{102}$$

$$C^{a}_{t,r,i}\left(pd_{t,r,i},pn_{t,i},pfx_{t}\right)\geq pa_{t,r,i}\quad\perp\quad a_{t,r,i}\geq 0,\tag{103}$$

$$C^{g}_{t,r}\left(pa_{t,r,agf},\ldots,pa_{t,r,srv}\right)\geq pgov_{t,r}\quad\perp\quad gov_{t,r}\geq 0,\tag{104}$$

$$C^{i}_{t,r}\left(pa_{t,r,agf},\ldots,pa_{t,r,srv}\right)\geq pinv_{t,r}\quad\perp\quad inv_{t,r}\geq 0,\tag{105}$$

$$C^{x}_{t,s}\left(pfix_{t,s},px_{t,s}\right)\geq pfx_{t}\quad\perp\quad xd\_fx_{t,s}\geq 0,\tag{106}$$

and

$$C^{m}_{t,s}\left(pfim_{t,s},pfx_{t}\right)\geq pm_{t,s}\quad\perp\quad ms_{t,s}\geq 0,\tag{107}$$

where $C^{y}_{t,r,s}$ is the unit cost function for production of s using new capital based on (5) and (12), $C^{y\_ex}_{t,r,s}$ is the unit cost function for production of s using extant capital based on (15), $C^{a}_{t,r,s}$ is the unit cost function for the Armington aggregate based on (1), $C^{g}_{t,r}$ is the unit cost function for the government good based on (36), $C^{i}_{t,r,s}$ is the unit cost function for the investment good based on (22), $C^{x}_{t,s}$ is the unit cost function for the open economy representation for exports (4), and $C^{m}_{t,s}$ is the unit cost function for the open economy representation of imports (3). A similar condition can be established for the "price" of full consumption

$$e_{t,r,h}\left(pa_{t,r,agf},\ldots,pa_{t,r,srv},tl_{t,r,h},tfica_{t,r,h},tc_{t,r}\right)\geq pcl_{t,r,h}\quad\perp\quad cl_{t,r,h}\geq 0,\tag{108}$$

where $e_{t,r,h}$ is the unit expenditure function for full consumption based on the intra-temporal preferences in (27). Following Section 2.7, the final zero-profit condition requires that for investors to hold capital the price must equal the present value of returns, such that

$$pk_{t,r}\geq pr_{t,r}+(1-\delta)\,pk_{t+1,r}\quad\perp\quad k_{t,r}.\tag{109}$$

A similar condition must also hold for the national new captial stock fund into which households invest, such that

$$pkh_{t}\geq prh_{t}+(1-\delta)\,pkh_{t+1}\quad\perp\quad pkh_{t}.\tag{110}$$

From Shepard's lemma the Hicksian demands for each input is the partial derivative of the unit cost function with respect to the price of the input times the level of the activity. As such, the input demands for profit maximizing firms using new capital, conditional on the equilibrium level of production, are

$$id_{t,r,ss,s} = \frac{\partial C^{y}_{t,r,s}}{\partial pa_{t,r,ss}} y_{t,r,s}, \tag{111}$$

$$kd_{t,r,s} = \frac{\partial C^{y}_{t,r,s}}{\partial pr_{t,r}} y_{t,r,s}, \tag{112}$$

$$ld_{t,r,s} = \frac{\partial C^{y}_{t,r,s}}{\partial pl_{t,r}} y_{t,r,s}, \tag{113}$$

and

$$res_{t,r,s} = \frac{\partial C^{y}_{t,r,s}}{\partial pres_{t,r,s}} y_{t,r,s}. \tag{114}$$

Similarly inputs to production using extant capital are defined as

$$id\_ex_{t,r,ss,s} = \frac{\partial C^{y\_ex}_{t,r,s}}{\partial pa_{t,r,ss}} y\_ex_{t,r,s}, \tag{115}$$

$$kd\_ex_{t,r,s} = \frac{\partial C^{y\_ex}_{t,r,s}}{\partial pr\_ex_{t,r,s}} y\_ex_{t,r,s}, \tag{116}$$

$$ld\_ex_{t,r,s} = \frac{\partial C^{y\_ex}_{t,r,s}}{\partial pl_{t,r}} y\_ex_{t,r,s} \tag{117}$$

and

$$res\_ex_{t,r,s} = \frac{\partial C^{y\_ex}_{t,r,s}}{\partial pres_{t,r,s}} y\_ex_{t,r,s}. \tag{118}$$

The inputs to the formation of capital and government consumption may be similarly defined as

$$g_{t,r,s} = \frac{\partial C^{g}_{t,r}}{\partial pa_{t,r,s}} gov_{t,r} \tag{119}$$

and

$$i_{t,r,s} = \frac{\partial C^{i}_{t,r}}{\partial pa_{t,r,s}} inv_{t,r}. \tag{120}$$

The inputs to the large open economy representation can also be defined as

$$xd_{t,s} = \frac{\partial C^{x}_{t,s}}{\partial px_{t,s}} xd\_fx_{t,s}, \tag{121}$$

$$fix_{t,s} = \frac{\partial C^{x}_{t,s}}{\partial pfix_{t,s}} xd\_fx_{t,s}, \tag{122}$$

$$ms\_fx_{t,s} = \frac{\partial C^{m}_{t,s}}{\partial pfx_{t}} ms_{t,s}, \tag{123}$$

and

$$fim_{t,s} = \frac{\partial C^m_{t,s}}{\partial pfim_{t,s}} ms_{t,s}. \tag{124}$$

Given the equilibrium level of full consumption, the demands for final discretionary consumption goods are

$$dcd_{t,r,h,s} = \frac{\partial e_{t,r,h}}{\partial pa_{t,r,s}} cl_{t,r,h}, \tag{125}$$

where leisure demand can be similarly defined as

$$leis_{t,r,h} = \frac{\partial e_{t,r,h}}{\partial pl_{t,r}} cl_{t,r,h}. \tag{126}$$

Total commodity demands are defined as,

$$cd_{t,r,h,s} = dcd_{t,r,h,s} + scd_{t,r,h,s}, \tag{127}$$

where $scd_{t,r,h,s}$ is exogenous. Imports and domestically-sourced use are defined conditional on the equilibrium level of the Armington aggregate as

$$d_{t,r,s} = \frac{\partial C^a_{t,r,s}}{\partial pd_{t,r,s}} a_{t,r,s}, \tag{128}$$

$$m_{t,r,s,dtrd} = \frac{\partial C^a_{t,r,s}}{\partial pn_{t,r}} a_{t,r,s}, \tag{129}$$

and

$$m_{t,r,s,ftrd} = \frac{\partial C^a_{t,r,s}}{\partial pm_{t,s}} a_{t,r,s}. \tag{130}$$

Exports are determined from the CET function in (2), such that

$$x_{t,r,s,dtrd} = \frac{y\_ex_{t,r,s} + y_{t,r,s}}{y0_{r,s}} \left(\frac{pn_{t,r}}{py_{t,r,s}}\right)^{te\_dx} \tag{131}$$

and

$$x_{t,r,s,ftrd} = \frac{y\_ex_{t,r,s} + y_{t,r,s}}{y0_{r,s}} \left(\frac{px_{t,s}}{py_{t,r,s}}\right)^{te\_dx}, \tag{132}$$

where the right hand side defines the optimal share of output supplied to the export markets based on the output transformation function in (2).

Given the Hicksian demands conditional on equilibrium activity levels, the market clearance conditions in Section 2.6 can be defined. If any of the conditions in (39)-(52) holds with strict inequality it would imply that supply exceeds demand in equilibrium, such that the price of that activity's output must be zero. This leads to a series of complementarity conditions, which define the market clearance conditions. The price of the Armington aggregate, $pa_{t,r,s}$, clears the goods market, such that

$$a_{t,r,s} \geq \sum_{ss}(id_{t,r,s,ss} + id\_ex_{t,r,s,ss}) + \sum_{h}(cd_{t,r,s,h}) + i_{t,r,s} + g_{t,r,s} \quad \perp \quad pa_{t,r,s} \geq 0. \tag{133}$$

The price of domestic output consumed domestically, $pd_{t,r,s}$, clears the domestic market, such that

$$\frac{y\_ex_{t,r,s} + y_{t,r,s}}{y0_{r,s}} \left(\frac{pd_{t,r,s}}{py_{t,r,s}}\right)^{te\_dx} \geq \frac{d_{t,r,s}}{d0_{r,s}} \quad \perp \quad pd_{t,r,s} \geq 0, \tag{134}$$

where the left hand side defines the optimal share of output supplied to the domestic market based on the output transformation function in (2). The price of labor, $pl_{t,r}$, (i.e., the wage rate) clears the labor market, such that

$$\sum_{h} l_{t,r,h} \geq \sum_{s} ld_{t,r,s} + ld\_ex_{t,r,s} \quad \perp \quad pl_{t,r} \geq 0. \tag{135}$$

The rental rate for sector specific extant capital, $pr\_ex_{t,r,s}$, clears the market for extant capital, such that

$$\frac{k\_ex_{t,r}}{k0_{r}} \left(\frac{pr\_ex_{t,r,s}}{pr\_ex\_agg_{t,r}}\right)^{te\_k\_ex} \geq \frac{kd\_ex_{t,r,s}}{kd0_{r,s}} \quad \perp \quad pr\_ex_{t,r,s} \geq 0 \tag{136}$$

where the left hand side defines the optimal share of extant capital supplied to sector s based on the extant transformation function in (19). The national weighted average rental rate households earn on their past investments in the extant capital stock is

$$prh\_ex_{t} \geq \frac{\sum_{r} pr\_ex\_agg_{t,r} k\_ex_{t,r}}{\sum_{r} k\_ex_{t,r}} \quad \perp \quad prh\_ex_{t} \geq 0. \tag{137}$$

The regional rental rate for new capital, $pr_{t,r}$, clears the market for new capital in that region, such that

$$k_{t,r} \geq \sum_{s} kd_{t,r,s} \quad \perp \quad pr_{t,r} \geq 0. \tag{138}$$

The price of regional new capital, $pk_{t,r}$, clears the investment market in that region, such that

$$k_{t-1,r}\,(1-\delta) + inv_{t-1,r} \geq k_{t,r} \quad \perp \quad pk_{t,r} \geq 0. \tag{139}$$

The rental rate for new capital investments clears the national market for new capital, ensuring the households savings are consistent with investment in new capital across the regions, such that

$$\sum_{r,h} kh_{t,r,h} \geq \sum_{r} k_{t,r} \quad \perp \quad prh_{t} \geq 0. \tag{140}$$

The price of foreign exchange, pfx_t, clears the foreign exchange market, such that

$$\sum_s ms_{t,s} + row_t \geq \sum_s xd_{t,s,ftrd} + inc\_row_t + tran\_row_t - gint\_row_t - curactbal_t \quad \perp \quad pfx_t \geq 0. \tag{141}$$

The price of internationally imported commodities, $pm_{t,s}$ clears the import market

$$\sum_r m_{t,r,s,ftrd} \geq ms_{t,s} \quad \perp \quad pm_{t,s} \geq 0. \tag{142}$$

The price of internationally exported commodities, $px_{t,s}$ clears the export market

$$\sum_r x_{t,r,s,ftrd} \geq xd_{t,s} \quad \perp \quad px_{t,s} \geq 0. \tag{143}$$

The price of the fixed factor for internationally imported commodities, $pfim_{t,s}$ clears the market

$$fim0_s \geq fim_{t,s} \quad \perp \quad pfim_{t,s} \geq 0. \tag{144}$$

The price of the fixed factor for internationally exported commodities, $pfix_{t,s}$ clears the market

$$fix0_s \geq fix_{t,s} \quad \perp \quad pfix_{t,s} \geq 0. \tag{145}$$

The price of commodities on the national market, $pn_{t,s}$, clears the market for national trade

$$\sum_r x_{t,r,s,dtrd} \geq \sum_r m_{t,r,s,dtrd} \quad \perp \quad pn_{t,s} \geq 0. \tag{146}$$

The rental rate for sector specific fixed factors, $pres_{t,r,s}$, clears the market for sector specific fixed factors, such that

$$\sum_h rese_{t,r,s,h} \geq res_{t,r,s} + res\_ex_{t,r,s} \quad \perp \quad pres_{t,r,s} \geq 0. \tag{147}$$

