Are We Building Too Much Natural Gas Pipeline? A comparison of actual US expansion of pipeline to an optimized plan of the interstate network

by

Thuy Doan, Matthias Fripp, and Michael J. Roberts

UHERO
THE ECONOMIC RESEARCH ORGANIZATION AT THE UNIVERSITY OF HAWAI'I

Working Paper No. 2022-2

April 25, 2022

UNIVERSITY OF HAWAI'I AT MANOA
2424 MAILE WAY, ROOM 540 • HONOLULU, HAWAI'I 96822
WWW.UHERO.HAWAII.EDU

Are We Building Too Much Natural Gas Pipeline?

A comparison of actual US expansion of pipeline to an optimized plan of the interstate network

Thuy Doan, Matthias Fripp, and Michael J. Roberts

University of Hawai'i at Mānoa

April 25, 2022

Abstract

Interstate natural gas transmission and storage infrastructure is facilitated using regulated, private transactions. Pipeline companies obtain long-term contracts from producers and wholesale purchasers, typically local distribution companies (LDCs). Historically, the Federal Energy Regulatory Commission (FERC) accepted these counterparty contracts as sufficient justification of need. Typically the LDCs are themselves regulated firms, which sometimes possess affiliations with pipeline companies. But with contracted costs largely passed through to retail customers via regulated prices, it is unclear whether contracting parties face sufficient competition or otherwise possess an incentive to find least-cost alternatives. To aid evaluation of past and future investments, we develop a national-level optimization model that can assess the need for new interstate pipeline and storage facilities. The model takes production and demand pathways as fixed and minimizes the infrastructure and operation costs of transport and storage in order to balance supply and demand on each day in each state. Transport of gas can be achieved using pipeline transmission of dry gas, or using truck or ship transport of liquefied natural gas (LNG), and optimal placement of liquefaction and gasification facilities. The model also accounts for international imports and exports of both dry gas and LNG. Three underground dry-gas storage facilities are considered, as well as LNG storage. We compare the model's optimized plan with observed outcomes as the sector grew rapidly with hydraulic fracturing. We find that the U.S. has built 38 percent more pipeline and 27 percent more underground storage than necessary, amounting to roughly $179 billion in excess investment. It would have been more economic to expand pipeline far less than observed and instead satisfy critical-peak demands for gas using LNG, plus necessary liquefaction and gasification facilities. Differences between optimized and observed investments vary across the interstate network, while flows between states and into and out of storage bear a close resemblance to observed outcomes.

Keywords: Natural gas pipeline, storage, optimization modeling, regulation

JEL Codes: C61, L52, Q49

1 Introduction

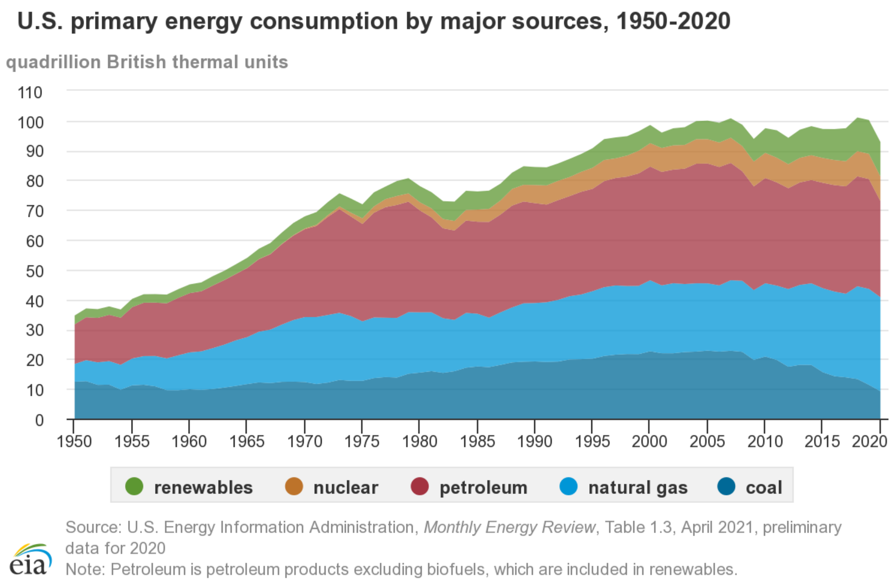
Due to technological advancements in natural gas extraction, natural gas production in the U.S. has increased substantially since 2007, and in 2020, became the largest source (39%) of U.S. energy consumption (Figure 1). The increase in supply contributed to a decline in natural gas prices, which in turn has contributed to increases in natural gas use by the electric power and industrial sectors. Natural gas has been considered a "bridge fuel" in the ongoing energy transition because it is cleaner and more efficient than coal, diesel, and other fossil fuels that is has displaced. With sharply lower fuel costs and stricter emission standards, gas-fired power plants have quickly replaced coal-fired plants since 2011 (EIA, 2019b). Gas-fired plants provide peak and balancing services to complement the variable electricity supply from wind and solar plants, because compared to coal and nuclear, gas plants can more easily ramp their power production up and down with fluctuating demand or renewable supply. American Gas Association and ICF (2022) and Gürsan and de Gooyert (2021) suggest that continued use of natural gas and its vast delivery infrastructure, if combined with an effort to reduce methane emissions and perhaps carbon capture, can increase the likelihood of successfully reaching net-zero targets while minimizing customer impacts.

Looking forward, a number of factors weigh on future demand for natural gas, in addition to decarbonization goals. These include growth of more efficient combined-cycle gas power plants, state-level efforts to grow renewable energy, and falling costs of batteries and other ways to store electricity, all of which will reduce demand from combustion turbines and other fast-ramping gas power plants. As a result of these and other factors, the Energy Information Administration (EIA) projects that growth of natural gas demand in the electricity sector will slow between 2020 and 2050 (EIA, 2021). In contrast, relatively low natural gas prices will drive growth in LNG exports and piped exports to Mexico, as well as growth of natural gas use by the industrial sector, especially chemicals and fertilizer.

Gas supply and demand are changing differently across regions of the U.S.. East and West Coast states are moving away from gas-fired power generation, while Midwest, Southern Mid Atlantic, and Southern regions continue to rely on gas in power generation. The Gulf Coast will see the greatest demand growth, driven largely by exports and concentration of chemical manufacturers. Natural gas production is projected to continue increasing in the Appalachian (Alabama to Maine) and Permian (Texas and New Mexico) basins. Most growth opportunities are connecting the Appalachian and Permian basins with rising demand in the Gulf Coast (EIA, 2021). These shifts in regional demand and supply have and will continue to require gas flows across the U.S. to change significantly (Figure 2).

Meanwhile, these shifts in supply and demand have driven expansion of pipelines and storage to facilitate interstate trade. These shifts, however, are transitory. By 2030 natural gas demand

Figure 1: U.S. Primary Source of Energy

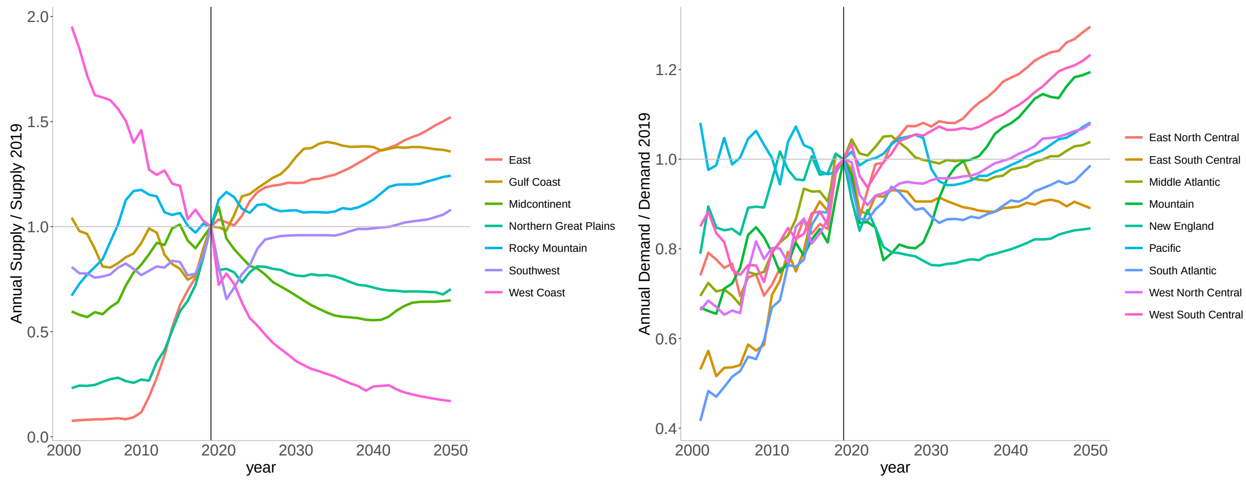


growth from the power sector is projected to slow and then decline. Lower future demand for gas-fired generations would, in combination with the electrification of heating loads in the residential and commercial sectors, imply a reduction in the total volumes of gas transported through the interstate pipelines. Pipelines risk becoming costly stranded assets if they are built without a serious look at how they fit with decarbonization goals. As revenue-regulated utilities, local distribution companies (LDCs) that entered a long-term contract with pipeline companies may pass procurement costs through to their end-use customers. This places the impact of any excess infrastructure costs on ratepayers. High costs of natural gas may lead more customers to electrify heat, which inevitably raises retail prices further for customers unable or unwilling to reduce natural gas use. The greatest burden is likely to be felt by low-income households who are least able to make the up-front investments required to improve energy efficiency or electrify (Davis & Hausman, 2022; Mohlin, 2021). Conducting a wide range literature reviews on direct and indirect effects of natural gas on the energy transition, Gürsan and de Gooyert (2021) argues that while natural gas can directly support renewable technologies in many functions, initial investments in natural gas could lock out emerging renewable technologies for extended periods. Over-investment in natural gas would amplify the complexity of the renewable energy transition, especially in the politics around affordable energy and the threat of losing fossil-fuel-related jobs. Their suggested policy is to specify the upper limits of fossil fuel capacity in the energy mix and a concrete allowable time frame to invest in fossil fuels, including natural gas.

Interstate natural gas pipeline additions and operations are subject to Federal Energy Regulatory Commission (FERC) licensing and rate-making authority. To obtain approval from FERC to build a new pipeline or expand a pipeline's capacity, a company is required to demonstrate

Figure 2: Shifting Natural Gas Regional Supply and Demand

(a) Supply (b) Demand



Note: This graph presents supply by production region and demand by division region as ratio of annual supply (demand) to the supply (demand) in 2019. The vertical black line indicates the year 2019. We aggregate annual-regional supply (demand) from historical monthly-state data in the period 2001-2019. Annual-regional supply (demand) projections in the period 2020-2050 is obtained directly from Annual Energy Outlook 2021 (EIA, 2021).

a market need for capacity expansion. Currently, proof of market need is established using long-term contracts between pipeline companies and gas shippers for gas transportation. These contracts act as transferable property rights for pipeline capacity and can be sold or released to other shippers on a secondary market. In operation, under cost-of-service regulation, pipelines are allowed to earn a reasonable return on their investment. The return includes a return on the pipeline's equity investment, an amount that allows recovery of interest on a pipeline's debt and equity financing, adjusted for income taxes net of the tax shield from debt (FERC, 1999). The simplicity of the regulation has greatly encouraged pipeline investments and has contributed to the speed of natural gas infrastructure development that has occurred with the fracking boom. This practice, however, has been criticized where both pipeline companies and shippers are affiliated entities, which creates an acute conflict of interest that is accentuated when the rate of return is above market rates. The reason is the inherent risk-shifting in such transactions, whereby pipeline developers stand to earn a return above risk and captive utility customers are levied with significant reservation costs regardless of whether their gas utility uses the pipeline capacity (Environmental Defense Fund, 2020; Mohlin, 2021; Tierney, 2019). Given both parties to such contracts are regulated, and costs are born by captive ratepayers, it is unclear whether there exists sufficient competition, or whether the contracting parties otherwise have an incentive to find least-cost alternatives.

Long-term contracts may also engender market power for those who own the pipeline shipping rights, especially in the interconnected natural gas and electricity markets. Marks et al. (2017)

shows that some gas distribution firms have utilized their contractual rights to schedule deliveries without actually delivering gas. This behavior blocks other firms from utilizing pipeline capacity, which limits gas supply to consumption regions like New England, which can drive up gas and electricity prices. This finding is consistent with McRae and Wolak (2019)'s findings on a reliability payment mechanism in the Colombian wholesale electricity market in which large generators who own reliability option contracts have the ability to unilaterally create a scarcity condition and increase market prices above the actual scarcity price.[1]

With regard to the regulated rate of return, von Hirschhausen (2008) estimated the weighted average cost of capital for US interstate pipeline projects between 1996 and 2003 at 11.6%, with returns ranging from 8.4% up to 12.64%, which is well above low-risk market interest rates. The use of rate-of-return regulation for price control may incentivize misallocation of economic resources (Averch & Johnson, 1962), or an excessive scale and cost of developed infrastructure (Stein & Borts, 1972), because the allowed rate of return generally exceeds the market interest rate, and involves minimal risk in practice. Helm and Thompson (1991) argues that the social costs of underinvestment are higher than the social costs of overinvestment and Baumol and Klevorick (2016) argues that the social loss of misallocation can be surprisingly small. Even if improved incentive mechanisms for public investment cannot be devised, there is a clear role for comprehensive models to guide investment planning and regulatory oversight.

In February 2021, FERC issued a Notice of Inquiry (FERC, 2021) seeking suggestions on what methodology and types of additional or alternative evidence FERC should examine to determine pipeline project need. This paper provides a model-based project evaluation method that will be useful to assess the need for a new interstate pipeline in the context of ongoing energy transition and decarbonization policies. In addition, the paper aims to contribute to the ongoing regulatory debate by examining the efficiency of the historical natural gas pipeline and storage investment in the U.S. Evaluation of infrastructure need ought to account for predicted future demand and supply as well as the substitution possibilities between new interstate pipelines, new underground storage, and liquefied natural gas (LNG) in order to meet growth and seasonal demand variation in a least-cost manner. In this study, we build a linear programming model that minimizes the total capital cost and operational cost of pipeline and storage to meet the domestic daily demand in each state of the contiguous U.S. The model determines the need for additional pipeline, storage, and LNG capacity to accommodate flows between supply

[1]McRae and Wolak (2019) study the design of capacity markets in the Colombian wholesale electricity market. These market-based incentives typically take the form of reliability option contracts that provide a implicit financial penalty to generation units that do not produce sufficient energy and a reward to generation firms that produce more than their firm capacity during critical conditions. This reliability payment mechanism aims to provide market-based incentives for plants to produce electricity during periods of system scarcity. The authors show that generation firms with market power may have the ability to create a scarcity condition by withholding their output. The incentive for generators to trigger scarcity conditions depends on the relative magnitudes of their "firm capacity" quantities and their forward contract quantities.

and demand regions. We then apply the model to examine how historic (2002-2019) interstate pipeline, storage, and LNG facility expansion compare to optimal expansion. Although computational optimization and equilibrium models are often used in economics, it is rare to see them systematically compared with observed outcomes, as we do here.

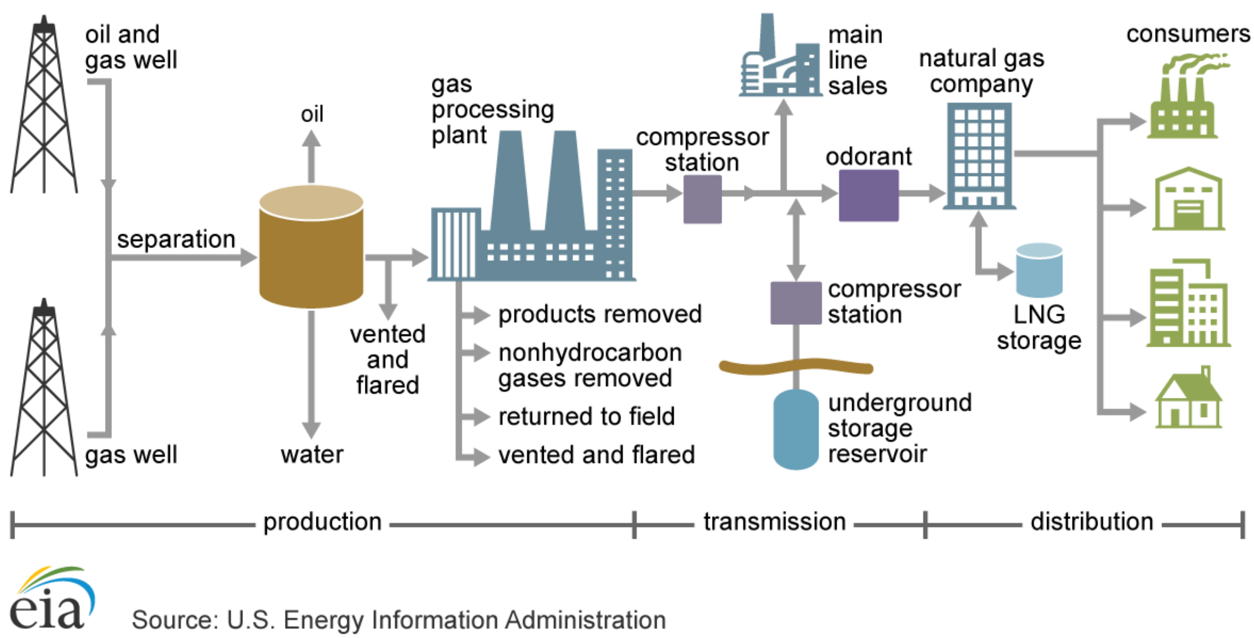
The results indicate that we have generally built too much pipeline and underground storage and slightly less LNG liquefication and gasification plants than optimal. However, this is not the case for all contiguous states. Pipelines have been overbuilt in some regions and underdeveloped in others. States in the Northeast region could have used more natural gas liquefaction and regasification plants. Overall, we find total capital cost spent on historical pipeline and storage projects in 2002-2019 was about 179 billion 2019-dollars higher than the least-cost plan devised by the model, not counting the costs associated with paying an excess rate of return on this investment, which will accumulate well into the future. The model solution suggests that building additional LNG facilities would have been more efficient than building more pipeline and storage in some states. Besides efficiency, a scenario in which we raise peak demand by 20% suggests that investing in LNG plants rather than pipelines also supports energy resilience, especially when the frequency of unprecedented extreme weather events is increasing as a result of climate change.

The rest of this paper is organized as follows: Section 2 provides background information of natural gas transmission and storage system in the U.S. and briefly sketches the trends in natural gas production and consumption. Section 3 presents our novel natural gas model along with its data inputs. We then verify the validity of the model in Section 4 and demonstrate model applications in Section 5. Finally, Section 6 summarises the main results and discusses further applications of the model.

2 Background of Natural Gas Transmission and Storage System

Natural gas is used in the residential and commercial sectors for heating, in the industrial sector for heating, power, and in the petrochemical industry as feedstock, and in electric power generation as fuel. Delivering natural gas from natural gas and oil wells to consumers requires many infrastructure assets and processing steps, and it includes several physical transfers of custody. Figure 3 presents the natural gas supply chain from wells to end-use consumers. EIA (2022b) characterizes three stages in natural gas system: Production, Transmission, and Distribution. Natural gas production can be complex and usually involves several processes to remove oil, water, and other impurities. Distribution is the final step, delivering natural gas to end-use customers. Large customers such as industrial and electric power plants usually receive natural gas directly from high capacity interstate and intrastate pipelines. Small users such as residential and commercial customers receive natural gas from their LDC. In this study, we focus mainly gas transmission, including transportation and storage, which is sometimes called the midstream industry. We also consider LNG storage, sometimes considered as part of the distribution system, because it can substitute for dry gas storage and pipeline. We take production and consumption quantities as exogenous.

Figure 3: Natural Gas System



Note: Natural gas system is often grouped into three categories: Production, Transmission, and Distribution. In this study, we focus on natural gas transmission which includes transportation and storage.

Natural gas transportation in the U.S. relies mainly on the steel pipeline network which are wide-diameter pipes that connect gathering systems in producing areas, natural gas processing plants, other receipt points, and the main consumer service areas. There are three types of natural gas transmission pipelines, including: (1) interstate pipelines that operate and transport natural gas across state borders; (2) intrastate pipelines that operate and transport natural gas within a state border; and (3) Hinshaw pipelines that receive natural gas from interstate pipelines and deliver it to consumers for consumption within a state border(EIA, 2022b). Our study considers the interstate pipeline network wherein each of the U.S. contiguous states is modeled as a node for pipeline deliveries and/or receipts.

