Learning and Investment under Demand Uncertainty in Container Shipping

Jihye Jeon†

February, 2018

Abstract

This paper investigates how firms invest under demand uncertainty focusing on the role of information in container shipping. I develop and estimate a dynamic oligopoly model that allows uncertainty about the demand process in addition to uncertainty about demand realizations: agents do not know the true parameters in the demand process, but form and revise their expectations based on information available at each decision-making moment. I find that the uncertainty about the demand process amplifies investment cycles through (i) leading firms to update beliefs more often and drastically as they experience demand volatility, and (ii) intensifying strategic incentives among firms.

KEYWORDS: Demand uncertainty, learning, dynamic games, investment, shipping

*I am indebted to my advisors John Asker, Luis Cabral, Robin Lee, and Ariel Pakes for their guidance and support. I am thankful to Victor Aguirregabiria, David Backus, Francesco Decarolis, Michael Dickstein, Myrto Kalouptsidi, Kei Kawai, Peter Newberry, Marc Rysman, and Laura Veldkamp as well as participants at the NYU Stern Applied Micro Seminar, the Harvard IO Workshop, IIOC 2016, and EARIE 2016 for helpful comments and suggestions. Financial support from the Center for Global Economy and Business is gratefully acknowledged. All errors are my own.

†Department of Economics, Boston University; jjeon@bu.edu.
1 Introduction

In many capital-incentive industries such as the oil, shipping, and chemical industries, firms invest in long-lived capital while facing highly volatile demand conditions. Thus, firms’ expectations about demand often play an important role. When world trade was booming in the mid-2000’s, container shipping companies ordered a large volume of new ships. Due to time-to-build a lot of these ships were delivered during the times of weak trade demand following the 2008 financial crisis.\(^1\) As a result, firms faced an oversupply of ships, and in turn fierce price competition and low profitability.

Many industry experts attribute industry excess capacity to the firms’ inability to forecast demand correctly.

The container-shipping industry has been highly unprofitable over the past five years. ... Some of the pain is self-inflicted: as in past cycles, the industry extrapolated the good times and foresaw an unsustainable rise in demand (Mckinsey Insights, 2014).\(^2\)

The problem is not limited to the 2008 crisis, as suggested by the CEO of one of the largest shipping companies:

It’s pretty clear that when we look back to the early part of 2011 when these ships were ordered, ours and everybody else’s view on growth was somewhat different than what it turned out to be (The Wall Street Journal, 2013).\(^3\)

This paper examines the role of information in investment cycles and overcapacity in the context of the container shipping industry where market power and strategic considerations are present. In order to capture agents’ information under a highly volatile environment, I develop a dynamic oligopoly model of firm investment that allows uncertainty about the demand process in addition to uncertainty about demand realizations. In this model, agents do not know the parameters governing the evolution of demand, but form and revise their expectations based on information available at each decision-making moment. My estimation strategy involves employing commonly-unavailable data on investment costs and scrap values to pin down firms’ learning process. I consider alternative models of agent information sets in order to understand implications of the uncertainty about the demand process. Counterfactual experiments shed light on the mechanisms through which learning affects firm investment behavior. Through the first set of counterfactuals regarding competition, I examine how strategic incentives interact with learning and agent beliefs in impacting

\(^1\)Shipping firms face a lag between the order and the delivery of a new ship. This lag is often called time-to-build and ranges from 2 to 4 years in this industry.


firm investment. Through the second set of counterfactuals with respect to demand volatility, I examine the informational channel through which demand volatility affects investment cycles.

This paper has three main contributions. First, it provides a dynamic oligopoly framework that incorporates agents’ changing beliefs and information sets about the aggregate demand process in a parsimonious and computationally tractable way. This framework is used to show that learning amplifies investment cycles through two channels and can help explain firm behavior in a setting subject to potential structural changes. Second, this paper sheds light on how the interaction between firms’ information and strategic incentives can lead to industry overcapacity and amplify boom-and-bust cycles of investment. Lastly, the paper shows that the modeling choice of firms’ expectations has policy implications. In particular, I show that policy based on a full-information model is more likely to block welfare-enhancing mergers.

Under the standard full-information assumption, firms may be uncertain about demand realizations due to the variance in the process, but know the true stochastic process of demand. Using this assumption is appropriate in many of the settings that we study where the underlying demand process can be easily observed and estimated by both agents in the model and the researcher. However, if this is not the case, for example, because the environment is highly volatile with potential regime shifts, the learning dynamics may arise such that agents put more weights on more recent observations (Durlauf et al. (2008)). Intuitively, as older information becomes less useful in predicting future demand after a structural change, a model in which an agent discount older observations can approximate agent beliefs in this setting. An alternative way to model agent beliefs in this setting would be to add richness to the specification of the full-information model to capture the complicated true data-generating-process. However, this approach increases the number of parameters in the model, potentially posing an identification issue and increasing the computational burden of solving the model. A learning model can instead provide a parsimonious and tractable way of incorporating agents’ changing beliefs in this setting.

