Incomplete Contracts in Commodities Trade: Evidence from LNG

Swati Dhingra, Ningyuan Jia, Gianmarco Ottaviano, Thomas Sampson, and Catherine Thomas

LSE and Bocconi

CEP and CEPR

PRELIMINARY. PLEASE DO NOT CITE.

September 19, 2024*

Abstract

This paper investigates supply chain resilience to demand shocks in the liquefied natural gas industry, where three quarters of global volumes are governed by long-term contracts between buyers and sellers. We find that contractual relationships adapt to unpredictable weather at destinations and are relatively efficient at mitigating the industry frictions arising from industry technology and geography. Sellers receive around one fifth of contract surplus, which is sufficient to compensate them for forgoing spot markets and for exerting relationship-specific investment to deliver shipments that match buyer demand. A counterfactual analysis that supplies all shipments to spot markets while holding quantities and final destinations constant suggests that adaptation within contracts contributed more than one third of LNG global industry value between 2009 and 2019.

*We thank Enrico Cavallotti and Edoardo Tolva for excellent research assistance. We are very grateful to Kpler for providing the detailed industry data. Heski Bar-Isaac, Jordi Blanes, Davin Chor, Rikard Forslid, Bob Gibbons, Erik Hurst, Rocco Macchiavello, and Saskia Meuchelböck, as well as seminar participants at Aarhus, ERWIT, LSE, Lucca IMT, MIT, and OASIS for very helpful comments. All conclusions are our own.

1 Introduction

For many goods and services in the modern economy, the sets of production and distribution activities in their supply chains are dispersed across both country borders and firm boundaries. Supply chains confront numerous trade frictions and are increasingly subject to disruptions of different linkages. This paper evaluates the efficiency of organizational form choices in a specific global supply chain: liquefied natural gas (LNG). The geography of supply and demand is such that almost all transactions are international. Liquefaction technologies enable large, discrete LNG shipments from producers to consumers and, because buyers are usually energy suppliers, demand is subject to shocks from unpredictable weather.

We focus on a characteristic feature of the LNG industry: while almost all transactions are at arm's length, producers sell to buyers both via bilateral long-term contracts as well as in dispersed spot markets. We show that the buyer-seller relationships governed by contracts increase LNG industry efficiency by adapting to meet varying buyer demand.

Much of the literature on contracts suggests the inability to adapt to a changing environment is a necessary contract cost, as the division of ex post quasi-rents must be clearly specified ex ante to enable relationship-specific investment (Klein, Crawford, and Alchian (1978), Williamson (1979), Masten and Crocker (1985)). However, our results suggest contracts in the LNG industry mitigate this trade off, remaining resilient to demand shocks while facilitating long-term capital investments. Buyers receive highly-valued contract shipments when there are demand shocks, ensuring sufficient shipment-level surplus that can be divided between buyers and sellers so that neither party prefers to deviate to spot markets. We argue that this is made possible by contract terms that are sufficiently incomplete.

LNG is a growing source of global energy, supplying around 10% of world demand.[1] Bilateral contracts govern around three quarters of this supply. On the supply side, sellers of LNG have made large capital investments in liquefaction plants and face very low marginal production costs. As a result, plants tend to produce at close to capacity and total output is unresponsive to spot prices. LNG is transported from sellers to buyers in large, specialized vessels that are often owned (or long-term leased) by one of the two parties. Buyers have also made large capacity investments, in regasification terminals, and tend to operate with some excess capacity. Contracts specify an annual quantity to be transacted between the buyer and seller at a price that is typically indexed to oil prices rather than to destination-specific LNG spot prices. Contracts are incomplete to the extent that they do not specify the arrival times or quantities of individual contracted shipments, which are bilaterally negotiated throughout the year

We focus on how contracts operate from week to week as buyers experience weather shocks. Data on shipment times and weather conditions between 2009 and 2019 in the largest LNG destinations show that the total quantity of contracted LNG arriving at a destination responds to local weather shocks. In weeks

[1] Natural gas provides around 23% of the world's energy, nearly 40% of which is supplied in its liquefied form (IEA, 2021).

when it is unexpectedly cold, relative to seasonal norms, buyers receive more LNG under contract. That is, sellers deliver LNG at times when it is particularly valuable to their contract partners.

Our analysis shows that delivered contract shipments are valued more than the shipments available to buyers on their local spot markets at the same time. Although LNG is itself a commodity, it is transported in large vessels that take time to travel the often long distances between sellers and buyers. The relatively small number of shipments in each spot market may not coincide with buyer demand at any point in time. We find that sellers exert relationship-specific effort within contractual transactions to deliver shipments in volumes and at times that match buyer demand.

The weeks buyers receive contracted shipments coincide with when the seller's outside option—selling on spot markets—is also more valuable because spot prices are positively correlated with weather shocks. Because seller effort is costly, and sellers forgo favorable spot market outside options when they fulfil contract requests, we show that sellers must anticipate extracting around one fifth of the overall estimated contract relationship value via the contract price in order for them to find contracts incentive compatible.

Contracts contribute to total industry gains from trade. A counterfactual industry-level analysis that converts all contract shipments to spot is welfare-reducing.[2] The value of the LNG industry over the time period studied was around 60% higher because of flexibility in contract shipment deliveries.

The shipment-level data used in this paper were made available to us by the company Kpler, a provider of data and analytic solutions for commodity markets. The data include information on the origin and destination of each LNG shipment, the volumes delivered, and whether the shipment was a contract or spot shipment. The data on LNG spot prices was downloaded from EIKON, and destination-specific weather data came from the World Bank Climate Knowledge Portal. We complement these sources with destination-country year data on total storage capacity from annual reports published by the industry body GIIGNL. We also use data on bilateral shipping times and costs between liquefaction plants and regasification terminals from Wood Mackenzie. Finally, origin-country year interest rate data come from the World Development Indicators.

The first empirical analysis is a reduced-form investigation of variation in volumes at the route-week level. Larger quantities flow to destinations when it is unseasonably cold. These findings are evidence that sellers try to send contract shipments to their relationship partners at times when buyers have the highest values. One implication of these results is that contracts are more valuable on routes where weather is less predictable. A second analysis confirms that buyers at destinations with more weather shocks are more

[2]Consistent with this conclusion, contractual relationships proved to be very resilient during the energy crisis in 2022 arising because of the war in Ukraine. Sellers continued to deliver contract shipments despite all-time high spot prices. Two notable exceptions were Eni and Gunvor in early 2022, who did not deliver anticipated contract shipments to Pakistan (S&P, 2022). While the total share of spot volumes had been increasing up to 2020, this trend has recently reversed, "LNG term contract volumes have leapt this year, as energy security becomes paramount worldwide" ((S&P, 2023).

likely to have contracts, controlling for total buyer and seller quantities and bilateral route distances.

We infer from these results that the observed contractual relationships are flexible enough to enable delivered quantities to respond to demand shocks. Consistent with this interpretation is selection of routes into having a contractual relationship when demand is less predictable. The long-term contracts governing relationships appear to operate as call options, where buyers are able to "call in" shipments when most valued. The prevalence of such contracts in the industry suggests sellers find it incentive compatible to maintain the relationships. In the appendix, we estimate lower bounds for the average value of contract shipments received relative to available spot shipments by finding the lowest value at which observed contract shipments are incentive compatible for buyers and sellers, given each party's spot market outside options.[3]

The reduced form findings motivate a structural model of the LNG industry that relates contract shipment timing to weather shocks. This model allows us to estimate shipment values and how value is shared between buyers and sellers, as well as contracting costs.

In the model, we allow buyer value for LNG to be a function of local weather. For contract shipments, buyer value is also a function of relationship-specific seller effort. The amount of effort the seller exerts depends on how much the effort contributes to buyer value, effort costs, and the share of contract surplus the seller extracts. In each period, buyers observe weather conditions and choose whether to request a contract shipment or to buy on the spot market. The spot value depends on an idiosyncratic shock drawn from a distribution with an expected value that is an increasing function of local spot market thickness in that period.[4] The model delivers a set of theoretical moments, derived from incentive compatibility constraints, that give the probability a contract buyer requests a contract rather than a spot shipment in a given week as a function of the weather and model parameters.

Because the model focuses on route-week-level buyer decisions, it does not explicitly take into account seller capacity constraints or any trade-offs the seller has to make between demand from multiple contract partners in any one period. In estimating the model, we select a subset of data in which these modeling assumptions are appropriate. We focus on route-years where a contract is in place, and use only route-week observations where we observe a contract shipment, and weeks with no contract shipment but both the buyer and seller are present in spot markets. This last criterion ensures sellers could have allocated more shipments to contract partners without being capacity constrained. The theoretical moments relate only

[3]The intuition is that the value premium must be sufficient to compensate sellers for not deviating from the contract when spot prices are high and compensate buyers for not deviating from the contract when spot prices are low. The estimated mean of the relative value for the contract shipments studied is $10.95 per cubic metre of LNG, around 6% of the mean spot price in the sample used here.

[4]Spot market thickness is measured as the relative size of the spot market in the region and how well the distribution of vessel sizes in the spot market fits the distribution of local contract demand.

to weeks where the buyer receives some positive quantities, either contract or spot shipments, so we also restrict the estimation sample to route-week observations where this is the case.

We estimate the model using non-linear least squares. Parameter identification is made possible by independent variation in the relationships between contract shipment timing and spot prices, holding weather constant, and between contract shipment timing and weather shocks, holding spot prices constant. Using the model estimates to predict the non-targeted moments of the elasticity of contract shipments with respect to weather shocks and spot prices gives values that are close to those in the data.

The results include distributions of estimates of the willingness to pay for the LNG in contract and spot shipments on each route-week in the sample. A cubic metre via a contract shipment is more than twice as valuable, on average, given the current size of the spot markets. There are three model mechanisms that explain this difference. Buyers tend to request contract shipments during weeks when they experience positive demand shocks; contract shipments are more likely to be the desired quantity; and sellers exert relationship-specific effort to further increase buyer value. The lower bound on the share of contract shipment surplus that sellers must extract via the contract price in order for contractual relationships to be incentive compatible is estimated to be 22%. We find that sellers exert relationship-specific effort to deliver contract shipments that buyers value more highly for around 90% of contract shipments.

We then turn to a counterfactual analysis that asks how much long-term contracts contribute to industry value compared to a world where all LNG is traded on local spot markets. We are able to do this by holding fixed total global supply and sellers' allocations of shipments to destinations in each week. We convert all arriving contract shipments to spot and add them to the spot shipments already arriving at the destination in a given week, giving us a counterfactual spot market in each destination. Because we have estimated the relationship between local spot market thickness and expected spot value, we can simulate the expected values buyers would have obtained in these enlarged counterfactual spot markets.[5]

We show that the efficiency of the LNG industry would have been significantly reduced in the absence of contracts between 2009 and 2019. The gains from contracts outweighed the fixed and variable costs of maintaining them, contributing around 39% of the total gains from trade in the industry over this period.

In the analyses in this paper, we take the existing contract network as fixed. While our reduced form evidence that sellers tend to have contracts with buyers in destinations with more unpredictable weather, consistent with selection into having a contract when it is most valuable, we do not model why some routes have contracts and others do not. We have also not modeled transport costs explicitly.[6] One other

[5]This approach allows us to sidestep the challenge of estimating counterfactual spot prices because we focus on the total surplus created by each shipment as a function of weather conditions, rather than how the unobserved counterfactual spot price divides this surplus. Setting up the counterfactual in this way does not take into account any allocative efficiency gains possible in the spot market if sellers were to divert spot shipments to locations where they anticipate higher prices.

[6]The counterfactual holds the distances traveled by each vessel fixed, so transport costs play no role in the welfare calcu-

important issue our analysis does not consider is seller market power. In theory, it could be that sellers choose to fulfil contract shipment requests in order to withhold quantities from spot markets and gain from the resulting higher spot prices. We do not allow sellers' supply decisions to include these strategic considerations. The main reason for this choice is that the routes between sellers and buyers tend to either be contract or spot. That is, if a seller has a contract with a buyer, it is unlikely that they are also selling on the spot market at that destination, and vice versa.[7]

This paper relates to several literatures. First, are industry-specific studies of the pros and cons of long-term quantity contracts. These include Joskow (1988), and Pirrong (1993), and the discussion in Lafontaine and Slade (2012). Zahur (2022) has studied the LNG industry in depth and has focused on the long-term contracts that we investigate here. He shows that contracts allow greater capital investments at the cost of allocative efficiency but does not examine whether contractual relationships create value through week-to-week adaptation to shocks. Our analysis is motivated by Zahur's observation that contracts are often signed around the time of the capital investments, which, due to the long life of capital, raises the issue we explore here of the efficiency implications of contract flexibility in delivery timing.[8]

In our setting, unpredictable weather causes exogenous demand shocks that we exploit to identify how supply chains respond. We establish that long-term contracts are sufficiently flexible to allow the supply side to adapt and add to overall industry value. This finding complements work showing that vertical integration can permit adaptation to unforeseen circumstances. Forbes and Lederman (2009) and Forbes and Lederman (2010) study vertical integration in the airline industry. They show that major airlines tend to own regional airline partners when adverse weather is more likely and where a route is more integrated into the major's network. Integrated routes also respond to weather shocks and perform better than non-integrated routes at the same airport on the same day. Costinot, Oldenski, and Rauch (2011) show that the boundaries of U.S. multinational firms reflect the extent to which sectors involve non-routine tasks. Vertical integration along the supply chain is more likely in sectors with high problem-solving intensity, consistent with ownership facilitating adaptation to changing circumstances. We find that long-term contracts enable adaptation in LNG supply chains in terms of delivery volumes and timing. We highlight that this finding implies contracts increase the resilience of the industry supply chain to demand shocks.[9]

The literature on relational contracts establishes how the prospect of future rents effectively deters

lations.

[7]While there are several large firms in this industry that have minority ownership stakes in multiple sellers, the majority ownership stake is often a state energy company with no cross-seller links.

[8]Because our counterfactual does not vary each shipment's destination, we do not explore allocative efficiency.

[9]Recent global events, perhaps most notably the Covid-19 pandemic, have prompted much political debate and study of the efficiency implications of potential supply-side disruptions within global value chains (see Grossman, Helpman, and Lhuillier (2023), Grossman, Helpman, and Sabal (2023), and the papers cited therein).

short term opportunism in a range of empirical settings (Macchiavello (2022)). Macchiavello and Morjaria (2015) use the dynamic incentive compatibility constraints implied by repeated seller-buyer interactions to estimate lower bounds on the value of relationships in the Kenyan flower industry, a setting of imperfect contract enforcement. We show that the future value of formal, but incomplete, contracts in the global LNG industry governs the short term actions of the sellers in the industry. Buyers, however, compare contemporaneous costs and benefits, interpreting contracts as call options to request shipments at times when they are particularly valuable.

One recent industry-specific study that examines the externalities that contracts exert on spot markets is Harris and Nguyen (2022).[10] They estimate the incentives that govern both relationship formation and the actions of buyers and sellers in U.S. trucking. Relationships form between trading partners with a high match quality, but the large number of contract relationships leads to thin spot markets, increasing the search costs for trading partners. Tolvanen, Darmouni, and Essig Aberg (2021) also find that frictions in spot markets provide an insurance role for buyers, in their analysis of transactions data from a large producer in the U.S. pulp industry. In our setting, we show that relationship surplus is an endogenous outcome reflecting demand shocks and seller incentives to exert costly relationship-specific effort to increase buyer value. Even though these costs are saved in the counterfactual that eliminates contracts, the increased thickness of the spot market is insufficient to compensate for the lost returns of seller effort.

Our interpretation of the relatively low buyer values for spot market shipments is that spot markets do not supply the quantities of LNG that buyers need at the ideal delivery times, even in the counterfactual thicker spot markets, due to frictions related to bulky shipping technology and the long distances between buyers and sellers. Brancaccio, Kalouptsidi, and Papageorgiou (2020) document how frictions arise in global shipping when country-level trade imbalances in dry bulk commodities lead to shipping price differences across markets. We note that LNG vessels are specialized and there is no two-way trade. In addition, vessels are very large relative to average weekly demand on a route, an indivisibility that amplifies the distance-related time lag between requesting and receiving a shipment. We infer that contracts were efficient in this industry because the total market size was too small, over the time period studied, for spot markets to overcome these geographical and technological frictions.

This paper proceeds as follows: Section 2 describes the industry and summarizes the data. Section 3 is the reduced-form analysis relating contract volumes to weather conditions. Section 4 presents the structural model, Section 5 discusses estimation, and Section 6 presents the results. Section 7 is the counterfactual analysis. Section 8 concludes.

[10] Kranton (1996) sets out theoretically how spot market externalities can be large enough or small enough to determine whether relationships are efficient.