In addition, the problem requires that households maximize intertemporal welfare in (23). The Karush-Kuhn-Tucker conditions for the welfare maximization problem are

$$\left(\frac{cl_{t,r,h}}{n_{t,r,h}}\right)^{-\eta} \geq \lambda_{t,r,h} pcl_{t,r,h} \quad \perp \quad cl_{t,r,h} \geq 0, \tag{148}$$

$$\beta_{t+1,r,h}\lambda_{t+1,r,h} \geq \beta_{t,r,h}\lambda_{t,r,h} \quad \perp \quad kh_{t+1,r,h} \geq 0, \tag{149}$$

and

$$
\begin{aligned}
pkh_{t+1}invh_{t,r,h} \geq & prh_t kh_{t,r,h} + (1 - tl_{t,r,h} - tfica_{t,r,h})\, pl_{t,r} te_{t,r,h} \\
& + prh\_ex_t kh\_ex_{t,r,h} + \sum_s pres_{t,r,s} rese_{t,r,s,h} \\
& + pfx_t \,(inc\_row_{t,r,h} + tran\_row_{t,r,h} + curactbal_{t,r,h})\, int\_share_{r,h} \\
& + cpi_t transfers_{t,r,h} + cpi_t \,(gint_{t,r,h} - deficit_{t,r,h} + icnadj_{t,r,h})\, g\_share_{r,h} \\
& + (tl_{t,r,h} - tl\_avg_{t,r,h})\, pl_{t,r} l_{t,r,h} \\
& - pcl_{t,r,h} cl_{t,r,h} - \sum_s (1 + tc_{t,r})\, pa_{t,r,s} scd_{t,r,s,h} \\
& \quad \perp \quad \lambda_{t,r,h} \geq 0.
\end{aligned}
\tag{150}
$$

where the level of labor supply is determined by the time constraint, such that

$$
te_{t,r,h} \geq leis_{t,r,h} + l_{t,r,h} \quad \perp \quad l_{t,r,h} \geq 0, \tag{151}
$$

and the new capital stock held by households is determine by investment according to

$$
kh_{t,r,h}\,(1 - \delta) + invh_{t,r,h} \geq kh_{t+1,r,h} \quad \perp \quad inv_{t,r,h} \geq 0. \tag{152}
$$

The terminal level of capital $kh_{T+1,r,h}$ is equal to $kht_{r,h}$, which is define below by the model closures.

The problem requires that the government budget constraint holds, as described in Section 2.5, such that

$$
\begin{aligned}
\sum_r pgov_{t,r} gov_{t,r} + & \sum_h cpi_t transferst, r, h + cpi_t gint_t + pfx_t gint\_row_t + cpi_t incadj_t \\
\geq \sum_r \sum_s \Bigg\{ & \; ty_{t,r,s} py_{t,r,s}\,(y_{t,r,s} + y\_ex_{t,r,s}) \\
& + tk_{t,r}\,[pr_{t,r} kd_{t,r,s} + pr\_ex_{t,r,s} kd\_ex_{t,r,s} + pres_{t,r,s}\,(res_{t,r,s} + res\_ex_{t,r,s})] \Bigg\} \\
+ \sum_r \sum_h & \Bigg[(tl_{t,r,h} + tfica_{t,r,h})\, pl_{t,r} l_{t,r,h} + tc_{t,r} pa_{t,r,s} cd_{t,r,s,h} \Bigg] + cpi_t deficit_{t,r,h} \\
& \perp \quad incadj_t \geq 0.
\end{aligned}
\tag{153}
$$

It must also be the case that the rest of world budget constraint associated with income from the fixed factors in the reduced form large open economy specification holds, such that

$$
pfx_t row_t \geq \sum_s \Bigg(pfix_{t,s} fix0_s + pfim_{t,s} fim0_s \Bigg) \perp \quad row_t \geq 0. \tag{154}
$$

Finally, we include the conditions to close the finite time approximation to the infinite time problem. As noted in Section 2.7, the post-terminal capital stock is determined by requiring that investment grows at the rate of aggregate consumption growth, such that

$$\frac{inv_{T,r}}{inv_{T-1,r}} \geq \frac{\sum_s y_{T,r,s} + y\_ex_{T,r,s}}{\sum_s y_{T-1,r,s} + y\_ex_{T-1,r,s}} \quad \perp \quad kt_r \geq 0. \tag{155}$$

The price is determined based on the law of motion for capital, such that

$$k_{T,r}\,(1-\delta) + inv_{T,r} \geq kt_r \quad \perp \quad pkt_r \geq 0, \tag{156}$$

where households' share of the post-terminal capital stock is assumed to be equivalent to their shares of the capital stock in the last period of the model

$$\frac{kh_{T,r,h}}{\sum_{r,h} kh_{T,r,h}} \geq kht_{r,h} \quad \perp \quad kht_{r,h} \geq 0 \tag{157}$$

and price of terminal capital for the households is equal to the average price of terminal capital

$$pkht \geq \frac{\sum_r pkt_r kt_r}{\sum_r kt_r} \quad \perp \quad pkht \geq 0. \tag{158}$$

The equations (101)-(158) define the equilibrium conditions of the model. The problem is formulated in the General Algebraic Modeling System (GAMS).[49] The model is solved using the PATH solver (Ferris and Munson, 2000). We set the numeraire to the price of foreign exchange, pfx_0, in the initial period.[50]

In this documentation all variables are defined in levels for ease of exposition and interpretation. The model in the code is mathematically equivalent to what is laid out in this documentation. However, in the implementation most variables are defined as indices relative to the benchmark value instead of in levels. This provides for a fairly well scaled problem with only limited need for scaling of equations and variables prior to the solve. This implementation does not affect the model solution, but does mean that some of the equations as implemented in the code may differ slightly from what is laid out in the documentation.

[49] GAMS Development Corporation. General Algebraic Modeling System (GAMS) Washington, DC, USA, 2014.

[50] SAGE solves for a set of *relative prices* through the selection of a numeraire, or a chosen price level used to denominate other prices in the model. Here, price changes are characterized relative to the foreign exchange rate in the initial period. A numeraire is required in this class of economic models to satisfy Walras' Law. A competitive general equilibrium is homogeneous of degree one in prices, meaning that the equilibrium level price vector scaled by a common factor is also a solution to the model. This indeterminacy is solved by fixing a single price to align the number of equations and variables in the model. In SAGE, we drop equation 141 for $t = 0$, though this condition is verified to hold post-solve. Notably, in intertemporal dynamic CGE models, only one price level in a single year is needed as a numeraire. For recursive dynamic formulations, a numeraire price can be assigned in every period if the model is homogeneous of degree one in prices within each period.

4.1 Multi-Year Timesteps

In this documentation the equations are written assuming an annual timestep for simplicity. However, solving the default version of the model at an annual timestep for a long enough time horizon to achieve convergence to the steady state is not computational feasible. Therefore, we use multi-year time steps to reduce the computational burden. To do so we adopt a step function approach to defining investment investment over a timestep. When defining the laws of motion for regional capital in equation (139) this translates to

$$k_{t-1,r}\left(1-\delta\right)^{interval_t} + interval_t inv_{t-1,r} \geq k_{t,r} \quad \perp \quad pk_{t,r} \geq 0, \tag{159}$$

where $interval_t$ is the number of years between model period t and $t+1$. Consistent with this adjustment the intertemporal no arbitrage condition in equation (109) needs to be updated to

$$pk_{t,r} \geq interval_t pr_{t,r} + \left(1-\delta\right)^{interval_t} pk_{t+1,r} \quad \perp \quad k_{t,r}. \tag{160}$$

Similar adjustment is required for the laws of motion in equations (152) and (156) and the intertemporal no arbitrage condition in equation (110).

Additional straight forward adjustments are needed to the exogenous temporal parameters in the model and the household discount factor relative to the annual timestep presentation in this documentation. These adjustments are made automatically by the model based on the simulation years selected. Section 6.5.1 provides details on setting the timesteps and time horizon when using the model.

4.2 Calculating Welfare Effects

Households' willingness to pay to avoid the costs of the policy requirements, that is the social costs associated with the policy, are estimated using equivalent variation (EV). EV is estimated as the amount of additional income households would require under baseline prices and still achieve the same level of welfare as simulated in the policy case. More specifically, EV is calculated as the difference between expenditures (on all goods including leisure) in the case where households face baseline prices but are constrained to the welfare level achieved in the policy case and expenditures in the baseline. This will be positive if the components of the policy modeled are on net an improvement in welfare and negative if they are on net a reduction in welfare.

The households' optimization problem described by (23)-(32) yields optimal levels of consumption and leisure given prices, taxes, transfers, and shadow prices on their budget constraint, $\lambda_{t,r,h}$. For simplicity of exposition, let $\mathbf{z}^{sim}_{r,h}$ be a vector of all prices, taxes, and transfers faced by household h in region r, where *sim* denotes the given simulation: *base* for baseline and *pol* for policy. The optimal levels of consumption and leisure are then given by

$$cd^{sim}_{t,r,s,h} = dcd_{t,r,s,h}\left(\mathbf{z}^{sim}_{r,h}, \boldsymbol{\lambda}^{sim}_{r,h}\right) + scd_{t,r,s,h} \tag{161}$$

and

$$leis^{sim}_{t,r,h} = leis_{t,r,h}\left(\mathbf{z}^{sim}_{r,h}, \boldsymbol{\lambda}^{sim}_{r,h}\right), \tag{162}$$

where $\boldsymbol{\lambda}^{sim}_{r,h}$ is a vector of shadow prices over the simulation's time horizon. Households' expenditures on full consumption are then defined as

$$expenditure\left(\mathbf{z}^{sim}_{r,h}, \boldsymbol{\lambda}^{sim}_{r,h}\right) = \sum_t \left(1 - tl^{sim}_{t,r,h} - tfica^{sim}_{t,r,h}\right) pl^{sim}_{t,r} leis^{sim}_{t,r,h} + \sum_s \left(1 + tc^{sim}_{t,r}\right) pa^{sim}_{t,r,s} cd^{sim}_{t,r,s,h}, \tag{163}$$

noting that since the prices are relative to the numeraire in the initial period they are already in present value terms.

From the first order conditions to the household optimization problem in (149), the evolution of the shadow price is defined as

$$\beta \lambda^{sim}_{t+1,r,h} = \lambda^{sim}_{t,r,h}, \tag{164}$$

where β is the discount factor defined in (24). Therefore, given a terminal value $\lambda_{T,r,h}$, the sequences of shadow prices can be determined. Given a vector of prices, taxes, and transfers and a vector of shadow prices, (161) and (162) define the paths of consumption and leisure. Based on those paths, (23) defines the households' welfare. Computing EV is therefore reduced to a problem of finding a value $\lambda^{ev}_{T,r,h}$ that, along with $\mathbf{z}^{base}_{r,h}$, leads to a level of welfare equal to the level in the policy simulation. Given this value, define

$$cd^{ev}_{t,r,s,h} = dcd_{t,r,s,h}\left(\mathbf{z}^{base}_{r,h}, \boldsymbol{\lambda}^{ev}_{r,h}\right) + scd_{t,r,s,h} \tag{165}$$

and

$$leis^{ev}_{t,r,h} = leis_{t,r,h}\left(\mathbf{z}^{base}_{r,h}, \boldsymbol{\lambda}^{ev}_{r,h}\right). \tag{166}$$

EV is then defined as

$$EV_{r,h} = expenditure\left(\mathbf{z}^{base}_{r,h}, \boldsymbol{\lambda}^{ev}_{r,h}\right) - expenditure\left(\mathbf{z}^{base}_{r,h}, \boldsymbol{\lambda}^{base}_{r,h}\right). \tag{167}$$

The value of EV defined in (167) is reported in the model output under the name `ev` and is based on the simulation years in the model and a finite time horizon. Two additional measures of EV are also standard outputs for a policy simulation. These include `ev_annual`, which linearly interpolates (163) between simulation years to estimate EV over all years covered by the policy simulation, and `ev_inf`, which extends `ev_annual` from a finite to an infinite time horizon based on the assumption that quantities and prices follow their steady state paths after the terminal period in the model. Measures of baseline expenditures on full consumption (consumption and leisure) are also output under the names `cl_base`, `cl_base_annual`, and `cl_base_inf`, with analogous temporal definitions, so that the user may compute EV as a share of baseline full consumption if desired.

5 Modeling Regulatory Requirements

Environmental regulations can vary over many dimensions. For example, EPA (2015) describe four categories of regulations commonly promulgated to address air pollution: single sector emission rate limits or technology standards; regional or state-implemented emission targets; multi-sector boiler or engine-level emission limits or technology standards; and federal product standards. Each of these categories has unique characteristics that may affect how the compliance requirements of the regulations and incentives created by the regulation should be modeled. Environmental regulations addressing additional pathways for pollution (e.g., land, water) have many similarities with the aforementioned categories but also have additional attributes that may be relevant for how they are modeled. Thus, the appropriate approach to introduce the requirements of a regulation into the model will depend on the specific details of the policy.