Underground storage is an important part of the natural gas delivery system to ensure balance between relatively steady supply and highly variable and seasonal demand. Demand for natural gas is usually higher during the winter due to heating demand in residential and commercial sectors while production and pipeline imports are relatively steady in the short term.[2] Demand can also spike during extreme cold or heat, since gas typically serves peaking power plants that respond to air conditioning and electric heating demand. Gas companies usually store natural gas during periods of low demand for times of peak demand, and as insurance against any unforeseen accidents, natural disasters, or other occurrences that may affect the production or delivery of natural gas (NaturalGas.org, 2013).

Natural gas is stored mainly in large underground reservoirs. The three main types of natural gas underground storage facilities are depleted fields, aquifers, and salt caverns. These types of storage differ in both development cost and number of cycles per year, which we discuss in more detail in section 3.4. *Depleted fields* are depleted natural gas or oil reservoirs, scattered throughout most U.S. regions with storage facilities. These reservoirs are large, but their delivery rates are relatively low, meaning the amount natural gas that can be extracted each day is limited. They typically require a long injection season with moderate withdrawals during winter months. *Aquifer* storage facilities are converted natural aquifers with water-bearing sedimentary rock formation overlaid with an impermeable cap rock. Aquifer storage typically requires larger base gas reserves and allow for less flexibility in injecting and withdrawing. The Midwest has the most aquifer storage. *Salt dome* storage facilities are naturally formed salt caverns shaped into a dome structure through leaching and dissolving the salt. Most salt caverns are located in the Gulf Coast states (South Central storage region) while a few exist in the Midwest and East regions. Salt caverns require very little base gas, and provide high deliverability rates relative to working gas capacity (Fang et al., 2016).

Two uses for natural gas in storage facilities are meeting base load requirements and meeting peak load requirements. *Base load storage* capacity is used to meet seasonal demand increases. Typically, the turn-over rate for natural gas in these facilities is a year; natural gas is generally injected during the summer (non-heating season), which usually runs from April through October, and withdrawn during the winter (heating season), usually from November to March. Instead, these facilities provide a prolonged, steady supply of natural gas. Depleted fields are the most common type of base load storage facility. *Peak load storage* facilities, on the other hand, are designed to have high-deliverability to meet sudden, short-term demand increases. These facilities cannot hold as much natural gas as base load facilities; however, they can deliver smaller amounts of gas more quickly, and can also be replenished in a shorter amount of time than base load facilities. Peak load facilities can have turn over rates as short as a few days or weeks. Salt caverns are the most common type of peak load storage facility, although aquifers may be used to meet these demands as well (NaturalGas.org, 2013).

[2] It takes many months of drilling before a new gas well can produce, and once it does, production from individual gas wells tends to follow a distinct and largely declining pattern over several years. Thus, adjusting supply in response to weather or seasonal variation is difficult. Empirically, aggregate production is smooth.

Underground natural gas storage, with the exception of on-grid salt cavern storage, thus is generally not economical for needle peaking which usually occurs for a day or two on an LDCs system. Mesko and Ramsey (1996) show that LNG is a better option for providing peaking services to LDCs in case pipeline capacity expansion are too costly. Adding pipeline capacity is an inefficient means of securing peaking service as LDCs may actually utilize this additional capacity for only a few days, or only a few hours in some years. In comparison with other alternative peaking sources such as linepack[3] and underground storage, the authors suggest that LNG is very competitive, particularly in the range of up to about 20 days of supply. A typical LNG value chain is composed of gas production, liquefaction, shipping, regasification (vaporization), and pipeline delivery (Figure 4). Natural gas companies and LDCs produce LNG by cooling natural gas to about -260°F at normal pressure. This process results in the condensation of the gas into liquid form, reducing its volume by 600 times. LNG then can be transported in specially-designed trucks, ISO containers, bunker barges, and ocean-going ships to areas where there is a demand for but an inadequate supply of natural gas or an limitation of pipeline capacity. Within continental U.S., LNG is shipped by trucks or rails. At the receipt point, a regasification plant heats LNG so that it expands back into the original gaseous state, for delivery into the natural gas pipeline system. In fact, the U.S. currently has more than 100 active LNG peak-shaving plants and other satellite facilities, The majority of these facilities locate in the Northeast, Upper Midwest, and Southeast (PHMSA - U.S. Department of Transportation, 2022; U.S. Department of Energy, 2013). LNG is an important part of the Northeast's supply and deliverability network since there is no underground storage located in New England[4] due to geologic unsuitability. LNG has been providing about 28% of design day supply in the winter for New England local gas utilities (Northeast Gas Assocition, 2022)

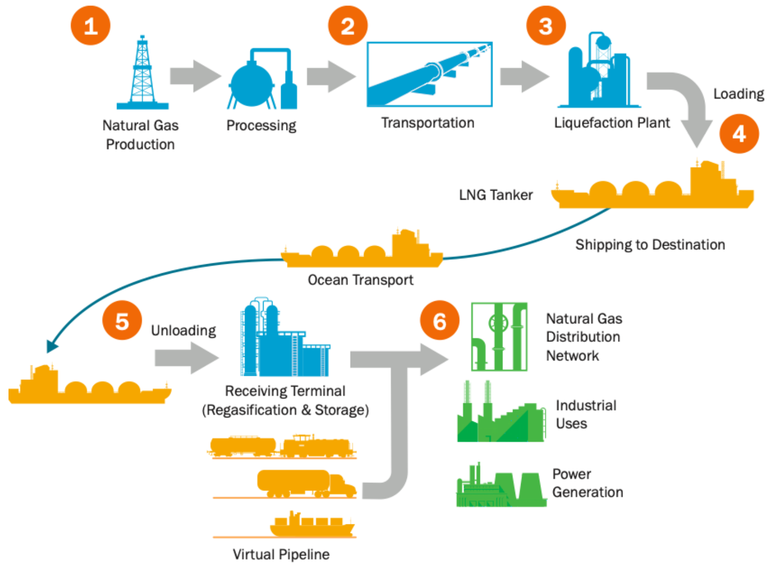
For better understanding of market forces affecting the natural gas transportation and storage infrastructures, we briefly characterize trends in natural gas production and consumption as well as the uses of natural gas infrastructures. Since the fracking boom after 2007, the increase in production contributed to a decline in natural gas prices, which in turn has contributed to increases in natural gas use by the electric power and industrial sectors (Figure 5a and 5b). Demand for natural gas during the summer months thus has been increasing (due to the demand for electricity to power air conditioners and the like.)

The rapid growth of natural gas production and consumption has led to a corresponding expansion of the U.S. pipeline system. Developers and operators have invested billions of dollars to connect major new production regions, such as the Marcellus (Pennsylvania) and Bakken (North Dakota) shale basins, to traditional gas markets. They have reversed and expanded

[3]Linepack refers to the volume of gas that can be stored in a gas pipeline.

[4]New England includes six northeastern states: Maine, Vermont, New Hampshire, Massachusetts, Connecticut, and Rhode Island.

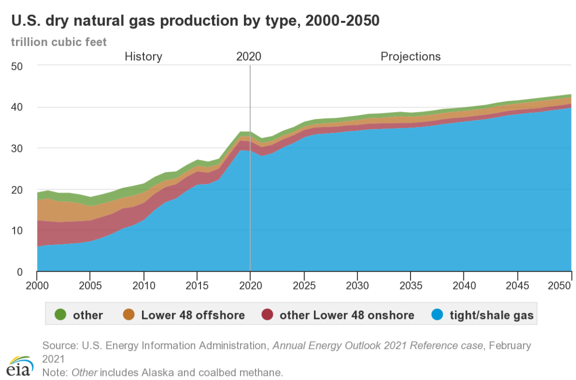
Figure 4: Liquefied Natural Gas Value Chain



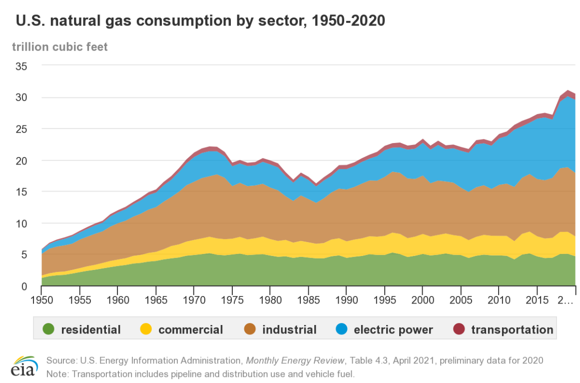
Source: U.S. Department of Energy

Figure 5: Production and Consumption Growth

(a) Growth in production



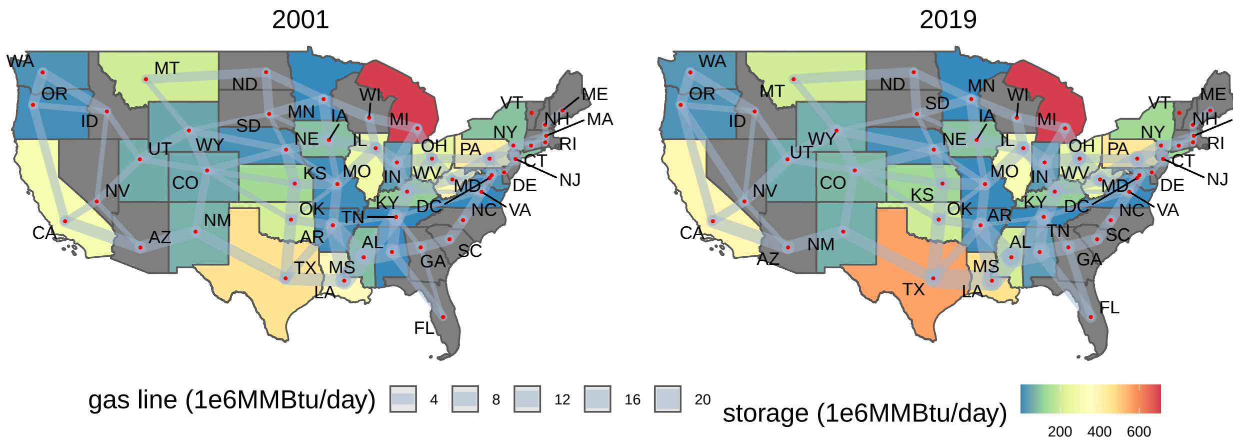
(b) Consumption Growth by End Use



existing pipelines, and developed entirely new, long-haul pipelines to reconfigure natural gas flows throughout North America (Diaz, 2021). Figure 7 shows interstate pipeline capacity has remarkably expanded within the production regions in South Central (Texas, Louisiana, Alabama, Mississippi) and between emerging production states (Pennsylvania, Ohio, and West Virginia) and their neighbor states. Similarly, underground storage has been expanded in most states during the period 2002 - 2019, except the New England due to geologic constraints. The storage utilization rate during winter reach 80-90% (Figure B.2 and Figure 6), especially states at the East and Northwest regions.

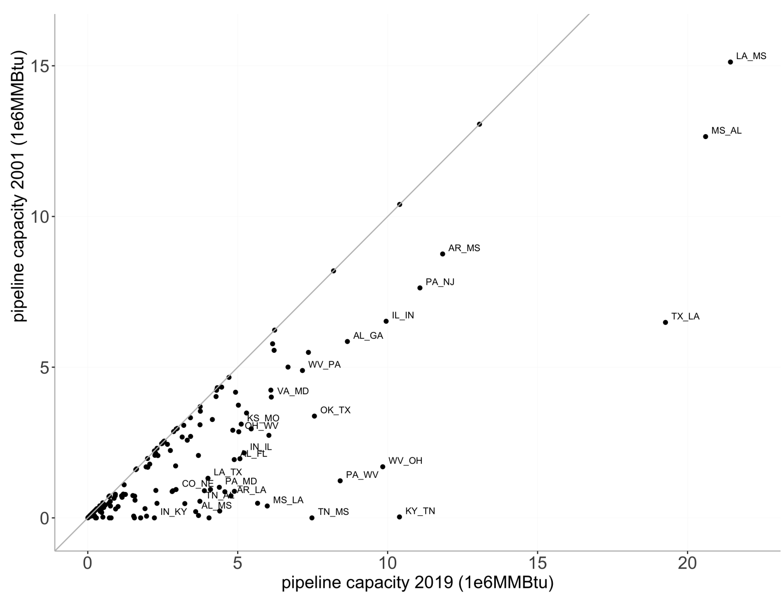
Given these characteristics of the U.S. natural gas system, the capacity expansion model pre-

Figure 6: Pipeline and Storage Capacity Expansion



Note: These maps illustrate the expansion of underground storage and interstate pipeline in 2019 versus 2001. The grey states have no underground storage (they may have above-ground LNG storage capacity). Segments connecting each pair of states indicate interstate pipelines, with the segment size indicating the pipeline capacity.

Figure 7: Interstate Pipeline Capacity Expansion



Note: This graph shows interstate pipeline capacity in 2019 plotted against capacity in 2001 for each directional connection. Labels are shown for connections that increased more than 2 million MMBtu/day. The label displays a pair of two-letter state abbreviations in which the first state is delivery point and the second state is receipt point. For instance, LA_MS represent the natural gas flow from LA (Louisiana) to MS (Mississippi) and MS_LA represent the gas flow from MS to LA.

sented in section 3 includes three main elements of natural gas transmission and storage infrastructures: interstate pipelines, underground storage, and LNG facilities.

3 A Capacity Expansion Model for Natural Gas Transportation and Storage

3.1 Existing Natural Gas Models

Along with growing natural gas uses in the U.S., several natural gas model have been recently built to either predict the natural gas market clearing prices and volumes or pipeline network expansions. The most recent model is EIA's Natural Gas Market Module (NGMM) of the National Energy Modeling System (NEMS) (EIA, 2018). EIA uses this model to project wellhead, spot, and market prices that balance *monthly* natural gas supply and demand through a simplified North American pipeline network. The objective is to maximize consumer plus producer surplus minus variable transportation costs. In the NGMM, storage capacity and storage withdrawals and injections are exogenous. The NGMM model solves for interstate pipeline capacity additions and transmission flows to meet peak demand in January and August each year. The NGMM represents natural gas markets in U.S., Canada, and Mexico at the regional - annual level.

Another popular model is ICF's Gas Market Model (GMM) (ICF Consulting Canada, 2015; U.S. Environmental Protection Agency, 2021), used by EPA and INGAA. ICF uses the GMM for generating the natural gas supply curves, which is a key input in its Integrated Planning Model (IPM®, an integrated wholesale power model). GMM is a linear programming model that incorporates a detailed representation of gas supply, demand, and an integrating pipeline transportation model to develop forecasts of gas supply, demand, prices, and flows in the North American gas market. Generating capacity and storage are given in GMM. Pipeline capacity expansions are input to GMM for the near–term and are endogenous for the longer term. GMM incorporates different regional analysis level for each variable and employs monthly data.

Commercial natural gas models include BRAC, Inc.'s Gas Pipeline Competition model (GPCM) and Deloitte's North American Gas Model (NAGM) (Busch, 2014; FERC, 2020). GPCM is a network partial equilibrium model of the North American natural gas market. This model is a tool for developing forecasts and scenarios for North American natural gas flows, price, and basis. It is used by many pipeline, storage, and consulting companies. In addition, the Federal Energy Regulatory Commission (FERC) and Sandia National Laboratories are licensees. The model solves for the peak and off-peak periods. The model decides pipeline capacity additions and storage withdrawals in the peak period. They then use the results to set pipeline capacity and storage injections in the off-peak period. In GPCM, transportation cost is a function of the flow. Storage injection and withdrawal costs and compressor fuel use are exogenous. On the other hand, Deloitte's NAGM simulates regional interactions of supply, transportation, and demand and determines market clearing prices, pipeline and storage flows, and pipeline capacity additions in the North American natural gas market (Deloitte, 2012). For more information about other natural gas models, see Busch (2014).

Overall, these natural gas models share some common features. They focus on market equilibrium and pipeline capacity expansion. Analysis is based on monthly or seasonal basis and at regional level. All of these models take storage capacity as exogenous and do not include the LNG sector. The LNG sector is an important part of the U.S. natural gas system, especially to serve peak demand in the Northeast region. Our model improves upon existing models by simultaneously optimizing capacity expansion of all three elements of natural gas infrastructures: pipelines, underground storage, and LNG facilities. We do this while co-optimizing pipeline flows, storage injections and withdrawals, as well as LNG flows, all on a daily basis, greatly improving upon the monthly time step in existing models. By doing so, we account for complement and substitution between pipelines, storage and LNG facilities in meeting local demand, and can account for critical peaks that occur with extreme weather. The geographic unit of analysis is the state, which is smaller than many other models. Each state serves as a node, and we solve for storage capacities, pipeline capacities between all pairs bordering states (separately for each direction), daily injections or withdrawals from storage, and daily transmission to and from each bordering state.

Because we currently focus on the efficiency of historic capacity expansion, we take supply and demand as exogenous, while optimizing gas flows and capacity additions. Extensions could endogenize supply, demand, and prices. Because cost of pipeline and storage presumably influences retail gas prices, and different prices would presumably feedback and influence quantities of gas, there may be small differences between our solution if demand and supply were endogenous. We believe these differences would be small given relatively inelastic demand for gas (Hausman & Kellogg, 2015) paired with the fact that transmission costs are a modest share of average price.[5]

[5]There is also some ambiguity about the influence of infrastructure on prices. Excess capacity would presumably increase average cost while lowering marginal cost during the times and places where capacity constraints bind in the optimized system but do not bind in the overbuilt system. The quantity effects would therefore depend on the structure of retail price tariffs, and whether excess fixed costs would be recaptured via fixed charges or volumetric prices. Excess capacity may also cause a larger difference between wholesale prices and retail prices, with the excess cost borne more by LDCs, while prices in the electricity and industrial sectors, which mostly engage with real-time wholesale markets, would see lower prices. These countervailing effects on quantities further suggest that endogenizing quantities would have minimal influence on optimal transmission infrastructure.

3.2 Model Structure

The model developed in this paper possesses a structure that is similar to state-of-the-art capacity expansion models used for integration of renewable energy and storage in electricity systems. The model simultaneously optimizes investment decisions and operations decisions in order balance supply and demand on all days in all states, while satisfying import and export demands. We build the model using a mixed-integer linear-programming software, SWITCH 2.0 (Fripp, 2016; Johnston et al., 2019), an open-source platform for optimal capacity planning that thus far has focused on electric power systems. SWITCH 2.0 has been employed in analyzing least-cost storage and transmission capacity expansion for California, Western North America, and Hawaii (Das et al., 2016; Fripp, 2012; Nelson et al., 2012). We modify the architecture to conform with gas networks. Unlike most implementations of capacity expansion models, we compare model solutions to realized historical investment and operations.

The optimization model minimizes the discounted total capital and operational cost of interstate pipeline, underground storage, and LNG facilities to meet the natural gas demand at each state in the 48 U.S. contiguous states and the District of Columbia. The main decision variables are necessary additional capacity of underground storage, LNG facilities, and state-to-state pipelines to accommodate the gas flow between supply and demand regions over the study period. In addition, the model computes optimal daily volume of natural gas injections into and withdrawals from underground storage, LNG facilities, as well as pipeline deliveries and receipts to meet the daily demand at each state. Constraints require that the total volume of natural gas from local production, net imports, net storage withdrawals, LNG regasification, and net pipeline receipts provides adequate natural gas for local consumption during on the daily basis. The amount of natural gas in underground storage, regasified from LNG, and transmitted through interstate pipeline are constrained by the capacity of underground storage, LNG facilities, and pipeline in each time period, respectively. Local production, consumption, and imports from and exports to other countries are taken as exogenous. Figure 8 illustrates variables of gas flows between states and across the gas value chain. Figure B.1 summarizes the objective function, decision variables, and constraints in the model. The mathematical optimization problem is as follows. Variables are indexed by states i and time t (monthly).

Capital Decision Variables

S_{it}: New storage capacity added.

Q_{it}: New LNG facility capacity added.

X_{jit}: New transmission build from state j to state i in period t.

Operations Decision Variables

w_{it}: Net storage withdrawal (negative implies injections) in state i and period t.

g_{it}: Net LNG volume used in regasification to produce dry natural gas (negative implies liquification) in state i and period t.

x_{jit}: State i net receipts from state j through pipeline in period t (negative values are deliveries).

Exogenous Variables and Parameters

δ_{it}: Demand in state i and period t.

η_{it}: Net imports (negative values are exports) from or to foreign countries in state i and period t.

β_{jit}: The fixed cost of expanding gas line capacity by X_{jit} MMbtu for transmitting gas from state j to state i.

γ_{it}: The fixed cost of expanding gas storage in state i in period t.

ρ_{it}: The fixed cost of expanding LNG facilities in state i in period t.

Endogenous Variables

Z_{it}: Total storage capacity in state i and time t

G_{it}: Total LNG capacity in state i and time t

L_{jit}: Total transmission capacity from state j to state i in time t

s_{it}: Amount of gas stored in state i and time t.