2 Industry and Data

2.1 Industry overview

Natural gas provides around 23% of the world's energy (IEA, 2021) and is supplied either via pipelines (62%) or LNG (38%).[11] In 2019, more than half of all LNG volumes were shipped to regasification plants in Japan, China, or South Korea, and a further 35 countries received positive shipments.[12] LNG is produced in dispersed, oil-rich, locations and liquefied by cooling it to $-160°$ centigrade then transported via ships to locations with regasification capacity. Large, long-lived, location-specific capital investments are required at both origins and destinations, which are typically located far apart from each other. Appendix A shows that the largest exporters in 2019, the last year of data used in this study, were Qatar, Australia, the USA, and Russia. Malaysia, Nigeria, and Indonesia, the three next largest suppliers, jointly supplied less than each of Qatar and Australia. Australia's destinations were concentrated in Japan and China, whereas Qatar's were more spread out.

Long-term sales and purchase agreements (SPAs) are contracts between sellers and buyers that govern the bulk of delivered volumes and are signed before or after capital investments are made (Zahur, 2022). Residual LNG volumes are traded on global spot markets. Contracts specify annual volumes and a pricing formula. Much of the detail in the contracts relates to their "take or pay" nature over a specified time period, a feature shared with pipeline gas contracts (Masten and Crocker, 1985).

In the majority of contracts, prices are indexed to the oil price index at the destination, which means that the seller bears risk related to the forgone opportunity of selling at high LNG spot prices. The buyer typically bears the quantity risk, in that there is an obligation to take the total quantity over the course of the year, whether they need it or not, or incur demurrage charges to compensate the seller. Contracts also contain various clauses related to unforeseen contingencies and the length of acceptable delivery windows.

SPAs do not specify the exact timing of individual contract shipments during the year. While parties typically operate under a three-month forward-looking schedule, shipments are governed by separate Confirmation Notices that describe timing, quantity, and other shipment-specific details. Confirmation Notices are agreed bilaterally prior to each shipment. Therefore, the governing contracts require sellers and buyers to arrange exactly when and with what quantity each shipment is delivered (or collected, when the buyer is responsible for the shipping).[13]

[11] Oil makes up 31%, and Coal 27%.

[12] Recent disruptions to pipeline gas in the EU due to the Russia-Ukraine war have led to many new LNG regasification investments. The import capacity in the EU and the UK is set to increase by one third by 2024, after expanding only modestly in the 10 years to 2022. https://www.eia.gov/todayinenergy/detail.php?id=54780.

[13] Although contractual agreements are proprietary, template SPAs are publicly available. For example, the company BP publishes its master sales and purchase agreement template on its website, https://www.bp.com/content/dam/bp/business-

2.2 Data

We have shipment-level data on 42,047 contract and spot market shipments from January 2009 to December 2019. The data were downloaded by year from the Kpler LNG platform. The shipments arrived at one of 130 destinations, which are regasification terminals spread across 38 countries.[14] 13 countries each imported more than 500 shipments in total over the 11 years in the data and Japan has been by far the largest importer historically, receiving 14,521 shipments in total to 36 terminals during the time period studied. South Korea had the second highest number of shipments, at 5,587, to only six terminals, some of which are among the world's largest. China has steadily increased imports, from around 100 shipments in 2009 to overtaking South Korea in 2019 as the second largest importer, with 864 shipments to 20 terminals. For each shipment, our data contain distance traveled, time taken, vessel capacity and type, as well as volume at origin and volume at destination.[15]

There are 35 origins in the data that we currently use. Of these, 16 are countries where liquefaction capacity is majority-owned by a state-owned company, with integrated oil companies typically owning minority stakes. The countries are: Algeria, Angola, Brunei, Egypt, Equatorial Guinea, Indonesia, Malaysia, Nigeria, Norway, Oman, Papua New Guinea, Peru, Qatar, Trinidad and Tobago, United Arab Emirates, and Yemen. The other 19 origins are separate liquefaction facilities in Australia (10), Russia (3), and the USA (6). We treat these plants as different producers because they have different owners and/or geographies: The Russian plant in Sakhalin is in the far east of the country, very near Japan, and the other two Russian plants, Yamal and Vysotsk, are in the west of the country. The Australian plants are also located on different sides of that country, as well as being owned by separate companies.

Out of the 4,550 possible origin-destination routes in each of the 11 years, there are 28,069 route-year pairs where the plant and terminal are both active for at least some weeks during the year because some facilities came into operation during the time period. The shipment-level data allow us to see the number of shipments traded on a given route in a given week, month, or year, and summing up the shipment volumes gives us the total volumes traded. Of all feasible route-years with active production facilities at both ends, only 5,432, or 19%, have any positive volumes, illustrating how sparse trade flows are in this industry.

In the empirical analysis, we restrict attention to routes ending at one of 10 countries with large LNG

sites/en/global/bp-trading-and-shipping/documents/bp-master-ex-ship-lng-sale-and-purchase-agreement-2019-edition%20(3)%202.pdf. A Confirmation Notice template for a specific shipment is given in the Appendix, Schedule 1, on pages 46 to 49.

[14] We exclude shipments to countries that are primarily large exporters.

[15] The volume arriving at the destination is always lower than the volume that left the origin. Throughout the paper, we measure quantities in cubic metres of LNG. To convert mmBtu into cubic metres: $1mmBtu = 0.048$ cubic metres of LNG. Equivalently, $1m^3 = 21.04mmBtu$. Source: http://www.lngplants.com/conversiontables.html

spot markets, where weekly spot price data is available. These are the regasification terminals in Belgium, China, France, India, Japan, South Korea, Spain, Taiwan, Turkey, and the United Kingdom.[16] These 10 countries receive 35,779, or 85% of all the contract and spot shipments over the time period. For the 5,432 active route-years with any positive flows during the year, 4,415 end in a destination in one of the 10 countries with a spot price.

Because the Kpler data indicate whether any given shipment is a contract or spot transaction, we can also compute the total contract and spot shipments and volumes on a route in a week, month, or year. Of the 4,415 active route-years ending in one of the countries with a spot price, 3,310 have some contracted volumes during that year, and 2,214 have some spot volumes.

We complement the shipment level data with data from several other data sources. We use location-specific LNG spot prices from EIKON, as shown in Figure 1. These are weekly data for the 10 different spot markets we study and we convert the mmbtu prices given there to price per cubic meter, which is the unit of quantity in the Kpler shipment-level data.

Country-specific temperatures are from the World Bank climate knowledge portal.[17] We select the mean monthly temperature in each country for each year.[18] We are interested in how well contracts allow parties to adapt to unforeseen circumstances, asking, in our case, how resilient the supply chain is to unexpected weather shocks. We follow the climatology literature (see, for example, Nicholson (1993)) and compute weather anomalies as the deviations from the destination's long-term mean, divided by its long-run standard deviation. To adjust for seasonality, we find this measure of the shock for each destination country-month-year. That is, we define the temperature shock as the mean temperature in a country that month minus the mean mean temperature across years in that country-month, divided by its standard deviation. We use the negative of this standardized shock, so that a positive value of the shock implies it is colder than expected in destination that month given the mean monthly temperature and typical variation in that month. Figure 2 shows the distribution of destination-month-year weather shocks.

We use data from Wood Mackenzie, a large energy sector data provider, on the journey time and shipping cost on each route. These data are more appropriate than distance measures to capture the trade frictions associated with a route between a buyer and a seller because they account for the indirect paths taken by sea, and the delay or cost associated with congested shipping routes, including any Suez and Panama canal crossings. We have access to the 2020 and 2021 cost matrices from Wood Mackenzie's LNG

[16] There is also a spot price series for the US, but we have excluded it as a destination as it is predominantly an exporter of LNG. We also exclude the LNG spot markets in Argentina, Brazil, and Mexico. This is because they are relatively small markets and also because demand appears not to vary with weather. Also note that we assume buyers in Taiwan can purchase on the China spot market due to its proximity.

[17] https://climateknowledgeportal.worldbank.org/download-data

[18] All the results are robust to using variation in the minimum monthly temperature in each year to measure weather shocks.

platform, and use the 2020 matrix as it is closest to the end of our shipment sample in 2019.

When estimating the structural model in Section 4, we also use data on the LNG storage capacity in each destination country in each year, downloaded from the GIIGNL annual reports found at the GIIGNL website[19]. We allow storage capacity to affect the responsiveness of demand to weather shocks.

2.3 Descriptive statistics

Table 1 summarizes the data used. The LNG quantity measure is number of shipments and volumes at the route-week level on active route-years, defined as route-years with some positive volumes. There are 192,911 observations at the week level on active route-years where we have a local spot market price for the destination country that week. Panel (a) of Table 1 shows the mean, standard deviation, and minimum and maximum values over all of these route-week observations, including the majority that have zero volumes. The mean number of shipments is 0.15 and the mean volumes are 20.95 thousand cubic meters. The largest number of shipments on any route-week is five. The average number of contract shipments and volumes is greater than the number of spot shipments and volumes.

Panel (b) of Table 1 summarizes the subset of route-week observations on active routes that have positive values of the relevant variable that week. The mean number of shipments conditional on there being any shipment is 1.15. This reflects the fact that on all active route-years, 98% of weekly observations with a local spot market price have zero or one shipment in that week. The mean vessel volume is 158.71 cubic meters. Contract shipments are, on average, both more frequent and larger than spot shipments.

Panel (c) of Table 1 summarizes destination-week-specific prices and temperatures. The mean spot price of LNG across all observations is 187.34 USD per cubic meter. The large standard deviation in spot prices of 84.80 can be seen also in Figure 1, where we see variation both across destinations and over time. Across all destination countries, the lowest mean monthly temperatures are in China and the highest are in India. The destination country with the smallest standard deviation in temperature for a given month across years is India and the country with the highest is Belgium. Panel (c) also summarizes route-week journey days and journey costs per cubic meter from the Wood Mackenzie data.[20] The mean number of days taken across all observations is 10.24 and the mean cost is 16.86 USD per cubic meter, which is around 9% of the spot price.

Panel (d) of Table 1 summarizes destination-week specific prices and temperatures for weeks with some

[19] GIIGNL is the industry association the International Group of Liquefied Natural Gas Importers. The annual reports are available here: https://giignl.org/resources2/.

[20] We were unable to find several origins and destinations in the Wood Mackenzie data, and there are some for which close substitutes were not easily available, these are the Yamal liquefaction plant in Northwest Siberia, Russia, and regasification terminals in Sweden, Finland, and in Okinawa, Japan, the last of which is in the East China Sea, far from other Japanese terminals.

positive volumes. The mean spot price of LNG across all route-week observations is 195.90 USD per cubic meter. It also shows journey times and costs. There are five routes in the data with positive flows where shipments take an average of less than a day, four of which are between Algeria and destinations in Spain, and one of which is between Sakhalin 2 in Russia and a destination in Northern Japan. The routes that take the most journey days are between Norway and either China, South Korea, or Japan.

The 189,982 weeks on active route-years where we have both spot price data and journey cost information (Table 1, panel (c)) come from 4,355 active routes with positive volumes. We look at whether traditional gravity-style variables predict which of the 19,556 potentially active routes have some positive flows during a year. Table 2 shows the results of three regressions of the indicator for a route-year being active on either journey days or transport cost per cubic meter. In columns 2 and 4, we include the total volume produced or received at the origin and destination in that year, and in columns 3 and 6, we include origin-year and destination-year fixed effects.

Column 1 shows that an increase in journey length of one day decreases the probability of the route-year being active by two percentage points, which is 10% of the mean probability. Column 2 includes the total volume of the origin and destination as controls, and column 3 includes origin and destination fixed effects. The coefficient on journey time remains very similar. Column 2 also shows that routes starting from origins or arriving at destinations with larger total quantities in a year are more likely to be active.

Columns 4 to 6 show the impact on the probability a route-year is active is lower when transport costs are higher. A one dollar increase in the per cubic meter transport cost decreases the probability of having any positive volumes on the route-year by one percentage point, which is 5% of the mean probability a route-year is active. Overall, Table 2 shows that active routes tend to be from larger origins to larger destinations that are close by and have low transport costs, consistent with gravity.

For the set of route-years with positive flows, the volumes tend to be either all contract or spot volumes. Figure 3 shows the bimodal share of volumes on a route-year that is under contract. Based on this figure, we construct an indicator variable for whether an active route-year is mostly under contract, defined as having at least 50% of total flows in the year being contract flows and investigate selection into being a contract route. 70% of the 4,415 active route years in the sample with destination spot price data are defined as contract routes under this definition.

3 Reduced form analysis

The data in Table 1 show there is large variation in the volumes across routes. We now turn to variation in volumes over time on any one route. We first ask whether this variation is related to fluctuations in destination-specific temperature level, as we expect demand to be larger during colder weeks. We consider only active routes, defined as those where plants at both the origin and destination are operational and

there is at least one shipment on that origin-destination route during the calendar year.

Defining Q_{lgt} as the volume of LNG on the route between liquefaction plant l and regasification terminal g in week t, in $'000$ cubic meters, and $temp_{gt}$ as the mean temperature during that month, we estimate versions of the following equation that includes various definitions of volumes.

$$Q_{lgt} = \gamma \times temp_{gt} + \Gamma_{lgy} + \epsilon_{lgt}, \tag{1}$$

where Γ_{lgy} are route-year fixed effects. Standard errors are clustered at the country-month-year level, which is the level at which the weather variable is defined, and there are often multiple destinations in a country.

Table 3 shows the results. In column 1, we see that a one-degree increase in the mean monthly temperature in the destination's country is associated with lower volumes on a given route, by 216 cubic meters, which is about one percent of the mean volumes. Column 2 includes an indicator for whether the route-year is a contract route or not. On spot routes, volumes are 79 cubic meters less on average and, on contract routes, the volume is 279 less when the temperature at the destination increases by one degree. In columns 3 and 4, the dependent variable is contract volumes and spot volumes, respectively. Contract volumes fall by 0.98% and spot volumes fall by 1.3% relative to their mean levels when the temperature is one-degree higher.

Much of the variation in mean temperature across months at a given destination is predictable, and therefore contractible, and the findings in Table 3 that greater volumes arrive in cold weeks is therefore unsurprising. We want to focus on responsiveness of supply to unexpected demand shocks. To do this, we use the measure of weather shock to temperature defined in Section 2.2, which is (the negative of) the deviation in the mean temperature in a given month from that month's average across years divided by its standard deviation at that destination. This allows us to explore whether more LNG arrives at destinations when it is unseasonably, or unexpectedly, cold.

We denote the weather shock variable w_{gt}. The distribution of this variable over active route-years is shown in Figure 2. In the first set of specifications, we regress total volumes on a given route-week, in $'000$ cubic meters, on the weather shock that month, as follows:

$$Q_{lgt} = \gamma w_{gt} + \Gamma_{gm} + \epsilon_{lgt}, \tag{2}$$

where Q_{lgt} is the total volume on active route lg in week t. Here, we include destination-month fixed effects, Γ_{gm}. These fixed effects control for the typical volumes on routes arriving at a destination in a given month of the year, across years, allowing us to focus on deviations from this mean that are associated with the weather shock. Standard errors are clustered at the destination country-month level.

The results of estimating equation (2) are shown in Table 4. Because a positive weather shock means that the weather is unpredictably cold at the destination that month, the positive and significant coefficient

in column 1 shows that more LNG arrives at the destination when it is colder than in the same calendar month in other years. The coefficient of 1.042 suggests that an increase of one standard deviation in the weather shock is associated with an increase in LNG volumes on a route week of 1.042 thousand cubic meters, or 5% of mean route-week volumes. Column 2 shows that this result is robust to including controls for the total volume supplied by the origin in a year and for route distance, in terms of journey days.

Column 3 of Table 4 interacts total volumes with an indicator for whether the route-year is a contract route. The coefficient on the indicator variable is positive and significant, showing that contract routes to a destination tend to have larger weekly volumes than spot prices. The coefficient on the weather shock remains positive but is now smaller and insignificant, showing that there is no significant increase in volumes on spot routes when it is unexpectedly cold. However, the interaction of the contract route indicator and the weather shock is positive and significant and larger than the coefficient in column 1. This suggests it is the volumes on contract routes that adjust to unexpected weather shocks. A one standard deviation increase in the weather shock in a given month is associated with a 5.9% larger volume on these routes. We might expect that origins producing larger quantities, that are also nearby, are more able to respond quickly to weather shocks, and if these routes are also contract routes then the increased responsiveness on contract routes could be due to this selection. Column 4 of Table 4 shows that the increase in contract route volumes in response to weather shocks remains when controlling for origin size and route distance.

The second set of reduced form specifications changes the dependent variable to be total volumes, contract volumes, or spot volumes on a route-week, controlling for the average quantity with route-month fixed effects. That is, we control for all route-specific factors, such as distance, and explore whether variation in flows in a calendar month across years varies with weather shocks. The estimated equation is:

$$Q_{lgt} = f(w_{gt})'\gamma + \Gamma_{lgm} + \epsilon_{lgt}, \tag{3}$$

where Q_{lgt} are the volumes in question, and Γ_{lgm} are the route-month fixed effects, rather than the destination-month fixed effects included in Table 4. The results are given in Table 5. In columns 1 to 3, $f(w_{gt}) = w_{gt}$ and γ is a scalar in equation (3), investigating whether there is a linear relationship between the weather shock and volumes. In columns 4 to 6 of Table 5, we show the results when $f(w_{gt}) = w_{gt} + w_{gt}^2$ in equation (3) so that γ is a vector that allows for a quadratic relationship. We include this specification because very large weather shocks could lead to very high spot prices and make sellers reluctant to forgoe these prices in order to fulfil contract shipments.