In practice, with the exception of federal product standards or prohibitions, environmental regulations are typically source-level technology standards, performance-based emission-rate limits, or workplace standards. In each case, sources are required to undertake abatement and monitoring activities in addition to their regular production activities. As a starting point for reflecting heterogeneity across regulatory approaches, the default version of the model has two built-in approaches for simulating abatement requirements on producers that may be calibrated to engineering or partial equilibrium estimates of compliance costs: a productivity shock and an explicit abatement activity. Each approach is discussed below, followed by an example that highlights the differences between them.

5.1 Compliance Requirements as a Productivity Shock

The production functions in the model, as described in equations (5)-(10), are of the calibrated share form in which the inputs are entered relative to the benchmark values. In other words, the production function in (5)-(10) can be described as

$$
\begin{aligned}
y_{t,r,s,v} = f_y\Bigg(& f_{mat}\left(\frac{id_{t,r,agf,s,v}}{id0_{r,agf,s}}, \dots, \frac{id_{t,r,srv,s,v}}{id0_{r,srv,s}}\right), \\
& f_{kle}\Bigg(f_{ene}\Bigg(\frac{id_{t,r,ele,s,v}}{id0_{r,ele,s}}, \\
& \qquad f_{en}\left(\frac{id_{t,r,col,s,v}}{id0_{r,col,s}}, \dots, \frac{id_{t,r,gas,s,v}}{id0_{r,gas,s}},\right)\Bigg), \\
& \qquad f_{kl}\left(\frac{kd_{t,r,s,v}}{kd0_{r,s}}, \frac{ld_{t,r,s,v}}{ld0_{r,s}}\right)\Bigg)\Bigg).
\end{aligned}
\tag{168}
$$

To generalize this discussion, (168) introduces the index $v \in (new, extant)$ to describe the vintage of capital used in the production function. Under the case $v = extant$, (168) represents the Leontief production function for production with extant capital implicitly described by (15)-(18).

In the code an additional parameter, $prod\_ind_{t,r,z,s,v}$, is introduced to allow for modeling of a productivity shock on input z. The implementation essentially redefines (168) as

$$
\begin{aligned}
y_{t,r,s,v} = f_y\Bigg(& f_{mat}\left(\frac{id_{t,r,agf,s,v}}{prod\_ind_{t,r,agf,s,v}id0_{r,agf,s}}, \ldots, \frac{id_{t,r,srv,s,v}}{prod\_ind_{t,r,srv,s,v}id0_{r,srv,s}}\right), \\
& f_{kle}\Bigg(f_{ene}\Bigg(\frac{id_{t,r,ele,s,v}}{prod\_ind_{t,r,ele,s,v}id0_{r,ele,s}}, \\
& \qquad f_{en}\left(\frac{id_{t,r,col,s,v}}{prod\_ind_{t,r,col,s,v}id0_{r,col,s}}, \ldots, \frac{id_{t,r,gas,s,v}}{prod\_ind_{t,r,gas,s,v}id0_{r,gas,s}}\right)\Bigg), \\
& \quad f_{kl}\left(\frac{kd_{t,r,s,v}}{prod\_ind_{t,r,k,s,v}kd0_{r,s}}, \frac{ld_{t,r,s,v}}{prod\_ind_{t,r,l,s,v}ld0_{r,s}}\right)\Bigg)\Bigg),
\end{aligned}
\tag{169}
$$

where in the baseline $prod\_ind_{t,r,z,s,v} = 1$, in which case (168) and (169) are equivalent.

The interpretation of this additional parameter is that increasing $prod\_ind$ for a specific input from 1 to $1 + \Delta$ and holding all other inputs fixed would require a $\Delta \times 100\%$ increase in the affected input to continue producing the baseline level. Based on this interpretation one can define an approach to calibrating the value of $prod\_ind$ to reflect the compliance requirements associated with a regulation. For example, suppose an engineering cost analysis estimates that a regulation impacting sector s will require additional expenditures on input z of $cost_{t,r}$ in year t and region r to produce the baseline level of output at new production sources, $y_{t,r,s,new}$. This may be represented by

$$
prod\_ind_{t,r,z,s,new} = 1 + \frac{cost_{t,r}}{(1+\tau_z)\, z0_{r,s}} \frac{y0_{r,s}}{y_{t,r,s,new}}, \tag{170}
$$

where $z0_{r,s}$ is the benchmark value of input z and τ_z represents any potential ad valorem tax on input z paid by producers (e.g., taxes on capital returns). This calibration would yield a situation where holding the output level and all other inputs fixed at their baseline levels, consistent with the setup in most engineering cost analyses, would require additional expenditures (gross of taxes) of $cost_{t,r}$ on input z. However, it should be noted that, after implementing the shock, firms may substitute away from the now less productive input towards other inputs. Based on the nature of the productivity shock these implicit substitution possibilities in the compliance activity are defined by the substitution elasticities in the regulated sector's production function.

5.2 Modeling Explicit Compliance Requirements

The model also allows for the explicit specification of input requirements for regulatory compliance. This is accomplished by extending the nesting structure of the production function depicted in Figures 4-6 to include a top level Leontief nest that combines production of saleable goods and services with pollution abatement activities. For the standard manufacturing and services production functions with new capital, this extended production function is presented in Figure 21. Production then requires both the traditional production activity and an abatement activity, which is itself a

Leontief function of inputs used in regulatory compliance.

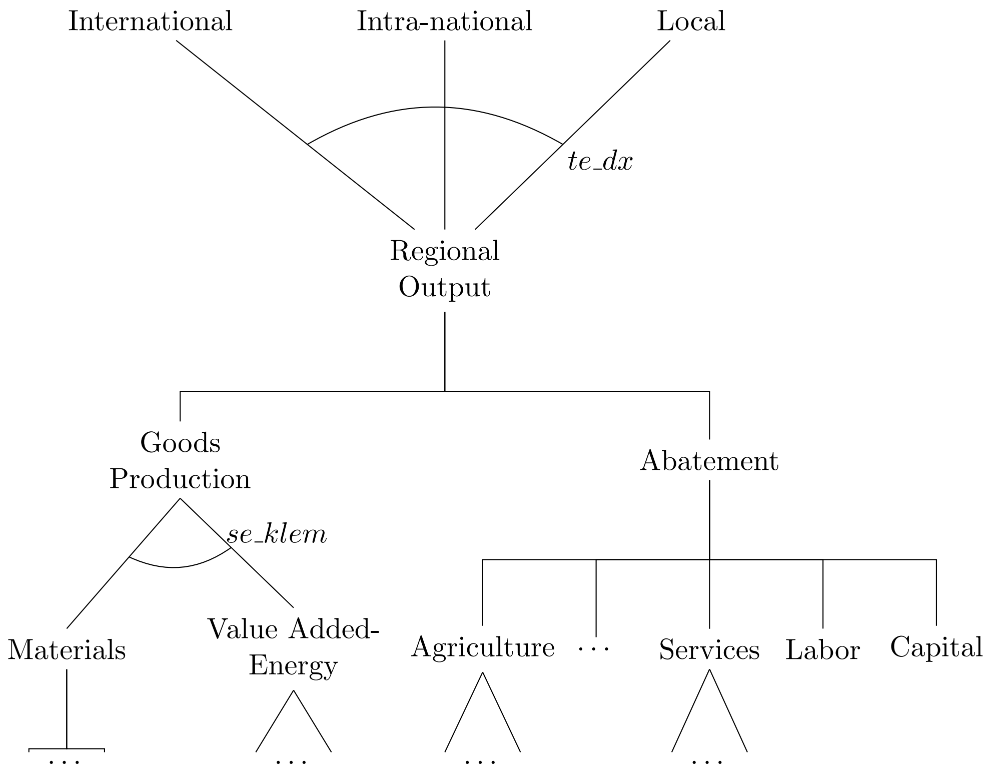


Figure 21: Manufacturing and Services Production Functions with Abatement

With the implementation of the extended production function to account for potential compliance activities, output in the manufacturing and service sectors that is not associated with a fixed factor resource is defined as

$$y_{t,r,s} = y0_{r,s} \min\left(\frac{klem_{t,r,s}}{klem0_{t,r,s}}, \frac{abate_{t,r,s}}{abate0_{t,r,s}}\right), \tag{171}$$

where $klem_{t,r,s}$ represents the traditional production activity defined in Section 2.2.1, such that

$$\begin{aligned} klem_{t,r,s} = klem0_{r,s} \Bigg[& cs\_klem_{r,s} \left(\frac{mat_{t,r,s}}{mat0_{t,r,s}}\right)^{\frac{se\_klem-1}{se\_klem}} \\ & + (1 - cs\_klem_{r,s}) \left(\frac{kle_{t,r,s}}{kle0_{t,r,s}}\right)^{\frac{se\_klem-1}{se\_klem}} \Bigg]^{\frac{se\_klem}{se\_klem-1}}. \end{aligned} \tag{172}$$

The abatement activity is defined as a Leontief function of intermediate inputs, labor, and new capital, such that

$$abate_{t,r,s} = abate0_{t,r,s} \min \left(\frac{id\_abate_{t,r,agf,s}}{id\_abate0_{t,r,agf,s}}, \ldots, \frac{id\_abate_{t,r,srv,s}}{id\_abate0_{t,r,srv,s}}, \frac{ld\_abate_{t,r,s}}{ld\_abate0_{t,r,s}}, \frac{kd\_abate_{t,r,s}}{kd\_abate0_{t,r,s}} \right), \tag{173}$$

where $id\_abate_{t,r,ss,s}$, $ld\_abate_{t,r,s}$, and $kd\_abate_{t,r,s}$ are inputs of commodity ss, labor, and capital for abatement activities, respectively. The "benchmark" values in (173) include time subscripts because the required level of abatement activities or the inputs associated with the abatement activity may change over time, for example due to a phase in of the regulation. With the extended production function that includes abatement activities, firms are assumed to maximize profits inclusive of the abatement inputs,

$$\begin{aligned} &(1 - ty_{t,r,s})\, py_{t,r,s} y_{t,r,s} - \sum_{ss} pa_{t,r,ss} \left(id_{t,r,ss,s} + id\_abate_{t,r,ss,s} \right) \\ &- (1 + tk_{t,r})\, pr_{t,r} \left(kd_{t,r,s} + kd\_abate_{t,r,s} \right) - pl_{t,r} \left(ld_{t,r,s} + ld\_abate_{t,r,s} \right), \end{aligned} \tag{174}$$

subject to the production function defined by (171)-(173) and (6)-(10). Similar extensions are implemented for production with extant capital and sectors associated with fixed factor resources. In the case of production associated with extant capital, abatement activities are still assumed to make use of new capital in (173).

The solution approach outlined in Section 4 is easily extended to accommodate the expanded production structure inclusive of abatement activities. While the default implementation represents abatement activities as Leontief technologies, alternative functional forms can be adopted if warranted. It is also possible to represent abatement activities as substitutes for emissions in a CES function where the elasticity is calibrated to match available estimates of marginal abatement cost curves following Kiuila and Rutherford (2013), allowing more complex regulatory designs to be modeled.

5.3 Difference Between Productivity Shock and Explicit Compliance Requirements

There are two main differences between modeling compliance requirements as a productivity shock versus a nesting structure that explicitly represents the abatement activity. First, the substitution possibilities allowed between inputs for compliance differ. The productivity shock implicitly assumes that compliance inputs have the same substitution elasticities as the underlying production technology for the regulated sector. Alternatively, explicit representation of abatement requirements, at least as defined in Section 5.2 and the default version of the model, does not allow flexibility in how the abatement requirements are met. Second, the explicit abatement requirement assumes that any capital inputs for compliance activities are always new capital investments regardless of whether the regulation affects new or existing sources of production. For the productivity shock, capital requirements associated with compliance activities at existing sources are implicitly assumed to be repurposed extant capital.[51]

To highlight the main differences between these approaches to modeling regulatory requirements the example `examples/regulatory_modeling_approach.R` simulates an identically-specified regulation under both approaches. For this example, as well as the others presented in Section 6, we use a hypothetical regulation in the primary metal manufacturing (*pmm*) sector loosely calibrated to an initial round of regulations that were promulgated about 20 years ago under section 112 of the Clean Air Act. Section 112 of the Clean Air Act (CAA) requires the EPA to list industrial categories of major sources of one or more hazardous air pollutants (HAPs) and to then establish a national emissions standard for those categories (also referred to as a NESHAP). Major sources of HAPs are defined as new or existing facilities that emit 10 tons or more annually of any single HAP or 25 tons or more annually of a combination of HAPs. A NESHAP is typically based on an assessment of the degree to which emission reductions have been achieved at the best performing facilities in a particular source category using existing abatement control techniques. This standard is referred to as a Maximum Achievable Control Technology or MACT floor because it specifies the minimum level of HAPs control required. Specifically, the Clean Air Act requires the NESHAP to reflect the maximum degree of reduction in HAP emissions that is achievable, taking into consideration the cost of achieving the emission reductions (as well as a few other factors). For existing sources, the MACT floor is the average emission rate of the least-emitting 12 percent of facilities within that industry at the time of promulgation.