Characterization

For our purpose of identifying infrastructure capacity to meet demand during the historic period of 2002-2019, in this model, demand δ_{it}, production y_{it}, and net imports η_{it} are assumed to be exogenous for each state i and time t. These in fact may vary with weather, month-of-year, and other market factors. Below we characterize the assumed flow of production.

Storage capacity at state i in time t is

$$Z_{it} = Z_{i,t-1} + S_{it} - R_{it} \tag{1}$$

LNG capacity at state i in time t is

$$G_{it} = G_{i,t-1} + Q_{it} - K_{it} \tag{2}$$

and directional gas line capacity from state i to state j at time t is

$$L_{jit} = L_{ji,t-1} + X_{jit} - T_{jit} \tag{3}$$

where R_{it}, K_{it}, and T_{jit} are, respectively, retirements of storage, LNG facilities, and line capacities in state i in time t.

Finally, underground storage and LNG storage, respectively, evolve over time by the relation:

$$s_{it} = s_{i,t-1} - w_{i,t} \tag{4}$$

where $w_{i,t} = s_{wid_{it}} - s_{inj_{it}}$ with $s_{wid_{it}}$ is storage withdrawals and $s_{inj_{it}}$ is storage injections.

and

$$g_{it} = g_{i,t-1} - g_{i,t} \tag{5}$$

Objective and Constraints

Let T denote number of time periods, θ be a discount factor satisfying $0 \leq \theta \leq 1$, and c_l, c_g, c_s respectively denotes variable unit cost of gas transmission, regasification, and storage withdrawal/injection. Given demand δ_{it}, net import η_{it}, and initial storage amount s_{i0}, the cost minimization problem is defined as follows.

$$\min_{S,G,X,s_{inj},s_{wid},g,x} TC = \sum_{i=1}^{49} \sum_{t=1}^{T} \theta^t (\sum_{j \neq i} \beta_{jit} X_{jit} + \gamma_{it} S_{it} + \rho_{it} G_{it} + c_l \sum_{j \neq i} x_{jit} + c_g g_{it} + c_s s_{inj_{it}} \tag{6}$$

subject to:

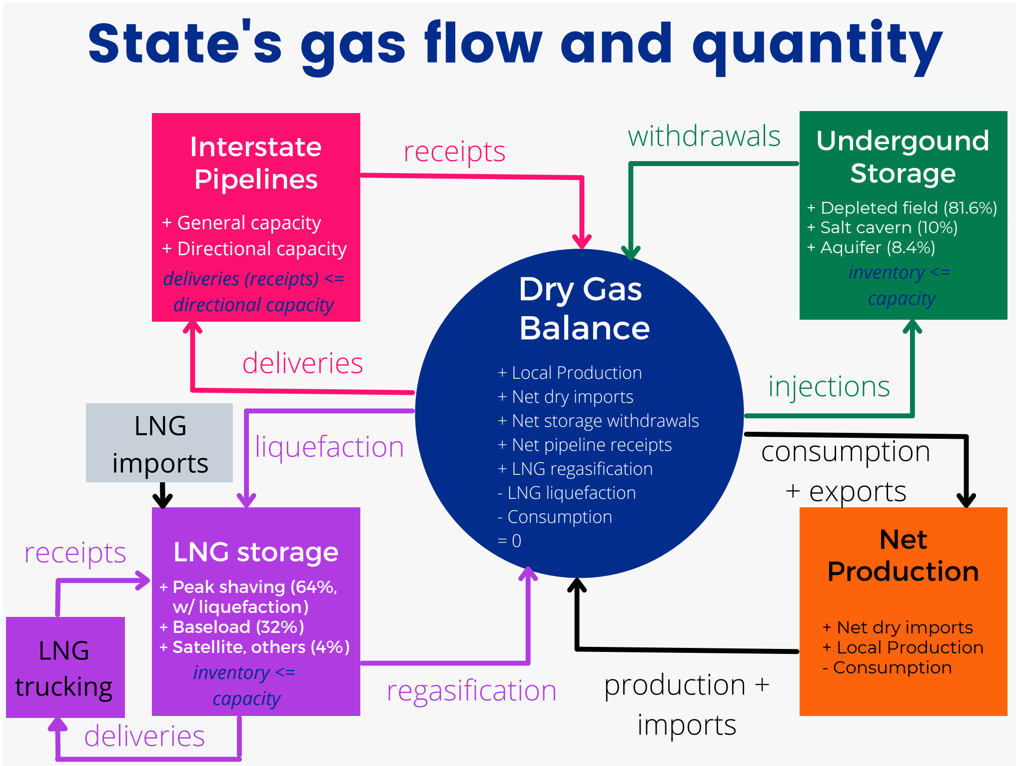
$$
\begin{aligned}
s_{it} + s_{wid_{it}} - s_{inj_{it}} &= s_{i,t-1}, && \text{given } s_{i0} && (7)\\
Z_{it} &= Z_{i,t-1} + S_{it} - R_{it} && \text{(Storage capacity)} && (8)\\
G_{it} &= G_{i,t-1} + Q_{it} - K_{it} && \text{(LNG capacity)} && (9)\\
L_{jit} &= L_{ji,t-1} + X_{jit} - T_{jit} && \text{(Pipeline capacity)} && (10)\\
s_{it} &\leq Z_{it} && \text{(Storage constraints)} && (11)\\
g_{it} &\leq G_{it} && \text{(LNG facilities constraints)} && (12)\\
x_{jit} &\leq L_{jit} && \text{(Pipeline constraints)} && (13)\\
\sum_{i=1}^{N} \sum_{j \neq i} x_{jit} &= 0 && \text{(U.S. Gas flow balance)} && (14)\\
\sum_{j \neq i} x_{jit} + w_{it} &= \delta_{it} - y_{it} - \eta_{it} && \text{(Balance constraints)} && (15)\\
s_{it} &\geq 0 && && (16)\\
S_{it} &\geq 0 && && (17)\\
X_{jit} &\geq 0 && && (18)\\
\forall i \in [1, 49] &\text{ and } \forall t \in [1, T]
\end{aligned}
$$

LNG facilities include liquefaction, regasification and storage. We denote a general term for LNG sector in the above mathematical problem for simplicity and to fix attention on dry natural gas flows. Our computational model in fact separately optimize capacity additions for LNG storage, liquefaction, and regasification as well as LNG amount to liquefy or regasify, and LNG volumes shipped from a state to another. Capacity and flow constraints applied on each element of LNG facilities are the same as in Equations 9 and 12.

3.3 Implementation

We use the model to examine the efficiency of historical pipeline expansion from 2002-2019 period, and assuming EIA projections of demand, supply, imports, and exports through 2050. To reduce the computational cost for running the model, we include a representative year for every five-year investment interval (Table 1) and account for every day in one year to reflect between day variation of gas supply-demand and interstate movement. The quantity in a representative year is the consumption weighted average over five years within each period. The model solution gives the least-cost expansion of pipeline, underground storage, and LNG facility capacity built each period. Finally, we contrast optimal outcomes and observed historic capacity

Figure 8: Model variables



Note: This diagram illustrate gas flows and quantities at each infrastructure element: interstate pipeline, underground storage, and LNG facility. The centered circle represents a state's net consumption, with the text describes the balance constraint in equation 15. The arrows show gas in-flow and out-flow from pipeline, storage, and LNG facilities to serve local demand in each state in each day. The black text and arrows represent gas amount and flows that are exogenous in the model.

in the period of 2002-2019[6] to examine whether the states have over-invested in pipelines given EIA-projected supply and demand to 2050.

Besides the baseline, overall least-cost scenario, we optimize capacity expansion in three other alternative scenarios (Table 2). We consider one scenario in which the extreme cold or hot weather leads to a sudden surge in local demand for natural gas while supply remain constrained. The winter of 2021 was considered the US's coldest winter within past 30 years and total natural gas consumption during the peak period from January through March increased by 17% relative to the average for that period from 2002-2019. We therefore set a demand buffer of *20%* higher than the original peak demand in both winter and summer each year. The model solution indicates the additional pipeline, storage, and LNG capacity needed to meet local critical peak demand under extreme weather conditions. The other two scenarios examine the efficiency of pipeline, storage, and LNG investment assuming the existing storage and LNG (pipeline) capacity that might be planed independently. That is, we fix one element of the

[6]the optimal capacity expansion in study periods from 2002 through 2021 is compared to historic capacity expansion in the period of 2002-2019

Table 1: Investment Periods in the Optimization Model

Period#	Investment period	Period start	Period end
1	2002	2002	2006
2	2007	2007	2011
3	2012	2012	2016
4	2017	2017	2021
5	2022	2022	2026
6	2027	2027	2031
7	2032	2032	2036
8	2037	2037	2041
9	2042	2042	2046
10	2047	2047	2050

Note: The study time span is 2002-2050, split into ten investment periods. Except the last period, each investment period covers five years.

system at observed levels (e.g., pipeline), and then solve for the least-cost capacity of the other two (storage and LNG). A comparison across the optimal results from these scenarios provide insight into the substitution of pipelines, storage, and LNG facilities, as well as the relative importance of storage, LNG facilities versus pipelines in ensuring energy resilience.

Table 2: Optimization Scenarios

Scenario name	Scenario description	Pipeline 2002-2021	Storage 2002-2021
Base	Gas demand, supply, cross-border imports and exports, and unit capital cost are exogenous.	Optimized	Optimized
Demand Buffer	Demand on the peak day in each winter and summer increase by *20%*. Supply remains the same as in Base scenario.	Optimized	Optimized
Fixed Pipeline	Pipeline capacity is predetermined at observed historic level in 2002-2019.	Exogenous	Optimized
Fixed Storage-LNG	Storage capacity is predetermined at observed historic level in 2002-2019.	Optimized	Exogenous

Note: The model always optimizes capacity additions of pipeline, storage, and LNG facilities in 2022 - 2050.

The model is based on several underlying assumptions:

1. The time that pipeline/ storage/ LNG projects start, stakeholders can predict correctly demand (consumption and net export), supply (production), and costs.

2. Pipeline, storage, and LNG capacity were built to meet the daily demand in each contiguous state.

3. Free disposal of excess supply, which may be unaccountable loss in transmission and storage, or inputs for processing LNG exports, or flaring.

4. Secondary market (for trading the right of transport and storage capacity uses) works efficiently as a supplement to the long-term take-or-pay pipeline and storage contracts.

3.4 Data

We obtain the historical natural gas data up to 2019, both market elements and infrastructure capacity, from the U.S. Energy Information Administration (EIA, 2022a). We employ projected data in the reference projection case from EIA Annual Energy Outlook 2021 (EIA, 2021) for regional natural gas supply, demand, and imports - exports. Assuming all states in each region will have the same growth rate in natural gas supply, demand, and trades, we compute the projected volume in 2020-2050 at state level by multiplying state's volume in 2019 by regional growth rate. Besides, we gather data for capacity of LNG facilities from the PHMSA - U.S. Department of Transportation (2022). The model input is daily data at the state level.

Production data

Production data input is a sum of dry natural gas production and supplemental gaseous fuels. EIA's data by state is available at monthly level since 2006 and at annual level since 1986. Meanwhile, aggregate national monthly data is available since 1986. We thus obtain monthly data for the period of 2002-2005 by proportionally distribute annual amount to each month of year using the ratio of national monthly quantity to annual quantity. Besides, total dry production in four Gulf coast states (Texas, Alabama, Louisiana, and Mississippi) includes the state's local production and the amount delivered from Gulf of Mexico. We then linearly interpolate monthly production into daily production while conserves the actual monthly mean.

Consumption data

Demand data is total dry natural gas consumption from all end-use sectors (residential, commercial, industrial, electric power, and vehicle fuel), obtained on a monthly basis from the Energy Information Administration. Most previous models similar to ours use this monthly data for their models, but it is important to consider flows on a shorter time scale to ensure that enough gas can be delivered to handle critical peak demands that occur with extreme weather events.

To obtain daily demand, we regress monthly demand on monthly heating degree days (HDD) and monthly cooling degree day (CDD) as specified in the Equation 19 below. The fit of this regression is strong and shows gas demand to increase significantly with both HDD and CDD, but more strongly with HDD. We then use the estimated relationship to predict daily demand by substituting daily CDD and HDD data into estimated relationship. Daily temperature observations were obtained from the 4-km gridded weather data developed by PRISM Climate Group (2021). To estimate continuous, daily CDD and HDD for each day, we assume that temperature follows a sine curve within each day between its minimum and maximum temperature using the methods developed by Schlenker and Roberts (2009). We then aggregate PRISM grid values to the state level by population-weighting individual grids using 1-km gridded population of the world from Socioeconomic Data and Applications Center (SEDAC) (2021). Monthly values for CDD and HDD simply aggregate the daily values.

Denoting q_{it} as average daily gas consumption in state i and time period t, we estimate a separate regression for each sector (residential, commercial, industrial, electric power, and vehicle fuel) using the specification:

$$\ln q_{it} = f^i(t) + h^i(t)HDD_{it} + g^i(t)CDD_{it} + \gamma X_{it} + u_{it} \tag{19}$$

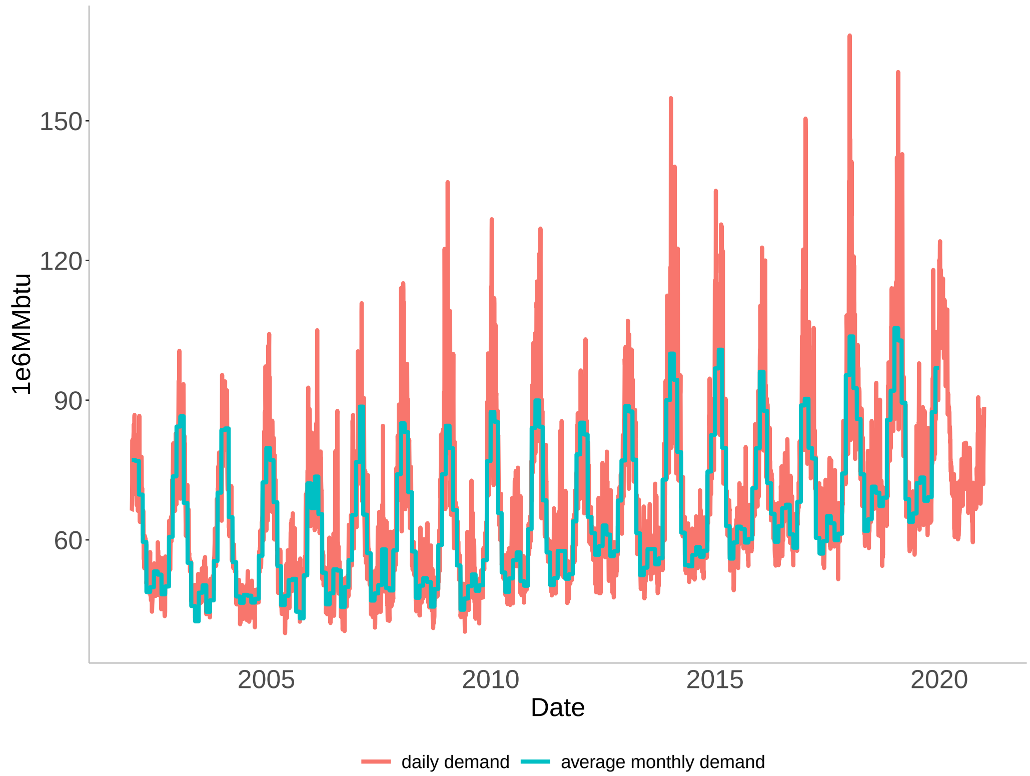
where $f^i(t)$, $h^i(t)$ and $g^i(t)$ are smooth functions of time estimated using a natural splines with three degrees of freedom (i.e., three knots each). By interacting the spline functions with CDD_{it} and HDD_{it}, we allow the sensitivity of demand to temperature to change smoothly over time (say, due to changes in mean energy efficiency). By estimating a separate spline for each state, we allow temperature sensitivity to also differ across states. Additional variables in the vector X_{it}, with estimating coefficient vector γ, are consumption-weighted average national HDD, consumption-weighted average national CDD, and one-month lags and leads of HDD and CDD at the state and national levels, and the share of days in the month that are weekends. Note that non-local and/or non-current weather can influence local consumption via prices; but the links are weak compared to local, current HDD and CDD (Hausman & Kellogg, 2015). Using cluster-robust standard errors by state-year, the weather variables are highly significant. The daily estimates from each sector, imputed by replacing monthly weather with daily weather and a zero-or-one indicator for weekend days, are then aggregated to state-by-day total demand values.

To illustrate how much daily downscaling matters, Figure 9 shows predicted aggregate gas use by day overlaid with monthly average use. Monthly values smooth out much of the variation across days within a month and does not reveal actual peak demand. Building the model at daily level better reflects the infrastructure requirements needed to meet peak demands.

Import and Export Data

We use EIA's data of *"Imports by point of entry"*, *"Exports by point of exit"*, and *LNG imports*

Figure 9: U.S. aggregate national daily demand versus average monthly demand



Note: This figure shows U.S. monthly natural gas demand (in blue) together with estimated daily natural gas demand (in orange) over time. Daily demand is estimated from daily weather using the link between monthly weather and monthly demand specified in Equation 19. Estimates also account for differences between weekends and weekdays.

terminal instead of imports and exports *by state* to account for the actual physical gas quantity at each state at a time and the need of storage capacity and pipeline capacity to transport the natural gas from the entry point to other states or from the production states to the exit point. We aggregate the volume of imports and exports over points of entry/exit or terminals within a state to obtain data at the state level. Export volume includes natural gas exported in different forms — pipeline, compressed natural gas, and LNG. Import volume includes natural gas exported through pipeline and compressed natural gas. We separately include LNG imports in the model because this is the main LNG supply source for the Northeast states. Main LNG import terminals are in Massachusetts, Maryland, Georgia, and Louisiana. LNG may be either regasified to dry natural gas at the import terminal and transmitted through pipeline or shipped by truck to other states and regasified at the destination.

We compute pipeline unit capital cost based on the median cost reported in the EIA Natural Gas Pipeline Projects data and state's cost multipliers reported by American Petroleum Institute (ICF International & API, 2017). Storage capital cost is based on the infrastructure report in ICF International and API (2017). LNG capital cost is approximated based on Mesko and

Ramsey (1996) and Katulak (2016). Lifetime for all infrastructures are assumed of 50 years. Fuel cost for compressor stations to boost the pressure needed for natural gas moving through the steel pipe is 0.04% per every 120 km. Compressor stations are located about every 50 miles to 100 miles along pipelines (American Gas Association, 2021; Clowney, 2003; EIA, 2007; Ulvestad & Overland, 2012). In the model, we use the average distance which is 75 miles, or approximately 120 kilometers. Gas fuel and loss is 2% of injection volume for underground storage, based on average level of fuel reimbursement in storage tariff across gas companies such as Kinder Morgan and Tallgrass Energy, which range between 1% to 2.6%. LNG liquefaction on average costs 9% and regasification costs 5% of feed gas, according to Chong et al. (2019) and Petrowiki (2018).

Table 3: Gas Underground Storage Operations

Type	Mean Cushion* to Working Gas Ratio$^{(1)}$	Injection Period (days)$^{(2)}$	Withdrawal Period (days)$^{(2)}$	Cycles per year	Capital cost$^{(3)}$ (2019$/MMBtu working-gas)
Aquifer	68%	200 - 250	100 - 150	1	39.8
Depleted field	48%	200 - 250	100 - 150	1	19.2
Salt cavern	34%	20 - 40	10 - 20	12	33.7

(1): average over all states in 2017-2020;

(2): Federal Energy Regulatory Commission (2004); (3): ICF International and API (2017)

(*): Cushion gas (or base gas) is volume of natural gas that must remain in the storage facility to provide the required pressurization to extract the remaining gas.

In 2020, depleted fields account for 81.6% of the total U.S. underground natural gas working storage capacity[7], while the rest is salt caverns (10%) and aquifers (8.4%). Developing an aquifer formation as a storage facility often is expensive due to the requirement for large volume of cushion gas. Therefore, aquifers had been used for natural gas storage only in states that geologic characteristics do not allow for developing other types of storage. Depleted field and aquifer storage have low cycle and are mainly used to store natural gas during summers and withdraw in winters. Meanwhile, salt caverns provide very high withdrawal and injection rates relative to their working gas capacity, and cushion gas requirements are relatively low (Table 3). Salt caverns can readily begin flowing gas on as little as one hour's notice, which is useful in emergency situations or during unexpected short term demand surges. Therefore, salt caverns are best suited for peak load storage. However, salt cavern storage facilities are not available in every state, primarily Gulf Coast and Northern states (NaturalGas.org, 2013). In this study, we set a parameter that allows specific state to build specific types of new storage facilities depends on their geologic characteristics. In particular, we allow salt caverns to be built in only

[7]Working gas is the volume of natural gas in the storage that can be extracted during the normal operation of the storage facility. This is the natural gas that is being stored and withdrawn; the capacity of storage facilities normally refers to their working gas capacity (NaturalGas.org, 2013)

those states where salt domes have been discovered. Depleted fields are allowed in those states that have drilled oil or gas wells. We assume that all states, except New England, allow aquifer storage.