One of the challenges in selecting appropriate informational assumptions and estimating a learning model is that the researcher does not directly observe agents’ beliefs. And as Manski (1993) points out, it is hard to identify information and model parameters simultaneously. This paper’s approach is to use data to identify the model of firm beliefs that can rationalize observed data patterns. Empirical papers on industry dynamics typically focus on recovering objects such as investment costs, entry costs, and exit values that can rationalize observed firm behavior while imposing a full-information structure (e.g. Collard-Wexler (2013)). For the container shipping industry, however, detailed data are available on ship sales prices and demolition prices. Hence, I rely on these price data to estimate investment costs and scrap values and instead focus on recovering the model of firm beliefs. The underlying logic of this approach is similar to that of Hortacsu &

---

4Note that this is different from perfect foresight where firms know future demand realizations exactly.
Puller (2008) that use marginal cost data to quantify how much firms’ actual bidding deviates from the optimal bidding predicted by their theoretical benchmark for the Texas electricity spot market. This paper similarly compares firms’ optimal investment behavior given the investment costs and scrap values observed in the data across different information structures and further estimates the model of firm beliefs.

Incorporating learning intensifies the computational burden of solving a dynamic model with many firms. Firms’ beliefs change over time, which means that equilibrium needs to be solved separately for each period in time. This paper addresses this challenge by adopting an equilibrium concept in which firms keep track of some summary statistics of rivals’ states instead of rivals’ detailed states based on the moment-based Markov equilibrium (MME) notion proposed by Ifrach & Weintraub (2016) and the experience-based equilibrium (EBE) notion by Fershtman & Pakes (2012). This approach vastly reduces the state space size, while still capturing strategic interaction among firms.

The estimates suggest that agents place 45% weight on a 10-year-old observation relative to the most current one. Endowing agents with the knowledge about demand parameters decreases total investment by 17% and the volatility of investment lower by 22%. More importantly, it reallocates investment across time leading firms to withhold investment during demand boom years and suffer less from overcapacity when faced with downturns in demand. This has a large impact on welfare: it increases producer surplus by 85%, and decreases consumer surplus by 3%.

The counterfactual experiments shed light on the mechanism through which the uncertainty about the demand process affects firm investment behavior. The first set of counterfactuals pertains to competition and industry consolidation. Its goal is to highlight how firms’ strategic incentives interact with agent beliefs and learning. Many theoretical studies have shown that strategic incentives such as business stealing and preemption can lead to overinvestment. The industry is moving towards consolidation as well. I perform a counterfactual experiment whereby the industry becomes monopolized. The total investment over the period of 2006 to 2014 drops by 34%, and the volatility of aggregate investment decreases by 22%. Producer surplus increases by $92 billion both from reduced shipbuilding costs and from higher prices, whereas consumer surplus drops by

---

5MME can be viewed as a special case of EBE. One interpretation of these two equilibrium concepts is that firms may have limited capacity to monitor or strategize over the relevant information of all rival firms, which justifies limiting agents’ information sets. An alternative interpretation of MME is that it is an approximation to Markov-perfect equilibrium (MPE).

6This estimate is very close to the estimates in the previous studies that estimate a constant-gain learning model based on aggregate survey data such as the Survey of Professional Forecasts or micro data on expectations (e.g. Malmendier & Nagel (2016), Milani (2007), and Orphanides & Williams (2005)). Doraszelski et al. (2016) also find that firms weight recent play disproportionately when forming expectations about competitors’ play.

7For example, see Mankiw & Whinston (1986) and Spence (1977)

8The top two firms in the industry formed a vessel-sharing agreement (the “2M Alliance”) in 2014. China’s two biggest shipping lines have also proposed to merge.
$42 billion in the Asia-Europe market. An alternative counterfactual of a merger between the top two firms decreases investment by 7.5%, and results in a producer surplus increase of $14 billion and a consumer surplus decrease of $1 billion. These results suggest that strategic incentives raise investment rates and amplify boom-and-bust investment cycles. The results also have policy implications as coordinated investment decisions may lead to a consumer surplus loss but a total welfare gain.\textsuperscript{9}

More importantly, I show that learning amplifies investment cycles through intensifying strategic motives. When uncertainty about the aggregate demand is introduced, reducing strategic interaction through a merger or monopolization leads to a larger decrease in the level and the volatility of investment as well as a larger increase in welfare. This is because during high demand periods in which firms have greater strategic incentives to steal business from and preempt rivals, learning also leads firms to collectively become more optimistic, which reinforces the strategic incentives. Consequently, policy that is based on a full-information model is more likely to block welfare-enhancing mergers.

The second counterfactual simulation sheds light on the informational channel through which demand volatility affects investment. Consistently with findings in previous empirical studies (e.g. Bloom (2009), Collard-Wexler (2013)), I find that an increase in demand volatility reduces investment. In addition, I find that under learning large fluctuations in demand lead firms to revise their beliefs more frequently and more drastically, which in turn amplifies boom-bust investment cycles.