For primary metal manufacturing, it was estimated that the abatement technology available to meet the initial emission limits for integrated iron and steel manufacturing and primary and secondary aluminum manufacturing would require capital investments equivalent to approximately 0.4% of those sector's capital stock and annual operating costs equivalent to approximately 2.0% of those sector's labor expenditures at the time. Since the goal is to provide a hypothetical scenario to test the behavior of the model and not to develop quantitative impact estimates for a specific policy, we make many simplifying assumptions to keep the example as clear as possible. For example, costs as a share of benchmark capital and labor inputs are assumed to be uniform across regions; costs scale with output over time; operating costs are assumed to be associated with labor only; new and existing sources are assumed to face the same compliance costs; the other primary metal production activities included in the default aggregated *pmm* sector are assumed to face similar compliance costs for abating HAP emissions; and the policy is assumed to begin in the second modeling period. The details of how this scenario is run are presented in Section 6.

For each modeling approach we consider three policy simulations in which the regulatory requirements apply to 1) all sources of production; 2) only production associated with extant capital; and 3) only production associated with new capital.[52] As previously noted, there are two potential

[51] For example, where part of an existing structure must be repurposed for compliance activities.

[52] In the default specification of abatement requirements in SAGE, abatement activities experience productivity growth equal to the economy-wide average as described in Section 3.4.1. However, to simplify the comparison between the two regulatory modeling approaches the scenarios presented here assume no productivity growth in the compliance activities.

Table 13: EV Comparison Across Modeling Approaches [Billion $]

Affected Sources	Productivity Shock	Abatement Requirement
All	-45.8	-45
Existing	-8.3	-8.4
New	-37	-36.1

differences between the approaches to modeling abatement requirements: differences in the substitution possibilities in the abatement activity, and the vintage of capital required for compliance. When only production associated with new capital is subject to the regulatory requirements, then new capital is required for compliance in both cases, and differences between the two approaches are driven by varying assumptions about substitution possibilities in the abatement activity. On the other hand, because production associated with extant capital is modeled as Leontief, if only production associated with extant capital is subject to the regulatory requirements, then neither modeling approach provides any substitution possibilities in the abatement activity and differences across the two approaches are due to the vintage of capital required for compliance.

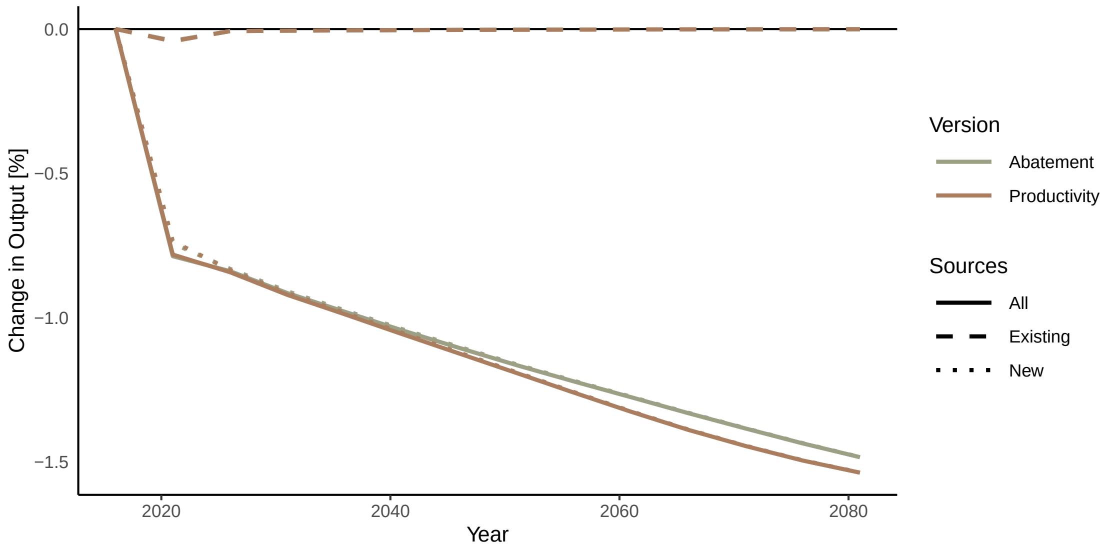
Figure 22 presents the simulated percent change in output and labor demand (inclusive of abatement requirements) for the regulated sector under the different scenarios. There is little difference in output between the two modeling approaches independent of which sources of production are subject to the abatement requirements. When the abatement requirements fall on existing sources and there is no difference in the assumptions about substitution possibilities in the abatement activity, the change in labor demand is approximately the same. In the other cases the labor intensity of production is lower in the case of the productivity shock relative to the explicit abatement requirement as firms substitute away from labor under this approach.

In a first best setting, we expect that the additional compliance flexibility assumed under the productivity shock approach would lower the social costs of the regulation, but in a second best setting the differential effect on the real wage rate leaves the direction of the difference ambiguous a priori. Table 13 presents estimates of EV under both approaches to represent the hypothetical regulation on the *pmm* sector (i.e., productivity shock and explicit abatement activity), varying which sources are affected (i.e., all, existing sources only, or new sources only).[53] For this example, regardless of which sources are affected, the EV for the two approaches to representing abatement requirements are within 0.5% of each other.

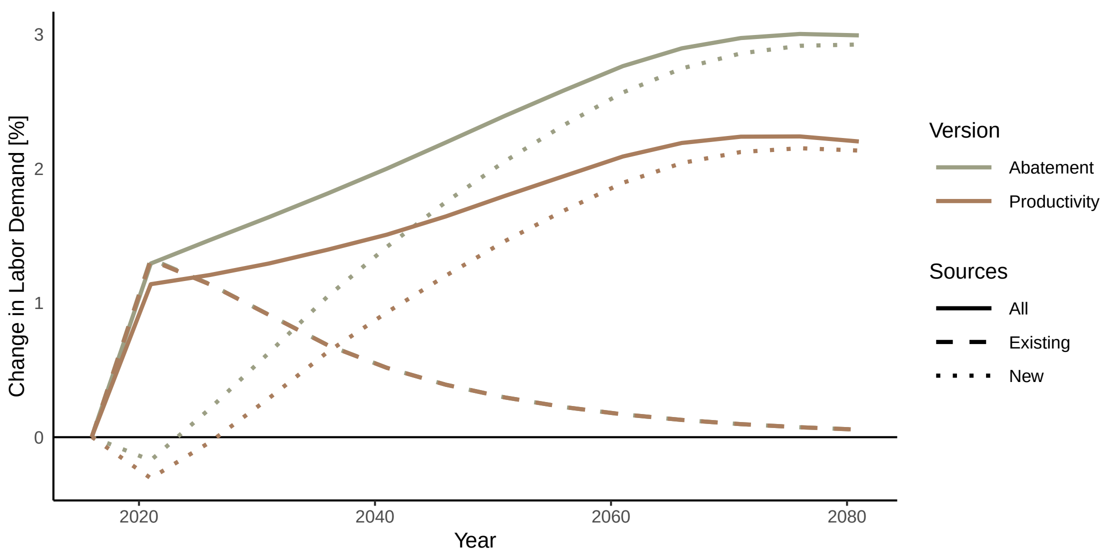
6 Using the Model

The core SAGE package is composed of 1) a build routine for constructing the model's database and 2) the modeling files for performing simulations. The programs are written to allow flexibility in how the datasets are constructed and provide options for including different modeling assumptions.

[53]The values presented represent the infinite time horizon approximation of EV discussed in Section 4.2 and contained in the output variable `ev_inf`.



(a) Change in Output in Regulated Sector



(b) Change in Labor Demand in Regulated Sector

Figure 22: Effect of Approach to Modeling Abatement Requirements

The build routine constructs a consistent set of value shares based on IMPLAN data and compiles all other exogenous data parameters (including elasticities, growth rates, population totals, tax rates, and oil and gas production data) to form the necessary inputs to run the model. The build routine is only run once for a given release to generate the model's datasets. Afterwards, conducting an analysis only requires running the model itself.

The SAGE model is run in a sequence and is intended to be used to compare the impacts of a policy shock against a specified reference case. The user must run the model to calculate the baseline level of all model variables, design a policy shock that alters the reference equilibrium point, and rerun the model to compare the resulting equilibrium with baseline values for computing the economy-wide impacts.

6.1 Directory Structure

The build routine, model, and all examples are designed to be run from the package's top directory level. The package is composed of the subdirectories in Table 14.

Table 14: Directory Structure

Subdirectory	Description
`build`	Subdirectory contains data and files for constructing the benchmark dataset. The build routine relies on a mixture of GAMS and R routines. The launching program is called `build_default_datasets.R`
`utilities`	Subdirectory for custom R routines and functions for compiling external data sources and running model code. This code is referenced throughout the build routine and modeling examples.
`data`	Subdirectory containing reconciled benchmark data and exogenous parameter files.
`model`	Subdirectory containing the core SAGE modeling file.
`diagnostics`	Subdirectory containing scripts for running model and data diagnostic tests.
`examples`	Subdirectory holding examples using the model.
`output`	Place holder subdirectory for generated model results.
`documentation`	Subdirectory containing documentation for the model. The documentation source is available in the latex file `documentation.tex`, while a typeset pdf version of the documentation is available in `documentation.pdf`.

6.2 Building the Dataset

The data compilation routines in the build stream includes programs written in both R and GAMS.[54] Some custom R routines are included (see the `utilities` subdirectory) to automate the download of external data sources and facilitate their subsequent compilation. External R packages used by SAGE but not currently available on the the system are automatically installed when the build stream is run, or may be installed separately by running `utilities/install_R_packages.R`.[55]

The build routine is controlled through the launch program, `build_default_datasets.R`, in the `build` subdirectory. The build script is designed to be run for the top-level directory of the

[54] R and GAMS must be included in the PATH environment variable. The current build stream was tested with GAMS 24.9 and R 3.5.

[55] Note that by default the package "fiftystater" is not installed, though this package is necessary for using the utility to create state choropleths with results or Figure 2. Instructions for how to install this package are included in `utilities/install_R_packages.R`.

SAGE package. Note that an internet connection is required for running the routine as it relies on external data sources for calibration.

The major steps in the build stream are as follow:

1. Initially, `build_parameters.R` compiles preference and technology parameters outside of the SAM (e.g., elasticities) from included and downloaded data files and creates `data/parameters.gms`.

2. `build_baseline.R` compiles the exogenous projections that help define the model's baseline and creates `data/baseline.gms`.

3. Additional external data sets needed for the creation of the SAM are downloaded and processed by `get_additional_economic_data.R`, `get_oil_and_gas_data.R`, and `get_population_data.R`.

4. Data on effective marginal tax rates are processed by `get_tax_data.R`. Note that aside from reconciling data from the Current Population Survey, `get_tax_data.R` submits the compiled data to the NBER TAXSIM model to derive weighted marginal tax rates.

5. IMPLAN data is extracted using `build/data/implan` by `read_implan_state.gms`, which partitions each state data file into its submatrix components.[56]

6. Resulting state-level GDX outputs are merged and fed into `build_benchmark.gms` to create the SAM and disaggregate the oil and gas extraction sectors.

7. The SAM is aggregated to the requested levels in `aggregate_benchmark.gms` based on the aggregation defined in `aggregation_file` at the top of the launching program.

8. The SAM is filtered and rebalanced in `balance_benchmark.gms`, imposing microconsistency on the dataset (i.e. data satisfying all needed accounting identities in the modeling framework) as well as other calibration assumptions on the dynamic structure of the model. The aggregated dataset is balanced and filtered using a least squares optimization framework with options as listed in Table 15.

9. Using the calibrated data, subsistence demands are estimated consistent with Section 3.3.6 as implemented in `calibrate_les.gms`. The build routine first generates the national data file to constrain the recalibration procedure when imposing regional disaggregations in the model.