In columns 1 and 4 of Table 5, the dependent variable is the total volumes on a route-week. Confirming the finding in column 4 of Table 4, column 1 shows that total volumes are greater when the temperature is unexpectedly low when controlling for route (and origin) fixed effects. Column 4 shows no evidence of a quadratic relationship. Columns 2 and 3 show that the linear relationship is significant only for contract volumes.

Tables 4 and 5 show that volumes are responsive to weather shocks on contract routes even controlling for origin and route-level factors. This suggests that contracts allow for adaptability under uncertainty and potentially even facilitate adaptation. If this were the case, we expect it to be efficient for destinations to have contracts with origins where adaptation to demand shocks is relatively low cost, for example, with nearby origins with large potential quantities available. Our data allow to test whether the routes most able to adapt to shocks are those most likely to be contract routes.

We explore whether the gravity-type variables origin size and route distance are related to being a contract route. In addition, we ask whether destinations with the most variable weather—those with the largest standard deviations in weather shocks—are more likely to have contracts. Table 6 includes an observation for each route-year, as in Table 2, the binary dependent variable is the contract route indicator and all columns include year fixed effects.

Columns 1 and 2 of Table 6 include year fixed effects and show that distance is insignificantly related to whether a route is a contract route or not. This could indicate that time is not of the essence for contract routes, perhaps because shipments can be organized in line with weather forecasts, which have a similar time horizon to the typical journey time on all active routes. These results show that larger origins and destinations are more likely to have contracts, consistent with size being related to the ability to supply, and to the magnitude of demand shock, when adverse weather hits. Column 3 includes origin-year fixed effects and shows that a given origin is more likely to have contracts with larger destinations.

Columns 4 to 6 look at the relationship between destination-year weather shocks and the probability a route has a contract. The key independent variable is mean standard deviation of the monthly weather shocks in the relevant destination. This variable is the mean of the twelve monthly standard deviations in weather shock across years. It does not vary by destination across years. For this reason, we cluster standard errors by year and destination in these estimations. Column 4 shows that destinations with greater weather variability are more likely to have contract routes. Column 5 shows that this result is robust to including gravity-type variables of journey days and origin-year volumes. Column 6 shows that destination weather variability is positively associated with having a contract controlling for journey days and also origin fixed effects.[21]

Considered together, the reduced form evidence shows that contract volumes respond to destination-specific demand shocks and that destinations that tend to have more shocks are particularly likely to have contract routes. Moreover, volumes on contract routes are responsive to weather shocks even controlling for selection on gravity-type variables into having a contract.

In the standard text of contracts between buyers and sellers in this industry, we see that shipment timing over the course of a year is left unspecified and must be bilaterally negotiated and agreed via

[21] In column six, standard errors are clustered at the year, destination and origin-year level.

shipment-specific confirmation notices. Our analysis suggests that buyers and sellers are able to adapt to agree shipments at times when they are most valuable for buyers. This insight, together with the condition that maintaining a contractual agreement must be incentive compatible for sellers given their spot market outside option, motivates the structural model of a contract that we set out in the following section.

4 Model

Building on the insight from the reduced form analysis in Section 3 showing that contract volumes respond to demand shocks, we model how a buyer decides whether to request a contract shipment as a function of the weather in a given week and the price at which LNG is available to the buyer on the local spot market. We take as given the presence of a contract on an active LNG trade route. While the reduced form results suggest demand uncertainty for buyer and the capacity of the seller are both positively related to selecting into having a contract, our objective here is to set up a model allowing us to estimate the value of observed LNG contracts during the time period studied.

We first describe the set up, then give details of the buyer's payoffs from a contract shipment and a spot shipment in turn. To derive the equilibrium shipment choice, we impose that the observed outcomes are incentive compatible in each period for the buyer and over the future duration of the contract for the seller.

4.1 Set up

Time is discrete and each period denoted t. There are two types of agents: L sellers, $l \in 1, 2, ..., L$, and G buyers, $g \in 1, 2..., G$. All agents are risk neutral. Each seller l produces an exogenous quantity of LNG in each t, and each buyer g has installed regasification capacity. Buyers and sellers are globally dispersed, and vessels containing LNG travel from sellers to buyers in each period.[22] All possible lg pairs are feasible trading routes.

There is a long-term contract in place on a subset of all lg routes. A contract specifies the quantity of LNG to be shipped from l to g over a time period that spans multiple t decision periods. For most contracts, the contracted quantity is annual and t in our model is a week. A contract also contains a formula that determines the contract price P^c_{lgt}.[23] Hence, at the start of each time period t, each l has a pre-determined set of contract buyers and each g has a predetermined set of contract sellers. Each contract

[22] We do not explicitly model the shipping stage of the value chain. Specialized LNG vessels are usually owned by or on long-leases to sellers or buyers and we assume shipping capacity is not constrained but that each available vessel is of a large fixed size.

[23] Each contract's price is indexed to the destination region's oil price, but the exact indexing rule is unobserved. Actual contract prices are also unobserved in our data.

is incomplete in that the exact timing and quantity of any one delivery are not contracted and shipments are agreed bilaterally across t in the contracting period.

Transactions on any lgt route-week can take one of two organizational forms, a contract shipment or a spot shipment. The timing is as follows:

- At the start of t, all parties observe weather conditions for all g, w_{gt}. All parties form expectations of the vector of gt-specific spot prices p^s_{gt} using information about historic supply, demand, transport costs, and weather shocks.
- Buyer g learns its demand for LNG in t, as a function of w_{gt}.[24]
- Buyer g makes a draw from the destination-specific spot market value distribution. The shape of this distribution reflects the expected market thickness of the relevant gt spot market.
- Buyer g makes one of three choices with regards to contract partner l: requests a contract shipment from l; does not request a contract shipment from l but instead purchases a shipment on the local spot market; does not request a contract shipment from l and also does not purchase a local spot market shipment.
- If g requests a contract shipment from l, l chooses whether or not to fulfil the request and also decides how much costly relationship-specific effort, $h_{lgt} \geq 0$, to exert to ensure the quantity and arrival time meet buyer g's specified requirements.
- Seller l delivers any remaining output to the spot market with the highest expected price.
- Destination spot prices p^s_{gt} are realized, equilibrating residual spot demand and supply in each destination.[25] Payments for contract and spot shipments are made.

4.2 Shipment values

4.2.1 Contract shipments

Buyer g's value for a contract shipment in t, V^c_{gt}, depends on w_{gt} and $h_{lgt} \geq 0$. We specify the functional form:

$$V^c_{gt}(w_{gt}, h_{lgt}) = v(w_{gt})(1 + h_{lgt})^{\mu} \tag{4}$$

In this "contract value production function", $\mu \leq 1$ is the intensity with which l's effort creates value for g. This parameter also captures any returns to scale in l's effort. Seller effort has a constant marginal cost, c. Each contract shipment also incurs a fixed cost F.

[24] For example, the buyer learns the ideal day(s) of delivery and ideal volume it needs, if any, in week t.

[25] Note that s is used to denote a spot shipment and not a seller.

An endogenous fraction β of contract value, net of fixed cost, goes to l and is built into the pre-agreed contract price index formula. Hence, the first order condition determining l's level of relationship-specific effort is:

$$\beta v(w_{gt})\mu(1+h_{lgt})^{\mu-1} = c, \tag{5}$$

and the seller's choice of effort (plus one) is:

$$(1+h_{lgt}) = v(w_{gt})^{\frac{1}{1-\mu}} \left(\frac{\beta\mu}{c}\right)^{\frac{1}{1-\mu}}. \tag{6}$$

Seller effort is increasing in effort intensity and in the share of contract shipment value going to the seller, and decreasing in marginal effort costs, c. Note that the production function implies potential complementarities between the responsiveness of buyer value to the gt demand shock, l's effort choice, and l's share of surplus. Variable effort costs, borne by l, are:

$$c(1+h_{lgt}) = \beta\mu v(w_{gt}) \left(v(w_{gt})^{\frac{\mu}{1-\mu}} \left(\frac{\beta\mu}{c}\right)^{\frac{\mu}{1-\mu}}\right). \tag{7}$$

The contract shipment surplus is defined as the value to the buyer less the fixed cost, in terms of the model parameters, the endogenous variable β, and the function $v(w_{gt})$:

$$S^c_{lgt} = v(w_{gt}) \left(v(w_{gt})^{\frac{\mu}{1-\mu}} \left(\frac{\beta\mu}{c}\right)^{\frac{\mu}{1-\mu}}\right) - F = v(w_{gt})^{\frac{1}{1-\mu}} \left(\frac{\beta\mu}{c}\right)^{\frac{\mu}{1-\mu}} - F \tag{8}$$

Note that S^c_{lgt} is deterministic conditional on the demand shock w_{gt} and model parameters.

4.2.2 Spot shipments

Buyer g's value for a spot shipment in t, V^s_{gt}, depends on w_{gt}. Reflecting the arms-length nature of the spot market, there are no seller-specific terms in this value. Because there is no bilateral negotiation, it is unlikely that g will be able to buy a spot market shipment that fits their exact preferences for delivery time and quantities. To model this form of shipment differentiation, we introduce a gt-specific shock to the spot market value, $\phi^s_{gt} \geq 1$, which serves to discount its value relative to the value of a contract shipment during the same period. The gt-specific shock to the spot value is ϕ^s_{gt}:

$$V^s_{gt} = \frac{v(w_{gt})}{\phi^s_{gt}}. \tag{9}$$

We let $\phi^s_{gt} \geq 1$ follow a Pareto distribution $\mathcal{F}_\phi$ with pdf $f_\phi = \frac{\lambda_{gt}}{\phi^{\lambda_{gt}+1}}$, and allow λ_{gt} to be an increasing function of the thickness of the local spot market in gt. This captures the idea that g is likely to find a better fit for its needs when there are relatively more spot shipments available and when those shipments' capacities are similar to the capacities of the contract shipments that are typically received.[26] The expected

[26] Appendix C describes how we measure spot market thickness.

value of $\left(\frac{\lambda_{gt}}{\phi^{\lambda_{gt}+1}}\right)$ is an increasing function of λ_{gt}. Larger realized draws of ϕ^s_{gt} mean that g can find only a poor match for its needs in the spot market in t.

The price g pays is the local spot market price, p^s_{gt}. Any seller in the spot market receives the spot market price less an lg-specific transport cost.

4.3 Equilibrium

An observed contract shipment on route lgt implies buyer g has found it incentive compatible to request the shipment and seller l has found it incentive compatible to fulfil the request. Following the evidence in Section 3, showing that contract volumes increase at times when there is an adverse weather shock, we assume that g's requests are consistent with per-period incentive compatibility.

To explain the fact that sellers choose to fulfil contract shipment requests even in times when their spot market outside alternative is particularly valuable, we assume the contractual relationship is dynamically incentive compatible for sellers given the future value they are able to extract from continuing relationships.

4.3.1 Buyer's incentive compatibility constraint

The state variables for buyer g in t are its pre-determined set of contract partners and the realizations of two shocks, the demand shock, w_{gt}, and thespot market value shock ϕ_{gt}. Having also observed all other demand shocks, g's expectation of the local spot price is $E(p^s_{gt})$. Requesting a contract shipment rather than no shipment is incentive compatible for g in t iff $(1-\beta)S^c_{lgt} \geq 0$, and requesting a contract shipment rather than making a spot purchase is incentive compatible for g iff:

$$(1-\beta)S^c_{lgt} \geq V^s_{gt} - E(p^s_{gt}). \tag{10}$$

From equations (8) and (9), this inequality is satisfied when ϕ_{gt} exceeds the threshold value ϕ^{s*}_{lgt}, which is given by:

$$\phi^{s*}_{lgt} = \frac{v(w_{gt})}{(1-\beta)\left[v(w_{gt})^{\frac{1}{1-\mu}}\left(\frac{\beta\mu}{c}\right)^{\frac{\mu}{1-\mu}} - F\right] + E(p^s_{gt})}. \tag{11}$$

Intuitively, the buyer prefers a contract shipment whenever the spot market shock (which discounts the spot shipment value) is sufficiently large. The threshold value ϕ^{s*}_{lgt} is greater when the demand shock is high. This is because the demand shock increases the value of all shipments relative to no shipment at all, but increases the value of a contract by more because the seller will choose to exert more relationship-specific effort.

Hence, whenever $(1-\beta)S^c_{lgt} \geq 0$, buyer g requests a shipment. The probability that g requests a contract shipment from l is therefore the probability $\phi_{gt} \geq \phi^{s*}_{lgt}$, which is:

$$\Pr(\phi^s_{gt} \geq \phi^{s*}_{lgt}) = 1 - F_\phi(\phi^{s*}_{lgt}) = (\phi^{s*}_{lgt})^{-\lambda}, \tag{12}$$

and the probability that g buys on the spot market is the probability that ϕ^s_{gt} is less than ϕ^{s*}_{lgt}.[27]

4.3.2 Seller's incentive compatibility constraint

We impose that the seller's equilibrium strategy satisfies a dynamic incentive compatibility constraint. Seller l fulfils a contract request if over the course of all future periods $t' > t$, the share of contract surplus it will receive, less the variable effort costs it will incur, outweighs the revenue it would earn from selling the same quantities in spot markets.[28] While the seller could sell in any spot market in t', the relevant spot revenues are those from selling in the market with the highest spot price net of seller-specific transport costs.

Seller l's incentive compatibility constraint assumes that buyer g is playing a trigger strategy, whereby if the seller does not send any or all requested contract shipment in any period t and, implicitly, if the buyer feels that the seller has reneged on the relationship governed by the contract, then the relationship ends and the seller must sell the future output assigned to the relationship to be best available spot market. Hence, the seller's incentive compatibility constraint summing over all future periods $t' > t$, can be written as follows:

$$\sum_{t' \geq t} \delta_{lt'}(\beta S^c_{lgt'} - c(1 + h_{lgt'})) \geq \sum_{t' \geq t} \delta_{lt'} E(max_{g \in G}(p^s_{gt'} - \tau_{lgt})), \tag{14}$$

where $g \in G$ indexes all spot markets, τ_{lgt} is the seller's transport cost to each spot market, and $\delta_{lt'}$ is the seller's discount factor.

5 Estimation

The model set out in Section 4 derives the conditions under which a contract shipment is observed on route-week lgt. We use the model to derive derive theoretical moments that relate the probability a contract shipment is observed to the variables observed in the data and model parameters. We then turn to a subsample of the data to estimate the model, choosing the subsample to match the assumptions made. The data provide a set of empirical moments detailing the route-weeks where contract shipments took place. Finally, we describe the estimation approach.

[27] Although we do not currently use it in the estimation, there is a second threshold value of ϕ^s_{gt}, above which the buyer prefers no shipment to a spot shipment, that is $\frac{v(w_{gt})}{\phi^s_{gt}} - E(p^s_{gt}) \leq 0$. This threshold value is:

$$\bar{\phi}^s_{gt} = \frac{v(w_{gt})}{E(p^s_{gt})}. \tag{13}$$

[28] Buyer g is more likely to request a shipment when it receives a more positive demand shock, which is exactly when l's outside option of selling on the spot market is likely to be of high value.

5.1 Theoretical moments

Equation 12 is the probability that buyer g requests a contract shipment from seller l in period t. Substituting equation (11) into this probability allows us to express it as a function of the weather shock, expected spot market price, and model parameters. The parameters of interest are those that relate the demand shock to shipment value (in the function $v(w_{lgt})$), seller effort intensity, μ, the share of seller surplus, β, the marginal costs of contract effort, c, and the contract shipment fixed costs, F_{lt}. We also focus on how the shape parameter of the Pareto distribution of spot value shocks, λ, relates to spot market thickness.

We first specify a functional form for the value function $v(w_{lgt})$. We allow this function to reflect how destinations in the data vary in their reliance on contemporaneous LNG supply as an energy source.[29] The responsiveness of demand to weather shocks depends on storage capacity relative to annual demand in a shipment value function that is a power function governed by three parameters, k, α_i, for $i = 0, 1$. We set $v(w_{lgt}) = k w_{gt}^{\alpha_0 R_{gt}^{\alpha_1}}$, where w_{gt} is the weather shock at g at time t and Rgt is the ratio of total annual demand relative to storage capacity. Our hypothesis is that a higher ratio of demand to storage capacity means destinations are less able to use inventories to smooth out demand shocks and so this ratio will be associated with a greater responsiveness of willingness to pay to a weather shock, implying $\alpha_1 > 1$.

Next, because the relationship between spot market thickness and the value of a spot market shipment will be of key importance in the counterfactual analysis, we estimate λ flexibly to be $\lambda = \lambda_0 (Th_{gt})^{\lambda_1}$, where Th_{gt} measures spot market thickness of the spot market faced by buyer g at time t.

Third, in order to allow the fixed costs of fulfilling a contract shipment to vary with deviations from the typical number of l's contract shipments, we specify $F = F_0 (N_{lt})^{F_1}$, where N_{lt} is the deviation from the mean number of contract shipments a seller fulfils in each period.