4 Model Validation

We check the credibility of the model by examining the extent to which the outputs: (1) generate coherent patterns of transmission and storage behavior in comparison to actual history; (2) examine how shadow values of the balance constraints vary over time and location in relation to actual bottlenecks on the system; and (3) compare model flows to historic pipeline capacity utilization rates. Utilization is measured as the ratio of the total amount of natural gas move through a pipeline and full operating capacity of the pipeline. Utilization rates typically run from zero to one.[8].A high capacity utilization rate indicates the possibility of pipeline bottlenecks under peak loads. A low utilization rate implies pipeline overcapacity. As a rule of thumb, to guarantee sufficient return on a pipeline project to investors, a pipeline requires a utilization rate of around 50%

First, we find that the natural gas system operates reasonably under both peak and low demand times. Resulting pipeline inflows and outflows, underground storage withdrawals and injections, LNG liquefaction and regasification amount, and LNG shipping between states behave logically and follow stylized facts in the natural gas market. Specifically, natural gas is injected into underground storage in summer and withdrawn from storage in winter. LNG is mainly consumed in New England states in winter since these states do not have underground storage due to geologic conditions and limited pipeline access through New York. Most consumption states rely on pipeline deliveries to meet local demand. Figures B.3 and B.4 show the similarity between optimal and observed pattern in how pipeline, underground storage, and LNG facilities are utilized to meet local demand at each state.

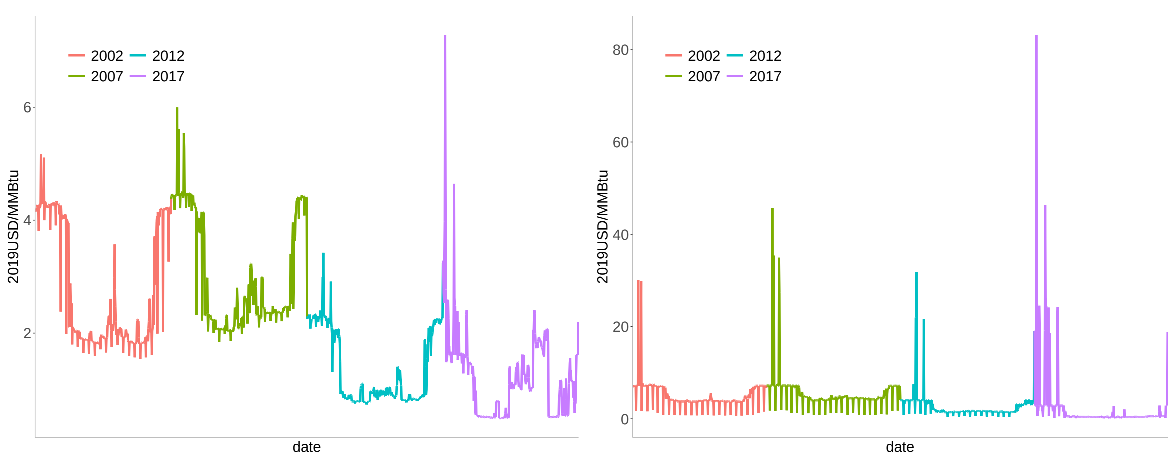
Second, shadow prices obtained from balance constraints in the optimization model well reflects the historic price volatility and seasonality. Figures 10a and 10b present consumption weighted average shadow price of the U.S. and New England region, respectively. The figures clearly display price seasonality with higher peak price in winter and lower peak price in summer. The lower peak in summer is especially visible in the period 2017 - when gas-fired power generation became more prevalent. On average across the U.S., shadow prices were highest in the period 2002 and 2007, reduced in the period 2012, and increased in the period 2017. This price variation is consistent with historical events during the periods 2002-2019. High price in period 2002 and

[8]It is normal that sometimes pipeline capacity utilization rate is greater than one. This happen when compressor stations are upgraded to push pipeline deliverability above initial design capacity.

2007 reflects the natural gas shortage in 2000-2008 and prices spiked due to a decline in the production and increase in demand for electricity generation and due to increase in oil price. Among four periods, only periods 2002 and 2007 experiences clear needle peak price in summer. This is consistent with price spike in August 2005 due to supply disruption after Hurricanes Rita and Katrina (FERC, 2005) and record high price of natural gas in July 2008, when Henry Hub prices peaked at $13.31/MMBtu (Congressional Research Service, 2011; FERC's Office of Enforcement, 2009). The low price in period 2012 is congruent with the fact that natural gas fracking technology led to new production started in the Northeast (Pennsylvania and West Virginia) and Ohio in 2009 and a steadily increase in supply since late 2014 (Xiarchos et al., 2020). While prices generally remained low in 2017, the winter price spike is compatible with the cold wave in the Northeast winter 2017/2018. In particular, New England, the states that experienced a severe cold spell in 2017-2018, have average shadow price of about $83/MMBtu in January of period 2017, close to the historic spot prices of $78.98/MMBtu at New England's natural gas and electricity market hub, Algonquin Citygate, at the same time (EIA, 2019a). Figure B.5 displays daily shadow prices over time by state (note that y-axis scale is different across states). Compared to all other states, the shadow price is much higher in New England states and Florida, where there is no underground storage. The model thus rationally proposes to build underground storage in Florida[9] (Figure 12) and more LNG plants in New York and New Jersey to serve New England's winter demand (B.12a.)

Figure 10: Average Shadow Prices
(consumption weighted)

(a) National Shadow Price (b) New England Shadow Price



Note: This figure shows the variation of shadow prices which is the marginal value (lambda) associated with relaxing the balance constraints in the Equation 15. The left panel displays consumption weighted average shadow prices in the U.S.. The right panel displays the weighted average shadow prices in New England area. Note that the scale of vertical axes is different between the two panels.

[9]Recall that New England's geologic characteristics do not allow for underground storage.

Finally and most importantly, the model's decision in adding new pipeline capacity is consistent with the observed actual needs for gas transmission through pipeline. When putting together the difference of the optimal pipeline capacity from the observed historic capacity and the actual observed pipelines' utilization rate (Figures 12 and 13), we find that those pipeline routes which the model decides to not build or build less capacity than in reality, the utilization rate is less than ten percent, or even equals zero on some pipelines. In contrast, those pipelines which the model chooses to build more or the same as observed levels have higher utilization rate. We discuss this in more details in section 5.1.

5 Applications

5.1 Are We Building Too Much Pipeline?

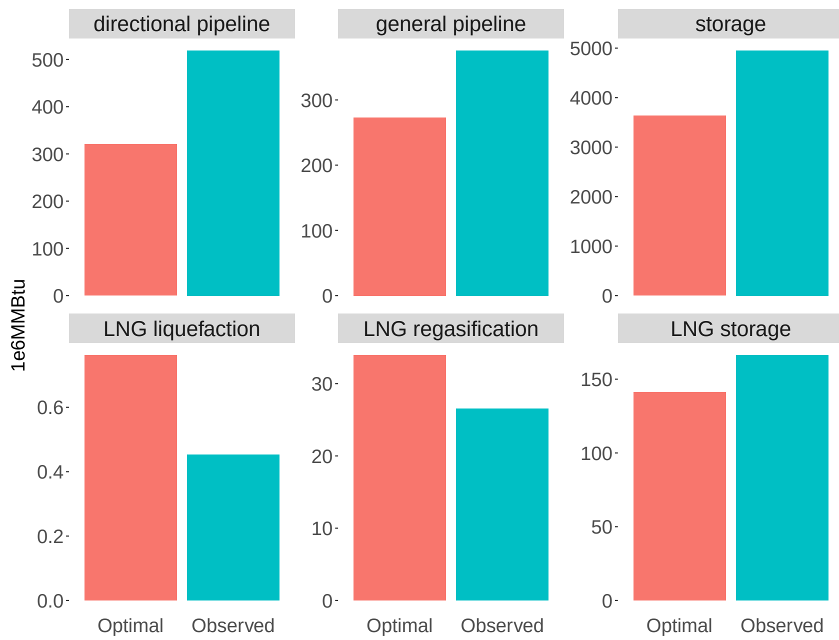
We find that on average, more pipelines and underground storage while less LNG facility were built than needed (Figure 11.) Specifically, the U.S. has overbuilt interstate pipeline connecting between production regions in South Central and Mid Atlantic, but under-invested in pipelines connecting production areas to consumption states such as pipeline from Pennsylvania and West Virginia to New England or from Alabama to Florida (Figure 12.) The lack of pipeline to New England is in part because New York environmental regulators has denied certification for several planned interstate pipeline projects since 2016. Meanwhile, as a large consumption state, Florida has always relied on pipeline transmission from Alabama and Georgia, and could benefit from underground storage. Figure 14 and Figure B.6 show that while the actual natural gas system heavily relies on underground storage in Michigan and Pennsylvania, the optimized least-cost scenario substantially utilizes underground storage in Texas, Louisiana, and Indiana. These states all historically possess large underground storage capacity. Moreover, we find more massive storage net withdrawals in the optimized least-cost scenario compared to observed values.

In addition, Massachusetts and Georgia could have expanded LNG regasification capacity to process LNG imports and serve New England's and Florida's demand. With restricted pipeline transmission, New England states rely on LNG to serve their peak demand in January and February. The model thus suggests to liquefy natural gas at New York and New Jersey[10], then delivers produced LNG to New England states such as Connecticut, Rhode Island, and New Hampshire (Figure B.11).

As a result of investing in too much in pipeline, the utilization rate of pipeline in some areas is quite low. Figure 13 shows observed capacity and utilization rates of pipeline in 2017.

[10]and small liquefaction requirements at Tennessee, Virginia, and Indiana

Figure 11: Optimal and Observed Total Capacity



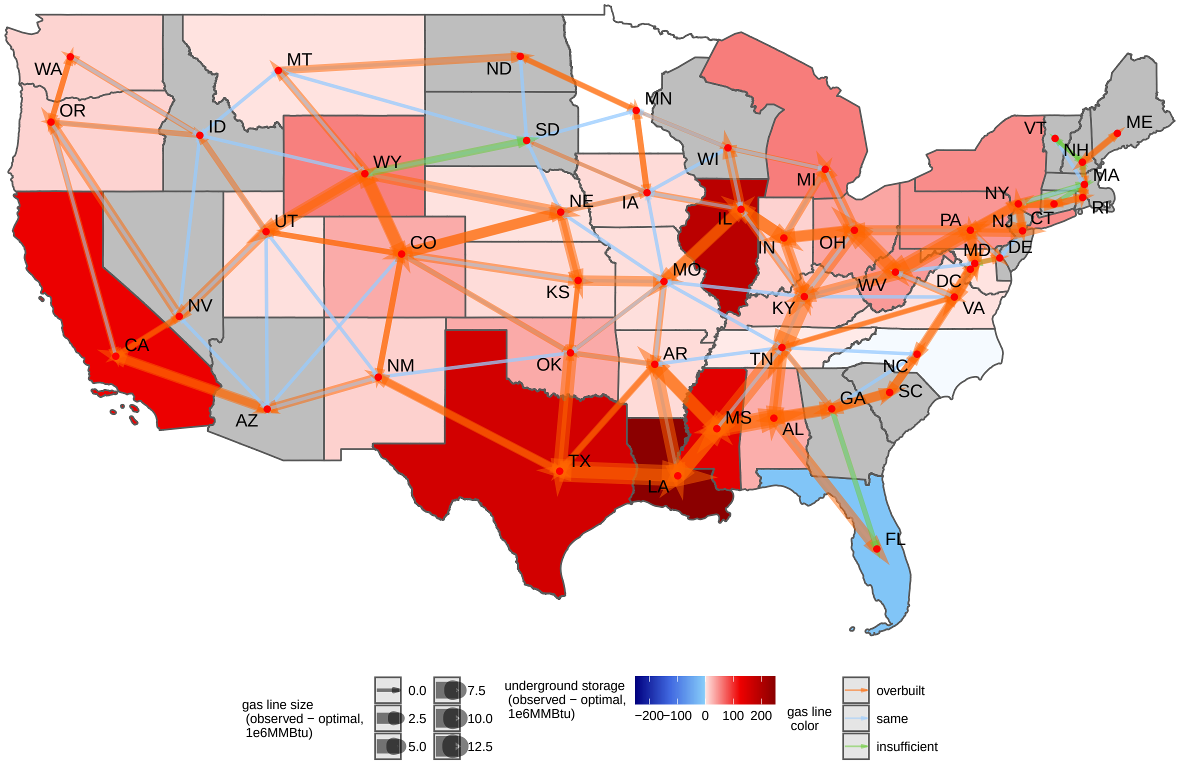
Note: The figure shows capacity of each infrastructure element in the estimated least-cost solution (optimal) juxtaposed with observed capacity. Optimal and observed capacity are represented in orange and blue, respectively. The first two panels show pipeline capacity, in which directional pipeline refers to specific capacity in each direction while general pipeline refers to the maximum capacity of both directions of the same interstate link. Some pipelines are bidirectional while some other pipelines operate (mainly) in just one direction.

The size of pipeline segments indicate pipeline capacity and the colors present pipeline utilization. Pipeline segments with pink color have utilization rate of zero to ten percent. These under-utilized pipelines mainly connect Mid-Atlantic states and East-North-Central states.[11] Contrasting Figure 12 and Figure 13, we find that the most under-utilized pipelines are pipelines that the model indicates are heavily overbuilt.

Actual gas infrastructure outlays from 2002 through 2019 exceeded optimal (estimated least cost) by about *179 billion* 2019-dollars (Figure 15). Given annual consumption of all sectors in the period 2002-2019 averaged 24,700 billion cubic feet, if the additional capital cost from overbuilt infrastructure is distributed along four investment periods (twenty years), natural gas unit cost will increase by about 36 cents per thousand cubic feet. According to EIA's data, an average household consumes about 70.8 thousand cubic feet a year. Thus, the capital cost for overbuilt pipeline and underground storage would translate into an additional 25.5 dollars per year on natural gas bills. Note that this simple calculation does not include any excess return allowed to pipeline companies. If the allowed rate of return exceeds the true cost of capital by two to six percent, the excess total cost grows by an additional $22 to $88.7 billion 2019-dollars,

[11] East North Central division includes five states: Illinois, Indiana, Michigan, Ohio and Wisconsin.

Figure 12: Observed and Optimal Underground Storage and Pipeline Capacities



Note: The map shows the difference between observed capacity and the least-cost optimized capacity in 2017 for both underground storage and interstate pipeline. The grey states do not have underground storage. The white states have no differences between observed and optimal storage capacity. Compared to optimal storage capacity, observed storage capacity is greater in red states and smaller in blue states. The segments connecting each pair of states indicates interstate pipelines, with the segment size indicating the difference between observed capacity and optimal capacity. The arrows show pipeline directions, and colors indicate whether a pipeline is over or under built: overcapacity is in orange and undercapcity is in green; blue indicates no difference between optimal and observed.

depending on the allowed rate of return and length of depreciation (Table 4).[12]

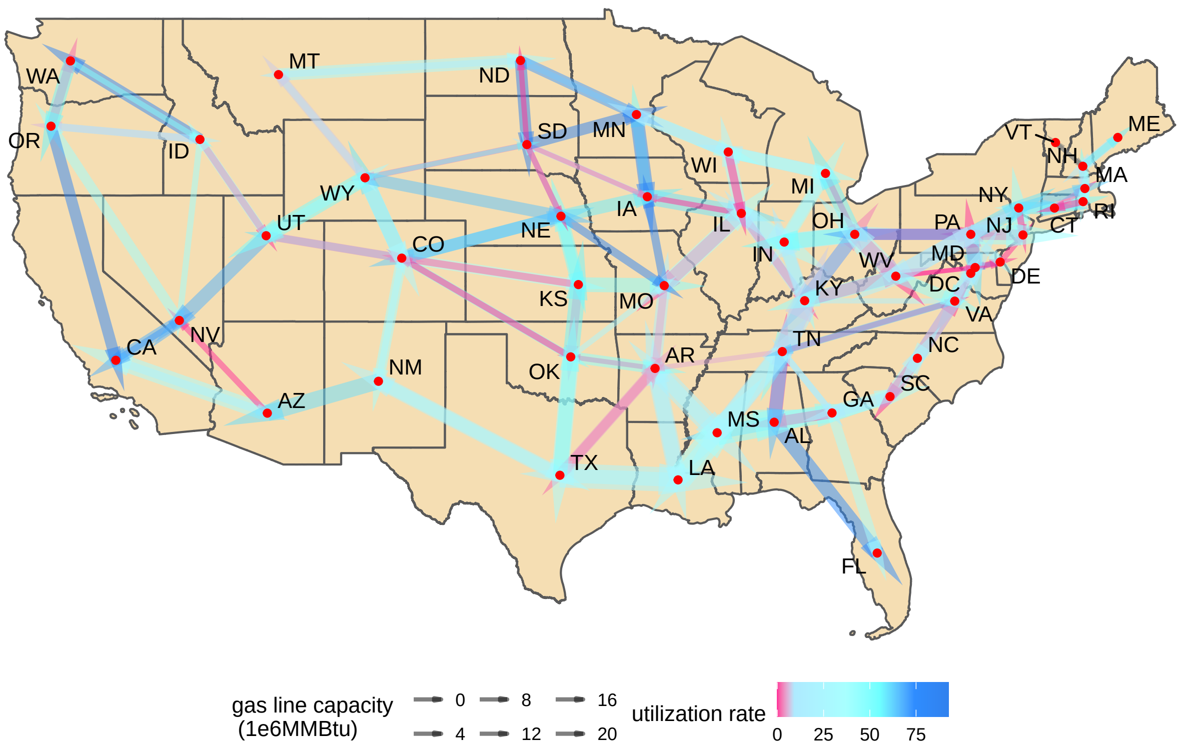
Table 4: Excessive Returns on Overbuilding Pipeline
(billion 2019 dollars)

Depreciable life (years)	Pipeline overbuilding cost	excessive rate 2%	excessive rate 4%	excessive rate 6%
30	130.75	29.12	58.23	87.35
35	130.75	29.57	59.15	88.72

Note: This table presents the total expense paid for excessive rate of returns on *overbuilding pipeline capacity* (i.e. not on total capacity expansion) in 20 years from 2002 to 2021. The excessive return cost of pipeline overcapacity is computed for different levels of excessive rate of return: 2%, 4%, and 6% and in cases the pipeline are depreciated in 30 or 35 years. The present value of excessive return is calculated using discount rate of 3%, base year 2019, and fixed instalment (straight line) depreciation method.