The build routine relies on both IMPLAN data and data from external sources (e.g., EIA and U.S. Census Bureau). The build stream requires that the state-level IMPLAN data files (`*.gms`) are stored in `build/data/implan` at the time of compilation.[57] All other data files are included with the

[56] The process for constructing a SAM from the IMPLAN dataset is based on the IMPLANinGAMS package (Rausch and Rutherford, 2009). The original version of this software can be found at: `http://www.mpsge.org/implan98.htm`.

[57] IMPLAN is a proprietary data set and is therefore, not included in the publically available version of SAGE. Given a licensed version of IMPLAN for SAGE's benchmark year, to build the SAGE data sets first follow the instructions in `build/data/implan/implan_data_instructions.txt` to add the necessary IMPLAN data files into the SAGE directory structure.

Table 15: Selected Options in Data Set Build Stream

Place	Option	Description	Default Value
Launch	`aggregation_file`	Name of the mapping file that characterizes the level of sector, region, and household aggregation. Mapping files are located in `build/aggregation_map`. Alternative mappings can be used to modify the dimensionality of the dataset.	default_aggregation.gms
	`aeo.year`	Year of the EIA's AEO to use in the calibration.	Most recent AEO that includes benchmark year
	`aeo.scenario`	AEO scenario to use in the calibration.	`paste0("REF",aeo.year)`
	`balanced_growth`	Binary flag to calibrate benchmark investment levels consistent with a balanced growth path.	1
Matrix Balancing	`filter_small`	Binary flag to filter out small numbers.	1
	`include_taxes`	Binary flag to allow tax rate adjustments in balancing.	0
	`threshold`	Filter threshold for smallest value allowed in benchmark dataset.	5e-4
	`frac_deviations`	Binary flag to minimize percent deviations rather than absolute deviations.	1

SAGE package in `build/data` or downloaded from the internet throughout the routine. Following the successful completion of these routines, the file `build/data/satellite_data_versions.csv` is generated, which includes versioning information for downloaded data from API (Application Programming Interface) requests, the current population survey, and the TAXSIM model. Important options for controlling the build stream are located in the launch program and the SAM filtering and rebalancing script (`build/balance_benchmark.gms`). These options are described in Table 15 and are listed at the top of the associated programs. Once the build stream is finished, the resulting balanced dataset and generated parameters file containing all elasticities and assumed dynamic parameters are stored in the `data` subdirectory at the top level of the package's directory structure.

6.3 Running the Model

The model itself is written in GAMS and located in `model/sage.gms`. The model is designed to be run from the package's top directory and requires the PATH solver to be installed and licensed. The model is designed to run a single scenario, either a baseline scenario or a policy scenario. Therefore, the general process will be to first solve the model for the baseline solution and then rerun the model to solve for the counterfacutal policy solution. `sage.gms` is written to minimize the need for user adjustments to core model code (i.e. equations and data declarations) when solving counterfactual scenarios. Instead, the model offers multiple points during the execution where additional code can be included to change the specification or behavior of the model (examples below).

The model can be run through the GAMS IDE, but is designed to be run from the command line to take advantage of command line arguments to specify options including additional code to be included during counterfactual simulations. Running the model is done with a command line call to GAMS: `gams model/sage.gms`. The available command line options are presented in Table 16 and may be applied when running the model with the syntax `gams model/sage.gms --option1=choice1 --option2=choice2 ....`

After each simulation all model variables are saved in both `.csv` and `.gdx` format. These files are written to the `output/` subdirectory. While these files are very similar in their contents, the CSV file contains additional output based on post-processing of the solution results, such as GDP, EV, etc. It is also worth noting that the quantity/activity variables in the GDX file are indices relative to the benchmark levels, while in the CSV file these indices have already been multiplied by the benchmark levels for the convenience of working with the output.

6.4 Solution Checks and Diagnostics

After each simulation the model performs a set of verification checks on the solution. The results of these post-solve diagnostics are reported in the GAMS listing file. This set of diagnostics includes checks that:

1. Nominal gross domestic product is the same when calculated based on expenditures and value added;
2. Accounting identities hold in the post-solve social accounting matrix; and
3. No arbitrage opportunities exist between the national new capital price for household investments and the regional capital prices.

Following a simulation, a new social accounting matrix is constructed based on the computed post-policy equilibrium. This constructed matrix serves to verify that all of the accounting closures hold. These accounting closures include a check on commodities (the value of production and imports less exports must equal demand), activities (the value of production must equal the costs of labor, capital, intermediate inputs, and tax obligations), households (the value of consumption, investment, and tax payments must equal factor income and transfers), government (the value of

Table 16: Command Line Options for the SAGE Model

Option	Description	Default Level
`benchmark_file`	File containing the benchmark dataset. The default aggregation is specified in Section 2.	`data/default_aggregation.gdx`
`putty_clay`	Binary flag for enabling the partial putty-clay specification. With a value of 0, capital is fully malleable.	1
`parameter_file`	File containing the elasticities and additional technology and preference assumptions.	`data/parameters.gms`
`baseline_file`	File containing the time steps and exogenous baseline assumptions.	`data/baseline.gms`
`gdx_baseline_file`	A gdx file containing the results of a previous model solve. May be used to set the starting values and/or define baseline prices to calculate equivalent variation.	
`balanced_start_values`	Binary flag to set the starting values based on a balanced growth path solution independent of whether a baseline file was provided.	0
`policy_file`	Optional file containing GAMS code to define the policy changes in the model. If NULL the baseline is run.	
`gdx_save`	Binary flag for saving model results in a GDX file. The resulting GDX file is stored in the file specified by the environment variable `gdx_results_file`.	1
`gdx_results_file`	Provides the location and output name of the GDX file where the model results will be stored.	`output/results.gdx`
`output_file`	Provides the location and output name of the CSV file where the model results will be stored.	`output/results.csv`
`prologue`	Optional file with GAMS code to be included *before* any data processing and model declaration. Useful for adjusting parameters.	
`epilogue`	Optional file with GAMS code to be included *after* the model solution and any post-processing. Useful for additional post-processing of results.	
`perturb_start`	Debugging option to additively and uniformly perturb initial starting values on $y_{t,r,s}$. Value specifies the size of the perturbation.	0

government purchases less transfers equals tax income), and the rest of the world (the value of imports equals exports plus an exogenously defined balance of payments deficit). While the listing file contains the numerical values for each of these checks, for convenience the log file reports if a given check has "FAILED" based on a selected tolerance, which has a default value of 10^{-4} (i.e., $100,000). If no error messages are reported, then all checks were passed.

6.4.1 Numeraire Test

The previous solution checks illustrated instances of hypothetical policy shocks that satisfied post equilibrium adding up conditions. One other form of model validation concerns test simulations where the outcome of the simulation is already known. One such test suggested by Dixon and Rimmer (2013) is to test the model's homogeneity assumptions. As described above, SAGE is homogeneous of degree one in prices and hence a numeraire is assigned to represent equilibrium prices as relative. The level of the numeraire is typically fixed to 1. Because the model relies on relative prices, perturbing the numeraire from 1 to $(1+\beta)$ should adjust all baseline price and value variables by the same perturbation factor (β) while quantity variables remain unchanged.

To test that this assumption holds in SAGE, the "policy" file, `diagnostics/numeraire_test.gms`, contains an adjustment to the numeraire by a perturbation factor of $\beta = 0.2$. "Policy" is written in quotes because the purpose of this test is to compare two separate baselines (not a policy case vs. a baseline) differed only by the magnitude of the numeraire, but can be functionally implemented with a policy file. Notably, the chosen value of β is arbitrary.[58] The example may be run from the command line in a similar fashion those already described or from the R script, `diagnostics/numeraire_test.R`.

6.5 Examples of Adjusting or Using the Model

The following sections present a information or examples for adjusting specifics of the model or conducting policy analysis.

6.5.1 Adjusting Timesteps and Horizon

The default parameterization of the model uses five year timesteps between 2016 and 2081. The simulation years for the model are defined by the set t in the baseline file. This default assumption can be changed in the default baseline file (`data/baseline.gms`). Alternatively, a new baseline file can be created with the updated simulation years and specified by the command line option `--baseline_file` when running the model. The default aggregation of the model does not support annual timesteps. Maintaining the end year of 2081 with annual time steps increases the dimensionality of the model and poses computational intractabilities. Use of annual time steps with a more aggregated version of the model (e.g., no regional representation) is feasible. The model does not require that the timesteps be evenly spaced.

[58] While this value is arbitrary, a large level of β may cause issues for the model's solver.

6.5.2 Example of a Hypothetical Regulation

The file `examples/sample_abatement_requirement.gms` contains a representation of the hypothetical regulation in the primary metal manufacturing (*pmm*) sector that is described in Section 5.3. As previously stated, the hypothetical scenario assumes compliance with the regulation requires capital investments equivalent to approximately 0.4% of the regulated sector's capital stock and annual operating costs equivalent to approximately 2.0% of the regulated sector's labor expenditures; cost shares are uniform across regions; costs scale with output across time; operating costs are associated with labor only; and new and existing sources face the same compliance costs.

As noted in Table 16, the model provides options for including GAMS code in the model at multiple points during its compilation. The `policy_file` command line option allows the user to define a file containing GAMS code that will be included right before the model's solve statement. The file `examples/sample_abatement_requirement.gms` (see Listing 1) defines compliance requirements for the hypothetical regulation in the *pmm* sector based on the explicit abatement requirement approach of Section 5.2 and is intended to be used with the `policy_file` option.[59]

Listing 1: `examples/sample_abatement_requirement.gms`

```
* the hypothetical regulation is assumed to affect the primary metal
* manufacturing sector and have an engineering cost estimate that compliance
* will require an additional 0.4% of baseline capital expenditures and 2% of
* labor expenditures. the requirements are assumed to begin in the second time
* period of the model. the requirements are assumed to be the same for
* production with new and extant capital.

ld_abate0(t,r,"pmm",v)$(ord(t) gt 1 or ord(t) eq card(t)) = 0.020*ld0(r,"pmm")*
    l_prod_agg(t);
kd_abate0(t,r,"pmm",v)$(ord(t) gt 1 or ord(t) eq card(t)) = 0.004*kd0(r,"pmm");
```

The variables `ld_abate0(t,r,s,v)` and `kd_abate0(t,r,s,v)` define the labor expenditures and capital stock required for compliance when producing the benchmark level of output, `y0(r,s)`, in period t, region r, and sector s with capital of vintage v. There is an analogous variable for intermediate inputs of commodity ss for compliance, `id_abate0(t,r,ss,s,v)`. Care needs to be taken to ensure that the abatement costs are entered in the correct format, that is, the compliance costs at sources of vintage v when output from those sources is at the benchmark level.

In this example, the labor inputs are multiplied by `l_prod_agg(t)`, which reflects the aggregate labor productivity growth in the economy. By default the model assumes that labor productivity in abatement activities will grow at this rate. Therefore, by increasing the costs at rate `l_prod_agg(t)` one is essentially negating that productivity growth and instead assuming that there is no productivity growth in the abatement activity. This is done only to keep this example the same as the one presented in Section 5.3 to maintain consistency throughout the examples.

[59] For an example that implements the compliance requirements as a productivity shock, see `examples/regulatory_modeling_approch.R`.

In this example, the average compliance expenditures per unit of output in a region are assumed to remain constant over time. In addition, the compliance requirements are assumed to begin in the second period, hence the conditional `ord(t) gt 1`. The second part of the conditional is to ensure that the example will work with the static version of the model, which by definition only has one time period.

To analyze the hypothetical regulation the baseline is first calculated, after which the model is run with the hypothetical regulation. This can be accomplished from the command line using the commands presented in Listing 2.

Listing 2: Running the Sample Abatement Requirement from the Command Line

```
gams model/sage.gms --gdx_results_file=output/baseline.gdx
                    --output_file=output/baseline_results.csv

gams model/sage.gms --gdx_baseline_file=output/baseline.gdx
                    --policy_file=examples/sample_abatement_requirement.gms
                    --output_file=output/regulation_results.csv
```

In calculating the baseline in the first model run, the command line option `gdx_results_file` defines the GDX file where the results of the model solve will be saved. The command line option `output_file` defines a CSV file where the baseline results will be stored. The GDX file is used to define the baseline in the policy run using the `gdx_baseline_file` command line option. In this case, the baseline is used to both set the starting values and provides the baseline prices for calculating EV. The `output_file` command line option defines a CSV file where the results of the model run with the abatement requirement will be saved. The two output files may be used to calculate the changes in variables between the two simulations. Policy impacts should only be compared to their corresponding baseline.

Once the post policy equilibrium solution is determined, the SAGE listing file (`sage.lst`) will include the diagnostic checks described above. The diagnostics help determine if the model solution satisfies the necessary closures.