Related Literature

This paper builds on an emerging field that studies uncertainty and agents’ beliefs in a learning framework. At the 2000 Ely Lecture, Hansen (2007) argued that the rational expectations approach endows agents with too much information and advocated putting econometricians and economic agents on comparable footing. Cogley & Sargent (2005) use a Bayesian learning model to study the role of the Federal Reserve’s changing beliefs in the monetary policy. Orlik & Veldkamp (2014) study uncertainty shocks in the Bayesian learning framework. This paper investigates how expectations formed through learning explain firm-level decisions and within-industry cycles of investment.

In the area of learning, the empirical literature in industrial organization has predominantly explored learning about firms’ private information (e.g. Jovanovic (1982)), learning about a new technology (e.g. Covert (2014)), or consumers’ learning about values of experience goods through experimentation (e.g. Dickstein (2011)). Doraszelski et al. (2016) examine learning about com-

\textsuperscript{9}US antitrust policy prohibits firms in the same business from colluding on investment decisions, while Japan allows cooperation among rivals along this dimension. O’Brien (1987) argues that Japan’s support for coordinated decision-making in investment is partially responsible for the country’s success in the steel industry.
petitors’ play and demand elasticity parameters in the context of the UK electricity market.

This paper is complementary to empirical studies on investment cycles, especially two papers on the bulk shipping industry: Kalouptsidi (2014) and Greenwood & Hanson (2015).\textsuperscript{10} The paper’s contribution is to introduce a new informational structure and strategic interaction. Kalouptsidi (2014) employs a fully rational model and uses second-hand ship prices to identify values of owning a ship non-parametrically. As the second-hand prices already reflect sellers’ and buyers’ beliefs about future demand, Kalouptsidi is indirectly incorporating firms’ beliefs in the estimation of values of owning ships. By contrast, this study models firms’ forecasting process explicitly. This approach will be useful in cases where the industry does not have active second-hand market or the second-hand market suffers from significant selection problems.\textsuperscript{11} Understanding how firms form expectations is interesting in its own right as well. Greenwood & Hanson (2015) introduce behavioral biases in persistence in earnings and long-run endogenous supply responses by rivals to explain bulk shippers’ investment behavior. In contrast, this study does not require biases in firm beliefs. In particular, firms’ beliefs about rivals’ actions are consistent with the rivals’ equilibrium strategies.

This paper makes a methodological contribution to the literature on the structural analysis of industry dynamics. Doraszelski & Pakes (2007)) provide an overview of this literature. Recent empirical papers include Ryan (2012), Collard-Wexler (2013), and Igami (2017). This paper adopts learning as the belief-formation process in a dynamic oligopoly framework. It shows that introducing an extra dimension of uncertainty can be crucial in analyzing firm behavior in an environment subject to structural changes. Incorporating this type of uncertainty also helps us understand the informational channel through which demand fluctuations can affect investment, which contributes to the body of empirical studies that quantify the effect of demand uncertainty on investment (e.g. Bloom (2009), Collard-Wexler (2013), and Kellogg (2014)).

The remainder of the paper is organized as follows. Section 2 describes the industry and the data. Section 3 presents the dynamic model of investment with learning for the shipping industry. Section 4 describes the estimation procedure and presents estimation results including the empirical implementation of the learning model. Section 5 examines alternative informational structures to understand the role of agents beliefs and learning. In this section, I also diagnose the models of firm beliefs base on the approach that is relatively free of structural assumptions of the model

\textsuperscript{10}Although the two shipping industries share many similar characteristics, there is stark difference in terms of market power with much higher concentration in the container shipping industry. Kalouptsidi (2014) assumes that each firm owns one ship only and develops a competitive model of the bulk-shipping industry. Also, container shippers operate according to fixed schedules, while bulk shippers operate on-demand services much like taxis.

\textsuperscript{11}Adverse selection may arise in the second-hand market if sellers privately observe the quality of the goods. If there is selection, the quality of goods traded in the second-hand market may be different from the quality of goods currently owned by firms. In this case, estimating the value of owning the goods from second-hand prices will lead to biased estimates.
using GDP forecast data. Section 6 discusses counterfactual experiments and section 7 concludes.

2 Industry and Data

2.1 Container Shipping Industry

The container shipping industry’s core activity is the transportation of containerized goods over sea according to fixed schedules between named ports. The containers come in two standard dimension (the twenty-foot dry-cargo container (TEU) or the forty-foot dry-cargo container (FEU)), which makes it easier to load, unload, and stack the cargo. The container ships transport a wide range of consumer goods and intermediate goods such as electronics, machinery, textiles, and chemicals. Container trade accounts for over 15% of global seaborne trade by volume and over 60% in value (Stopford (2009)).

Container shipping is a capital-intensive industry. Companies can invest in capital by purchasing new vessels. The price of building a ship fluctuates depending on the conditions of the shipbuilding and shipping markets at the time of the order, including freight rates, the strength of trade demand, the size of the order book, and expectations. Container carriers also rely on chartered vessels, which are leased out by third parties. Chartered vessels account for approximately 50 percent of the total container ship capacity operated by the largest 20 firms. The majority of charter contracts for container ships are time charters which involve the hiring of a vessel for a specific period of time. The average contract length is 7-10 months (Reinhardt et al. (2012)). The charterer has operational control of the ships, while the ownership and management of the vessel are left in the hands of the shipowner. Firms can also scrap old ships which cannot be operated profitably. The demolition prices depend on the demand for scrap metal and the availability of ships for scrap.