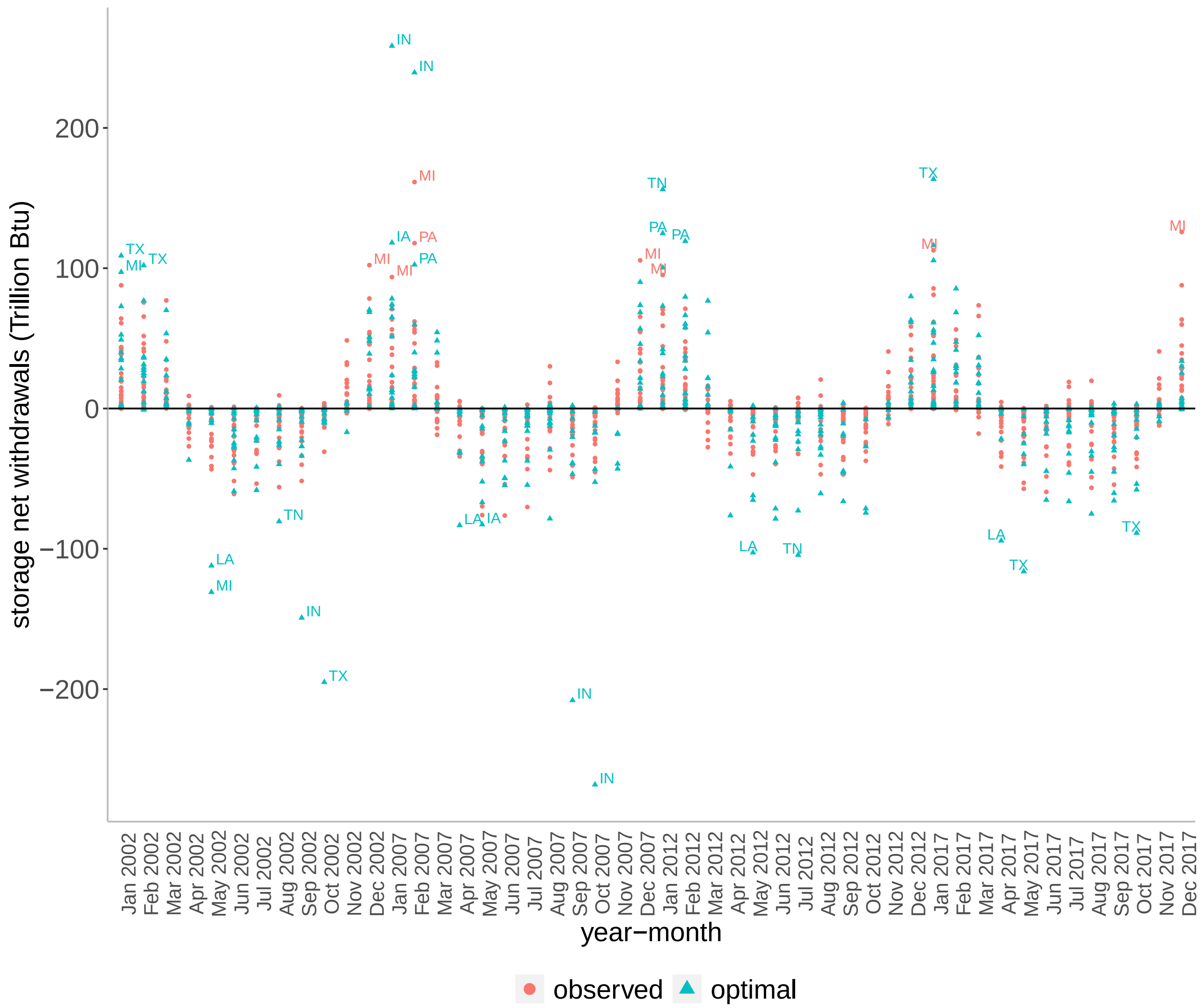
[12]This additional cost only considers possible excess return on excess pipeline investment, and excludes any excess return on investments in storage, LNG, or capital expenditure up to the optimal level.

Figure 13: Observed Pipeline Utilization Rate in 2017



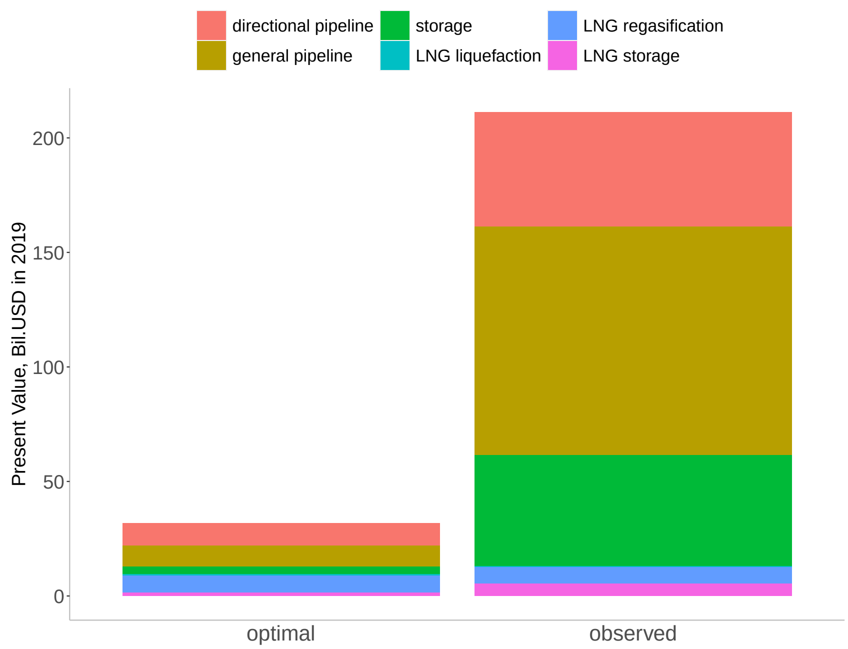
Note: This map presents capacity and utilization rate of pipeline in reality in the period 2017. The size of pipeline segments indicate pipeline capacity, the arrows show pipeline direction, and the colors present pipeline capacity utilization. Pipeline segments with pink color have utilization rate of zero or less than ten percents.

Figure 14: Observed and Optimal Net Withdrawals from Underground Storage



Note: This figure shows average monthly underground storage net withdrawals in each month during 2002, 2007, 2012, and 2017. Each dot represents a state's amount of storage net withdrawals in the least-cost solution or the actual observed value. Negative value indicates storage net injections. The figure shows state labels only where absolute value of storage net withdrawals is greater than or equal to 60 trillion Btu. The orange solid circle points indicate observed storage net withdrawals and the blue triangle points indicate optimal (estimated least cost) values.

Figure 15: Optimal and Observed Capacity Expansion



Note: This figure indicates capital cost of capacity expansion in the estimated least cost solution versus observed in reality. The capital cost displayed here is the sum of present value of fixed cost spent in each period (2002, 2007, 2012, and 2017) calculated using discount rate of 3% and base year 2019.

5.2 Are LNG Facilities a Viable Alternative to Pipeline?

We examine the substitution of underground storage, LNG facilities, and pipeline by considering the scenario "Fixed storage and LNG", in which we keep capacity of storage and LNG facilities at the observed level and optimize pipeline capacity. In the "Demand Buffer" scenario, we examine how much additional infrastructure capacity needed in case peak demand in winter and summer increase by 20%. With the "Fixed Pipelines" scenario, we restrict pipeline capacity to the amount observed and optimize underground and LNG capacity. Transmission flows and storage/ LNG quantities are optimized under all scenarios. Due to limitation in computing capacity, for the comparison across the base case and three additional scenarios, we includes only three periods - 2007, 2012, and 2017. The optimal outcomes are similar in this three-period model and the full model that covers the period of 2002-2050.

Results, summarized in Figure 17 and Figure 18, show that, compared to all other three scenarios (except the "Fixed Pipelines"), the U.S. has built too much pipeline. Interestingly, when restricting the capacity of underground storage and LNG facilities at the observed level, a comparison between "Fixed Storage and LNG" and "Observed" suggests that even if we keep the storage and LNG capacity as it is, we do not need much more pipeline capacity (the top lelf panel in Figure 18). The bottom left panel in Figure 18 on LNG liquefaction capacity provides important implications. The "Fixed Pipeline" scenario results in 50% decrease in LNG liquefaction capacity compared to the base scenario. This suggests that pipeline and LNG facility are substitutes. Moreover, during extreme weather events (hot and cold) that causes local demand to increase by up to 20% (Figure 18), the least-cost method of serving these demand is to increase LNG liquefaction capacity rather than more pipeline capacity.

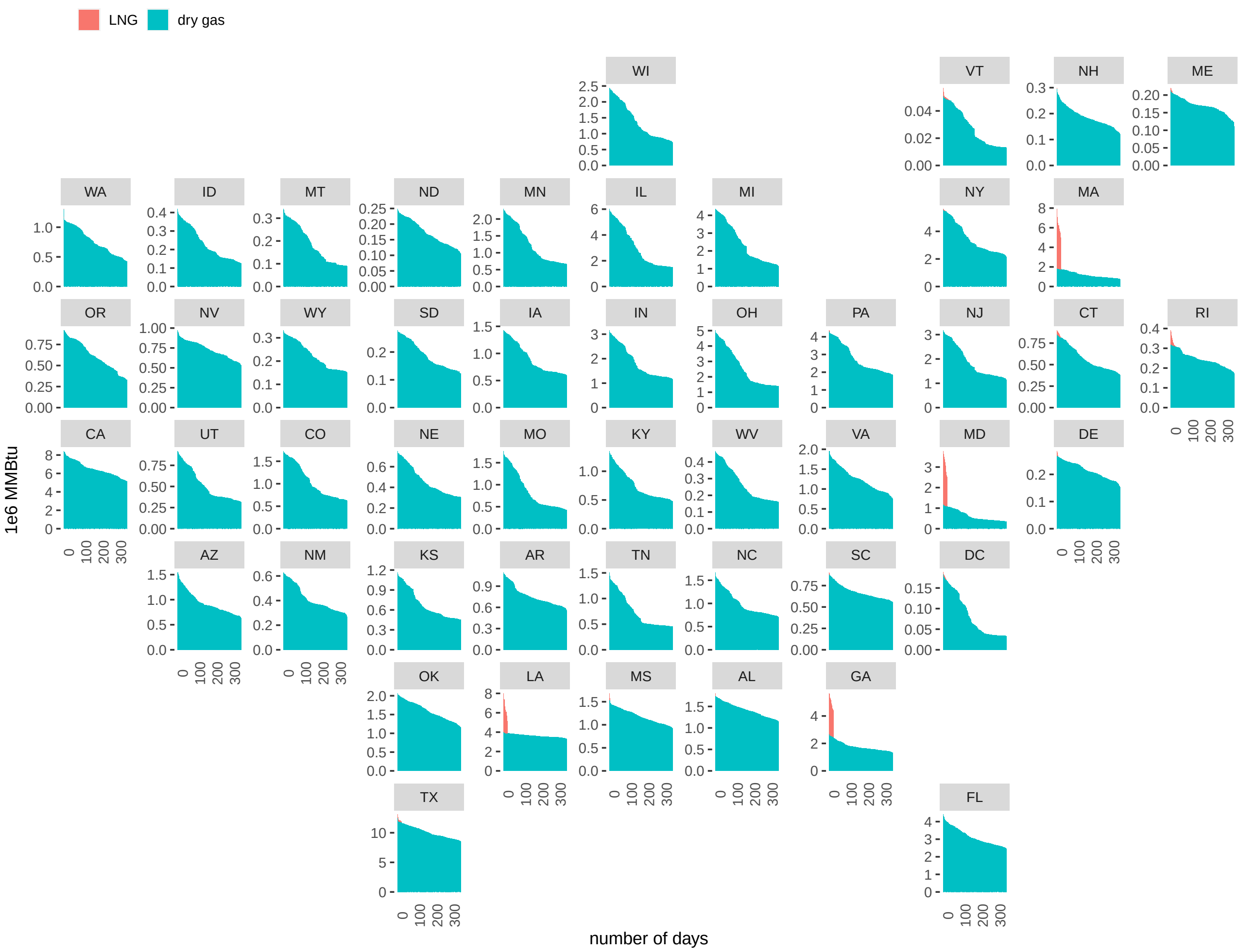
Figure 16 displays the load duration curve for natural gas demand in each state. The load duration curve shows the frequency distribution of demand days, ordered from highest demand to lowest. The figure shows where LNG consumption, colored in orange, helps to satisfy highest-demand days. Dry gas consumption which colored in blue may be from local production, withdrawals from storage, interstate pipeline net receipts, or pipeline imports. This curve represent the load requirements occur during an average year in the period of 2002-2019. In most states, the peak load arises over just few days a year. By closely examining each state, we find Washington, Pennsylvania, Delaware, Mississippi, Alabama, and South Carolina all use a small amount of LNG few days in a year. New England has more substantial use of LNG, about one month, and is clearly visible on the graph. New England states receive a large amount of regasified LNG that is transmitted through pipeline from the import terminals in Massachusetts B.10a.

6 Conclusion

This study uses an optimization model to perform a retrospective analysis of natural gas capacity expansion in the U.S., including interstate pipelines, underground storage, and LNG facilities. Our principal objective is to evaluate whether historical regulatory processes encouraged an efficient level and disposition of investment, and more generally, to develop a tool that may be used to guide investment going forward. Our main results indicate that the U.S. has built substantially too much pipeline and underground storage, amounting to an excess of 179 billion 2019-dollars of excess capital investment. The overbuilding costs an average household roughly 25.5 dollars per year. An overarching concern is that this substantial overinvestment will become more cumbersome in time as natural gas demand declines with electrification of heat and renewable energy displaces natural gas energy sources. Our findings suggest that LNG facilities may be an alternative for long-haul pipelines to serve critical peak demands, and might be more economical means of satisfying any transportation needs associated with near-term growth in gas demand. LNG facilities are also important for New England during its coldest winter month.

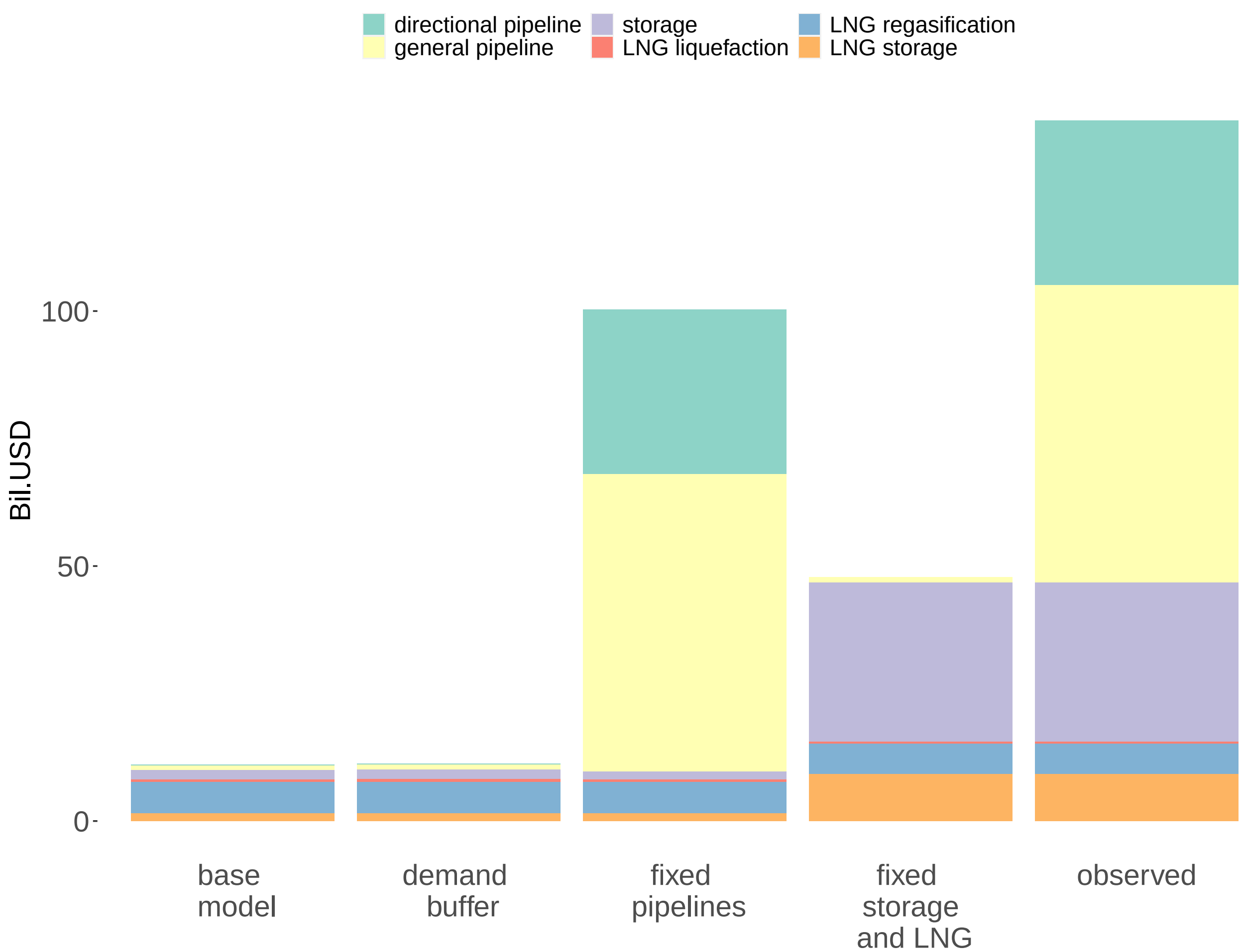
All of the data and code used to develop this model will be made open source in order to ease replication and extensions by others. It may also be a useful tool for regulatory authorities, such as the FERC, the Environmental Protection Agency, and the Department of Energy, as well as

Figure 16: Natural Gas Load Duration Curve



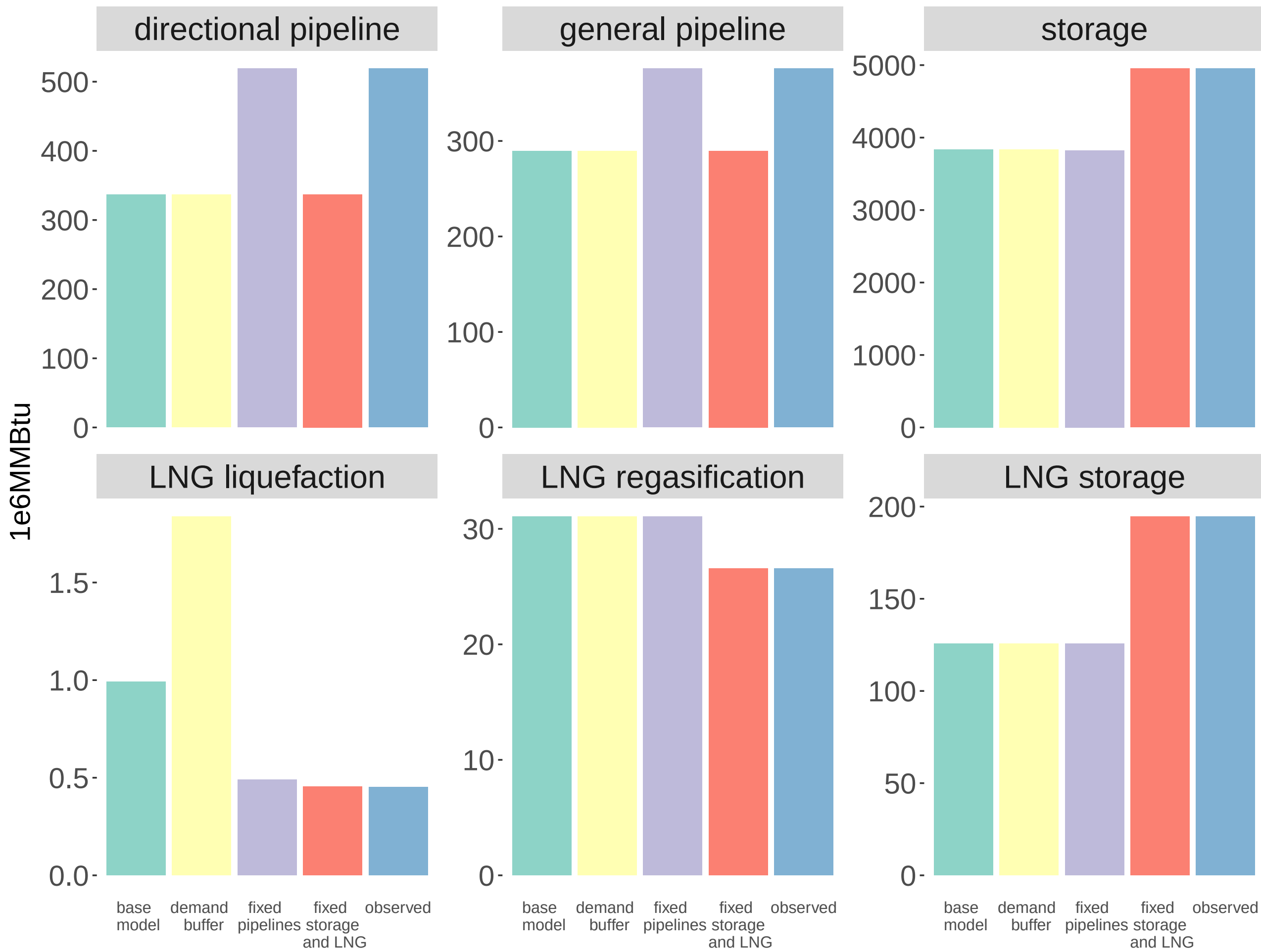
Note: This plot shows natural gas load duration curves for each state over for an average year during the period of 2002-2021 under the estimated least cost. The horizontal axis indicates number of days each load requirement occurs in a year. The blue area represents the amount of natural gas supplied by local production, net pipeline receipts, and underground storage net withdrawals. The orange area shows the amount of natural gas supplied by LNG regasification.

Figure 17: Optimal and Constrained Optimal Cost of Additional Capacity Relative to Observed Capital Cost.



Notes: This graph shows the overall optimal (the overall least-cost base model) and capital costs under alternative scenarios that increase peak demand by 20 percent (demand buffer) or hold some elements of the system (fixed pipelines or fixed storage and LNG), all in relation to observed capital expenditure. Due to limited computing capability, for the comparison in this figure, we include only three periods - 2007, 2012, and 2017, in each scenario. The outcomes from base case are similar between solving the model in the periods 2007-2017 and the full study period, 2002-2050.