The model's use of command line options and compile time code inclusions allows the model to be easily run from scripts. The modeling package includes a series of R utilities that are located in the file `utilities/R_utilities.R`, which provides functions to run the model and process the results from R. The file `examples/basic_example.R` shows how this hypothetical abatement requirement may be run and results processed from an R script. An analogous example of how such routines can be built in GAMS is included in `examples/basic_example.gms`.

6.5.3 Example of a Regulation with Phased In Requirements

The previous example focused on a hypothetical regulation where the compliance requirements were assumed to begin in the second period of the model and where the average compliance expenditures per unit of output in a region are assumed to remain constant over time. It is possible to model situations where the average compliance costs per unit of output are expected to vary over time. One example of this regulatory context is when a regulation is phased in over time. This section describes the implementation differences from previous examples to accommodate temporal variation in compliance cost model inputs.

The policy file `examples/sample_phased_in_abatement_requirement.gms` contains a modified representation of the hypothetical regulation described in Section 6.5.2 where the average compliance costs per unit of output in the second period of the model are 50% of their expected value once the regulation is fully implemented starting in the third period of the model. This scenario assumes compliance with the regulation requires capital investments equivalent to approximately 0.2% of the regulated sector's capital stock and annual operating costs equivalent to approximately 1.0% of the regulated sector's labor expenditures in the first year of compliance (second model period). Subsequent years of compliance (third model period and thereafter) requires capital investments equivalent to approximately 0.4% of the regulated sector's capital stock and annual operating costs equivalent to approximately 2.0% of the regulated sector's labor expenditures. All other aspects of the hypothetical regulation match the example described in Section 6.5.2.

Listing 3 presents the policy file for this phased in version of the example regulation. The variables `ld_abate0(t,r,s,v)` and `kd_abate0(t,r,s,v)` define the labor expenditures and capital stock required for compliance when producing the benchmark level of output, `y0(r,s)`, in period t, region r, and sector s with capital of vintage v. In this case, there is a definition for the two phases of the hypothetical regulation: the first phase denoted by `ord(t) eq 2` (the model's second period) and the second phase denoted by `ord(t) gt 2` (all model periods after the second). The second part of the conditional for the second phase costs is to ensure that the example will work with the static version of the model, which by definition only has one time period and is assumed to be associated with the full implementation costs in this example.

Listing 3: `examples/sample_phased_in_abatement_requirement.gms`

```
* this is a phased in version of the sample regulation in
* examples/sample_abatement_requirement.gms the hypothetical regulation is
* assumed to affect the primary metal manufacturing sector and have an
* engineering cost estimate that compliance will require an additional 0.4% of
* baseline capital expenditures and 2% of labor expenditures when fully
* implemented. the requirements are assumed to begin in the second time period
* of the model but with only half the costs compared to once they become fully
* implemented in the third period of the model. the requirements are assumed to
* be the same for production with new and extant capital.

* costs in the regulation's first phase (second model period)
ld_abate0(t,r,"pmm",v)$(ord(t) eq 2) = 0.010*ld0(r,"pmm")*l_prod_agg(t);
kd_abate0(t,r,"pmm",v)$(ord(t) eq 2) = 0.002*kd0(r,"pmm");

* costs in the regulation's second phase (third model period and thereafter)
ld_abate0(t,r,"pmm",v)$(ord(t) gt 2 or ord(t) eq card(t)) = 0.020*ld0(r,"pmm")*
    l_prod_agg(t);
kd_abate0(t,r,"pmm",v)$(ord(t) gt 2 or ord(t) eq card(t)) = 0.004*kd0(r,"pmm");
```

This example may be run from the command line in a similar fashion to Listing 2 by updating the name of the policy file. The R script `examples/policy_phase_in_example.R` runs a comparison between the phased in policy and the immediate implementation version in Section 6.5.2. Implementing a phased in regulation as a productivity shock would be accomplished in a similar fashion by defining the productivity shock (represented by the parameter `prod_ind`) for each phase using a conditional argument for `t`.

6.5.4 Example of a Regulation in a Large vs. Small Open Economy

In this section, we illustrate the importance of the large open economy assumption in SAGE for the aforementioned hypothetical regulatory scenario. A model switch has been integrated into the code to convert the model to a small open economy where the United States is treated as a price taker in the world market for commodities with perfectly elastic demand for its international exports and perfectly elastic supply for international imports. This assumption can be toggled from the command line as illustrated in Listing 4 (note that previous commands for running the baseline and policy case still apply here).

Listing 4: Running the Model as a Small Open Economy

```
gams model/sage.gms --loe=0
```

We run the hypothetical environmental regulation for all sources in the primary metal manufacturing sector (*pmm*). As before, we assume that compliance with the regulation requires additional capital and labor inputs to production. Functionally, the large open economy formulation imposes a non-zero slope in the international demand for U.S. export curve and supply of international imports into the U.S. market curve. The steepness of these curves are evaluated in Table 9. Figure 23 reports the percent change in imports and exports across all model years for each sector in the economy. Because estimated elasticities are high for imports, the percent change across the two formulations are almost equivalent. Both model formulations estimate increases in the imports for primary metals due to import substitution from relatively more expensive domestic output. The story is different for exports. Estimated elasticities are smaller in absolute value for exports and therefore the large open economy modeled results diverge more significantly from the small open economy results. Imposing a downward sloping demand curve for exports mutes quantity impacts in the export market and therefore percent changes in the large open economy formulation are smaller than in the small open economy case.

While quantity changes are muted in the large open economy case, export and import prices differentially change for each commodity. This contributes to larger social costs estimated in the large open economy framework relative to the simpler alternative as reported in Table 18.

6.5.5 Example of Sector Specific Consumption Taxes

To further augment the section on sensitivity simulations, we illustrate model results for hypothetical sector specific consumption tax scenarios. The simulations estimate the equivalent variation of

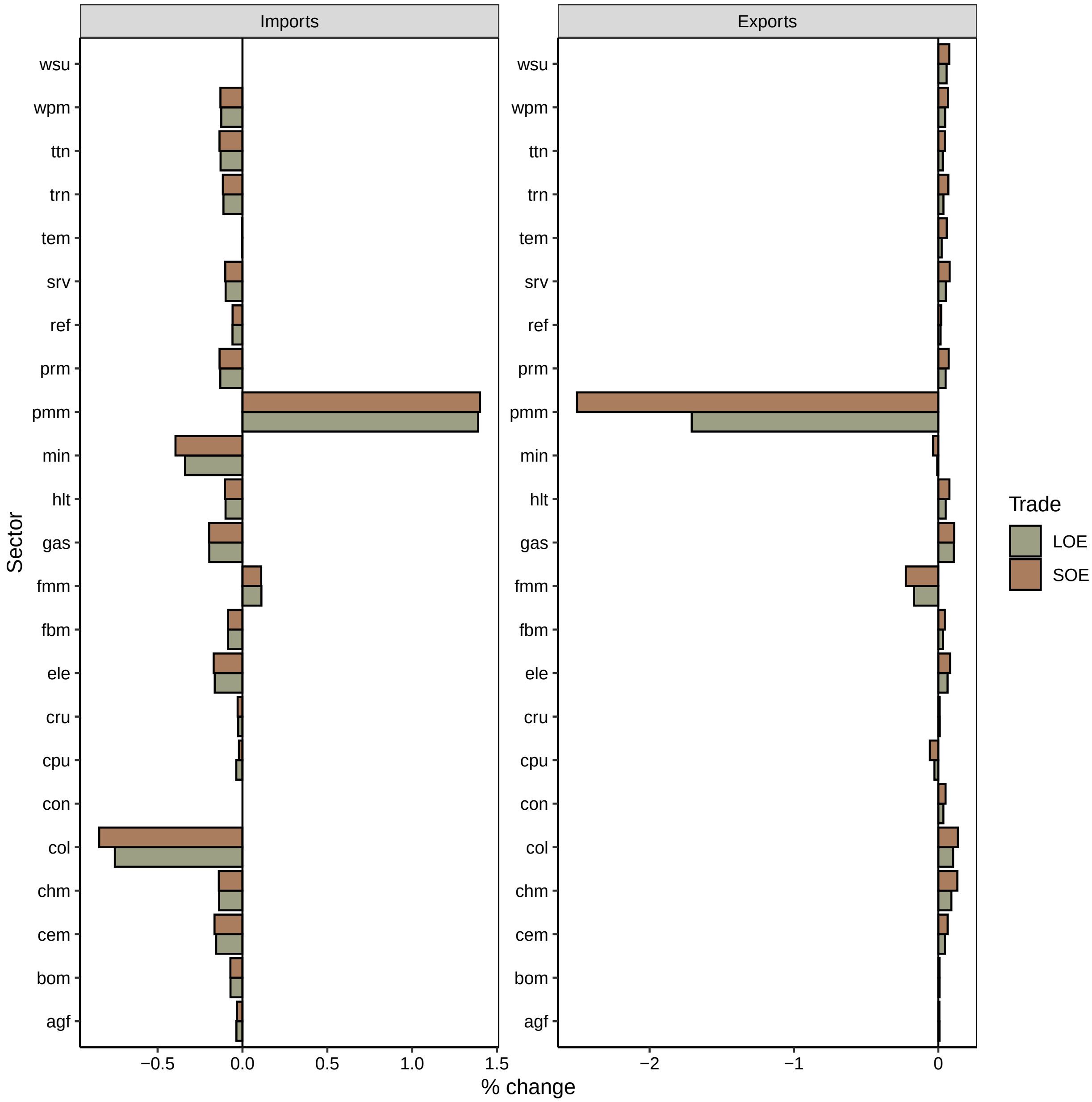


Figure 23: Percent Change in Imports and Exports

Table 18: EV Comparison Across Trade Assumptions [Billion $]

Large Open Economy	Small Open Economy
-44.98	-41.67

sector specific consumption taxes that raise an additional $10 billion in total government revenues in each time step of the model based on the output value in the baseline (this revenue target represents approximately 2% of government revenue from sales and excise taxes in the benchmark year).

This suite of simulations is run via the R program, `examples/consumption_tax.r` that relies on a policy file called `examples/consumption_tax.gms` (also reported in Listing 5). The policy file selects the taxed sector as controlled by the R script and imposes a hypothetical addition to the existing consumption tax for the selected sector equal to the specific amount of government revenue raised in each period.

Listing 5: `examples/consumption_tax.gms`

```
* implements consumption tax change
*
*   s_tax:           taxed sector set by defining s_tax("xxx") = YES;
*   shock_size:      initial year value of the shock in billions
*
set
  s_tax(s)       taxed sector;

parameter
  shock_size     value of the shock in billions
  tax            addition to tax rate;

* there are no regulated sectors until one is set
s_tax(s) = no;

* load the shock specification
$include examples/tax_definition.gms

* tax amount
tax(t,s_tax)$sum((rr,h), pa.l(t,rr,s_tax)*cd.l(t,rr,s_tax,h)*cd_base(t,rr,s_tax,h)
    )
  = shock_size/sum((rr,h), pa.l(t,rr,s_tax)*cd.l(t,rr,s_tax,h)*cd_base(t,rr,s_tax,
      h));

* augment existing consumption tax
tc(t,r,s_tax) = tc(t,r,s_tax) + tax(t,s_tax);

display tax;
```

Figure 24 reports the calculated social costs as characterized by changes in equivalent variation. These results are for diagnostic purposes only and should not be interpreted as the results of a specific policy. The social cost is dependent, in part, on the use of the revenue, where in these examples the recycling occurs through the models standard government budget constraint.