The industry is vulnerable to sharp swings in global trade demand, but it is hard for firms to respond quickly to supply-demand imbalances in the short run. There is a gap between the time of placing a new order and the time of receiving the ordered ships due to time-to-build ranging from 2 to 4 years. Moreover, whereas bulk shippers can easily move their idle ships into lay-up, container shippers are limited to do so due to their pre-announced schedules (Stopford (2009)). When firms cannot fill their ships due to the oversupply of ships, they engage in fierce price competition in order to attract more customers. Hence, the ability to make correct forecasts about future demand and invest accordingly is important in this industry.

---

12 The construction of new ships happens at shipyards. There are approximately 300 major shipyards and many smaller ones globally.

13 The freight cost is the most important criterion for customers, although other factors such as transit time, schedule reliability, and frequency of departure matter as well (Reinhardt et al. (2012)).
Figure 1: Total investment and investment costs

Notes: This figure shows the volume of new orders and the average price of building new ships from 2001:Q1 to 2014:Q4.

Figure 1 shows the industry-level quarterly quantity of ship orders and the price of those orders for 2001 to 2014. Investment is concentrated in the times of high shipbuilding prices. Although the price is on average 42% higher compared to the 2009-2014 period, the volume of new orders is higher by more than 60% in the 2006-2008 period.

2.2 Data

This project uses two main datasets on the container shipping industry. The first dataset combines data collected from two sources: MDS Transmodal, a U.K.-based research company, and Clarksons Research, a U.K.-based ship-broking and research company. This dataset covers quarterly information from 2006 to 2014. The key information includes: (1) quantities and prices of container trade by trade route; (2) firm-level information on the number and the capacity of ships that each firm owns, charters, and has in its order book as well as the capacity deployed in each of the routes the firm operates on; and (3) industry-level charter rates, scrap prices, and shipbuilding prices.

The prices of building a new ship and the number of ships in the industry order book are available by size category (2500 TEU, 3700 TEU, 6700 TEU, 8800 TEU, 10000 TEU, and 13500 TEU). I first obtain per TEU shipbuilding prices for each size category and construct the weighted average of these prices. The average scrap value is constructed in a similar way.
Estimating firms’ beliefs for the sample period from 2006 to 2014 requires historical price and quantity data that go further back than 2006, ideally from the inception of the industry. The first dataset on firm-level investment and capital is therefore supplemented with the historical price and quantity data compiled from the Review of Maritime Transport published by the United Nations that goes back to 1997. It contains information on the average freight rates and cargo flows on major routes. The volume of trade is available at the yearly level in this dataset, although the price level is available at the quarterly level. The quarterly quantity of container trade are imputed based on the data on the value of trade by origin-destination pair from the IMF Direction of Trade Statistics database.

Figure 2: Prices on major trade routes

Notes: This figure shows quarterly average prices of shipping an unit of trade goods (TEU) on major container trade routes from 1997 to 2014. The shaded area covers the period on which this paper’s main analysis lies from 2006 to 2014.

Figure 2 shows the average prices on major trade routes (front-haul and back-haul separately) from 1997 to 2014. These routes together account for approximately 55% of all interregional container trade by trade volume and approximately 60% of deployed ship capacity. Shipping firms can adjust their capacity across different routes relatively easily and the network of services is constantly changing to meet the needs of trade (Stopford (2009)). For this reason, Kalouptsidi

---

15 Although this is roughly the start date of the official public data on the aggregate price and quantity of container trade, firms may have longer historical data and use them in forming expectations. Section 4.4 discusses my empirical strategy in estimating firms’ beliefs given the truncated nature of the price and quantity data.

16 The imputation assumes that the quarterly container trade volume is proportional to the value of trade in each year.
Table 1: Descriptive statistics