The threshold value of the shock above which a contract shipment is preferred to a spot shipment, ϕ_{lgt}^{s*}, from equation (11), is now:

$$\phi_{lgt}^{s*} \equiv \frac{k w_{gt}^{\alpha_0 R_{gt}^{\alpha_1}}}{(1-\beta) S_{lgt}^{c} + E(p_{gt}^{s})}.$$

Hence, the theoretical moments giving the probability buyer g requests a contract shipment in lgt, conditional on $S_{lgt}^{c} \geq 0$, are:

$$\Pr(\phi_{gt}^{s} \geq \phi_{lgt}^{s*}) = \begin{cases} \left(\frac{k w_{gt}^{\alpha_0 R_{gt}^{\alpha_1}}}{(1-\beta) S_{lgt}^{c} + E(p_{gt}^{s})} \right)^{-\lambda} & S_{lgt}^{c} \geq 0 \\ 0 & S_{lgt}^{c} < 0 \end{cases} \tag{15}$$

where the term S_{lgt}^{c} is the surplus from a contract shipment and, under the functional form assumptions,

[29] It some countries, it is a core energy source and in others it serves as a "peak shaver". Related, countries vary in their storage capacity relative to total demand. Investments in storage capacity are an endogenous response to the contribution LNG makes to mean energy supply and to its variation.

equation (8) becomes:

$$S^c_{lgt} = \left(k w_{gt}^{\alpha_0 R_{gt}^{\alpha_1}}\right)^{\frac{1}{1-\mu}} \left(\frac{\beta\mu}{c}\right)^{\frac{\mu}{1-\mu}} - F,$$

and the theoretical moments in expression (15) are functions of all model parameters of interest. In particular, the theoretical moments depend on eight parameter values $\theta = \{k, \alpha_0, \alpha_1, \lambda_0, \lambda_1, F_0, F_1, \frac{\beta}{c}\}$. We note that β and c cannot be separately identified from these moments, as they enter expression (15) similarly. Both these parameters relate to the seller's incentives to fulfil a contract shipment, β being the share of contract shipment surplus going to the seller and c being the seller's marginal cost of relationship-specific effort. To derive estimates for each of these parameters, we combine the theoretical moments in expression (15) with l's dynamic incentive compatibility constraint, given in inequality (14), to derive lower bounds for β, which permits an estimate of c.

The theoretical moments in expression (15) give the probability of a contract shipment on lgt as a non-linear function of all the parameters. Appendix E establishes that each parameter is separately identified. Some intuition for parameter identification can be found by observing that if w_{gt} is held constant, then the numerator and first term in the denominator of the theoretical moments in expression (15) are fixed, implying that the buyer's payoff from a contract shipment is fixed. Variation in the probability of observing a contract shipment is therefore determined only by variation in buyer payoff from a spot shipment, which varies with the expected spot price and the responsiveness of spot shipment value to spot market thickness, as set out in equation (9). This variation allows us to identify $\lambda = \{\lambda_0, \lambda_1\}$, relating spot market value to market thickness.[30] Next, the given the functional form assumptions for V^o_{gt}, $o \in \{c, s\}$, and holding $E(p^s_{gt})$ fixed, variation in the probability of observing a contract shipment comes only from the relative responsiveness of contract shipment value to variation in w_{gt}. This allows us to estimate the parameters in the V^o_{gt} function, k, α_0, and α_1, along with μ and a term that is a function of both β and c.[31]

5.2 Data selection for model estimation

Section 4 models decisions about contract shipments at the lgt level given the existence of a contract on route lg at time t, and the ability of supplier l to fulfil buyer g's contract shipment request if one is made. It hence abstracts from several key aspects of the buyer and seller's broader problems. For the purposes of model estimation, we select a subset of the available lgt data to limit the impact of these simplifications.

[30] As described in the set up, we assume the observed spot price is the expected spot price, hence we have data on expected spot price variation.

[31] We have estimated the model including all parameters and also simpler versions of the model that set $F_1 = 0$ and then $\alpha_1 = 0$. In these specifications, the fixed cost is independent of the seller's total number of contracts fulfilled or the responsiveness of buyer demand to weather shocks is invariant to the storage capacity in the destination country, respectively. Our base case results used for counterfactual analysis will be the model specification that best matches the non-targeted moments.

First, we restrict attention to lgt observations on lg routes where there is a contract in place. To do this we consider those active route-years where at least 50% of the observed shipments are contracted. As shown in Figure 3 and described in Section 2.3, the distribution of contracted shipments is concentrated at zero or one. Second, we focus only on weeks where one of two conditions holds: either we observe at least one contracted shipment on that route-week, or both the seller and buyer are present in the spot market. The second condition ensures that the seller either decided to fulfil all contract shipment requests from the buyer or chose to not fulfil a request and sold on the spot market instead. It also ensures that the buyer had some positive demand in that week. This ensures that the contract shipment surplus defined in equation (8) is positive, which is a necessary condition for the probability a contract shipment is requested by g to be expressed as equation (12). When contract shipment surplus is positive, the buyer's decision is only whether or not to request a contract shipment from a contract seller origin rather than a spot shipment, conditional on having positive demand.[32]

One comment relevant to estimation sample selection is that we use the full set of observations described in Table 1 and in Section 2.3 to construct the gt-specific measure of spot market thickness, Th_{gt}. This variable allows us to estimate how the idiosyncratic shock to spot shipment value varies with the nature of the local spot market. Our preferred measure has two components. We consider the relative size of the spot market in the region in that period compared to the destination's contract demand in that period. If buyers have total demand for a large number of shipments, and much of it is fulfilled via contract, then their spot market options are more limited. We make an adjustment to this measure by multiplying it by the extent of overlap between the distribution of vessel capacities requested by the buyer under contract and the distribution of vessel capacities available on the region-specific spot market. This adjustment accounts for the fact that even if the regional spot market is relatively large, a buyer might not be able to find a vessel that fits their demand specifications at any point in time.[33]

We also use the full set of observations described in Table 1 and in Section 2.3 to construct N_{lt}, which is the number of contracts seller l fulfils in t relative to the typical number. This variable is included in the estimation in the specification that includes the F_1 parameter to allow the fixed costs of contracts to reflect any seller congestion in contract supply in t.

Finally, the discount factor used is $\delta_{lt'} = \frac{1}{1+r_{lt'}}$, where $r_{lt'}$ is the seller-time-specific real interest rate in period t'. This is relevant for estimating the seller's incentive compatibility constraint given in equation (14).[34] Because buyer g makes decisions that are incentive compatible in each time period, estimating the

[32]Our model suggests weeks where buyers receive both contract and spot shipments are times when the buyer would have preferred to receive more contract shipments, but their additional requests were unfulfilled by the seller(s). We implicitly assume that, when they are asked, sellers $l \in L$ send sufficient contract shipments on a route to avoid relationship breakdown.

[33]Appendix C gives a detailed description of how Th_{gt} is constructed.

[34]These country-year interest rates are taken from the World Development Indicators.

model parameters does not require making any assumptions about buyers' discount factors.

5.3 Empirical moments and estimation

There are $30,050$ lgt observations that satisfy the sample restrictions in subsection 5.2. That is, they are on a contract route, where there is either at least one contract shipment on the route, or both the buyer and seller are in the spot market that period.

The empirical moments are the share of all observed shipments that are contract rather than spot. We aggregate the selected route-week observations to the origin-destination country level and find the share of all shipments in a given week to destinations in a given country that are contract shipments. This aggregation gives us $16,818$ empirical moments that we can relate to data on weather and spot price shocks, which are are the destination country-month and destination country-week level, respectively. We call this variable C_{lgt}, with the slight abuse of notation that buyer g now denotes the destination country rather than the destination plant.

In this sample of $16,818$ empirical moments, the share of contract shipments C_{lgt} is positively correlated with the weather shocks, w_{gt}, and with the local spot price, p^s_{gt}, in regressions that include country destination month fixed effects. Although local spot price variation is correlated with weather shocks, there is sufficient independent variation that including both in a regression on C_{lgt} gives positive and significant coefficients for both variables. We will use the coefficients in these three regressions as non-targeted moments in our estimation as a robustness check for the parameter estimates. The regression is:

$$C_{lgt} = \gamma_0 + \gamma_1 w_{gt} + \gamma_2 p^s_{gt} + \varepsilon_{lgt}. \tag{16}$$

We use the weather shock as defined in Section 2.2 and also normalize the distribution of local spot prices, so that the magnitudes of the coefficient estimates are comparable. The regression results are shown in Table 7. The estimated coefficients in the specification with both weather shocks and spot prices on the left hand side, as in equation (16), are $\hat{\alpha}_1 = 0.043$ and $\hat{\alpha}_2 = 0.056$, and both are significant at the 1% level.

With both the theoretical moments given in equation (15) and the empirical moments C_{lgt} in hand, we employ non-linear least squares to estimate the model parameters $\boldsymbol{\theta} = \{D, k, \alpha_0, \alpha_1, \lambda_0, \lambda_1, F_0, F_1, \frac{\beta}{c}\}$, where D is a constant term that can be interpreted as the probability of a contract shipment when the weather shock and spot price are equal to zero. We find the parameter values that solve the following minimization problem:

$$\min_{\boldsymbol{\theta}} \sum_{t=1}^{T} \sum_{g=1}^{G} \sum_{l=1}^{L} \left((\phi^{s*}_{lgt})^{-\lambda_{gt}} - C_{lgt} \right)^2, \tag{17}$$

where $t = (1, ..., T)$ are the weeks in the data, $g = (1, ...G)$ are destination country buyers, and $l = (1, ..., L)$ are sellers.

6 Results

6.1 Parameter estimates

Estimating equation (17) by non-linear least squares gives model parameter estimates. In this section, we present the model specification that best fits the non-targeted moments. This is the version with both α_1 and F_1 set equal to zero. Appendix F presents two model extensions, including the specification with all model parameters.

Parameter	Description	Estimate	Confidence Interval
D	Constant	0.39	[0.387, 0.394]]
k	Value function normalization	10.07	[9.997, 10.135]
α_0	Value function elasticity	0.08	[0.075, 0.081]
α_1	Value function elasticity storage term	0.00	
μ	Effort intensity	0.51	[0.508, 0.512]
c	Marginal effort cost	1.05	[1.035, 1.068]
F_0	Contract fixed cost normalization	1.85	[1.827, 1.878]
F_1	Contract fixed cost elasticity	0.00	
λ_0	Spot market thickness normalization	0.73	[0.710, 0.741]
λ_1	Spot market thickness elasticity	2.29	[2.241, 2.331]

All estimates are significantly different from zero. The positive value function elasticity parameter α_0 shows that the value g receives from a shipment is higher when there is a positive weather shock, that is, when it is unexpectedly cold. The positive estimates of λ_0 and λ_1 show that the value g receives from spot market shipments is positively associated with Th_{gt}, our measure of local spot market thickness at time t. The coefficient μ of 0.51 shows that relationship-specific seller effort contributes positively to contract shipment value. The positive coefficient F_0 suggests there are fixed costs associated with each contract shipment.

To find the β parameter estimate, we return to inequality (14):

$$\sum_{t' \geq t} \delta_{lt'}(\beta S^c_{lgt'} - c(1 + h_{lgt'})) \geq \sum_{t' \geq t} \delta_{lt'} E(max_{g \in G}(p^s_{gt'})),$$

which delivers an estimate of β for each lgt. The distribution of these estimates is given in Figure 4. Since we need our estimate of β to satisfy all of these constraints, we take it as the maximum of the distribution in Figure 4, which is 0.22, with a confidence interval $[0.221, 0.229]$.

The results give that the seller exerts relationship-specific effort, that is, $h > 0$, for 89% of contract shipments. The seller incurs additional variable costs in these cases. This suggests that the seller's share of the contractual relationship surplus is sufficient to compensate them for foregoing spot markets and also

exerting effort for these shipments.[35] In the model, h is an increasing function of the weather shock, and there is a strong positive correlation between estimated h_{lgt} values and w_{gt}. That is, sellers exert effort within contractual relationships when their contract partners face unexpectedly cold weather.

The parameter predictions allow us to predict the responsiveness of C_{lgt}, the share of contract shipments, to observed variation in weather and spot prices using equation (16). Table 8 gives the estimated moments using model parameters. These non-targeted moments are close to the actual data, with the model predicting more variation with each variable than we see in the data. Appendix F shows that the model specification that includes α_1 or F_1 produces similar coefficient estimates for the other parameters to the model specification described here, but the non-targeted moments are less well matched.

6.2 Shipment value estimates

The estimated values of the observed contract shipments are shown in Figure 5. The far right histogram in the figure presents calculated shipment values from equation (4) given the functional form assummption for $v(w_{gt})$ made in Subsection 5.1:

$$V^c_{gt}(w_{gt}, h_{lgt}) = kw^{\alpha_0}_{gt}(1 + h_{lgt})^{\mu}$$

The middle histogram subtracts estimated fixed costs, F_{lt}, giving the distribution of S^c_{lgt}. The left (purple) histogram is the surplus of contract shipments net of estimated variable seller effort costs, $c(1 + h_{lgt})$. The mean of the left (purple) distribution is 6.68 and the standard deviation is 0.73. These values can be compared to the normalized distribution of the local spot price per mmbtu, the mean of which is 0.33. That is, we estimate that the mean value of a contract shipment is 20 times higher than the mean spot price over the period studied.

The expected spot market shipment value can be calculated from local weather and the estimate of $\lambda_{gt} = \lambda_0 Th^{\lambda_1}_{gt}$, where Th_{gt} is the observed thickness of the local spot market.[36] It is:

$$E(V^s_{gt}) = E\left(\frac{kw^{\alpha_0}_{gt}}{\phi^s_{gt}}\right) = kw^{\alpha_0}_{gt}\frac{\lambda}{\lambda + 1} \tag{18}$$

Figure 6 shows the calculated expected spot values together with the net contract surplus net of variable effort costs (purple histogram) from Figure 5.

In Figure 6, the blue histogram gives the expected values of spot shipments when at least one spot shipment is observed in lgt, weighted by the actual number of spot shipments observed in the data. The mean of the distribution is 3.10 and the standard deviation is 0.67. The estimated mean value of a spot

[35] We note that the seller receives only a β share of the additional surplus generated through its effort, and so is suboptimal.

[36] The derivation is: $E(V^s_{gt}) = E\left(\frac{kw^{\alpha_0}_{gt}}{\phi^s_{gt}}\right) = kw^{\alpha_0}_{gt}\frac{\lambda_{gt}}{\lambda_{gt}+1} = kw^{\alpha_0}_{gt}E(\frac{1}{\phi^s_{gt}}) = kw^{\alpha_0}_{gt}\int \frac{1}{\phi^s_{gt}}\frac{\lambda_{gt}}{(\phi^s_{gt})^{\lambda_{gt}+1}}d\phi^s_{gt} = kw^{\alpha_0}_{gt}\lambda_{gt}\int \frac{1}{(\phi^s_{gt})^{\lambda_{gt}+2}}d\phi^s_{gt} = -kw^{\alpha_0}_{gt}\frac{\lambda_{gt}}{\lambda_{gt}+1}(\phi^s_{gt})^{-(\lambda_{gt}+1)}|^{\infty}_1 = kw^{\alpha_0}_{gt}\frac{\lambda_{gt}}{\lambda_{gt}+1}$.

shipment is hence only 46% of the mean value of a contract shipment, but is still around nine times larger than the normalized mean spot price. Around 37% of the gains from trade in the current spot market goes to the seller and around 63% goes to buyers. Note that we estimate $\beta = 0.22$, which is the share of surplus going to sellers in the contract relationships, suggesting that the sellers extract a larger share of the gains from trade in spot shipments compared to their of the surplus in contract shipments. However, because the magnitude of the surplus is larger for contract shipments, the sellers have around 30% larger gains from trade in contract shipments.

7 Counterfactual Analysis

We now ask how much long-term contractual relationships contribute to industry value relative to a situation where the same vessels arrived in each destination in each period and were transacted in the spot market. That is, the counterfactual is larger spot markets in each lgt where contract shipments are currently observed. The trade-off is that all shipments become "anonymous" arm's length transaction, and sellers exert no relationship-specific effort and incur zero contract fixed and variable costs.

We are able to estimate this counterfactual because the λ parameter estimates from Section 6 allow us to calculate the expected spot shipment value as a function of the observed weather shock and the counterfactual spot market thickness, Th_{gt}, using equation (18). The measure of Th_{gt} typically becomes equal to 1 because the share of regional supply that is spot is now 100%. In addition, because the vessel sizes that were demanded under contract by each buyer are now a subset of the distribution of vessel sizes available via spot, the distributions of vessel sizes become equivalent (see Appendix C).

One appeal of this approach is that we do not need to simulate counterfactual spot prices. We also avoid needing to address the seller's allocation problem, that is, where they would send vessels in the absence of contracts. A related benefit is that our counterfactual estimates are invariant to transport costs, which are currently excluded from the analysis, but which remain constant in the counterfactual.

For this analysis, we estimate the counterfactual values of all contract and spot shipments, including those that did not satisfy the estimation sample restrictions in Subsection 5.2. This gives us $31,201$ contract shipments and $6,198$ spot shipments, so $37,399$ shipments in total.