6.5.6 Additional Examples

Other examples are included in the `examples` subdirectory and are listed in Table 19. These relatively simple examples are designed to demonstrate basic features of the modeling framework and their general impact on simulation results. The examples are intended to be run from the top-level directory of the SAGE package. Most of the routines listed in Table 19 use the R programming

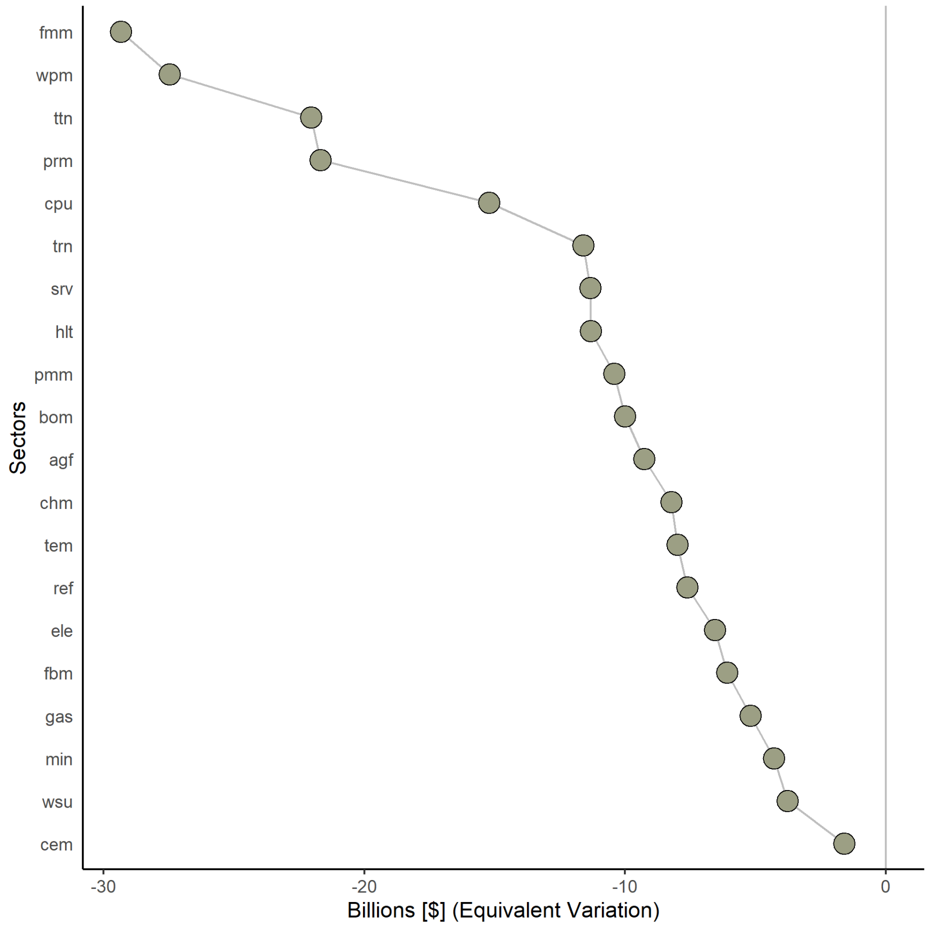


Figure 24: Social Costs of Equivalent Hypothetical Sector Specific Consumption Tax Increases

language to conduct the simulations and process the results.[60]

In addition to the simpler examples, there is a more extensive suite of simulations contained in the file `examples/scenario_analysis.R`, which conducts sensitivity analyses around hypothetical regulations implemented as abaement requirements in the spirit of those considered in Marten et al. (2019). Note that the results in Marten et al. (2019) are based on SAGE v. 1.0.7 and a different implementation of the hypothetical regulations, such that `examples/scenario_analysis.R` is not intended to replicate the quantitative results of that paper.[61]

[60]The scripts may be run from a development environment with an R backend or from the command line using Rscript (assuming it is available).

[61]For code to replicate the specific results of Marten et al. (2019), see the Dataverse site for the Journal of the Association of Environmental and Resource Economists.

Table 19: Additional Simulation Examples

File Name	Description
`static_vs_dynamic.R`	Compares the results from the sample abatement requirement using a dynamic vs. static version of the model.
`putty-clay_vs_putty-putty.R`	Compares the results of the sample abatement requirement under the default partial putty-clay capital framework vs the case of fully malleable capital under the putty-putty assumption.
`national_vs_regional.R`	Compares the results from the sample abatement requirement from the dynamic model with and without regional delineation.
`regulatory_modeling_approach.R`	Compares the output of the sample abatement requirement when modeled as a productivity shock vs. an explicit abatement requirement per unit of output. Produces the comparisons in Figure 22.
`open_economy.R`	Compares the output of the sample abatement requirement when assuming a large vs. small open economy in international trade. Produces the comparisons in Figure 23.
`scenario_analysis.R`	Runs sensitivity analyses around hypothetical regulations similar to those considered in Marten et al. (2019). Uses the file `examples/productivity_shock.gms` as the policy file to define a variety of hypothetical regulations as productivity shocks.
`consumption_tax.R`	Runs sensitivity analyses for hypothetical sector specific consumption tax increases as reported in Figure 24. Uses the file `examples/consumption_tax.gms` as the policy file to define the hypothetical scenarios.
`plot_results.R`	Illustrates custom plotting functions designed to organize and display results from SAGE outputs. For function code, see `R_utilities.R`.

References

Aguiar, M. and Bils, M. (2015). Has consumption inequality mirrored income inequality? *American Economic Review*, 105(9):2725–56.

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

Arora, V. (2014). Estimates of the price elasticities of natural gas supply and demand in the United

States. Munich Personal RePEc Archive, MPRA Paper No. 54232.

Arora, V. and Cai, Y. (2014). US natural gas exports and their global impacts. *Applied Energy*, 120:95–103.

Arrow, K. J., Cropper, M., Gollier, C., Groom, B., Heal, G. M., Newell, R. G., Nordhaus, W. D., Pindyck, R. S., Pizer, W. A., Portney, P., et al. (2013). How should benefits and costs be discounted in an intergenerational context? the views of an expert panel. *The views of an expert panel (December 19, 2013). Resources for the future discussion paper*, (12-53).

Babiker, M. H., Maskus, K. E., and Rutherford, T. F. (1997). Carbon taxes and the global trading system. University of Adelaide, Center for International Economic Studies, Seminar Paper 97-08.

Balistreri, E. J. and Rutherford, T. F. (2001). Dynamic general equilibrium analysis at the state level: Assessing the economic implications of the Kyoto Protocol. Presented at the 4th Annual Conference on Global Economic Analysis, Purdue University, June 27-29.

Baumeister, C. and Hamilton, J. D. (2019). Structural interpretation of vector autoregressions with incomplete identification: Revisiting the role of oil supply and demand shocks. *American Economic Review*, 109(5):1873–1910.

Beck, T., Jolly, L., and Loncar, T. (1991). Supply response in the Australian black coal industry (no. 91.1). Canberra, Australia: Australian Board of Agricultural and Resource Economics.

Beckman, J., Hertel, T., and Tyner, W. (2011). Validating energy-oriented CGE models. *Energy Economics*, 33(5):799–806.

Bjørnland, H., Nordvik, F., and Maximilian, R. (2017). Supply flexibility in the shale patch: Evidence from North Dakota. Norges Bank Research Working Paper 9.

Boetel, B. L., Hoffmann, R., and Liu, D. J. (2007). Estimating investment rigidity within a threshold regression framework: The case of US hog production sector. *American Journal of Agricultural Economics*, 89(1):36–51.

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

Böhringer, C., Schneider, J., and Asane-Otoo, E. (2016). Trade in carbon and the effectiveness of carbon tariffs. University of Oldenburg, Department of Economics, Discussion Papers in Economics, No. V-388-16.

Borges, A. M. and Goulder, L. H. (1984). Decomposing the impact of higher energy prices on long-term growth. In Scarf, H. E. and Shoven, J. B., editors, *Applied General Equilibrium Analysis*. Cambridge University Press.

Brockway, P. E., Heun, M. K., Santos, J., and Barrett, J. R. (2017). Energy-extended CES aggregate production: Current aspects of their specification and econometric estimation. *Energies*, 10(2).

Brown, J. P., Fitzgerald, T., and Weber, J. G. (2016). Capturing rents from natural resource abundance: Private royalties from U.S. onshore oil & gas production. *Resource and Energy Economics*, 46:23 – 38.

Brown, S. P. and Huntington, H. G. (2003). Terms of trade and OECD policies to mitigate global climate change. *Federal Reserve Bank of Dallas Economic and Financial Policy Review*, 2(1):1–25.

Buhr, B. L. and Kim, H. (1997). Dynamic adjustment in vertically linked markets: the case of the US beef industry. *American Journal of Agricultural Economics*, 79(1):126–138.

Burniaux, J.-M. and Truong, T. P. (2002). GTAP-E: an energy-environmental version of the GTAP model. Purdue University, Center for Global Trade Analysis, GTAP Technical Paper No. 16.

Cai, Y., Newth, D., Finnigan, J., and Gunasekera, D. (2015). A hybrid energy-economy model for global integrated assessment of climate change, carbon mitigation and energy transformation. *Applied Energy*, 148:381–395.

Caldara, D., Cavallo, M., and Iacoviello, M. (2018). Oil price elasticities and oil price fluctuations. *Journal of Monetary Economics*, 103:1–20.

Capros, P., Van Regemorter, D., Paroussos, L., Karkatsoulis, P., Fragkiadakis, C., Tsani, S., Charalampidis, I., and Revesz, T. (2013). GEM-E3 model documentation. European Commission Joint Research Centre, JRC Technical Report 26034 EN.

Caron, J., Karplus, V. J., and Schwarz, G. A. (2017). Modeling the income dependence of household energy consumption and its implications for climate policy in china.

Caron, J. and Rausch, S. (2013). A global general equilibrium model with US state-level detail for trade and environmental policy analysis–technical notes. MIT Joint Program on the Science and Policy of Global Change, Joint Program Technical Note 13.

CBO (2017). International comparisons of corporate income tax rates. U.S. Congressional Budget Office. `https://www.cbo.gov/system/files/115th-congress-2017-2018/reports/52419-internationaltaxratecomp.pdf`.

CBO (2018). The distribution of household income, 2014. U.S. Congressional Budget Office. `https://www.cbo.gov/system/files/115th-congress-2017-2018/reports/53597-distribution-household-income-2014.pdf`.

CEA (2017). Discounting for public policy: Theory and recent evidence on the merits of updating the discount rate. Council of Economic Advisers Issue Brief.

https://obamawhitehouse.archives.gov/sites/default/files/page/files/201701_cea_discounting_issue_brief.pdf.

Chateau, J., Dellink, R., and Lanzi, E. (2014). An overview of the OECD ENV-Linkages model: Version 3. OECD Environment Working Papers, No. 65.

Coughlin, C. C. and Novy, D. (2013). Is the international border effect larger than the domestic border effect? Evidence from US trade. *CESifo Economic Studies*, 59(2):249–276.

Dahl, C. and Duggan, T. E. (1996). US energy product supply elasticities: A survey and application to the US oil market. *Resource and Energy Economics*, 18(3):243–263.

Dissou, Y., Karnizova, L., and Sun, Q. (2015). Industry-level econometric estimates of energy-capital-labor substitution with a nested CES production function. *Atlantic Economic Journal*, 43(1):107–121.

Dixon, P. B. and Rimmer, M. T. (2013). Validation in computable general equilibrium modeling. In *Handbook of Computable General Equilibrium Modeling*, volume 1, pages 1271–1330. Elsevier.

EIA (2001). Model documentation: Coal market module of the National Energy Modeling System. DOE/EIA-M060(2001).

EIA (2011). Performance profiles of major energy producers 2009. Report Number: DOE/EIA-0206(2009).

EPA (2010). Guidelines for preparing economic analysis. U.S. Environmental Protection Agency. Washington, DC.

EPA (2015). Economy-wide modeling: Social cost and welfare white paper. https://yosemite.epa.gov/sab/SABPRODUCT.NSF/0/07E67CF77B54734285257BB0004F87ED.

Farzin, Y. H. (1992). The time path of scarcity rent in the theory of exhaustible resources. *The Economic Journal*, 102(413):813–830.

Feenberg, D. and Coutts, E. (1993). An introduction to the TAXSIM model. *Journal of Policy Analysis and Management*, 12(1):189–194.

Feenstra, R. C., Inklaar, R., and Timmer, M. P. (2015). The next generation of the Penn World Table. *American Economic Review*, 105(10):3150–82.

Ferris, M. C. and Munson, T. S. (2000). Complementarity problems in GAMS and the PATH solver. *Journal of Economic Dynamics and Control*, 24(2):165–188.

Fitzgerald, T. and Rucker, R. R. (2016). US private oil and natural gas royalties: Estimates and policy relevance. *OPEC Energy Review*, 40(1):3–25.

Foley, P. and Clark, J. (1981). US copper supply: An economic/engineering analysis of cost—supply relationships. *Resources Policy*, 7(3):171–187.

Golombek, R., Hagem, C., and Hoel, M. (1995). Efficient incomplete international climate agreements. *Resource and Energy Economics*, 17(1):25–46.

Greene, D. L. and Liu, C. (2015). U.S. oil dependence 2014: Is energy independence in sight? *Energy Policy*, 85:126–137.

Ha, S. J., Lange, I., Lecca, P., and Turner, K. (2012). Econometric estimation of nested production functions and testing in a computable general equilibrium analysis of economy-wide rebound effects. University of Stirling, Division of Economics, Stirling Economics Discussion Papers 2012-08.