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Industry-level data (2006-2014)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shipbuilding price ($1000/TEU)</td>
<td>11.62</td>
<td>2.22</td>
<td>8.69</td>
<td>15.76</td>
</tr>
<tr>
<td>Scrap price ($1000/TEU)</td>
<td>2.62</td>
<td>0.55</td>
<td>1.50</td>
<td>3.81</td>
</tr>
<tr>
<td><strong>Market-level data (1997-2014)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asia to Europe Quantity (1 mil. TEU)</td>
<td>2.37</td>
<td>1.10</td>
<td>0.70</td>
<td>3.98</td>
</tr>
<tr>
<td>Price ($1000/TEU)</td>
<td>1.51</td>
<td>0.28</td>
<td>0.80</td>
<td>2.09</td>
</tr>
<tr>
<td>Europe to Asia Quantity (1 mil. TEU)</td>
<td>1.08</td>
<td>0.39</td>
<td>0.51</td>
<td>1.76</td>
</tr>
<tr>
<td>Price ($1000/TEU)</td>
<td>0.78</td>
<td>0.10</td>
<td>0.57</td>
<td>1.07</td>
</tr>
<tr>
<td>Asia to North America Quantity (1 mil. TEU)</td>
<td>2.57</td>
<td>0.78</td>
<td>1.12</td>
<td>3.92</td>
</tr>
<tr>
<td>Price ($1000/TEU)</td>
<td>1.67</td>
<td>0.21</td>
<td>1.27</td>
<td>2.20</td>
</tr>
<tr>
<td>North America to Asia Quantity (1 mil. TEU)</td>
<td>1.20</td>
<td>0.41</td>
<td>0.63</td>
<td>2.14</td>
</tr>
<tr>
<td>Price ($1000/TEU)</td>
<td>0.89</td>
<td>0.16</td>
<td>0.68</td>
<td>1.43</td>
</tr>
<tr>
<td>Europe to North America Quantity (1 mil. TEU)</td>
<td>0.78</td>
<td>0.15</td>
<td>0.48</td>
<td>1.05</td>
</tr>
<tr>
<td>Price ($1000/TEU)</td>
<td>1.32</td>
<td>0.16</td>
<td>0.93</td>
<td>1.77</td>
</tr>
<tr>
<td>North America to Europe Quantity (1 mil. TEU)</td>
<td>0.55</td>
<td>0.14</td>
<td>0.32</td>
<td>0.76</td>
</tr>
<tr>
<td>Price ($1000/TEU)</td>
<td>0.99</td>
<td>0.21</td>
<td>0.67</td>
<td>1.60</td>
</tr>
<tr>
<td><strong>Firm-level data (2006-2014)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Capacity of owned ships (1m TEU)</td>
<td>0.30</td>
<td>0.25</td>
<td>0.04</td>
<td>1.47</td>
</tr>
<tr>
<td>Capacity of ships in order book (1m TEU)</td>
<td>0.18</td>
<td>0.13</td>
<td>0.00</td>
<td>0.64</td>
</tr>
<tr>
<td>Capacity of chartered ships (1m TEU)</td>
<td>0.31</td>
<td>0.29</td>
<td>0.01</td>
<td>1.55</td>
</tr>
<tr>
<td>Capacity of ships deployed in Asia-Europe market (1m TEU)</td>
<td>0.22</td>
<td>0.19</td>
<td>0.04</td>
<td>0.99</td>
</tr>
</tbody>
</table>

Notes: There are 36 industry-level, 216 market-level, and 612 firm-level observations. Other routes include Asia to North America, North America to Asia, North America to Europe, and Europe to North America routes.

(2014) considers the shipping market to be a global one and considers a single demand component in her model. In my application, I account for demand in the Asia-Europe (A-E) market separately from demand in other major markets on the grounds that A-E demand was the main driver of the building boom in the mid-2000’s. Many of the ships that were ordered in this period were built specifically for the A-E routes and could not be easily redeployed to other routes due to the size restrictions. The A-E traffic mainly goes through the Suez Canal whose restriction on the ship size is approximately 18,000 TEU. The traffic on the next largest trade route – the Asia-North America route – goes through the Panama Canal (if going to or from the East Coast of North America) which has a smaller size restriction of approximately 5,000 TEU (expanded to approximately 14,000 TEU in June 2016) and the West Coast ports are also limited in their ability to handle large ships. Over 90% of ships in the orderbook in my sample were above this size restriction, indicating that many of these newly ordered ships are too large to fit through the

17For example, Maersk ordered eight E-class container ships (of size 14,770 TEUs) from 2006 to 2008 all of which were intended to be operated on the A-E route.
The analysis focuses on firms that deployed over 80,000 TEU of ships in the Asia-Europe market on average in the 2006 to 2014 period. These firms account for more than 95 percent of the total capacity of ships deployed in the Asia-Europe market. This results in a quarterly panel of 17 firms from 2006 to 2014. There is no entry into or exit by these firms in this period. Table 1 provides summary statistics of this dataset. On average, firms in the sample own 300,000 TEU in capacity, charter 310,000 TEU and have an order book of 180,000 TEU.

Figure 3: The distribution of firm size

Notes: This figure shows the capacity owned by each firm as a percentage of total industry capacity, where the capacity is averaged over the period of 2006:Q1 to 2014:Q4.

Figure 3 shows the distribution of firm size based on the average owned capacity over the period from 2006 to 2014. The market structure is quite concentrated with more than 40% of the total capacity concentrated on the top three firms in contrast to the bulk shipping industry which consists of a large number of small ship-owning firms (Kalouptsidi (2014)). While there is considerable size variation among the top two firms, the rest of the firms are similar in size.

Before describing the model, I look for preliminary evidence of changes in firms’ investment policy. It is inherently difficult to test whether firms are adjusting their beliefs about demand as they get new information since the beliefs are not directly observed. Instead, I search for suggestive evidence based on the difference in the predictions that a full-information model and a learning

---

Notes: Kalouptsidi (2014) shows that the largest fleet share is 3% for Handysize bulk carriers.
The Herfindahl index for the industry is 970 when these 17 firms are accounted for.
model make. A learning model generally predicts that even after controlling for all payoff-relevant variables, firms’ beliefs, hence, firms’ actions will be different before and after experiencing large demand shocks. By contrast, under full information new demand realizations do not contain any new information. So firms’ perceived probabilities of transitioning to different demand states from a given state stay fixed over time and firms’ actions in a given state also remain unchanged. Hence, I examine whether firm behavior changes significantly after firms experience large demand shocks while controlling for potentially important variables. In particular, I test for a structural break in the firm’s investment policy function and find evidence for such a break. The details of the test, results, and interpretation are provided in appendix B.