The predicted values of these $37,399$ shipments sold in the enlarged spot markets is given in the red histogram in Figure 7. For comparison, the blue histogram is the distribution of actual spot shipment values given the observed much smaller number of spot shipments, as shown in Figure 6. Figure 7 also includes the observed shipment surplus for contracts net of variable costs (the purple histogram). The red histogram reflects higher counterfactual spot values than in the blue histogram because buyers expect to receive a spot shipment that is a closer fit to their needs in larger spot markets. This counterfactual distribution has a mean value of 3.70 and a standard deviation of at 0.39, which is smaller than the

standard deviation in the observed spot value distribution. This is because there is no variation in Th_{gt} in the counterfactual.

Because the counterfactual holds all transport costs fixed, and varies only the value created for buyers, it is possible to compare total welfare in the observed and counterfactual cases.

The total change in welfare can be decomposed into three parts, where the variables with asterisks denote the number of counterfactual shipments, Q_s^*, and the expected value of the distribution of the expected value of those counterfactual shipments, $E(V_s^*)$:

$$\begin{aligned}\Delta W &= \Delta W_s + \Delta W_c \\ &= Q_s^* E(V_s^*) - (Q_s E(V_s) + Q_c(V(h) - F - c(1+h)) \\ &= (Q_s + Q_c)E(V_s^*) - (Q_s E(V_s) + Q_c(V(h) - F - c(1+h))) \\ &= Q_s(E(V_s^*) - E(V_s)) + Q_c(E(V_s^*) - V(h)) + Q_c(F + c(1+h)) \end{aligned} \tag{19}$$

The first term in equation (19) is the percentage increase in the value of spot shipments in the counterfactual larger spot market. This change is 1.1%. The second term is the percentage reduction in the value of the shipments that are currently transacted via contracts, and is estimated to be 78.4%. This loss can be attributed to the loss in shipment value from the fact that sellers no longer match these shipments to buyers' exact needs and exert no relationship-specific costly effort to ensure the quality of the match. The third term in equation (19) is the welfare gain coming from the savings in the seller's variable effort costs and fixed costs associated with contract shipments. This saving is equivalent to 14.9 for variable cost savings and 23.16% for fixed cost savings of industry value over the time period studied.

The net impact of the removal of contracts is a reduction in industry value of 39.29%. We infer that total industry size remains too small for spot markets to deliver shipments to buyers that fit their demand even if all industry supply were transacted on global spot markets rather than via contracts.

8 Conclusion

Long-term contracts dominate the LNG industry and the nature of these relationships shapes industry value. This paper brings insights from recent advances in the theory of incomplete contracts to bear on detailed transaction-level data. Our main contribution is the finding that these contracts allow transactions to flexibly adjust to unexpected demand circumstances. We also quantify the significant role played by relationship-specific seller effort in creating value for the buyer. Adaptation, often enabled by non-contractible effort, increases industry efficiency relative to a counterfactual of LNG spot markets characterized by frictions from technology and geography. The analysis reveals that the value of the global LNG industry would have been less than two thirds of its actual size between 2009 and 2019 in the absence of adaptation within long-term contracts.

Figures

Figure 1: LNG spot prices by destination country over time, USD

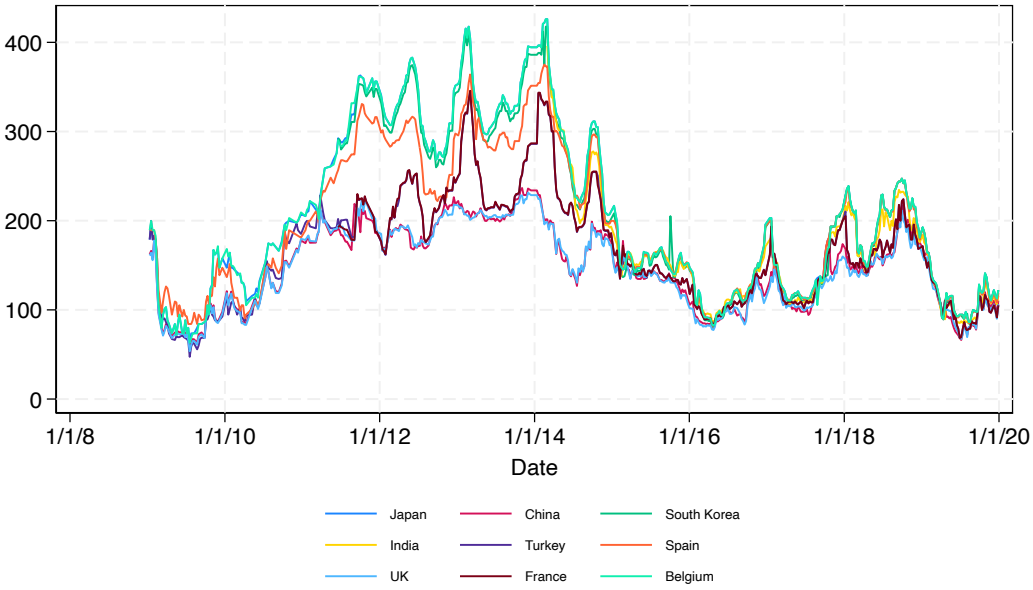
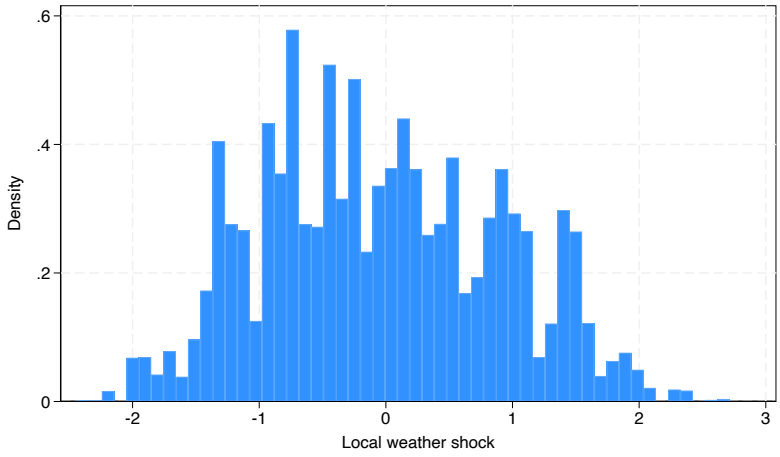


Figure 2: Destination-month temperature shocks, standardized



B.5.

Figure 3: Share of route-year volumes under contract

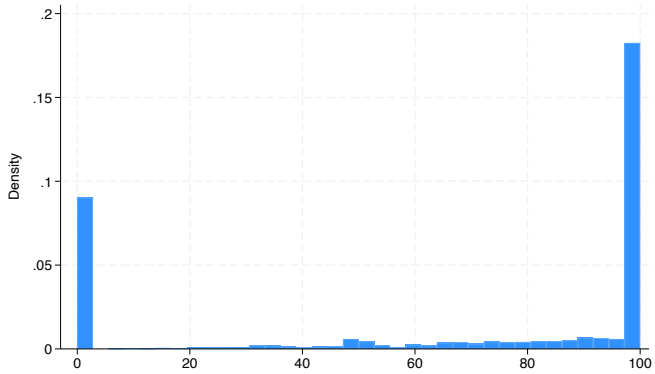


Figure 4: Estimated lower bounds on β

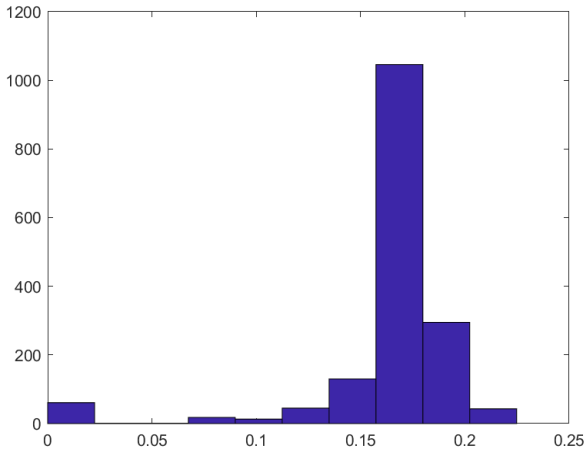


Figure 5: Estimated value of contract shipments

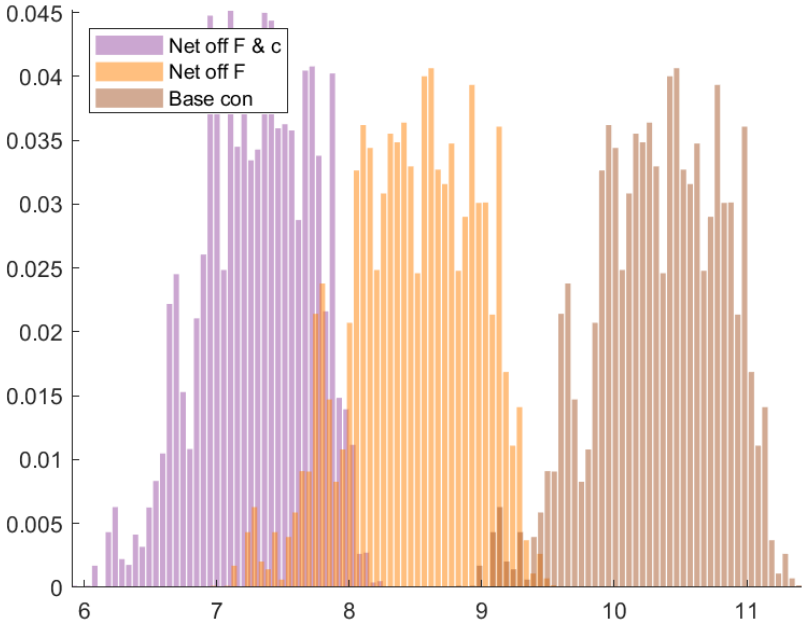


Figure 6: Expected value of spot and contract shipments

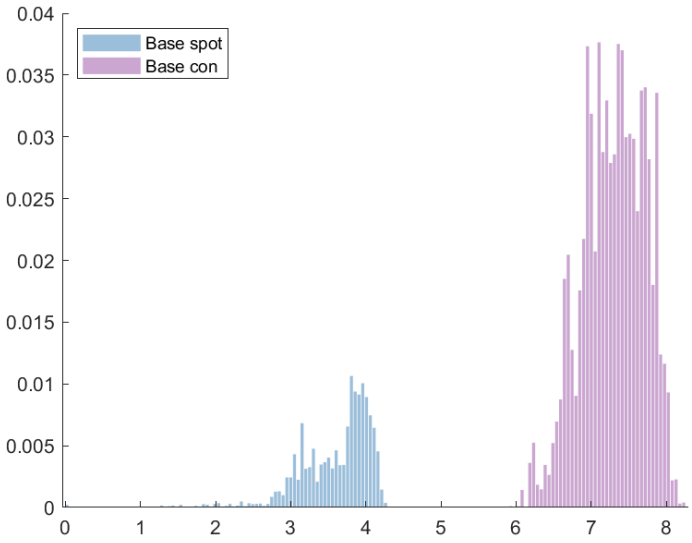
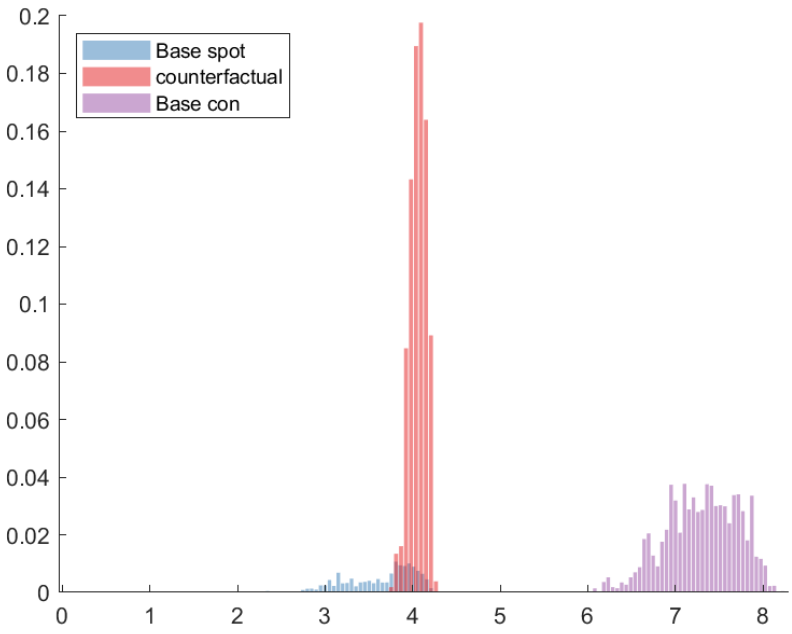


Figure 7: Observed and Counterfactual Shipment Values



Tables

Table 1: Summary statistics for active route-weeks, 2009 to 2019

	Obs	Mean	Std.dev	Min	Max
(a) All weeks					
Total shipments	192,911	0.15	0.42	0	5
Total volumes, $'000m^3$	192,911	20.95	60.77	0.00	999.98
Contract shipments	192,911	0.13	0.39	0	5
Contract volumes, $'000m^3$	192,911	17.72	56.84	0.00	999.98
Spot shipments	192,911	0.02	0.16	0	3
Spot volumes, $'000m^3$	192,911	3.23	21.83	0.00	437.62
(b) Weeks with positive flows of relevant variable					
Total shipments	25,466	1.15	0.42	1	5
Total volumes, $'000m^3$	25,466	158.71	78.15	0.15	999.98
Contract shipments	21,474	1.15	0.43	1	5
Contract volumes, $'000m^3$	21,474	159.21	80.62	0.15	999.98
Spot shipments	4,424	1.03	0.17	1	3
Spot volumes, $'000m^3$	4,424	140.79	37.61	0.87	437.62
(c) All weeks					
Local spot price, USD per m^3	192,911	187.34	84.80	47.34	426.06
Local mean temp, °c	192,911	12.34	8.94	-8.99	31.25
Local temp shock, °c	192,911	-0.04	0.93	-2.44	2.61
Journey days	189,982	10.24	6.49	0.24	28.79
Journey cost, USD per m^3	189,982	16.86	9.39	3.37	45.66
(d) All weeks with positive total volumes					
Local spot price, USD per m^3	25,466	195.90	88.42	47.34	426.06
Local mean temp, °c	25,466	11.63	8.99	-8.99	31.25
Local temp shock, °c	25,466	0.01	0.93	-2.44	2.61
Journey days	25,247	8.47	5.31	0.24	28.79
Journey cost, USD per m^3	25,247	14.30	7.54	3.37	45.66

Table 2: Probability route active in a year.

	(1)	(2)	(3)	(4)	(5)	(6)
Journey days	-0.02***	-0.01***	-0.02***			
	(0.00)	(0.00)	(0.00)			
Volume at origin		0.04***			0.04***	
		(0.00)			(0.00)	
Volume at dest		0.18***			0.18***	
		(0.00)			(0.00)	
Cost per m^3				-0.01***	-0.01***	-0.01***
				(0.00)	(0.00)	(0.00)
Constant	0.44***	0.22***	0.47***	0.46***	0.24***	0.49***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Observations	19556	19556	19556	19556	19556	19556
R^2	0.079	0.235	0.374	0.082	0.238	0.374
Fixed effects	year	year	year	year	year	year

Observations are at the route-year level.

Standard errors in parentheses, clustered by year.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3: Route-week volumes

	(1)	(2)	(3)	(4)
	Total	Total	Contract	Spot
Mean temp	-0.216***	-0.079***	-0.173***	-0.043***
	(0.016)	(0.020)	(0.013)	(0.007)
Mean temp * contract route		-0.200***		
		(0.025)		
Constant	23.610***	23.650***	19.855***	3.755***
	(0.237)	(0.238)	(0.192)	(0.119)
Observations	192910	192910	192910	192910
R^2	0.361	0.361	0.389	0.087
Fixed effects	route-year	route-year	route-year	route-year

Observations are route-week volumes on active route-years.

Volumes are in ′000 cubic meters.

Standard errors in parentheses, clustered at destination country-month level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4: Route-week volumes

	(1)	(2)	(3)	(4)
	Total	Total	Contract vol	Spot vol
Weather shock	1.042***	0.946***	0.113	0.379
	(0.184)	(0.184)	(0.243)	(0.272)
Total vol origin		3.066***		2.811***
		(0.273)		(0.270)
Journey days		-0.908***		-0.878***
		(0.047)		(0.060)
Contract route			19.714***	14.243***
			(0.847)	(0.662)
Weather shock * contract route			1.125***	0.681**
			(0.297)	(0.321)
Constant	21.001***	18.751***	7.240***	9.423***
	(0.009)	(0.970)	(0.588)	(1.093)
Observations	192909	189980	192909	189980
R^2	0.067	0.123	0.086	0.132
Fixed effects	dest-month	dest-month	dest-month	dest-month

Observations are route-week volumes on active route-years.

Volumes are in ′000 cubic meters.