Figure 18: Optimal and Constrained Optimal Cost of Additional Capacity Relative to Observed Capital Cost by Sector.



Note: This graph shows the total capacity of each sector of infrastructure at the end of period 2017 in each scenario listed in Table 2: *Demand buffer* scenario is the same as base model except for one peak day in each summer and winter, daily demand increases by 20%; In *Fixed pipeline* scenario, pipeline capacity additions are predetermined while optimizing capacity of underground storage and LNG facilities; *Fixed storage and LNG* scenario optimizes pipeline capacity while taking underground storage and LNG facilities as given.

state Public Utilities Commissions (PUCs). While other factors and models are surely useful and important to consider in pipeline approval processes, the transparency afforded of having an open-source model and data ought to foster more enlightened engagement between stakeholders and regulators. The modular structure of the Switch platform should enable extensions and adjustments to assumptions. For example, a regulatory authority (say FERC or a state PUCs) might evaluate a set of privately proposed projects of pipeline/storage/LNG capacity expansion and identify which combination of projects, rather than a single project, ought be approved. While Kiss et al. (2016) have conducted such evaluation for cross-border natural gas projects in Europe using their model, to our knowledge no such analysis has been conducted for the U.S. natural gas system.

There are a few directions for further consideration. First, while we find excess investment to date, it is possible that prudent investment going forward given past overinvestment might offset some of the excess investment incurred thus far. In particular, while we find underinvestment along some network connections, and modest underinvestment in gasification and liquefication facilities, these may be less necessary in the future given historic overinvestment in other parts of the system. To assess unrecoverable excess investment, it would be useful to consider a forward looking model that takes past investment as given. Second, assessment of investments going forward ought to consider a wider range of alternative futures besides those in EIA's current projections. Under rapid electrification and decarbonization, there is distinct possibility that pipeline and storage infrastructure will be utilized less than EIA currently projects. While we do not explicity consider uncertainty in this perfect-foresight model, option values associated with irreversible investment would likely favor less investment in the near term, not more (Dixit et al., 1994).

In any case, our findings suggest that pipeline companies have substantially over-invested in recent years. If this general result does not have a clear alternative explanation, it casts some doubt upon conventional wisdom following from Baumol and Klevorick (2016) that cost-of-service regulation, even with an excessively high allowed rate of return, imposes minimal social costs. Regulators might use models such as the one we develop here to better assess and regulate investment needs, or consider alternative mechanisms to better incentivize prudent investment.

In future work, we plan to endogenize natural gas demand, supply, and LNG imports, and integrate the model explicitly with a national, Switch-based power system model of North America. We will then use the extended model to project optimal capacity investments of the integrated gas-electricity system, looking particularly at least-cost pathways under alternative future decarbonization scenarios.

References

American Gas Association. (2021). How Does the Natural Gas Delivery System Work? https://www.aga.org/natural-gas/delivery/how-does-the-natural-gas-delivery-system-work-/

American Gas Association, & ICF. (2022). *Net-Zero Emissions Opportunities for Gas Utilities* (tech. rep.). American Gas Association. https://www.aga.org/news/news-releases/new-pageamericas-natural-gas-industry-will-be-essential-to-achieve-a-net-zero-emissions-future/

Averch, H., & Johnson, L. L. . (1962). Behavior of the Firm Under Regulatory Constraint. *The American Economic Review*, *52*(5), 1052–1069.

Baumol, W. J. ., & Klevorick, A. K. . (2016). Input Choices and Rate-of-Return Regulation. *The Bell Journal of Economics and Management Science*, *1*(2), 162–190.

Busch, L. K. (2014). *Review of Natural Gas Models in Support of U.S. Energy Information Administration Natural Gas Transmission and Distribution Module (NGTDM) Redesign Effort* (tech. rep.). leidos. https://www.eia.gov/outlooks/documentation/workshops/pdf/Review%5C%20of%5C%20Natural%5C%20Gas%5C%20Models.pdf

Chong, R., Dediu, D., & Gupta, R. (2019). Setting the bar for global LNG cost competitiveness. *McKinsey and Company*, 1–5. https://www.mckinsey.com/industries/oil-and-gas/our-insights/setting-the-bar-for-global-lng-cost-competitiveness

Clowney, S. L. (2003). *Role of Compressor and Pump Stations* (tech. rep.). https://primis.phmsa.dot.gov/rd/mtgs/121103/Facilities/RD_Forum_12-11-03_Facilities_PRCI_Compressor.pdf

Congressional Research Service. (2011). *Natural gas markets: An overview of 2008* (tech. rep.).

Das, P., Chermakani, D., & Fripp, M. (2016). *Development of SWITCH-Hawaii Model : Loads and Renewable Resources* (tech. rep.). Hawaii Natural Energy Institute.

Davis, L., & Hausman, C. (2022). *Who Will Pay for Legacy Utility Costs?* (Energy Institute WP 317R), Energy Instittute at Haas. https://haas.berkeley.edu/wp-content/uploads/WP317.pdf

Deloitte. (2012). Natural Gas Models. https://fossil.energy.gov/ng_regulation/sites/default/files/programs/gasregulation/authorizations/2012/applications/sierra_exhibits_12_101_lng/Ex._28_-_Deloitte_Natural_Gas_Models.pdf

Diaz, M. N. (2021). *U . S . Energy in the 21 st Century : A Primer* (tech. rep.). Congressional Research Service. https://crsreports.congress.gov/product/pdf/R/R46723

Dixit, R. K., Dixit, A. K., & Pindyck, R. S. (1994). *Investment under uncertainty*. Princeton University Press.

EIA. (2007). *Natural gas compressor stations on the interstate pipeline network: Developments since 1996* (tech. rep.). http://www.eia.gov/pub/oil_gas/natural_gas/analysis_publications/ngcompressor/ngcompressor.pdf

EIA. (2018). *Natural Gas Market Module of the National Energy Modeling System: Model Documentation 2018* (tech. rep.). US Energy Information Administration. https://www.eia.gov/outlooks/aeo/nems/documentation/ngmm/pdf/ngmm(2018).pdf

EIA. (2019a). *Liquefied natural gas imports played a key role in reducing price spikes in New England this winter* (tech. rep.). U.S. Energy Information Administration. https://www.eia.gov/naturalgas/weekly/archivenew_ngwu/2019/04_18/

EIA. (2019b). *More than 60% of electric generating capacity installed in 2018 was fueled by natural gas* (tech. rep.). U.S. Energy Information Administration. https://www.eia.gov/todayinenergy/detail.php?id=38632

EIA. (2021). *Annual Energy Outlook 2021* (tech. rep.). U.S. Energy Information Administration. https://www.eia.gov/outlooks/aeo/

EIA. (2022a). Natural Gas data. https://www.eia.gov/naturalgas/data.php#storage

EIA. (2022b). Natural gas explained. https://www.eia.gov/energyexplained/natural-gas/

Environmental Defense Fund. (2020). On Petitions for Review of Orders of the Federal Energy Refulatory Commission. https://www.edf.org/sites/default/files/documents/EDF_Filed_Brief_20-1016_6.26.20.pdf

Fang, H., Ciatto, A., & Brock, F. (2016). *U.S. Natural Gas Storage Capacity and Utilization Outlook* (tech. rep.). ICF International. https://www.energy.gov/sites/prod/files/2017/01/f34/U.S.%20Natural%20Gas%20Storage%20Capacity%20and%20Utilization%20Outlook_0.pdf

Federal Energy Regulatory Commission. (2004). *Current State of and Issues concerning Underground Natural Gas Storage* (tech. rep.).

FERC. (1999). *Cost-of-Service Rates Manual* (tech. rep.). Federal Energy Regulatory Commission.

FERC. (2005). *High Natural Gas Prices: the Basics* (tech. rep.). Federal Energy Regulatory Commission. https://doi.org/10.1136/bmj.324.7337.621/a

FERC. (2020). Impacts of COVID-19 on the Energy Industry. https://www.ferc.gov/sites/default/files/2020-07/Panel-3_2.pdf

FERC. (2021). Notice of Inquiry: Certification of New Interstate Natural Gas Facilities. https://www.federalregister.gov/documents/2021/02/24/2021-03808/certification-of-new-interstate-natural-gas-facilities#footnote-2-p11269

FERC's Office of Enforcement. (2009). *State of the Markets 2008* (tech. rep.).

Fripp, M. (2012). Switch: A planning tool for power systems with large shares of intermittent renewable energy. *Environmental Science and Technology*, *46*(11), 6371–6378. https://doi.org/10.1021/es204645c

Fripp, M. (2016). Switch Power System Planning Model. https://switch-model.org

Gürsan, C., & de Gooyert, V. (2021). The systemic impact of a transition fuel: Does natural gas help or hinder the energy transition? *Renewable and Sustainable Energy Reviews*, *138*(September 2020). https://doi.org/10.1016/j.rser.2020.110552

Hausman, C., & Kellogg, R. (2015). *Welfare and Distributional Implications of Shale Gas*, Brookings Papers on Economic Activity.

Helm, D., & Thompson, D. (1991). Privatised Transport Infrastructure and Incentives to Invest. *Journal of Transport Economics and Policy*, *25*(3), 231–246. http://www.jstor.org/stable/20034963

ICF Consulting Canada. (2015). *Supply and Demand Market Assessment in Support of AC LNG Export Licence* (tech. rep.). ICF International. https://docs2.cer-rec.gc.ca/ll-eng/llisapi.dll/fetch/2000/90466/94153/552726/2784947/2784950/2784685/Appendix_1_-_ICF_Consulting_Canada_-_Supply_and_Demand_Market_Assessment_-_A4L5I9.pdf?nodeid=2784687&vernum=-2

ICF International, & API. (2017). U.S. Oil and Gas Infrastructure Investment Through 2035, 155. https://www.api.org/$%5Csim$/media/Files/Policy/Infrastructure/API-Infrastructure-Study-2017.pdf

Johnston, J., Henriquez-Auba, R., Maluenda, B., & Fripp, M. (2019). Switch 2.0: A modern platform for planning high-renewable power systems. *SoftwareX*, *10*, 100251. https://doi.org/https://doi.org/10.1016/j.softx.2019.100251

Katulak, F. (2016). Is it time to rethink gas storage and pipelines? https://www.russoonenergy.com/content/it-time-rethink-gas-storage-and-pipelines

Kiss, A., Selei, A., & Tóth, B. T. (2016). A top-down approach to evaluating cross-border natural gas infrastructure projects in Europe. *Energy Journal*, *37*(Special Issue 3), 61–79. https://doi.org/10.5547/01956574.37.SI3.akis

Marks, L., Mason, C. F., Mohlin, K., & Zaragoza-Watkins, M. (2017). *Vertical Market Power in Interconnected Natural Gas and Electricity Markets* (No. 6687). https://doi.org/10.2139/ssrn.3072237

McRae, S. D., & Wolak, F. A. (2019). Market power and incentive-based capacity payment mechanisms.

Mesko, J., & Ramsey, J. (1996). *The Use of Liquefied Natural Gas For Peaking Service* (tech. rep.). INGAA Foundation, Inc. https://www.ingaa.org/File.aspx?id=21698

Mohlin, K. (2021). *The U . S . Gas Pipeline Transportation Market for the Energy Transition: An Introductory Guide with Research Questions for the Energy Transition* (tech. rep.). Enviromental Defense Fund.

NaturalGas.org. (2013). *Storage of natural gas*. Retrieved January 19, 2022, from http://naturalgas.org/naturalgas/storage/

Nelson, J., Johnston, J., Mileva, A., Fripp, M., Hoffman, I., Petros-Good, A., Blanco, C., & Kammen, D. M. (2012). High-resolution modeling of the western North American power system demonstrates low-cost and low-carbon futures. *Energy Policy*, *43*, 436–447. https://doi.org/10.1016/j.enpol.2012.01.031

Northeast Gas Assocition. (2022). *The Role of LNG in the Northeast Natural Gas (and Energy) Market* (tech. rep.). northeastgas.org. https://www.northeastgas.org/about_lng.php

Petrowiki. (2018). Liquified natural gas (LNG). https://petrowiki.spe.org/Liquified_natural_gas_(LNG)

PHMSA - U.S. Department of Transportation. (2022). Liquefied Natural Gas (LNG) Annual Data - 2010 to present. https://www.phmsa.dot.gov/data-and-statistics/pipeline/gas-distribution-gas-gathering-gas-transmission-hazardous-liquids

PRISM Climate Group. (2021). PRISM Climate Data. Retrieved January 3, 2021, from https://prism.oregonstate.edu

Schlenker, W., & Roberts, M. J. (2009). Nonlinear temperature effects indicate severe damages to us crop yields under climate change. *Proceedings of the National Academy of sciences*, *106*(37), 15594–15598.

Socioeconomic Data and Applications Center (SEDAC). (2021). Gridded Population of the World (GPW), v4. https://sedac.ciesin.columbia.edu/data/set/gpw-v4-admin-unit-center-points-population-estimates-rev11/data-download

Stein, J. L., & Borts, G. H. (1972). Behavior of the Firm Under Regulatory Constraint. *The American Economic Review*, *62*(5), 964–970. https://www.jstor.org/stable/181521

Tierney, S. F. (2019). *FERC's Certification of New Interstate Natural Gas Facilities - Revising the 1999 Policy Statement for 21 st Century Conditions* (tech. rep.). Analysis Group. https://www.analysisgroup.com/globalassets/content/insights/publishing/revising_ferc_1999_pipeline_certification.pdf

Ulvestad, M., & Overland, I. (2012). Natural gas and CO2 price variation: Impact on the relative cost-efficiency of LNG and pipelines. *International Journal of Environmental Studies*, *69*(3), 407–426. https://doi.org/10.1080/00207233.2012.677581

U.S. Department of Energy. (2013). *Liquefied Natural Gas: Understanding the Basic Facts* (tech. rep.). https://doi.org/10.2514/3.5665

U.S. Environmental Protection Agency. (2021). *Chapter 8. Natural Gas, "Documentation for EPA's Power Sector Modeling Platform v6 - Summer 2021 Reference Case"* (tech. rep.). https://www.epa.gov/airmarkets/documentation-epas-power-sector-modeling-platform-v6-summer-2021-reference-case

von Hirschhausen, C. (2008). Infrastructure, regulation, investment and security of supply: A case study of the restructured us natural gas market. *Utilities Policy*, *16*(1), 1–10. https://doi.org/https://doi.org/10.1016/j.jup.2007.08.001

Xiarchos, C. H. I. M., Coupal, R., Kelsey, T. W., & Krannich, R. S. (2020). Shale Gas and Oil Development: A Review of the Local Environmental, Fiscal, and Social Impacts.

Economics of Energy and Environmental Policy, *9*(2). https://doi.org/10.5547/2160-5890.9.1.chit

Appendices

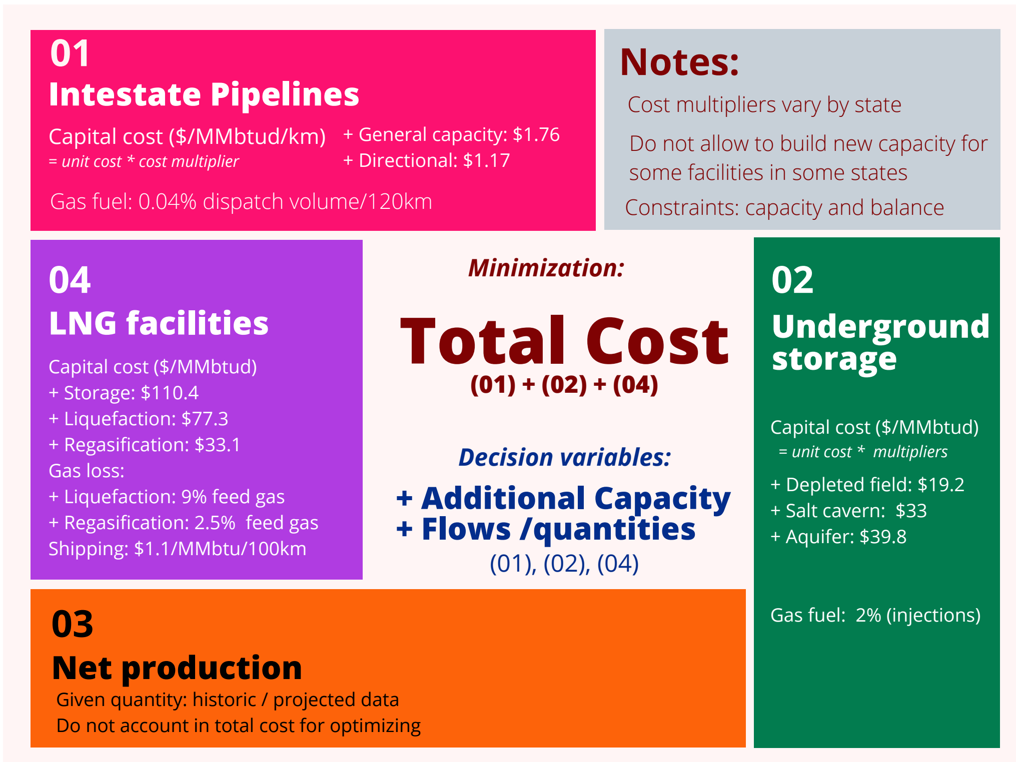
A Tables

Table A.1: State Two-letter Abbreviations

	State Name	State Abbreviation		State Name	State Abbreviation
1	Alabama	AL	25	Mississippi	MS
2	Alaska	AK	26	Missouri	MO
3	Arizona	AZ	27	Montana	MT
4	Arkansas	AR	28	Nebraska	NE
5	California	CA	29	Nevada	NV
6	Colorado	CO	30	New Hampshire	NH
7	Connecticut	CT	31	New Jersey	NJ
8	Delaware	DE	32	New Mexico	NM
9	District of Columbia	DC	33	New York	NY
10	Florida	FL	34	North Carolina	NC
11	Georgia	GA	35	North Dakota	ND
12	Hawaii	HI	36	Ohio	OH
13	Idaho	ID	37	Oklahoma	OK
14	Illinois	IL	38	Oregon	OR
15	Indiana	IN	39	Pennsylvania	PA
16	Iowa	IA	40	Rhode Island	RI
17	Kansas	KS	41	South Carolina	SC
18	Kentucky	KY	42	South Dakota	SD
19	Louisiana	LA	43	Tennessee	TN
20	Maine	ME	44	Texas	TX
21	Maryland	MD	45	Utah	UT
22	Massachusetts	MA	46	Vermont	VT
23	Michigan	MI	47	Virginia	VA
24	Minnesota	MN	48	Washington	WA
			49	West Virginia	WV

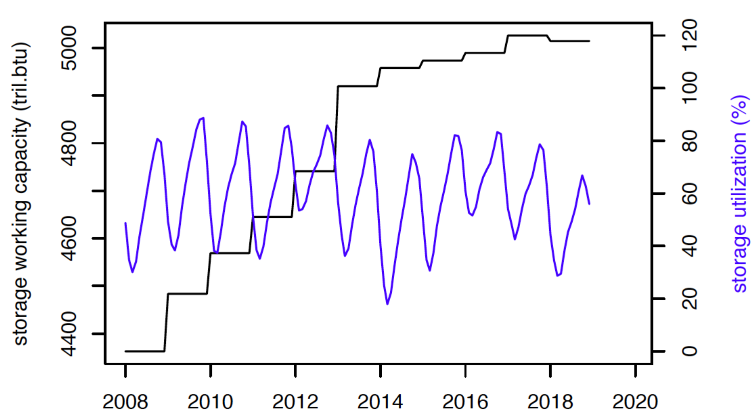
B Figures

Figure B.1: Model components



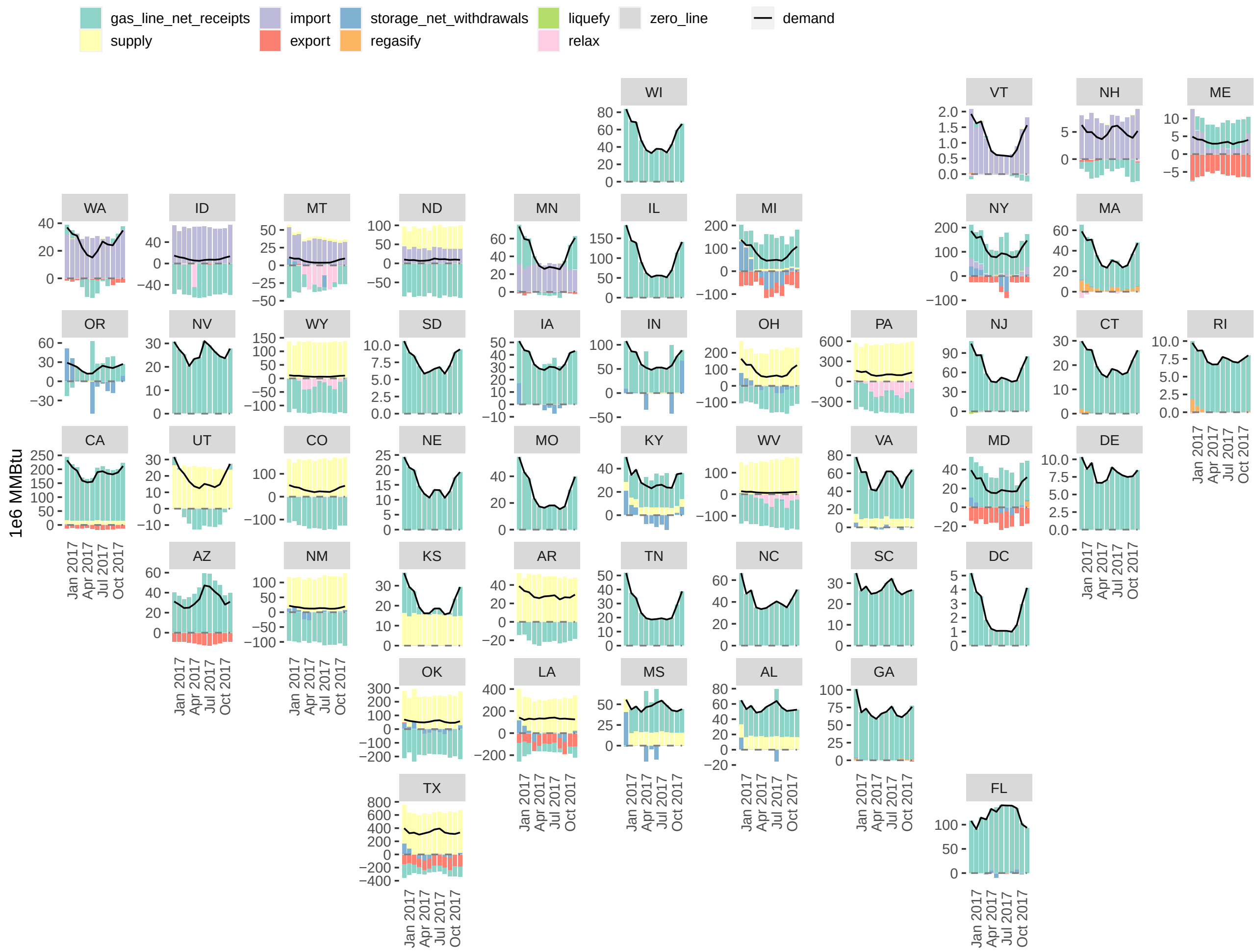
Note: This diagram summarises the objective function, decision variables, and constraints along with unit-cost parameters in the base least-cost model presented in section 3. Gas flows, quantities, and capacity in black texts are exogenous in the model.