3 Model

This section presents the model for the container shipping industry. The model builds on the dynamic oligopoly framework developed by Ericson & Pakes (1995) and the learning literature in macroeconomics. Firms’ beliefs about demand change over time as firms re-estimate the parameters of the demand process using up-to-date information available to them. In each period a firm decides whether to invest in new ships and whether to scrap existing ships based on its own capital and order book levels, and rivals’ aggregate capital and order book levels as well as its beliefs about future demand. In the product market competition stage, firms decide on how much capacity to charter (lease from a third-party chartering company) and how much capacity to station in each market. I start by describing the model of firm beliefs in section 3.1. Section 3.2 presents firms’ dynamic problem, and section 3.3 demand for container shipping services and product market competition. Section 3.4 provides a definition of equilibrium.

3.1 A Model of Firms’ Beliefs about Demand

This section proposes an adaptive learning model of firms’ expectations about demand. Section 5.1 presents details of the alternative models considered in this paper including a full-information model, a Bayesian learning model, and a full-information model with time-varying demand volatility. Under adaptive learning agents form expectations about demand based on information available to them in each period. They operate like econometricians who estimate the parameters of the model based on best information at their disposal and make forecasts using their estimates.

For reasons discussed in section 2.2, I consider demand in the Asia-Europe market separately from demand on other major markets. Agents contemplate a first-order autoregressive model for

---

20The payoff-relevant variables are defined to be those variables that are not current controls and affect the current profits of at least one of the firms as in Ericson & Pakes (1995) and Maskin & Tirole (2001).
demand in the Asia-Europe market, denoted by $z_t$, as the following:

$$z_t = \rho^0 + \rho^1 z_{t-1} + \omega_t$$

$$= \rho' x_t + \omega_t$$

(1)

where $\omega_t \sim N(0, \sigma^2_t)$, $\rho = [\rho^0, \rho^1]'$, and $x_t = [1, z_{t-1}]'$. Similarly, the model for demand in other major markets, henceforth called the “outside market” ($\tilde{z}_t$), is given as:

$$\tilde{z}_t = \tilde{\rho}^0 + \tilde{\rho}^1 \tilde{z}_{t-1} + \tilde{\omega}_t$$

$$= \tilde{\rho}' \tilde{x}_t + \tilde{\omega}_t$$

(2)

where $\tilde{\omega}_t \sim N(0, \tilde{\sigma}^2_t)$, $\tilde{\rho} = [\tilde{\rho}^0, \tilde{\rho}^1]'$, and $\tilde{x}_t = [1, \tilde{z}_{t-1}]'$. In the full-information model, the parameters in the demand model, $\{\rho^0, \rho^1, \sigma, \tilde{\rho}^0, \tilde{\rho}^1, \tilde{\sigma}\}$ are known to the agents. By contrast, under adaptive learning agents revise expectations by re-estimating these parameters in each period based on demand realizations up to time $t$, $\{z_{\tau}, \tilde{z}_{\tau}\}_{\tau=0}^t$. At each $t$, firms’ beliefs about demand can be described by the estimates of the AR($1$) parameters, denoted as $\eta_t = (\rho^0_t, \rho^1_t, \sigma_t, \tilde{\rho}^0_t, \tilde{\rho}^1_t, \tilde{\sigma}_t)$.

Firms are assumed to have homogenous beliefs about the aggregate demand. The prices and quantities of container trade are public information periodically published in trade journals and other publications. Moreover, swings in global trade demand common to all firms are the main source of demand shocks in this industry. The model also assumes that agents use their current beliefs in forecasting demand. This assumption has two behavioral interpretations. The first interpretation is that agents believe current beliefs to be the correct or the best forecasts for future demand. The alternative interpretation is that agents use current beliefs in forecasting as these approximate future beliefs. Appendix D.1 implements one way of relaxing this assumption and discusses implications for estimation results and challenges in the implementation.

Let $X_t = [x_0, x_1, \ldots, x_t]'$ and $R_t = \frac{X_t'X_t}{t}$. The expectations at time $t$ regarding the Asia-Europe

\footnote{21 Allowing correlation in demand in these two markets is straightforward and does not change the results quantitatively and qualitatively.}

\footnote{22 On a practical level there are no publicly available data that provide information on firm-level demand to my knowledge, which are necessary to allow heterogeneous firm beliefs. Nevertheless, heterogeneity in firms’ beliefs would arise if firms experienced different demand shocks, for example, through different customer pools. How firms form heterogeneous beliefs and how they affect firm decisions and industry dynamics are interesting topics of study for future work.}