Standard errors in parentheses, clustered at destination country-month level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5: Route-week volumes

	(1)	(2)	(3)	(4)	(5)	(6)
	Total	Contract	Spot	Total	Contract	Spot
Weather shock	0.745***	0.648***	0.097	0.754***	0.662***	0.092
	(0.175)	(0.156)	(0.075)	(0.175)	(0.154)	(0.076)
Weather shock sq				-0.099	-0.166	0.067
				(0.184)	(0.164)	(0.079)
Constant	20.988***	17.757***	3.232***	21.074***	17.900***	3.174***
	(0.008)	(0.007)	(0.004)	(0.159)	(0.140)	(0.068)
Observations	192859	192859	192859	192859	192859	192859
R^2	0.350	0.369	0.110	0.350	0.369	0.110
Fixed effects	route-year	route-year	route-year	route-year	route-year	route-year

Observations are route-week volumes on active route-years.

Volumes are in ′000 cubic meters.

Standard errors in parentheses, clustered at destination country-month level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6: Probability active route mostly contract in a year.

	(1)	(2)	(3)	(4)	(5)	(6)
Journey days	-0.004	-0.004	-0.012		-0.004	-0.015***
	(0.003)	(0.003)	(0.009)		(0.003)	(0.003)
Total vol orirgin		0.014***			0.018***	
		(0.002)			(0.002)	
Total vol dest		0.025***	0.055***		0.008	0.031*
		(0.006)	(0.013)		(0.016)	(0.016)
Weather variability				0.329***	0.327***	0.242**
				(0.079)	(0.080)	(0.082)
Constant	0.719***	0.650***	0.756***	0.413***	0.387***	0.614***
	(0.026)	(0.025)	(0.093)	(0.069)	(0.083)	(0.084)
Observations	5059	5059	5059	4221	4161	4149
R^2	0.015	0.033	0.201	0.031	0.062	0.295
Fixed effects	year	year	year, origin	year	year	year, origin-year

The dependent variable is the probability an active route is a contract route.

Weather variability is the mean standard deviation in month level temperatures across years.

Standard errors in parentheses, clustered at the year level in all columns, also at the origin level in column three, at the dest level in columns 4 and 5 and also at the origin-year level in column 6.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 7: Non-targeted moments share of contract shipments.

	(1)	(2)	(3)
	prob	prob	prob
Weather shock	0.062***		0.043***
	(0.012)		(0.013)
Local spot price (normalized)		0.063***	0.056***
		(0.011)	(0.011)
Constant	0.366***	0.376***	0.351***
	(0.009)	(0.006)	(0.010)
Observations	19588	16818	16818
R^2	0.002	0.002	0.003

Observations are the share of shipments that are via contract on origin-destination country routes in weeks that satisfy sample selection criteria. Robust standard errors in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 8: Non-targeted moments

Coefficient	Data	Model predictions
Weather shock	0.043**	0.070***
	(0.013)	(0.003)
Local spot price (normalized)	0.056**	0.084***
	(0.011)	(0.002)

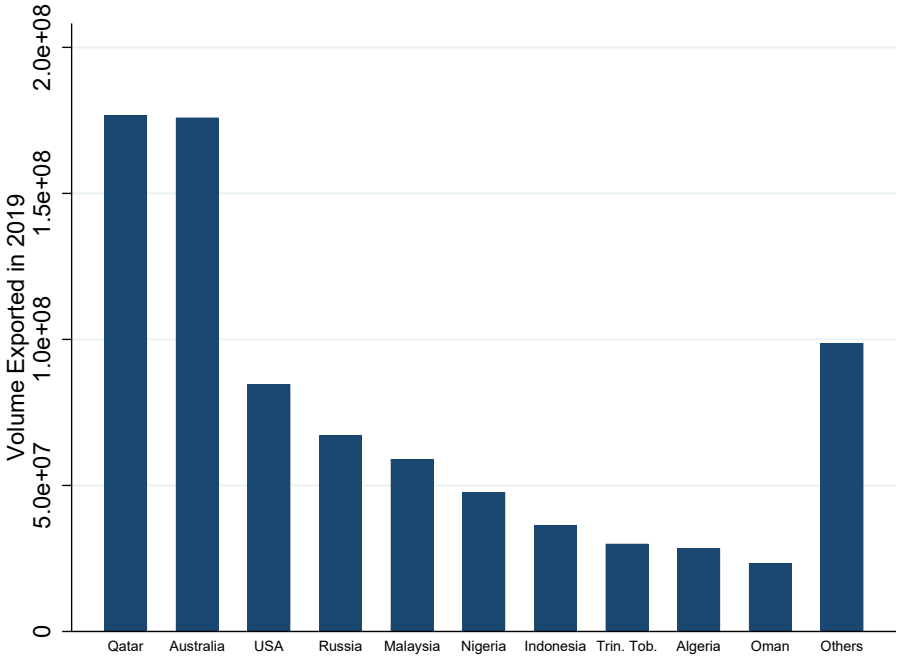
Appendix

A Global supply and demand in 2019

A.1 Supply

The shipment-level data we use, which runs from the start of 2009 to the end of 2019, contains shipments from 19 countries. Qatar has been the largest exporter historically, with over 500 shipments in each of the 11 years. Australia has increased exports over the time period, and exported more than 500 shipments in the last four years in the data. After Qatar and Australia, the largest exporter countries are Malaysia, Indonesia, Algeria, and Nigeria. Figure A.1 focuses on 2019, and shows total volumes exported by source country in cubic metres. The USA, Russia, and Trinidad and Tobago are more prominent in 2019 than in the period as a whole.

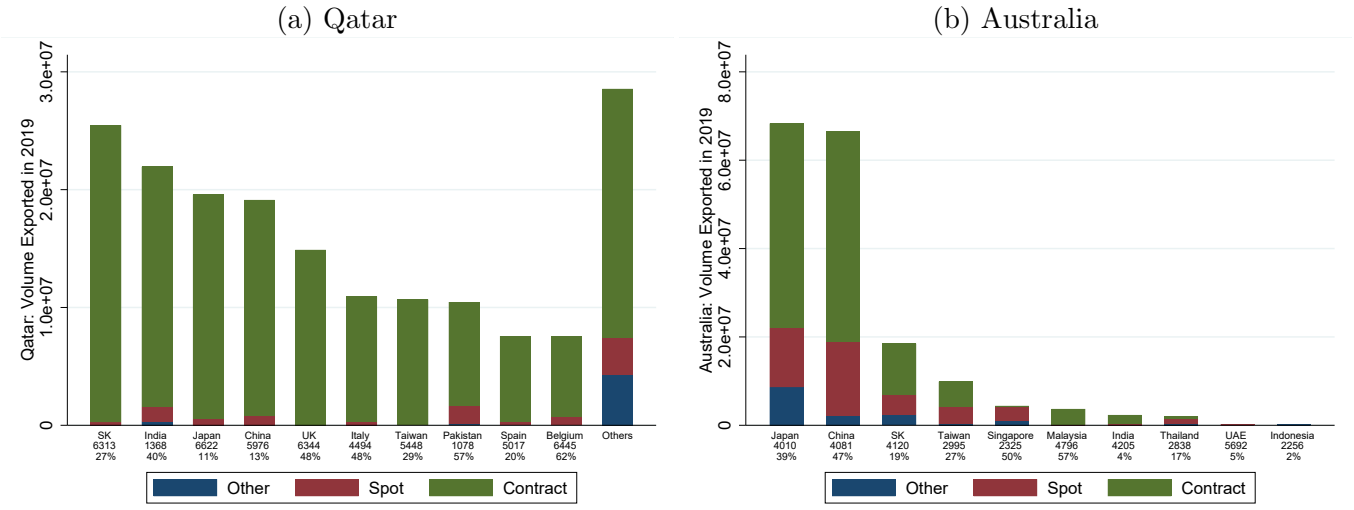
Figure A.1: Largest exporter countries in 2019



Because the LNG industry is a costly shipping technology for a commodity product, we might think that both sellers and buyers have incentives to minimize the distance travelled. The overall flows confirm this, in general. For example, Figure A.2 shows that the two largest destinations for Australia are the relatively close Japan and China, where Australia supplies 39% and 47% of these countries' LNG, respectively. Qatar supplies less of these countries' demand, at 11% and 13%. The figures below the country name show the average mileage travelled on routes between the source and destination countries. The LNG travelling to Japan and China from Australia travels much shorter distance than the LNG from Qatar.

However, there are exceptions. For example, Qatar supplies 27% of South Korea's LNG and 29% of

Figure A.2: Volume by Destination, 2019



Taiwan's despite LNG having to travel over 5,000 miles to get there. Australia is much closer to both destinations and supplies only 19% and 27%, respectively.[37] We suspect deviations from gravity relate to the staggered timing at which liquefaction and regasification capacity was installed in different countries and the agreed long term contracts between buyers and sellers.

Figure A.2 also offers some insight about how shipments are organized. The bulk of LNG exports from Qatar in 2019 were contracted quantities, meaning very little LNG from Qatar ended up in the spot market. For Australia, spot volumes were relatively more common.

Figure A.3 focuses on 2019, and shows total volumes imported by destination country.

Figure A.4 shows the volumes arriving in Japan in 2019 by source country. The figures below the country names are the average distance travelled and the share of that country's LNG exports arriving in Japan. There are some interesting comparisons. Australia sends Japan 39% of its output but is much further away than Malaysia, which sends a similar share of its total output to Japan.

B Lower bounds on buyer values for contract shipments

This appendix estimates lower bounds on the value of receiving LNG in a contract shipment rather than a spot shipment. It takes as given the thickness of the spot market and status quo spot prices. Our starting point is the analysis in Section 3, showing that buyers request contract shipments in periods when they have positive demand, which is also when spot prices tend to be high. We ask how much more valuable

[37] We have some data on production costs. Qatar has a much lower operating production cost than the installations in Australia (see supply functions in earlier update), so the delivered cost from Qatar may still be much lower than from Australia.

Figure A.3: Largest importer countries in 2019

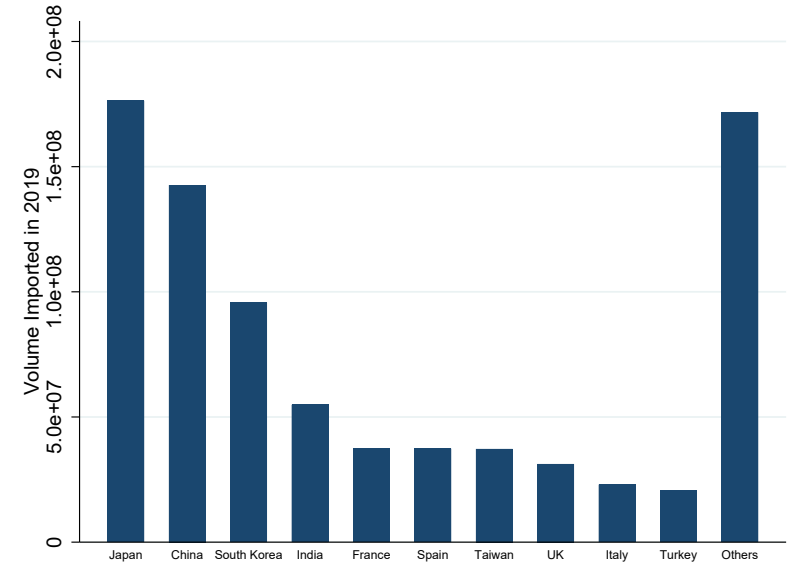
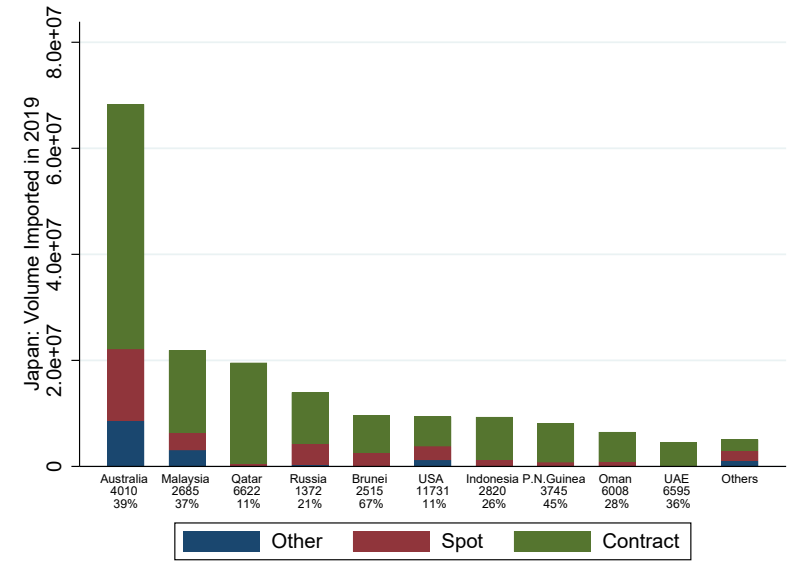


Figure A.4: Japan imported volumes by Origin, 2019



contract shipments must be to buyers in order to be consistent with contract prices that are sufficiently high to ensure both buyers and sellers find it incentive compatible to maintain the contract .

We introduce the notation that will also be used in the model in Section 4. When there is a contract in place on a given seller-buyer route, the downstream buyer, g, has a value V_{lgt} for one cubic metre of LNG arriving from seller l in a contract shipment at time t. The value g has for a spot shipment in t is V^s_{gt}, which is not l-specific.

The per cubic metre prices for LNG arriving via contract and spot are denoted p^c_{lgt} and p^s_{gt}, respectively. A seller can deliver shipments at time t to any of their contract partners or sell on any spot market, so the relevant spot price for the seller is the highest available spot price in t, net of the destination-specific transport costs. The object of interest in this section is $V^c_{lgt} - V^s_{gt}$ and the challenge is that only p^s_{gt} and bilateral transport costs, τ_{ij} are observed in the data (using Wood Mackenzie data on bilateral transport costs.)

We start by considering the buyer's surplus from a contract shipment, $V^c_{lgt} - p^c_{lgt}$, and appeal to revealed preference: whenever the buyer receives a contract shipment, it must create more value for them than would a spot shipment at the local prevailing spot price. That is, buyer incentive compatibility requires:

$$V^c_{lgt} - p^c_{lgt} \geq V^s_{gt} - p^s_{gt},$$

or, equivalently,

$$V^c_{lgt} - V^s_{gt} \geq p^c_{lgt} - p^s_{gt}. \tag{20}$$

Turning next to the seller's incentive compatibility constraint, we argue that sellers must, on average, receive at least as much as what they could earn selling the same quantities on spot markets over the contract relationship. For simplicity, in this section, we let the seller be forward-looking only over the month of a relevant contract shipment. The seller's dynamic incentive compatibility constraint implies that the contract prices they receive in expectation, net of transport costs to g, must exceed the expected highest net spot price in t in the periods t when contract shipments are delivered:

$$E(p^c_{lgt} - \tau_{ij}) \geq E(max_G(p^s_{gt} - \tau_{lg})), \tag{21}$$

where expectations are taken over all spot markets, G, over the contracted time period. We note that this is clearly a lower bound on the surplus required to align sellers' incentives in the relationship since it does not account for any variable or differential fixed costs incurred on contract shipments (other than the transport costs).

If both parties' expectations are consistent, we can combine the incentive compatibility constraints (20) and (21) to give:

$$E\left(V^c_{lgt} - V^s_{gt}\right) \geq E(p^c_{lgt}) - p^s_{gt} \geq E(max_G(p^s_{gt} - \tau_{lg})) + \tau_{lg} - p^s_{gt}. \tag{22}$$

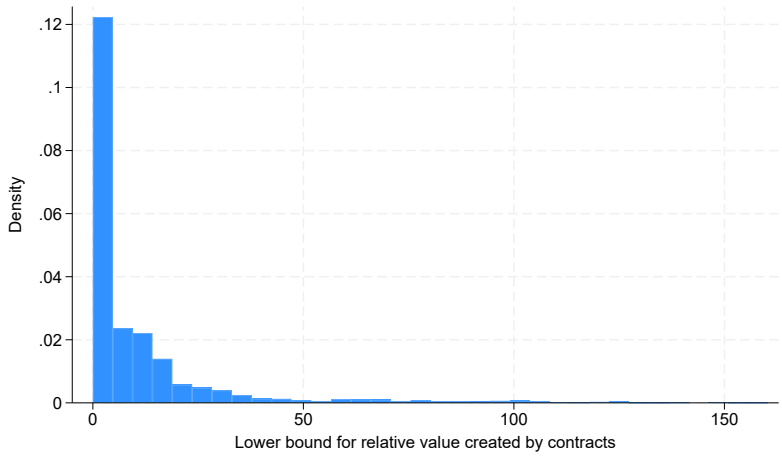
The right had side of inequality (22) can be calculated using data on spot prices in all g and bilateral transport costs τ_{lg} at all destination at times when contracts are received by buyer g.

The data yield an estimate of a lower bound on $E\left(V^c_{lgt} - V^s_{gt}\right)$ per cubic metre of LNG received under contract by destination g in month t, which is the contract period under consideration in this section. We find estimates at the contract shipment level and then find the mean for each route-year. $3,070$ route-years. Mean contract shipment value per shipment is \$10.97 per cubic meter, which is 5.9% of the mean spot price net of transport costs for the lgt observations where we have contract shipments and all the required data.