Figure B.2: Working Gas Underground Storage Capacity and Utilization Rate



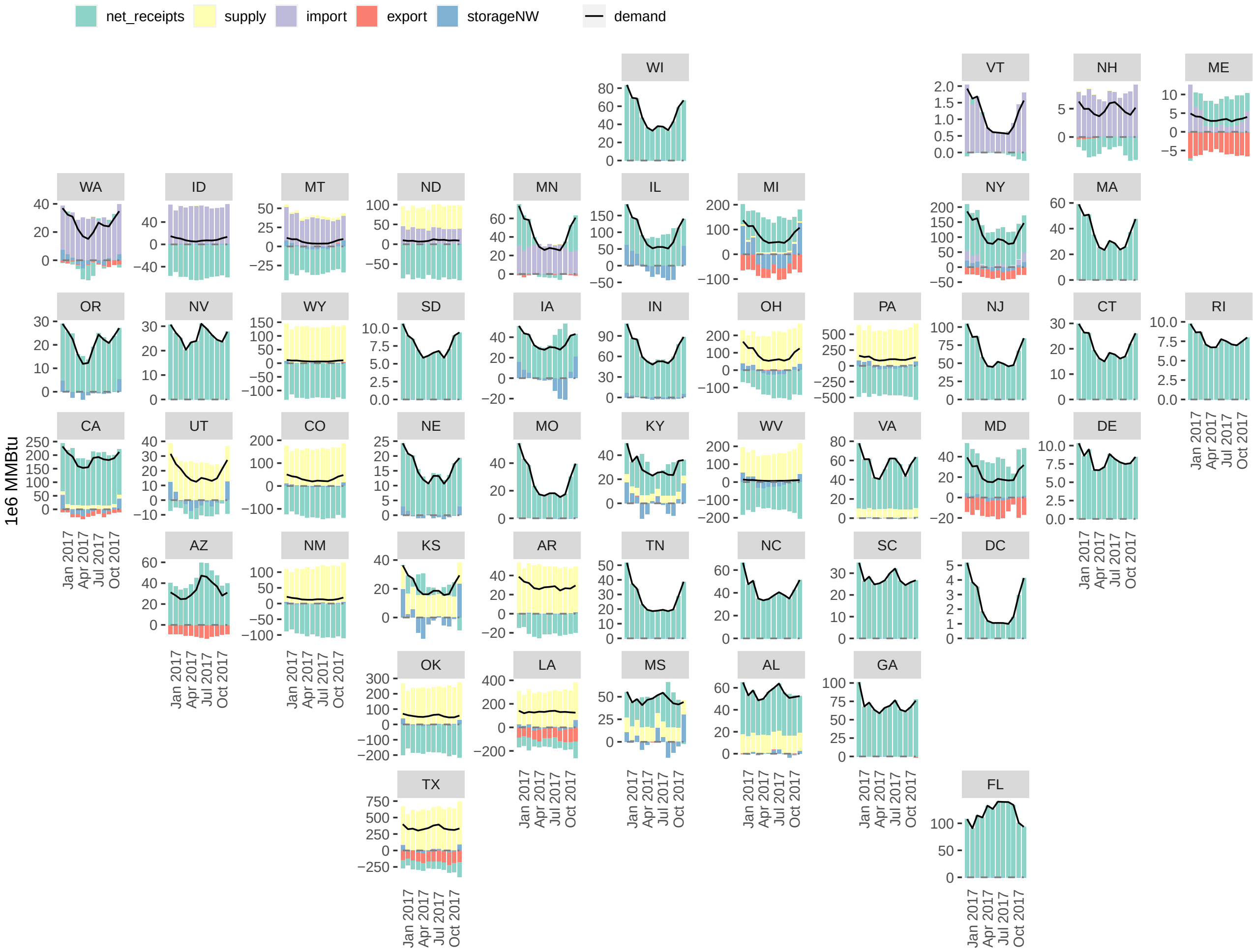
Note: The figure shows total capacity and average utilization rate of underground natural gas storage in the U.S. over time. The black line and the left vertical axis show national underground storage capacity. The blue line and the right vertical axis display storage capacity utilization rate, which is the ratio of working gas to storage capacity.

Figure B.3: *Optimized* Natural Gas Supply - Demand Balance by State - Month, Period 2017



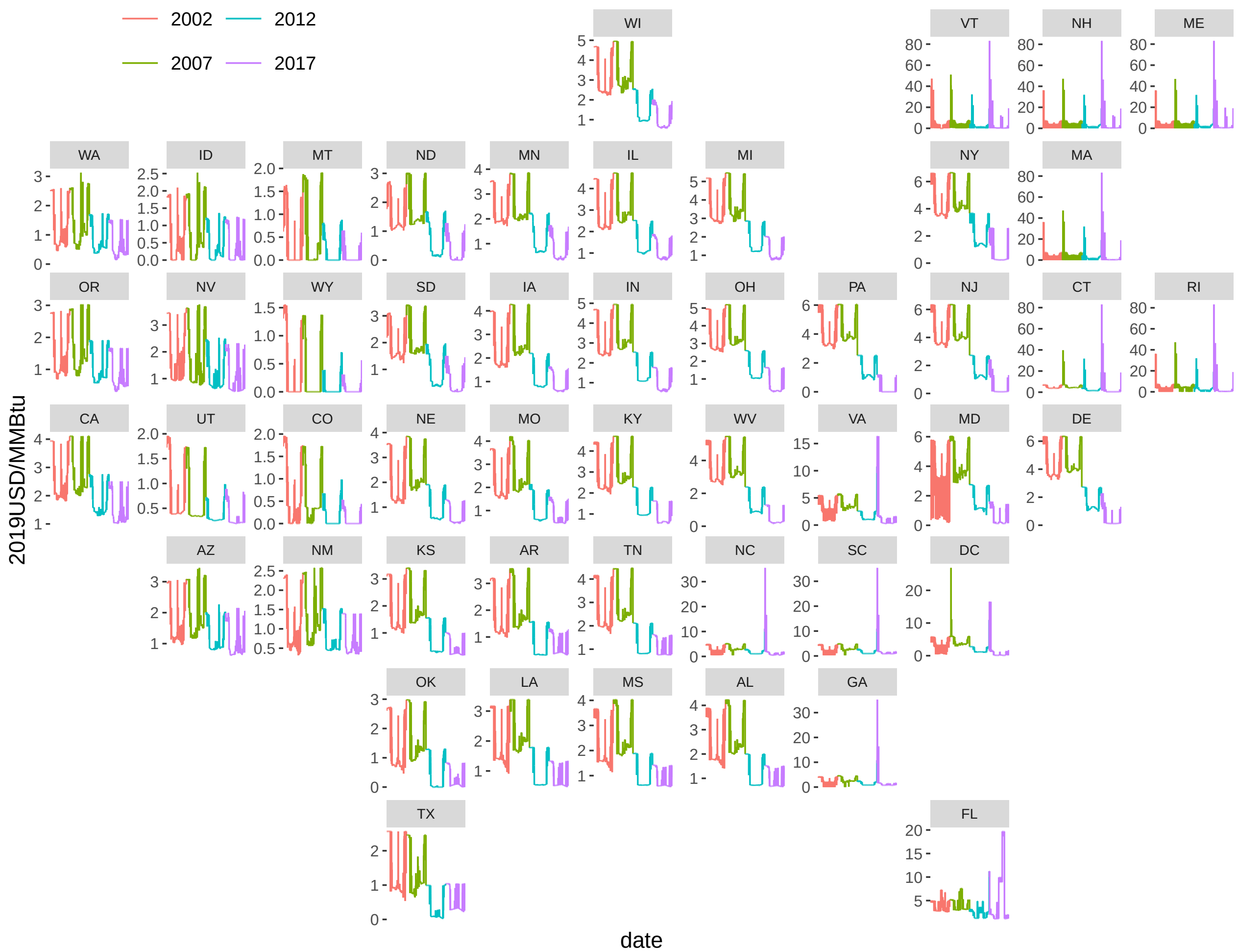
Note: This plot presents how each state supplies their local natural gas demand over months in an average year in the period 2017-2021 under the overall least-cost scenario. Each state may receive natural gas from pipeline net receipts, local supply (i.e. production), net imports, underground storage net withdrawals, and/or gas liquefied from LNG. "Relax" amount, colored in pink, is the excess supply volume which is assumed to be free disposal of.

Figure B.4: *Observed* Natural Gas Supply - Demand Balance by State - Month, Period 2017



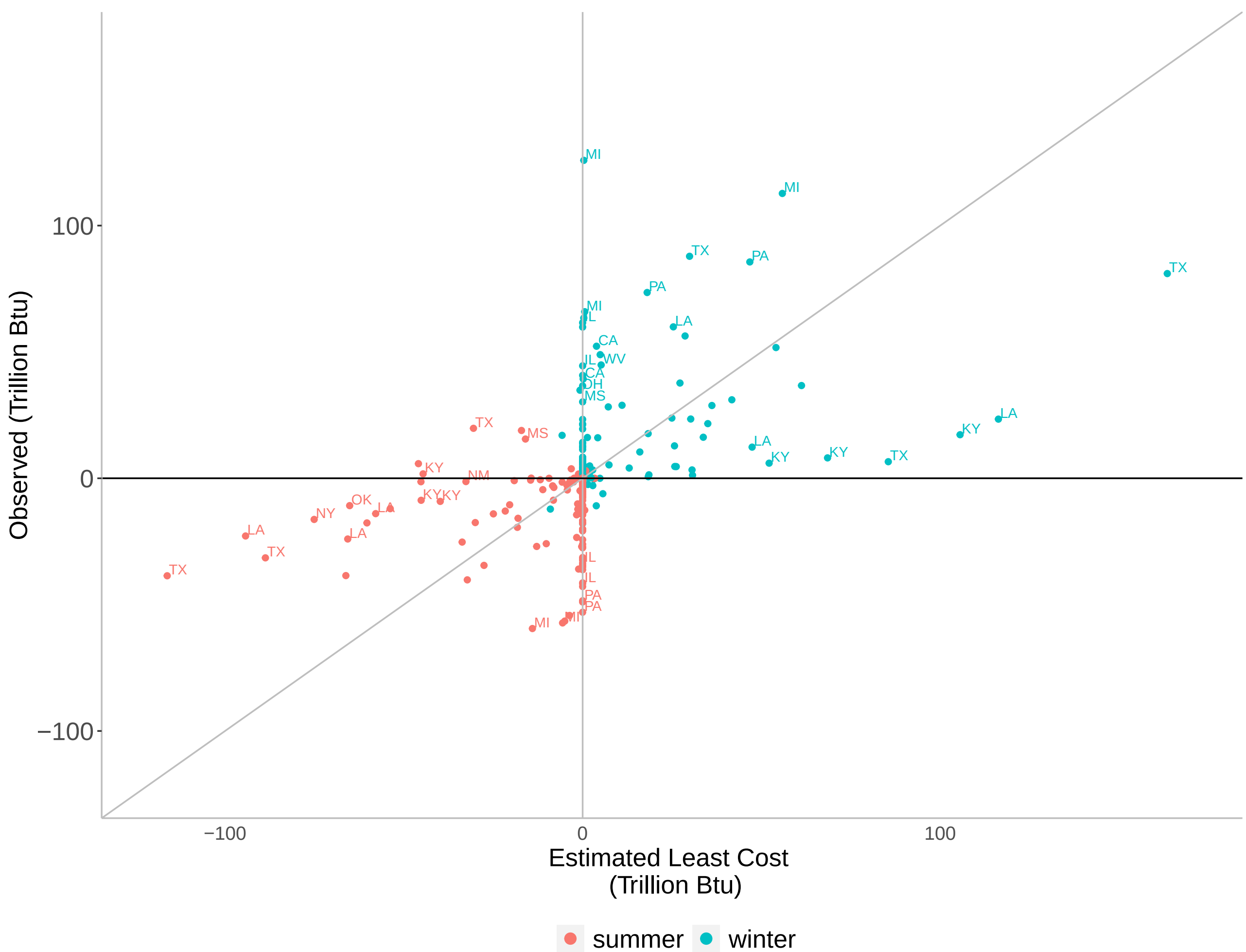
Note: This plot presents how each state actually supplied their local natural gas demand over months in an average year in the period 2017-2019 in reality. Each state may receive natural gas from pipeline net receipts, local supply (i.e. production), net imports, and/or underground storage net withdrawals. Since data for natural gas liquefied from LNG and monthly pipeline net receipts is not available, we assume that net required demand is supplied by pipeline net receipt. Particularly, pipeline-net-receipts = demand - (supply + net imports + storage net withdrawals). The LNG liquefaction volume is small for all states but New England. Thus, this approximation marginally affects, if any, the precision of observed states' supply-demand pattern.

Figure B.5: Shadow Prices by State



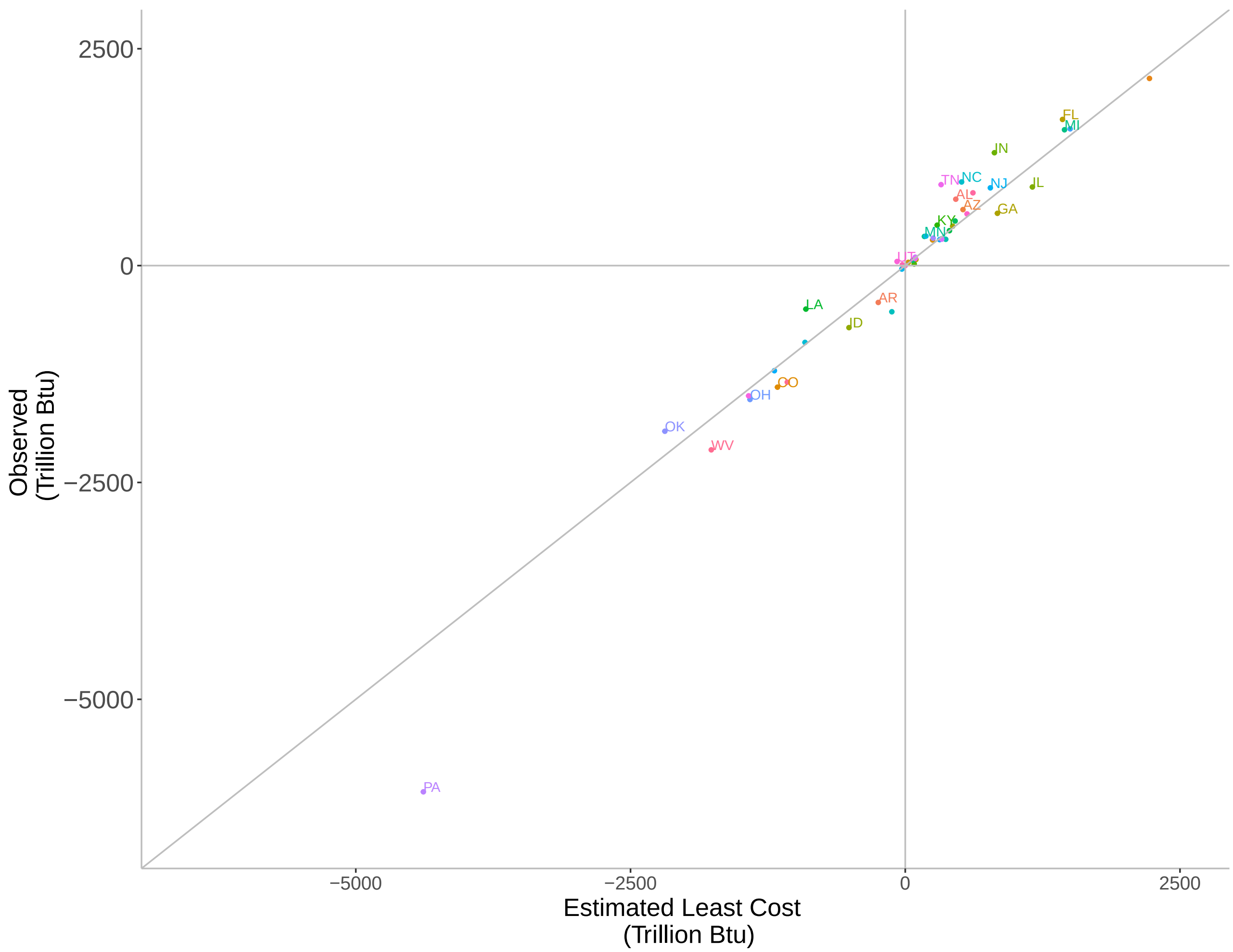
Note: This figure shows by state the variation of shadow prices which is the marginal value (lambda) associated with relaxing the balance constraints in the Equation 15. Note that the scale of vertical axis is different across panels.