\footnote{23 This implies that agents do not internalize the possibility of learning in the future. Suppose information is endogenous to agents’ decisions, for example, because agents are making consumption decisions for experience goods for which quality is difficult to observe in advance. In this case, the assumption of myopic learning rules out experimentation, while allowing agents to internalize learning in the future may encourage experimentation. In this paper’s setting, since information about the aggregate trade demand is exogenous to agents’ actions, there is no room for experimentation regardless of the assumption on learning.}
market demand under adaptive learning can be written recursively as

\[
\rho_t = \rho_{t-1} + \lambda_t (R_t)^{-1} x_t \left( z_t - \rho'_{t-1} x_t \right) 
\]  

(3)

\[
R_t = R_{t-1} + \lambda_t (x_t x'_t - R_{t-1}) 
\]  

(4)

where \( \lambda_t \) is the weight parameter that governs how responsive the estimate revisions are to new data (Evans & Honkapohja (2012)). Figure 4 plots relative weights placed on observations for different values of \( \lambda_t \). For example, if \( \lambda_t = \frac{1}{t} \), agents put equal weight on all observations in their information set. If \( \lambda_t \) is some constant between 0 and 1, weights geometrically decline with the age of the observation such that agents assign heavier weights to more recent observations. This would be a natural way to form expectations if agents were concerned about the possibility of structural changes (Evans & Honkapohja (2012)). A larger value of \( \lambda_t \) leads to heavier discounting of older observations. For example, when \( \lambda_t = 0.03 \), agents put a 30\% weight on a 10-year-old observation relative to the most current observation, while when \( \lambda_t = 0.02 \), agents put a 45\% weight on a 10-year-old observation.
3.2 Firms’ Dynamic Problem

Time is discrete with an infinite horizon and is denoted by \( t \in \{0, 1, 2, \ldots \} \). There are \( n \) incumbent firms and the set of incumbent firms is denoted by \( N = \{1, 2, \ldots, n\} \). Firms are heterogeneous with respect to their firm-specific state, \( x_{it} = (k_{it}, b_{it}) \), where \( k_{it} \) is the capacity of ships owned by firm \( i \) and \( b_{it} \) is the backlog, or the capacity of firm \( i \)’s order book.\(^{24}\) The underlying industry state is \( s_t = ((x_{it})_i, d_t) \) where \((x_{it})_i\) is the list of all incumbents’ firm-specific states and \( d_t = (z_t, \tilde{z}_t) \) includes the demand states of the Asia-Europe market and the outside market.

The timing of events is as follows: (1) Firms observe their current state as well as their private cost shocks associated with investing and scrapping. They update their beliefs about demand. (2) Firms make investment and scrapping decisions. (3) Firms choose how much capacity to charter and how much capacity to deploy in the Asia-Europe market and the outside market. They engage in period competition and receive period profits. (4) The dynamic decisions are implemented and the delivery and depreciation outcomes are realized. The industry evolves to a new state.

Computing a Markov perfect equilibrium (in which each incumbent firm follows a Markov strategy that is optimal when all competitors follow the same strategy) is subject to the curse of dimensionality. As the number of incumbent firms grows, the number of states grows more than exponentially.\(^{25}\) To address this challenge, I consider an alternative equilibrium concept which can be viewed in the context of the moment-based Markov Equilibrium (MME) of Ifrach & Weintraub (2016), or more broadly the experience-based equilibrium (EBE) of Fershtman & Pakes (2012).

In MME, firms keep track of and condition their strategies on the detailed state of strategically important firms (dominant firms) and a few moments of the distribution describing non-dominant firms’ states, instead of the detailed state of all incumbents. This reduces the size of the state space thereby alleviating the computational burden. My application allows firms to keep track of their own firm-specific states, the sum of all incumbents’ states, and the aggregate demand states. Firms’ strategies thus depend on the firm-specific state, \( x_{it} = (k_{it}, b_{it}) \), and the moment-based industry state defined as \( \hat{s}_t = (\sum_i x_{it}, d_t) \). MME strategies are not necessarily optimal, however; there may be a profitable unilateral deviation to a strategy that depends on the detailed state of all firms. This is because the moment-based state may not be sufficient statistics to predict the future evolution of the industry. In appendix D.2, I consider a version that allows richer information by adding a dominant firm’s state into the moment-based industry state and show that the model predictions are robust to this change.

\(^{24}\)The owned capacity space denoted by \( \mathcal{K} \) is discretized into 19 points such that \( \mathcal{K} = \{k_0, k_1, k_2, \ldots, k_{18}\} \) and the order book capacity space denoted by \( \mathcal{B} \) into 7 points such that \( \mathcal{B} = \{b_0, b_1, \ldots, b_6\} \). \( \mathcal{K} \) and \( \mathcal{B} \) are both discretized in 100,000 TEU increments such that \( k_0 = 0 \) TEU, \( k_1 = 100,000 \) TEU, and so on, and \( b_0 = 0 \) TEU, \( b_1 = 100,000 \) TEU, and so on.