While we have made some strong assumptions to derive these lower bounds on the value contract shipments create for buyers under prevailing spot prices, the logic follows the simple intuition that the relative expected value from a contract is sufficient to compensate the seller for not switching to the spot market when the spot price is high together with the amount required to compensate the buyer for not switching to the spot market when the price is low at times of similar demand.

Figure B.5 is a histogram of the estimated lower bounds of the relative value per cubic metre of LNG arriving at the destination in contract shipments compared to available spot shipments for the $1,158$ buyer-months in this analysis.

Figure B.5: Relative value of contract shipments



A regression of route-year mean contract shipment value on 11 year fixed effects has an adjusted R-squared of 0.06. Including only fixed effects for the 82 buyers that appear in this data sample gives an adjusted R-squared of 0.36. A regression with 556 fixed effects for the interaction of buyer and year fixed effects gives an adjusted R-squared of 0.75.[38]

The estimated lower bounds can be interpreted as the value of a contract shipment over a unilateral

[38] A similar exercise grouping time periods by destination-specific demand shock and assuming the difference between contract shipment value and contract price is at least as large as the difference between the highest and lowest prices paid for spot shipments under similar demand conditions yields similar estimates of relative contract value lower bounds.

deviation to a spot shipment in the same week, conditional on the vector of global spot prices at that time and the matrix of bilateral transport cost. In order to derive estimates of the relative value, rather than lower bounds, we need a structural model that delivers estimates of LNG demand and the relationship between spot market size and buyer spot value, as well as the nature of seller effort in contractual relationships and the division of contract surplus. This is the purpose of the Section 4 where we exploit exogenous weather shocks as well as local spot price variation.

C Measuring spot market thickness, Th_{gt}

For each destination-week, the measure Th_{gt} is the product of two terms.

First, we find the share of spot shipments in the region (Europe or Asia) in that month, which is the total number of spot shipments delivered to all destinations in the region divided by the total number of spot and contract shipments delivered in the region. The higher is this share, the more likely any buyer in the region could purchase spot shipments at the time they demand it.

Second, we turn to the match of spot shipment capacities with those demanded via contracts. We find the distribution of spot vessel sizes that have been delivered in the region over the last three months. We then add to these shipments the contract shipments delivered to the destination that month to give the capacity distribution of typical region spot shipments together with the destination's contract capacity demand. We use the Kolmogorov-Smirnov statistic to tell us how similar these two distributions are. If they are very similar, we infer that the destination's demand contract capacities could be catered to by the region's spot market. The more similar the distributions the thicker the local spot market, in terms of shipment capacity availability.

In the counterfactual, by construction, both of these terms equal one. Hence, the spot market in each destination-week achieves the maximum spot market thickness that is ever observed in the actual data. The buyer values for this enlarged spot market are still discounted due to the value of ϕ^s_{gt} in V^s_{gt} being determined by the estimated λ_0 and λ_1 parameters delivered in Section 6.

D Model identification proof

We employ Nonlinear Least Squares estimation to identify the model parameters, denoted $\boldsymbol{\theta} = \{k, \alpha_0, \alpha_1, \lambda_0, \lambda_1, F_0, F_1,$ Specifically, we find the parameter values that solve the following minimization problem:

$$\min_{\boldsymbol{\theta}} \sum_{t=1}^{T} \sum_{g=1}^{G} \sum_{l=1}^{L} \left((\phi^{s*}_{lgt})^{-\lambda_{gt}} - C_{lgt} \right)^2 .$$

This appendix shows that this highly nonlinear model with numerous parameters is identified. Silvey (1970) state that a model is identifiable if no two different sets of parameter values result in the same

model. One obvious cause of non-identifiability, therefore, is parameter redundancy, when a model can be respecified in terms of a smaller set of parameters (Catchpole and Morgan (1997)).

Cole, Morgan, and Titterington (2010) provided a comprehensive framework for detecting parameter redundancy. Applied to our model, it starts with defining the vector $\boldsymbol{\kappa} = [\kappa_{111}, \kappa_{121}, ...\kappa_{lgt}.., \kappa_{LGT}]^T$ with $\kappa_{lgt} = (\phi_{lgt}^{s*})^{-\lambda_{gt}}$. If the rank of $\boldsymbol{C} \equiv \frac{\partial \boldsymbol{\kappa}}{\partial \boldsymbol{\theta}}$, denoted r, is smaller than the number of parameters q in $\boldsymbol{\theta}$, then the model is parameter redundant and will be non-identifiable. This scenario aligns with the concept of rank-deficient nonlinear least squares, which typically does not yield a unique solution (Eriksson and Gulliksson, 2004). Conversely, if $r = q$, the model is of full rank, not parameter redundant, and will be at least locally identifiable.[39].

Given our two functional forms of interest, $v(w_{gt}) = k w_{gt}^{\alpha_0 R_{gt}^{\alpha_1}}$ and $\lambda_{gt} = \lambda_0 T H_{gt}^{\lambda_1}$, with $k > 0, \alpha_0 > 0, F_0 > 0, \lambda_0 > 0, \mu > 0$, where R_{gt} is the ratio of annual demand to storage capacity, w_{gt} is the normalized weather shock, and Th_{gt} is spot market thickness, we can express κ_{lgt} as follows:

$$
\begin{aligned}
\kappa_{lgt} &= \left(\frac{v(w_{gt})}{(1-\beta) v(w_{gt})^{\frac{1}{1-\mu}} (\frac{\beta\mu}{c})^{\frac{\mu}{1-\mu}} - (1-\beta)F + E(P_{gt})} \right)^{-\lambda_{gt}} \\
&= \left(\frac{(1-\beta) v(w_{gt})^{\frac{1}{1-\mu}} (\frac{\beta\mu}{c})^{\frac{\mu}{1-\mu}} - (1-\beta)F + E(P_{gt})}{v(w_{gt})} \right)^{\lambda_{gt}} \\
&= \left((1-\beta)(\frac{\beta\mu}{c})^{\frac{\mu}{1-\mu}} (k w_{gt}^{\alpha_0 R_{gt}^{\alpha_1}})^{\frac{\mu}{1-\mu}} - (1-\beta)F \frac{1}{k} w_{gt}^{-\alpha_0 R_{gt}^{\alpha_1}} + E(P_{gt}) \frac{1}{k} w_{gt}^{-\alpha_0 R_{gt}^{\alpha_1}} \right)^{\lambda_0 T h_{gt}^{\lambda_1}},
\end{aligned}
$$

which can be rewritten as:

$$
\kappa_{lgt} = \left(G(k w_{gt}^{\alpha_0 R_{gt}^{\alpha_1}})^{\frac{\mu}{1-\mu}} - \tilde{F}\tilde{k} w_{gt}^{-\alpha_0 R_{gt}^{\alpha_1}} + E(P_{gt}) \tilde{k} w_{gt}^{-\alpha_0 R_{gt}^{\alpha_1}} \right)^{\lambda_0 T h_{gt}^{\lambda_1}},
$$

where

$$
\begin{aligned}
\tilde{k} &= \frac{1}{k}, \\
\tilde{F} &= (1-\beta)F, \\
G &\equiv (1-\beta) \left(\frac{\beta\mu}{c} \right)^{\frac{\mu}{1-\mu}}.
\end{aligned}
$$

Because an increase in the marginal seller effort cost c can be offset by a decrease in the seller share of surplus β, c cannot be separately identified. Therefore, the parameters that need to be checked for identifiability are $\theta = [\tilde{k}, \alpha_0, \alpha_1, \tilde{F}, G, \mu, \lambda_0, \lambda_1]$.

[39] See Cole, Morgan, and Titterington (2010), Theorem 2a

The first derivative matrix for κ_{lgt}, $\boldsymbol{C} \equiv \frac{\partial \boldsymbol{\kappa}}{\partial \boldsymbol{\theta}}$, is given by

$$\boldsymbol{C} = \frac{\partial \boldsymbol{\kappa}}{\partial \boldsymbol{\theta}} = \begin{bmatrix} \frac{\partial \kappa_{111}}{\partial \theta_1} & \frac{\partial \kappa_{121}}{\partial \theta_1} & \dots & \frac{\partial \kappa_{LGT}}{\partial \theta_1} \\ \frac{\partial \kappa_{111}}{\partial \theta_2} & \frac{\partial \kappa_{121}}{\partial \theta_2} & \dots & \frac{\partial \kappa_{LGT}}{\partial \theta_2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \kappa_{111}}{\partial \theta_9} & \frac{\partial \kappa_2}{\partial \theta_9} & \dots & \frac{\partial \kappa_{LGT}}{\partial \theta_9} \end{bmatrix}$$

If the model is not parameter redundant, then the rank of the matrix $\boldsymbol{D}$ should be equal to the number of parameters, in which case the matrix has full row rank. To demonstrate this, we derive the first derivatives contained within the matrix:

$$\frac{\partial \kappa_{lgt}}{\partial \tilde{k}} = \lambda_{gt} \kappa_{lgt}^{\frac{\lambda_{gt}-1}{\lambda_{gt}}} \left(-G \frac{\mu}{\mu - 1} (k w_{gt}^{-\alpha_0 R_{gt}^{\alpha_1}})^{\frac{1}{\mu-1}} - \tilde{F} + E(P_{gt}) \right) w_{gt}^{-\alpha_0 R_{gt}^{\alpha_1}}$$

$$\frac{\partial \kappa_{lgt}}{\partial \alpha_0} = -\lambda_{gt} \kappa_{lgt}^{\frac{\lambda_{gt}-1}{\lambda_{gt}}} (E(P_{gt}) - \tilde{F}_0 G S_{lt}^{F_1} \tilde{k} w_{gt}^{-\alpha_0 R_{gt}^{\alpha_1}} + G \frac{\mu}{\mu - 1} (\tilde{k} w_{gt}^{-\alpha_0 R_{gt}^{\alpha_1}})) \tilde{k} w_{gt}^{-\alpha_0 R_{gt}^{\alpha_1}} R_{gt}^{\alpha_1} \ln w_{gt}$$

$$\frac{\partial \kappa_{lgt}}{\partial \alpha_1} = -\lambda_{gt} \kappa_{lgt}^{\frac{\lambda_{gt}-1}{\lambda_{gt}}} (E(P_{gt}) - \tilde{F}_0 G S_{lt}^{F_1} \tilde{k} w_{gt}^{-\alpha_0 R_{gt}^{\alpha_1}} + G \frac{\mu}{\mu - 1} (\tilde{k} w_{gt}^{-\alpha_0 R_{gt}^{\alpha_1}})) \tilde{k} w_{gt}^{-\alpha_0 R_{gt}^{\alpha_1}} \alpha_0 R_{gt}^{\alpha_1} \ln w_{gt} \ln R_{gt}$$

$$\frac{\partial \kappa_{lgt}}{\partial F_0} = \lambda_{gt} \kappa_{lgt}^{\frac{\lambda_{gt}-1}{\lambda_{gt}}} G S_{lt}^{F_1} \tilde{k} w_{gt}^{-\alpha_0 R_{gt}^{\alpha_1}}$$

$$\frac{\partial \kappa_{lgt}}{\partial F_1} = \lambda_{gt} \kappa_{lgt}^{\frac{\lambda_{gt}-1}{\lambda_{gt}}} \tilde{F}_0 G S_{lt}^{F_1} \tilde{k} w_{gt}^{-\alpha_0 R_{gt}^{\alpha_1}} \ln G S_{lt}$$

$$\frac{\partial \kappa_{lgt}}{\partial G} = \lambda_{gt} \kappa_{lgt}^{\frac{\lambda_{gt}-1}{\lambda_{gt}}} (\tilde{k} w_{gt}^{-\alpha_0 R_{gt}^{\alpha_1}})^{\frac{\mu}{\mu-1}}$$

$$\frac{\partial \kappa_{lgt}}{\partial \mu} = -\lambda_{gt} \kappa_{lgt}^{\frac{\lambda_{gt}-1}{\lambda_{gt}}} (G (\tilde{k} w_{gt}^{-\alpha_0 R_{gt}^{\alpha_1}})^{\frac{\mu}{\mu-1}} \frac{\ln(\tilde{k} w_{gt}^{-\alpha_0 R_{gt}^{\alpha_1}})}{(\mu - 1)^2})$$

$$\frac{\partial \kappa_{lgt}}{\partial \lambda_0} = \kappa_{lgt} \ln \kappa_{lgt}^{\frac{1}{\lambda_{gt}}} T h_{gt}^{\lambda_1}$$

$$\frac{\partial \kappa_{lgt}}{\partial \lambda_1} = \kappa_{lgt} \ln \kappa_{lgt}^{\frac{1}{\lambda_{gt}}} \lambda_0 T h_{gt}^{\lambda_1} \ln T h_{gt}$$

The matrix $\boldsymbol{C}$ is of full row rank if the only solution to the following equation, satisfied simultaneously for all lgt-level data, is $\beta^i = 0, \forall i \in \boldsymbol{\theta}$.

$$\beta^{\tilde{k}} \frac{\partial \kappa_{lgt}}{\partial \tilde{k}} + \beta^{\alpha_0} \frac{\partial \kappa_{lgt}}{\partial \alpha_0} + \beta^{\alpha_1} \frac{\partial \kappa_{lgt}}{\partial \alpha_1} + \beta^{F} \frac{\partial \kappa_{lgt}}{\partial F} + \beta^{G} \frac{\partial \kappa_{lgt}}{\partial G} + \beta^{\mu} \frac{\partial \kappa_{lgt}}{\partial \mu} + \beta^{\lambda_0} \frac{\partial \kappa_{lgt}}{\partial \lambda_0} + \beta^{\lambda_1} \frac{\partial \kappa_{lgt}}{\partial \lambda_1} = 0 \qquad (23)$$

where β^i represents the coefficients associated with each parameter i in the parameter set $\boldsymbol{\theta}$. This condition requires that each parameter i in $\boldsymbol{\theta}$ does not contribute linearly to the model outcomes, which ensures that $\boldsymbol{C}$ is of full row rank by demonstrating the linear independence of its rows.

Upon substituting the first derivatives into the matrix $\boldsymbol{D}$ and collecting terms, we obtain the following expression:

$$\begin{aligned}
0 =&(\beta^{\tilde{k}} - \beta^{\alpha_0}\tilde{k}R_{gt}^{\alpha_1}\ln w_{gt} - \beta^{\alpha_1}\tilde{k}\alpha_0 R_{gt}^{\alpha_1}\ln w_{gt}\ln R_{gt})H_{lgt}^v \\
&+(\beta^{F_0} + \beta^{F_1}\tilde{F}_0\ln GS_{lt})H_{lgt}^F \\
&+(\beta^C - \beta^\mu(G\frac{\ln(\tilde{k}w_{gt}^{-\alpha_0 R_{gt}^{\alpha_1}})}{(\mu-1)^2}))H_{lgt}^G \\
&+(\beta^{\lambda_0}\frac{1}{\lambda_0} + \beta^{\lambda_1}\ln Th_{gt})H_{lgt}^\lambda
\end{aligned} \tag{24}$$

where $H_{lgt}^v \equiv (E(P_{gt}) - \tilde{F}_0 GS_{lt}^{F_1}\tilde{k}w_{gt}^{-\alpha_0 R_{gt}^{\alpha_1}} + G\frac{\mu}{\mu-1}(\tilde{k}w_{gt}^{-\alpha_0 R_{gt}^{\alpha_1}})^{\frac{1}{\mu-1}})w_{gt}^{-\alpha_0 R_{gt}^{\alpha_1}}$, $H_{lgt}^F \equiv GS_{lt}^{F_1}\tilde{k}w_{gt}^{-\alpha_0 R_{gt}^{\alpha_1}}$, $H_{lgt}^G \equiv (\tilde{k}w_{gt}^{\alpha_0 R_{gt}^{\alpha_1}})^{\frac{\mu}{\mu-1}}$, and $H_{lgt}^\lambda \equiv \kappa_{lgt}^{\frac{1}{\lambda_{gt}}}\ln\kappa_{lgt}^{\frac{1}{\lambda_{gt}}}$.

Parameter redundancy in a model can be categorized into two types: intrinsic parameter redundancy, which is inherent to the model structure regardless of the data, and extrinsic parameter redundancy, which arises due to the characteristics of the dataset used. For instance, consider a scenario where $w_{gt} = 1\,\forall gt$. In such a case, the parameters β^{α_0} and β^{α_1} can assume any values, while all other parameters $\beta^i = 0\,\forall i \neq \alpha_0, \alpha_1$, thus satisfying equation (23) and leading to parameter redundancy induced by this specific dataset. This prevents the identification of β^{α_0} and β^{α_1}.