Figure B.6: Observed and Optimal Net Withdrawals from Underground Storage in 2017



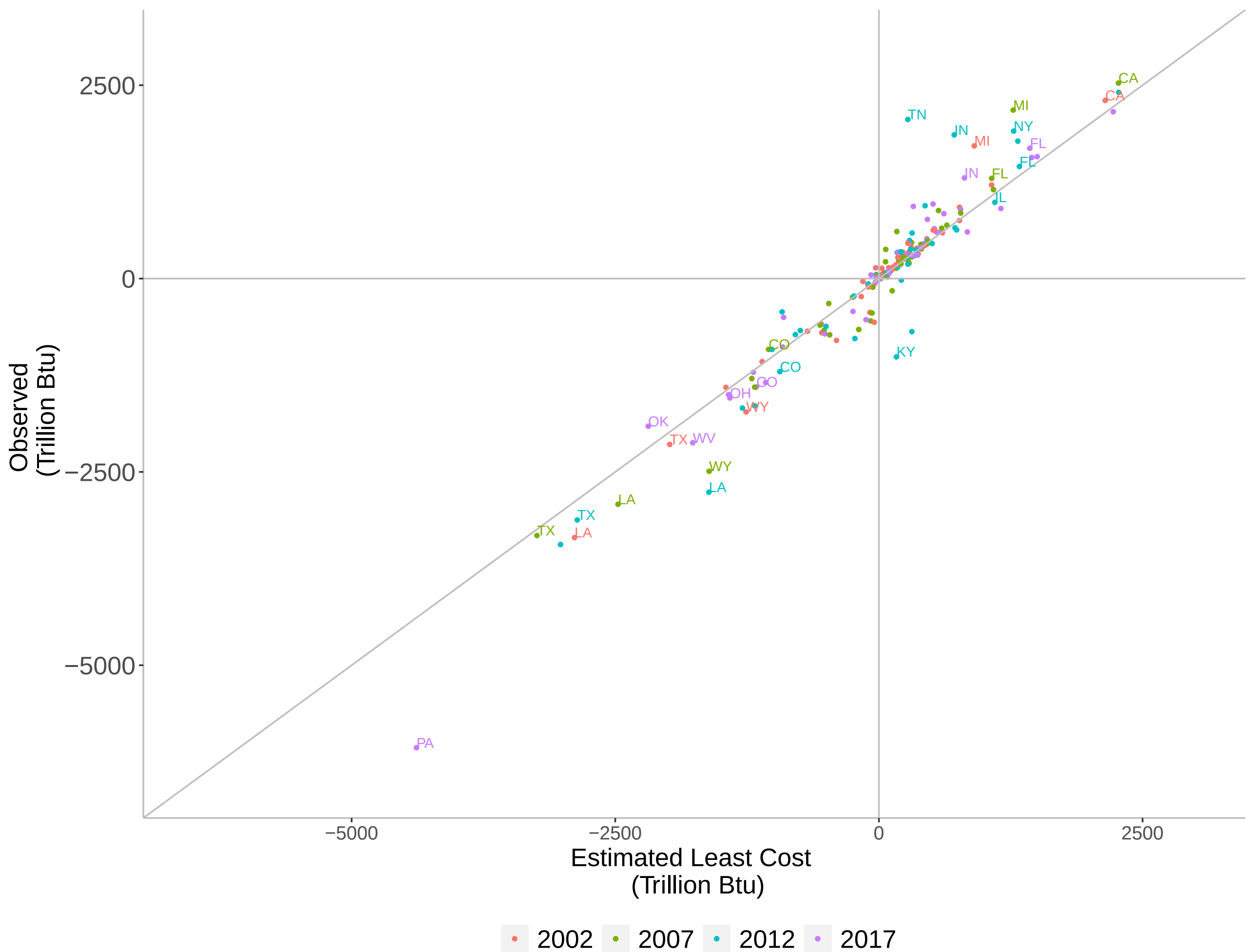
Note: This figure displays average monthly underground storage net withdrawals over the period 2017-2021. Each dot indicates a state's observed storage net withdrawals in a month (the vertical axis) in relation to the least cost solution (the horizontal axis). Negative value indicates storage net injections. The figure only shows state abbreviations when the difference between observed and optimized values is greater than 30 trillion Btu. The orange color represents summer months (April-October) while the blue color shows winter months (November-March).

Figure B.7: Observed and Optimal Annual Pipeline Net Receipts, 2017



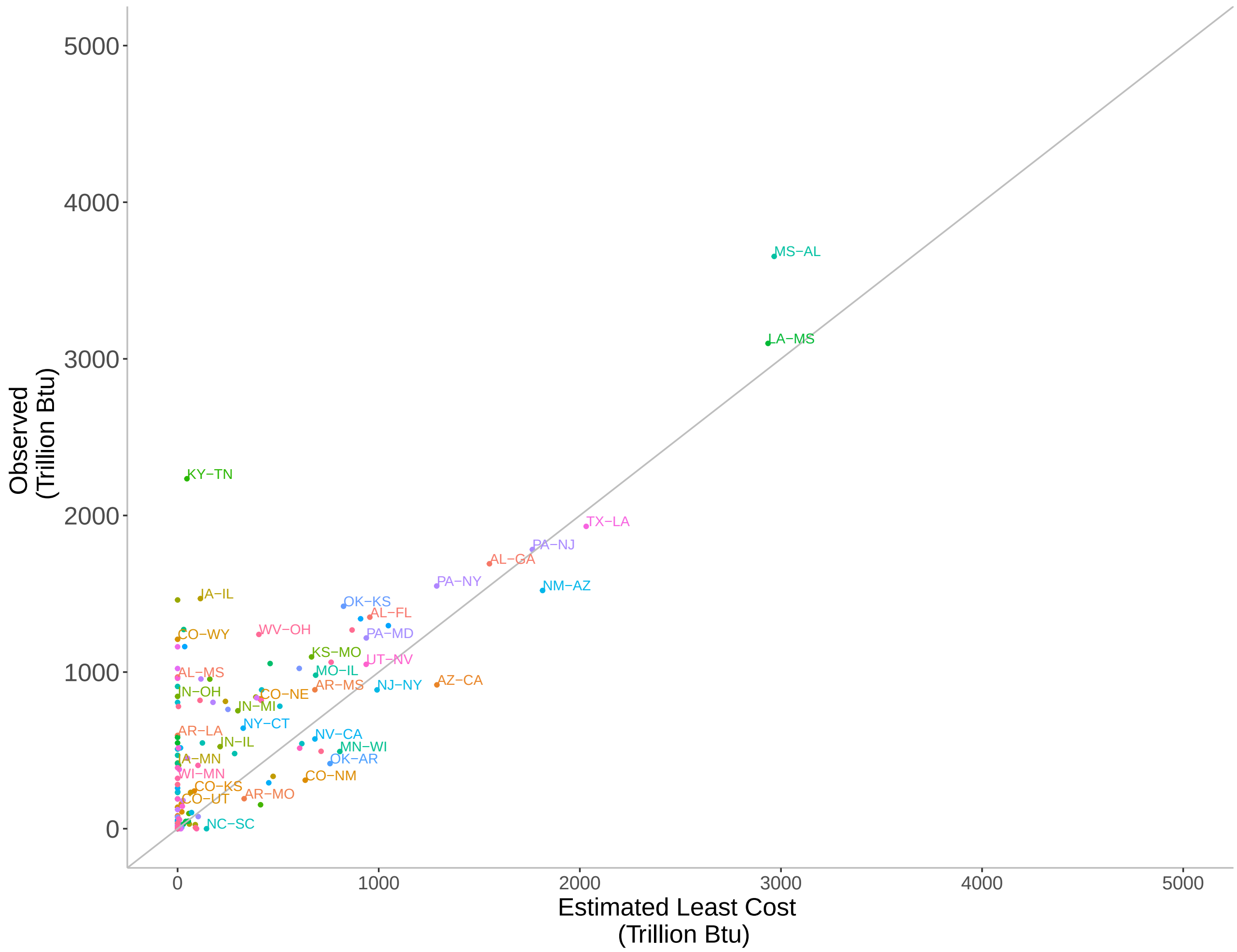
Note: This figure shows average annual pipeline net receipts in the period of 2017-2021. Each dot represents a state's pipeline net receipts in estimated least cost solution versus observed value. Negative value indicates pipeline net deliveries. The figure only shows labels for states for which the difference between observed value and optimized value is greater than 75 trillion Btu, which is the median difference between the two values across all state-periods. Labels are two-letter state abbreviations of corresponding states. Dots and labels are colored by state.

Figure B.8: Observed and Optimal Annual Pipeline Net Receipts by Period



Note: This figure shows average annual pipeline net receipts in each state-period. Each point represents a state's pipeline net receipts in estimated least cost solution versus observed value. Negative value indicates pipeline net deliveries. The figure only shows labels for state that (i) difference between observed value and optimized value is greater than 75 trillion Btu, which is the median difference between the two values across state-periods; and (ii) the absolute value of pipeline net receipts is 1000 trillion Btu or above. Labels are two-letter state abbreviations of corresponding states. Dots and labels are colored by period.

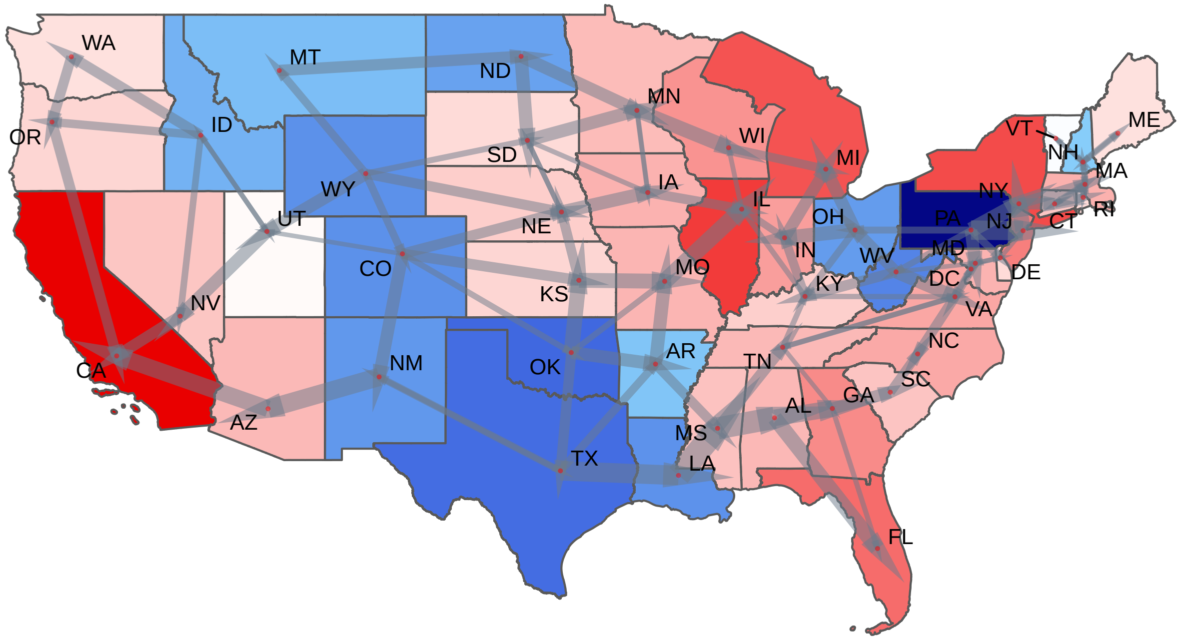
Figure B.9: Observed and Optimal Annual Pipeline Deliveries, 2017



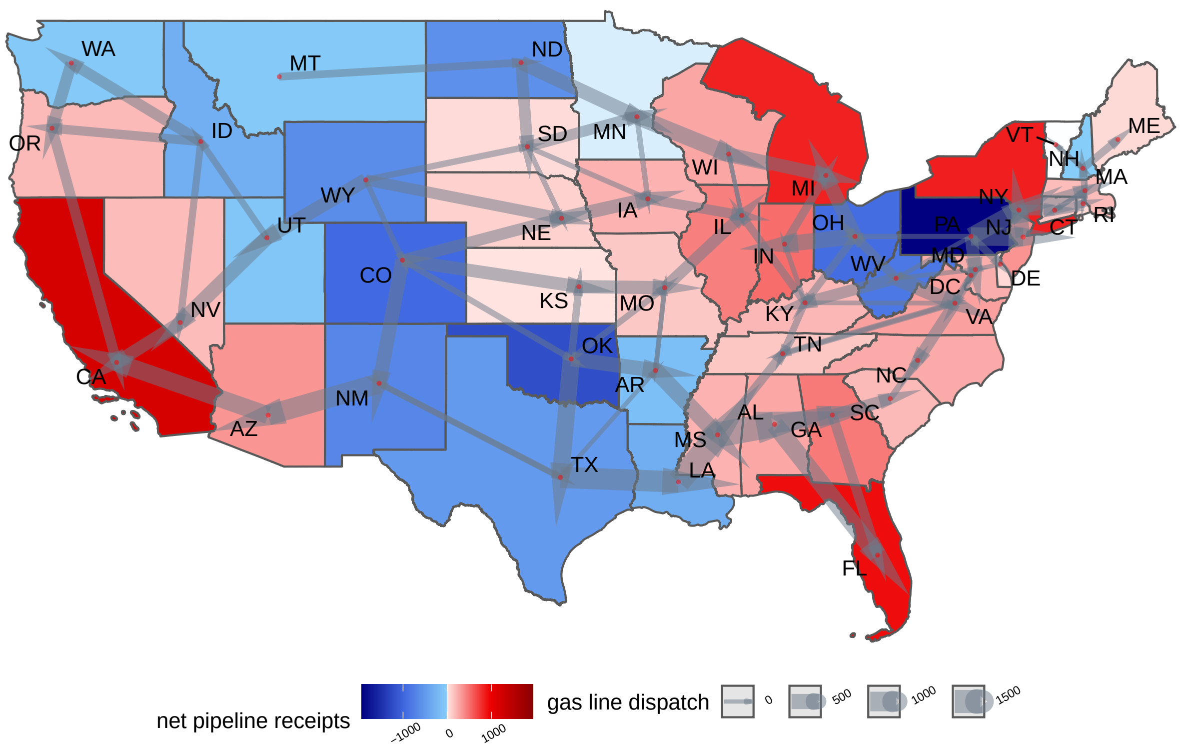
Note: This figure shows average annual pipeline deliveries from a state to another state in the period of 2017-2021. Each dot represents an inter-state pipeline, showing interstate gas flow in estimated least cost solution versus observed value. The figure only shows labels for interstate pipeline that difference between the observed gas flow and optimized flow is greater than 121 trillion Btu, which is the median difference between the two values across pipeline-periods. Labels display pairs of two-letter state abbreviations in which the first state is delivery point and the second state is receipt point. For instance, LA_MS represent the natural gas flow from LA (Louisiana) to MS (Mississippi) and MA_LA represent the gas flow from MS to LA. Dots and labels are colored by individual directional pipelines.

Figure B.10: Optimal Pipeline Net Receipts and Flows

(a) *Winter*

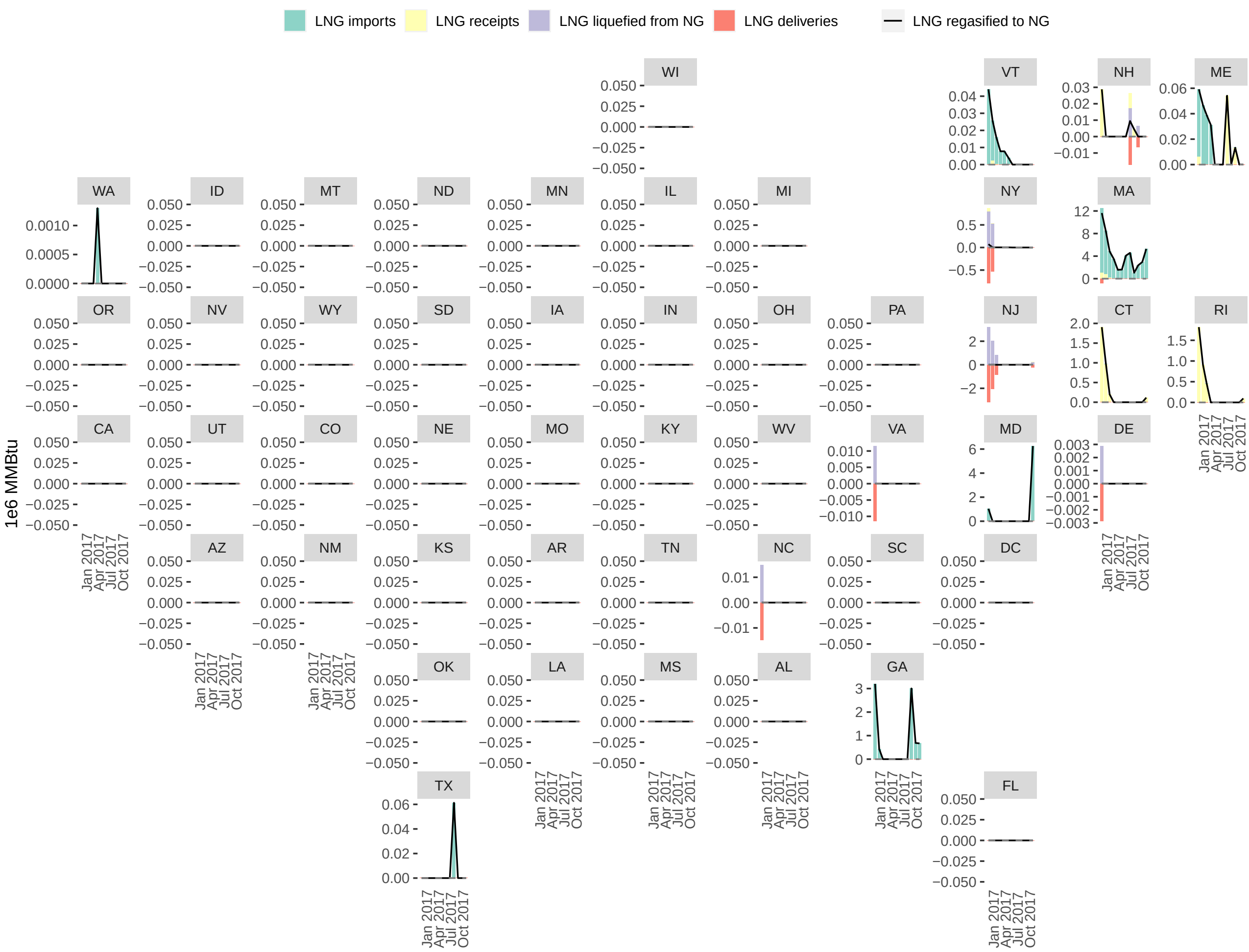


(b) *Summer*



Note: This map shows interstate pipeline net receipts in each state and gas flows from each state to neighboring states under the overall least-cost scenario. Pipeline interstate net receipts equal total volume receipts less total volume deliveries. Each state is colored by relative amount of pipeline net receipts: blue indicates negative net receipts, red indicates positive net receipts, while white means zero net receipts. The segments connecting each pair of states presents pipeline gas flow. The segment size indicates relative amount of gas transported through pipeline. The segment arrows show direction of gas flow.

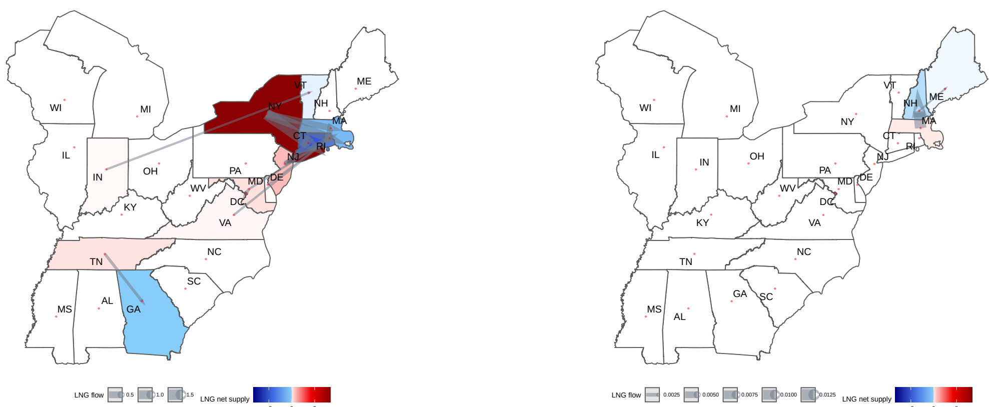
Figure B.11: LNG Supply - Demand Balance, Period 2017-2021



Note: This plot presents how each state supplies their LNG needs for regasification over months in an average year under the overall least-cost scenario. Each state may receive LNG from imports, truck shipping net receipts (i.e. LNG receipts less LNG deliveries), natural gas liquefaction.

Figure B.12: Optimal LNG Net Supply and Flow

(a) *Winter* (b) *Summer*



Note: This map shows interstate LNG net supply in each state and LNG flows from each state to other states under the overall least-cost scenario. This partial U.S. map includes all states with LNG activities. LNG net supply equal total volume imports plus volume liquefied from dry gas less regasified volume. Each state is colored by relative amount of LNG net supply: blue indicates negative net supply, red indicates positive net supply, while white means zero net supply - often is the case that states completely do not use LNG. The segments connecting each pair of states presents LNG shipping routes. The segment size indicates relative amount of LNG transported. The segment arrows show direction of LNG flow.

Figure B.13: Observed and Optimal LNG Facility Capacity

Note: This plot shows by state the observed capacity of LNG facilities, including above-ground storage, liquefaction, and regasification plants, in comparison to the least cost solution. The bar graph is colored by type of LNG facilities.

Figure B.14: Observed and Optimal Underground Storage Capacity

Note: This plot shows by state the observed capacity of underground storage versus the least cost solution. The bar graph is colored by type of underground storage: depleted field, aquifer, and salt cavern.