Haggerty, M., Lawson, M., and Pearcy, J. (2015). Steam coal at an arm's length: An evaluation of proposed reform options for US coal used in power generation. Manuscript, Headwaters Economics and Montana State University.

Havranek, T., Horvath, R., Irsova, Z., and Rusnak, M. (2015). Cross-country heterogeneity in intertemporal substitution. *Journal of International Economics*, 96(1):100–118.

Hendricks, N. P., Smith, A., and Sumner, D. A. (2014). Crop supply dynamics and the illusion of partial adjustment. *American Journal of Agricultural Economics*, 96(5):1469–1491.

Hertel, T., Hummels, D., Ivanic, M., and Keeney, R. (2007). How confident can we be of CGE-based assessments of free trade agreements? *Economic Modelling*, 24(4):611–635.

Hertel, T. W., McDougall, R., Narayanan, B., and Aguiar, A. H. (2008). Behavioural parameters. In Narayanan, B. and Walmsley, T. L., editors, *Global Trade, Assistance, and Production: The GTAP 7 Data Base*. Center for Global Trade Analysis, Purdue University.

Hertel, T. W. and Tsigas, M. (2002). Primary factor shares. In Dimaranan, B. V. and McDougall, R. A., editors, *Global Trade, Assistance, and Production: The GTAP 5 Data Base*. Center for Global Trade Analysis, Purdue University.

Higgs, P. J. and Powell, A. A. (1990). Forecasting small-area agricultural incomes using a CGE model of the Australian economy. *The Annals of Regional Science*, 24(1):43–59.

Ho, M., Britz, W., Delzeit, R., Leblanc, F., Roson, R., Schuenemann, F., and Weitzel, M. (2020). Modelling consumption and constructing long-term baselines in final demand. *Journal of Global Economic Analysis*, 5(1):63–108.

Holt, M. T. and Aradhyula, S. V. (1998). Endogenous risk in rational-expectations commodity models: A multivariate generalized ARCH-M approach. *Journal of Empirical Finance*, 5(2):99–129.

Holt, M. T. and McKenzie, A. M. (2003). Quasi-rational and ex ante price expectations in commodity supply models: An empirical analysis of the US broiler market. *Journal of Applied Econometrics*, 18(4):407–426.

Horridge, M., Meeraus, A., Pearson, K., and Rutherford, T. F. (2013). Solution software for computable general equilibrium modeling. In *Handbook of computable general equilibrium modeling*, volume 1, pages 1331–1381. Elsevier.

Huntington, H. G. (1992). Inferred demand and supply elasticities from a comparison of world oil models. In Sterner, T., editor, *International Energy Economics*. International Studies in Economic Modelling, Springer, Dordrecht.

Iqbal, M. Z. and Babcock, B. A. (2018). Global growing-area elasticities of key agricultural crops estimated using dynamic heterogeneous panel methods. *Agricultural Economics*, 49(6):681–690.

ITC (2011). Magnesium from China and Russia. Investigation Nos. 731-TA-1071-1072 (Review) Publication 4214.

ITC (2014a). Gray portland cement and cement clinker from Mexico. Investigation No. 731-TA-451 (Final) Publication 2305.

ITC (2014b). Silica bricks and shapes from China. Investigation No. 731-TA-1205 (Final) Publication 4443.

Jomini, P., Zeitsch, J. F., McDougall, R., Welsh, A., Brown, S., Hambley, J., and Kelly, J. (1991). SALTER: A general equilibrium model of the world economy, Volume 1, Model structure, database and parameters. Industry Commission, Canberra.

Jordan, B., Lange, I., and Linn, J. (2018). Coal demand, market forces, and US coal mine closures. RFF Working Paper 18-13.

Jorgenson, D. W., Goettle, R. J., Ho, M. S., and Wilcoxen, P. J. (2013). *Double dividend: Environmental taxes and fiscal reform in the United States*. MIT Press.

Kaiser, H. M. (2012). An economic analysis of the national pork board checkoff program. Report prepared for the National Pork Board.

Kapombe, C. M. and Colyer, D. (1998). Modeling US broiler supply response: A structural time series approach. *Agricultural and Resource Economics Review*, 27(2):241–251.

Keller, W. J. (1976). A nested CES-type utility function and its demand and price-index functions. *European Economic Review*, 7(2):175–186.

Kemfert, C. (1998). Estimated substitution elasticities of a nested CES production function approach for Germany. *Energy Economics*, 20(3):249–264.

Kim, H. and Moschini, G. (2018). The dynamics of supply: US corn and soybeans in the biofuel era. *Land Economics*, 94(4):593–613.

Kiuila, O. and Rutherford, T. F. (2013). Piecewise smooth approximation of bottom–up abatement cost curves. *Energy Economics*, 40:734–742.

Koesler, S. and Schymura, M. (2015). Substitution elasticities in a constant elasticity of substitution framework—Empirical estimates using nonlinear least squares. *Economic Systems Research*, 27(1):101–121.

Krichene, N. (2002). World crude oil and natural gas: A demand and supply model. *Energy Economics*, 24(6):557–576.

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

Lau, M. I., Pahlke, A., and Rutherford, T. F. (2002). Approximating infinite-horizon models in a complementarity format: A primer in dynamic general equilibrium analysis. *Journal of Economic Dynamics and Control*, 26(4):577–609.

Lecca, P., Swales, K., and Turner, K. (2011). An investigation of issues relating to where energy should enter the production function. *Economic Modelling*, 28(6):2832–2841.

Lin, C.-Y. C. and Wagner, G. (2007). Steady-state growth in a Hotelling model of resource extraction. *Journal of Environmental Economics and Management*, 54(1):68 – 83.

Liu, J., Arndt, C., and Hertel, T. W. (2004). Parameter estimation and measures of fit in a global, general equilibrium model. *Journal of Economic Integration*, 19(3):626–649.

Lofgren, H., Harris, R. L., and Robinson, S. (2002). *A standard computable general equilibrium (CGE) model in GAMS*, volume 5. Intl Food Policy Res Inst.

Markandya, A. and Pedroso-Galinato, S. (2007). How substitutable is natural capital? *Environmental and Resource Economics*, 37(1):297–312.

Marsh, J. M. (2003). Impacts of declining US retail beef demand on farm-level beef prices and production. *American Journal of Agricultural Economics*, 85(4):902–913.

Marten, A. and Garbaccio, R. (2018). An applied general equilibrium model for the analysis of environmental policy: Sage v1.0 technical documentation. NCEE Working Paper 2018-05.

Marten, A. L., Garbaccio, R., and Wolverton, A. (2019). Exploring the general equilibrium costs of sector-specific regulations. *Forthcoming in Journal of the Association of Environmental and Resource Economists*.

Mathiesen, L. (1985). Computational experience in solving equilibrium models by a sequence of linear complementarity problems. *Operations research*, 33(6):1225–1250.

McClelland, R. and Mok, S. (2012). A review of recent research on labor supply elasticities. Congressional Budget Office, Working Paper 2012-12.

Narayanan, B., Aguiar, A., and McDougall, R. (2016). An overview of the gtap 9 data base. *Journal of Global Economic Analysis*, 1(1):181–208.

Nordhaus, W. D. (2007). A review of the stern review on the economics of climate change. *Journal of Economic Literature*, 45(3):686–702.

Paltsev, S., Reilly, J. M., Jacoby, H. D., Eckaus, R. S., McFarland, J., Sarofim, M., Asadoorian, M., and Babiker, M. (2005). The MIT Emissions Prediction and Policy Analysis (EPPA) Model: Version 4. Joint Program Report Series, Report 125.

Ramsey, F. P. (1928). A mathematical theory of saving. *The Economic Journal*, 38(152):543–559.

Rausch, S., Metcalf, G. E., Reilly, J. M., and Paltsev, S. (2011). Distributional impacts of a U.S. greenhouse gas policy: A general equilibrium analysis of carbon pricing. In Metcalf, G. E., editor, *US Energy Tax Policy*. Cambridge University Press.

Rausch, S. and Rutherford, T. F. (2009). Tools for building national economic models using state-level implan social accounts. Unpublished manuscript. Available online at: http://www.cepe.ethz.ch/people/profs/srausch/IMPLAN2006inGAMS.pdf.

Roberts, M. J. and Schlenker, W. (2013). Identifying supply and demand elasticities of agricultural commodities: Implications for the US ethanol mandate. *American Economic Review*, 103(6):2265–95.

Ross, M. T. (2005). Documentation of the Applied Dynamic Analysis of the Global Economy (ADAGE) Model. RTI International Working Paper 05-01.

Ross, M. T. (2014). Structure of the Dynamic Integrated Economy/Energy/Emissions Model: Electricity Component, DIEM-Electricity. Duke University, Nicholas Institute for Environmental Policy Solutions, Working Paper NI WP 14-11.

Rucker, R. R., Burt, O. R., and LaFrance, J. T. (1984). An econometric model of cattle inventories. *American Journal of Agricultural Economics*, 66(2):131–144.

Russo, C., Green, R. D., and Howitt, R. E. (2008). Estimation of supply and demand elasticities of California commodities. Agriculture and Resource Economics Working Papers, Department of Agricultural and Resource Economics, UCD, UC Davis.

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.

SAB (2017). SAB advice on the use of economy-wide models in evaluating social costs, benefits, and economic impacts of air regulations. EPA-SAB-17-012. U.S. Environmental Protection Agency. Washington, DC.

SAB (2020). Technical review of epa's computable general equilibrium model, sage. EPA-SAB-20-010. U.S. Environmental Protection Agency. Washington, DC.

Sands, R., Malcolm, S., Suttles, S., Marshall, E., et al. (2017). Dedicated energy crops and competition for agricultural land. *Economic Research Report-Economic Research Service, USDA*, (223).

Sarmiento, C. and Allen, P. G. (2000). Dynamics of beef supply in the presence of cointegration: A new test of the backward-bending hypothesis. *Review of Agricultural Economics*, 22(2):421–437.

Serletis, A., Timilsina, G. R., and Vasetsky, O. (2010a). Interfuel substitution in the United States. *Energy Economics*, 32(3):737–745.

Serletis, A., Timilsina, G. R., and Vasetsky, O. (2010b). International evidence on sectoral interfuel substitution. *The Energy Journal*, 31(4):1–29.

Shideed, K. H. and White, F. C. (1989). Alternative forms of price expectations in supply analysis for US corn and soybean acreages. *Western Journal of Agricultural Economics*, 14(1836-2016-150808):281.

Shoven, J. B. and Whalley, J. (1984). Applied general-equilibrium models of taxation and international trade: An introduction and survey. *Journal of Economic literature*, 22(3):1007–1051.

Simola, A. (2015). Intensive margin of land use in CGE models–Reviving CRESH functional form. *Land Use Policy*, 48:467–481.

Stern, D. I. (2012). Interfuel substitution: A meta-analysis. *Journal of Economic Surveys*, 26(2):307–331.

Stern, N. (2007). *The Economics of Climate Change: The Stern Review*. Cambridge University Press.

Stone, R. (1954). Linear expenditure systems and demand analysis: an application to the pattern of british demand. *The Economic Journal*, 64(255):511–527.

Su, X., Zhou, W., Nakagami, K., Ren, H., and Mu, H. (2012). Capital stock-labor-energy substitution and production efficiency study for China. *Energy Economics*, 34(4):1208–1213.

Sue Wing, I. (2001). *Induced technical change in computable general equilibrium models for climate-change policy analysis*. PhD thesis, Massachusetts Institute of Technology.

Sue Wing, I. (2006). The synthesis of bottom-up and top-down approaches to climate policy modeling: Electric power technologies and the cost of limiting US CO2 emissions. *Energy Policy*, 34(18):3847–3869.

Sue Wing, I., Daenzer, K., Fisher-Vanden, K., and Calvin, K. (2011). Phoenix Model Documentation.

Van der Mensbrugghe, D. (1994). GREEN-the reference manual. OECD Economics Department Working Papers, No. 143.

Van der Werf, E. (2008). Production functions for climate policy modeling: An empirical analysis. *Energy Economics*, 30(6):2964–2979.

Young, A. T. (2013). US elasticities of substitution and factor augmentation at the industry level. *Macroeconomic Dynamics*, 17(04):861–897.

Yuan, M., Rausch, S., Caron, J., Paltsev, S., and Reilly, J. (2019). The mit us regional energy policy (usrep) model: The base model and revisions. *Joint Program Technical Note TN*, 18.

Zahniser, S., Hertz, T., Dixon, P., and Rimmer, M. (2012). The potential impact of changes in immigration policy on U.S. agriculture and the market for hired farm labor: A simulation analysis. U.S. Department of Agricuture, Economic Research Service, Economic Research Report Number 135.