Given a sufficiently large and varied dataset, our analysis is concentrated on intrinsic parameter redundancy, under the assumption of an ideal dataset. This entails that the functions $H_{lgt}^v, H_{lgt}^F, H_{lgt}^G, H_{lgt}^\lambda$ are linearly independent, given our parameter assumptions that $k > 0, \alpha_0 > 0, F_0 > 0, \lambda_0 > 0, \mu > 0$. Consequently, it must follow that each coefficient in front of $H^i, i \in \{v, F, G, \lambda\}$, is equal to 0:

$$\begin{aligned}
\beta^{\tilde{k}} - \beta^{\alpha_0}\tilde{k}R_{gt}^{\alpha_1}\ln w_{gt} - \beta^{\alpha_1}\tilde{k}\alpha_0 R_{gt}^{\alpha_1}\ln w_{gt}\ln R_{gt} &= 0 \\
\beta^{F_0} + \beta^{F_1}\tilde{F}_0\ln GS_{lt} &= 0 \\
\beta^G - \beta^\mu(G\frac{\ln(\tilde{k}w_{gt}^{-\alpha_0 R_{gt}^{\alpha_1}})}{(\mu-1)^2}) &= 0 \\
\beta^{\lambda_0}\frac{1}{\lambda_0} + \beta^{\lambda_1}\ln Th_{gt} &= 0
\end{aligned} \tag{25}$$

These four conditions must hold for any lgt, within a perfect dataset. Consequently, the solution to these equations is $\beta^i = 0\,\forall i \in \boldsymbol{\theta}$. This establishes the proof for the identifiability of our nonlinear least squares model.

Equations (24) and (25) also explain the variations utilized to separately identify each parameter. From equation (24), variations in $H_{lgt}^v, H_{lgt}^F, H_{lgt}^G$, and H_{lgt}^λ enable the separate identification of parameters associated with the value function, fixed costs, returns to scale, and market thickness, respectively. Similarly,

equation (25) indicates that variation in $R_{gt}^{\alpha_1} \ln w_{gt}$ allows for the separate identification of α_0 from $\tilde{k}$, and variation in $\ln R_{gt}$ allow for the identification of α_1 from α_0.

The same intuition also applies to the other parameters.

E Estimating simple versions of model to demonstrate identification

The full model is that in each period t, g calls in a contract shipment i iff $(1-\beta_u)S_{lgt} \geq E(U_{lt}^s)$, that is, if $(1-\beta_u)S_{lgt} - E(U_{gt}^s) \geq 0$. S_{udtc} is the surplus from a contract shipment:

$$S_{lgt} = v(w_{gt})^{\frac{1}{1-\mu}} (\frac{\beta\mu}{c})^{\frac{\mu}{1-\mu}} - F_{lt}$$

$$E(U_{gt}^s) = \frac{v(w_{gt})}{\phi_{gt}^s} - E(P_{gt})$$

Therefore, $(1-\beta_u)S_{lgt} \geq E(U_{lt}^s)$ yields

$$(1-\beta_u)S_{lgt} - \frac{v(w_{gt})}{\phi_{gt}^s} + E(P_{gt}) \geq 0$$

$$\phi_{lgt}^s \geq \frac{v(w_{gt})}{(1\beta_u)S_{lgt} + E(P_{gt})} \equiv \phi_{lgt}^{s*}$$

$$\phi_{lgt}^s \geq \frac{v(w_{gt})}{(1-\beta_u)S_{lgt} + E(P_{gt})} \equiv \phi_{lgt}^{s*}$$

The probability that g requests a contract shipment from l is the probability give in expression (12):

$$\Pr(\phi_{lgt}^s \geq \phi_{lgt}^{s*}) = (\phi_{lgt}^{s*})^{-\lambda}$$

We estimate equation (16) in the main text to give the two non-targeted moments relating to this probability. The two moments are the weather shock and spot price elasticities in the data. Equation (16) is:

$$C_{ldt} = \alpha + \theta_w w_{dt} + \theta_p p_{dt} + \epsilon_{ldt}$$

The estimated coefficients are (see Table 7):

$$\theta_w = 0.043$$

$$\theta_p = 0.056$$

$$\alpha = 0.351$$

We estimate different model specifications, successively adding parameters and evaluate how well each model specification fits the non-targeted moments.[40]

[40]Throughout this Appendix, $Thgt = D10m^3. * spot_{ctm}/(contract_{ctm} + spot_{ctm})$. The estimates will be redone to ensure consistency with the model specifications estimated in the main text and in Appendix F

E.1 Model 1, with only k and λ_0

This gives the on-targeted moments of $\theta_p = 0.015, \alpha = 0.494$. Equation (16) simplifies to:

$$\begin{aligned} C &= (\frac{E(P_{gt})}{v(w_{gt})})^{\lambda} \\ &= (\frac{E(P_{gt})}{k})^{\lambda} \end{aligned}$$

The estimated model parameters are:

$$\lambda = 0.0141$$

$$k = 8.19e + 20$$

Note that this model specification gives an unreasonably large estimate of k. The λ estimate shows that the spot value responds positively to spot market thickness.

E.2 Model 2, adding α_0 to k and λ_0

This gives the non-targeted moments $\theta_w = 8.457$, $\theta_p = 0.435, \alpha = -8.203$. Equation (16) simplifies to:

$$\begin{aligned} C &= (\frac{E(P_{gt})}{v(w_{gt})})^{\lambda} \\ &= (\frac{E(P_{gt})}{kw_{gt}^{\alpha_0}})^{\lambda} \end{aligned}$$

The estimated model parameters are:

$$\lambda = 9798.877$$

$$k = 7330.472$$

$$\alpha_0 = -9999$$

These parameter value are all large in (absolute) value.

E.3 Model 3, setting $\beta = 0.3$, with k, λ_0, α_0, and also adding μ and c, and then also F_0.

Adding μ and c gives the non-targeted moments of $\theta_w = 0.049$, $\theta_p = 0.001, \alpha = 0.467$. Equation (16) simplifies to:

$$\begin{aligned} C &= (\frac{(1-\beta_u)S_{lgt} + E(P_{gt})}{v(w_{gt})})^{\lambda} \\ &= ((1-\beta_u)(kw_{gt}^{\alpha_0})^{\frac{\mu}{1-\mu}}(\frac{\beta\mu}{c})^{\frac{\mu}{1-\mu}} + \frac{E(P_{gt})}{kw_{gt}^{\alpha_0}})^{\lambda} \end{aligned}$$

The estimated model parameters are:

$$\lambda = 2.009$$

$$k = 7355.060$$

$$\alpha_0 = 1.152$$

$$\mu = 0.032$$

$$c = 31.410$$

Now also including F_0 gives the non-targeted moments of $\theta_w = 0.049$, $\theta_p = 0.000, \alpha = 0.467$ Equation (16) simplifies to:

$$\begin{aligned} C &= (\frac{(1-\beta_u)S_{lgt} + E(P_{gt})}{v(w_{gt})})^{\lambda} \\ &= ((1-\beta_u)(kw_{gt}^{\alpha_0})^{\frac{\mu}{1-\mu}}(\frac{\beta\mu}{c})^{\frac{\mu}{1-\mu}} - \frac{(1-\beta_u)F_0}{kw_{gt}^{\alpha_0}} + \frac{E(P_{gt})}{kw_{gt}^{\alpha_0}})^{\lambda} \end{aligned}$$

The estimated model parameters are:

$$\lambda = 1.992$$

$$k = 881878.753$$

$$\alpha_0 = 4.737$$

$$\mu = 0.008$$

$$c = 93.175$$

$$F_0 = 29.739$$

E.4 Model 4, with $\beta = 0.3$, $k, \lambda_0, \alpha_0, \mu, c, F_0$, and D

This gives the non-targeted moments $\theta_w = 0.044$, $\theta_p = 0.000, \alpha = 0.47033$. That is, it still gives a spot price elasticity that is almost zero. Equation (16) simplifies to:

$$\begin{aligned} C &= (D\frac{(1-\beta_u)v(w_{gt})^{\frac{1}{1-\mu}}(\frac{\beta\mu}{c})^{\frac{\mu}{1-\mu}} - (1-\beta_u)F_0 + E(P_{gt})}{v(w_{gt})})^{\lambda} \\ &= D^{\lambda}((1-\beta_u)v(w_{gt})^{\frac{\mu}{1-\mu}}(\frac{\beta\mu}{c})^{\frac{\mu}{1-\mu}} - \frac{(1-\beta_u)F_0}{v(w_{gt})} + \frac{E(P_{gt})}{kw_{gt}^{\alpha_0}})^{\lambda} \end{aligned}$$

The estimated model parameters are:

$$\lambda = 0.380$$

$$k = 263.555$$

$$\alpha_0 = 0.158$$

$$\mu = 0.487$$

$$c = 0.207$$

$$F_0 = 67.030$$

$$D = 0.002$$

Note that including the constant term D, which can be interpreted as the share of contract shipments when w_{lgt} and p_{gt} are both equal to zero, gives a much more reasonable estimate of k.

E.5 Model 5, as in Model 4, and further adding α_1 and F_1.

This is the full model, with the exception of setting $\beta = 0.3$. It gives the non-targeted moments $\theta_w = 0.044$ $\theta_p = 0.223, \alpha = 0.343$. Note that this specification gives a predicted price elasticity that is far higher than in the data. Equation (16) is:

$$C = (D\frac{(1-\beta_u)v(w_{gt})^{\frac{1}{1-\mu}}(\frac{\beta\mu}{c})^{\frac{\mu}{1-\mu}} - (1-\beta_u)F_0 CS_{lt}^{F_1} + E(P_{gt})}{v(w_{gt})})^{\lambda_0 Th_{gt}^{\lambda_1}}$$
$$= D^{\lambda}((1-\beta_u)v(w_{gt})^{\frac{\mu}{1-\mu}}(\frac{\beta\mu}{c})^{\frac{\mu}{1-\mu}} - \frac{(1-\beta_u)F_0 CS_{lt}^{F_1}}{v(w_{gt})} + \frac{E(P_{gt})}{kw_{gt}^{\alpha_0 R_{gt}^{\alpha_1}}})^{\lambda_0 Th_{gt}^{\lambda_1}}$$

The estimated model parameters are:

$$\lambda = 8.440$$

$$k = 15.558$$

$$D = 0.212$$

$$\alpha_0 = 0.007$$

$$\mu = 0.294$$

$$c = 0.016$$

$$F_0 = 96.14$$

$$\alpha_1 = 0.094$$

$$\lambda_1 = 0.929$$

$$F_1 = 0.301$$

Note that the main results in Section 6 flexibly estimated β, but leave out F_1 and α_1. We show the results of adding F_1 and α_1 separately in Appendix F.[41]

F Results from alternative specifications

F.1 Including F_1

In this version of the model, we include the parameter F_1, the contract fixed cost elasticity, in addition to the parameters discussed in Section 6. This parameter allows the seller's fixed cost of a contract shipment to vary with the number of contract shipments that a seller fulfils in week t relative to its usual number. This variable is N_{lt}, and is described in Subsection 5.1.

[41] We have not yet computed the full model including both F_1 and α_1.

Parameter	Description	Estimate	Confidence Interval
D	Constant	0.02	[X, X]
k	Value function normalization	15.13	[X, X]
α_0	Value function elasticity	0.15	[X, X]
α_1	Value function elasticity storage term	0.00	
μ	Effort intensity	0.24	[X, X]
c	Marginal effort cost	0.79	[X, X]
F_0	Contract fixed cost normalization	15.69	[X, X]
F_1	Contract fixed cost elasticity	0.54	[X, X]
λ_0	Spot market thickness normalization	1.37	[X, X]
λ_1	Spot market thickness elasticity	0.97	[X, X]

The estimated β in this specification is 0.20. While this specification matches the non-targeted moment of the contract probability elasticity with respect to the weather shock, predicting γ_w of 0.04, which is the same as in the data. It predicts $\gamma_p = 0.15$, whereas in the data, it is 0.05. In this specification, the counterfactual suggests a welfare loss of 11.98% of industry value if all shipments were transacted on the spot market. (CONFIDENCE INTERVALS ARE UNDER CONSTRUCTION)

F.2 Including α_1

The third set of results are for the model specification that includes α_1 in addition to the parameters discussed in Section 6. The α_1 parameter describes how the value function elasticity to weather shocks varies with the storage capacity installed in destination country g.

Parameter	Description	Estimate	Confidence Interval
D	Constant	0.15	[X, X]
k	Value function normalization	9.81	[X, X]
α_0	Value function elasticity	0.16	[X, X]
α_1	Value function elasticity storage term	0.02	
μ	Effort intensity	0.34	[X, X]
c	Marginal effort cost	1.74	[X, X]
F_0	Contract fixed cost normalization	4.99	[X, X]
F_1	Contract fixed cost elasticity	0.00	[X, X]
λ_0	Spot market thickness normalization	0.26	[X, X]
λ_1	Spot market thickness elasticity	2.75	[X, X]

The estimated β in this specification is 0.69. This specification does not match either non-targeted

moment well. It predicts γ_w of 0.004, compared to the value in the data of 0.04. It predicts $\gamma_p = 0.01$, whereas in the data, it is 0.05. In this specification, the counterfactual suggests a welfare loss of 18.45% of industry value if all shipments were transacted on the spot market. (CONFIDENCE INTERVALS ARE UNDER CONSTRUCTION)

References

Brancaccio, Giulia, Myrto Kalouptsidi, and Theodore Papageorgiou. 2020. "Geography, transportation, and endogenous trade costs." *Econometrica* 88 (2):657–691.

Catchpole, Edward A and Byron JT Morgan. 1997. "Detecting parameter redundancy." *Biometrika* 84 (1):187–196.

Cole, Diana J, Byron JT Morgan, and DM Titterington. 2010. "Determining the parametric structure of models." *Mathematical biosciences* 228 (1):16–30.

Costinot, Arnaud, Lindsay Oldenski, and James Rauch. 2011. "Adaptation and the boundary of multinational firms." *The Review of Economics and Statistics* 93 (1):298–308.

Eriksson, Jerry and Mårten Gulliksson. 2004. "Local results for the Gauss-Newton method on constrained rank-deficient nonlinear least squares." *Mathematics of Computation* 73 (248):1865–1883.

Forbes, Silke J and Mara Lederman. 2010. "Does vertical integration affect firm performance? Evidence from the airline industry." *The RAND Journal of Economics* 41 (4):765–790.

Forbes, Silke Januszewski and Mara Lederman. 2009. "Adaptation and vertical integration in the airline industry." *American Economic Review* 99 (5):1831–1849.

Grossman, Gene M, Elhanan Helpman, and Hugo Lhuillier. 2023. "Supply Chain Resilience: Should Policy Promote International Diversification or Reshoring?" *Journal of Political Economy* forthcoming.

Grossman, Gene M, Elhanan Helpman, and Alejandro Sabal. 2023. "Resilience in Vertical Supply Chains." Tech. rep., National Bureau of Economic Research.

Harris, Adam and Thi Mai Anh Nguyen. 2022. "Long-term Relationships and the Spot Market: Evidence from US Trucking." .

Joskow, Paul L. 1988. "Asset specificity and the structure of vertical relationships: empirical evidence." *The Journal of Law, Economics, and Organization* 4 (1):95–117.

Klein, Benjamin, Robert G Crawford, and Armen A Alchian. 1978. "Vertical integration, appropriable rents, and the competitive contracting process." *The journal of Law and Economics* 21 (2):297–326.

Kranton, Rachel E. 1996. "Reciprocal exchange: a self-sustaining system." *The American Economic Review* :830–851.

Lafontaine, Francine and Margaret Slade. 2012. "Inter-firm contracts: Evidence." *Handbook of organizational economics* :958–1013.

Macchiavello, Rocco. 2022. "Relational contracts and development." *Annual Review of Economics* 14:337–362.

Macchiavello, Rocco and Ameet Morjaria. 2015. "The value of relationships: evidence from a supply shock to Kenyan rose exports." *American Economic Review* 105 (9):2911–2945.

Masten, Scott E and Keith J Crocker. 1985. "Efficient adaptation in long-term contracts: Take-or-pay provisions for natural gas." *The American Economic Review* 75 (5):1083–1093.

Nicholson, Sharon E. 1993. "An overview of African rainfall fluctuations of the last decade." *Journal of climate* :1463–1466.

Pirrong, Stephen Craig. 1993. "Contracting practices in bulk shipping markets: a transactions cost explanation." *The Journal of Law and Economics* 36 (2):937–976.

Silvey, S. 1970. "Statistical inference Penguin Books." *Middlesex, England* .

S&P. 2022. "Gunvor backs out of LNG deliveries to Pakistan for April-June deliveries." https://www.spglobal.com/commodityinsights/en/market-insights/latest-news/lng/032822-gunvor-backs-out-of-lng-deliveries-to-pakistan-from-april-june-officials. Accessed: 2023-09-12.

———. 2023. "Demand for term LNG contracts firms amid supply security concerns." https://www.spglobal.com/commodityinsights/en/market-insights/latest-news/lng/071223-demand-for-term-lng-contracts-firms-amid-supply-security-concerns. Accessed: 2023-09-12.

Tolvanen, Juha, Olivier Darmouni, and Simon Essig Aberg. 2021. "Pulp Friction: The Value of Quantity Contracts in Decentralized Markets." *Available at SSRN 4381156* .

Williamson, Oliver E. 1979. "Transaction-cost economics: the governance of contractual relations." *The journal of Law and Economics* 22 (2):233–261.

Zahur, Nahim Bin. 2022. "Long-term contracts and efficiency in the liquefied natural gas industry." *Available at SSRN 4222408* .