\(^{25}\)There are 17 active firms in my application. Even a simple specification with a single state variable that can take up to 5 different values would result in over a billion of states.
Firms make an investment decision \((ι_t ∈ \{0, 1\})\) and a scrapping decision \((δ_t ∈ \{0, 1\})\) in order to maximize expected discounted profits.\(^{26}\) I denote the strategy profile as \(μ_{it} = (ι_{it}, δ_{it})\). Each investing firm pays an investment cost. The investment cost consists of a part common to all firms which is a function of the aggregate state, \(κ(\hat{s}_t)\), and a privately observed part of the cost, \(ε_{it}^ι \sim N(0, (σ^ι)^2)\). If a firm decides to scrap its ships or if there is depreciation such that it can sell its ships for scrap metal, the firm receives a scrap value. The scrap value is the sum of the value common to all firms, \(φ(\hat{s}_t)\), and an iid private value distributed as \(ε_{it}^δ \sim N(0, (σ^δ)^2)\). Deprecation occurs with a probability proportional to the firm’s current capital amount, given as \(ζk_{it}\) for some constant \(ξ^2\).\(^{27}\) I denote as \(ν(δ_{it}, x_{it})\) the expected amount of capital reduction from depreciation or scrapping before the realization of the depreciation outcome such that \(ν(δ_{it}, x_{it})\) is one if \(δ_{it} = 1\) and \(ζk_{it}\) otherwise. The value function of a firm after observing its private shocks and before making investment and scrapping decisions can be written as

\[
V_{η_t}(x_{it}, \hat{s}_t) = \max_{ι_{it}, δ_{it}} π(x_{it}, \hat{s}_t) - ι_{it}(κ(\hat{s}_t) + ε_{it}^ι) + ν(δ_{it}, x_{it})(φ(\hat{s}_t) + ε_{it}^δ) + βE[V_{η_t}(x_{it+1}, \hat{s}_{t+1}|x_{it}, \hat{s}_t)]
\]

where \(η_t\) is the vector of parameters summarizing firms’ beliefs in period \(t\) about future demand. The value function is a function of \(η_t\) as it depends on how firms perceive the demand state to evolve. Note that the problem is still stationary due to the assumption that firms use the current period’s beliefs \((η_t)\) in forecasting future demand (see appendix D.1 for the discussion on relaxing this assumption.) This means that the continuation value is a function of \(η_t\) only, and not \(η_{t+1}, η_{t+2}, \) and so on.

The current model does not allow for persistent heterogeneity in the investment costs and scrap values across firms. The analysis of transaction-level pricing data on investment and demolition confirms that there is no significant firm heterogeneity at least in the observed transaction prices of investment and scrapping. The model incorporates firm heterogeneity in other areas, however, since it may be important given the persistent concentration of market power. First, the cost of chartering ships from a third party is allowed to depend on firm size, since larger firms may have greater bargaining power over charterers. Second, the marginal cost of production depends on the capacity of firm’s deployed ships. The detailed specification of these cost functions is given in

---

\(^{26}\)Firms are restricted to invest and/or scrap up to only one unit (100,000 TEU) per period. In the data there are no observations of a capital reduction by more than one unit and there are only three instances of an investment of more than one unit. Capping the maximum investment level to one unit for each firm reduces the action space thus alleviating the computational burden.

\(^{27}\)If a firm scraps its vessels, there is no depreciation in the same period such that the maximum reduction in \(k_{it}\) is one unit. This assumption is made since the data do not provide any observations of a capital reduction by more than one unit. The interpretation of this assumption can be that when a firm decides to scrap its vessels, it chooses the oldest vessels that are about to deprecate on their own. This assumption can be easily relaxed.
section 3.3.

State Transitions

When a firm invests, the order book capacity increases by one unit when there is no delivery at \( t \) and stays constant if there is delivery. A firm’s own capacity is determined by scrapping decision, depreciation and delivery outcomes. The transition of the firm-specific state is described as:

\[
\begin{align*}
    k_{it+1} &= k_{it} + \tau_{it} - \min(\delta_{it} + \psi_{it}, 1) \\
    b_{it+1} &= b_{it} + \mu_{it} - \tau_{it}
\end{align*}
\]

where \( \tau_{it} \) is delivery and \( \psi_{it} \) is depreciation. The probability of delivery is a linear function of the firm’s order book capacity such that the delivery happens with the probability of \( \xi b_{it} \) for some constant \( \xi \). Similarly, the probability of depreciation is \( \zeta k_{it} \) such that it linearly increases in the capital stock. The perceived evolutions at time \( t \) of the aggregate demand states for the Asia-Europe market and the outside market follow first-order autoregressive processes as the following:

\[
\begin{align*}
    z_t &= \rho_0 + \rho_1 z_{t-1} + \omega_t \\
    \tilde{z}_t &= \tilde{\rho}_0 + \tilde{\rho}_1 \tilde{z}_{t-1} + \tilde{\omega}_t
\end{align*}
\]

where \( \omega_t \sim N(0, \sigma_t^2) \) and \( \tilde{\omega}_t \sim N(0, \tilde{\sigma}_t^2) \). This process is described in more detail in section 3.1. The parameters in the AR(1) model, \( \eta_t = (\rho_0, \rho_1, \sigma_t, \tilde{\rho}_0, \tilde{\rho}_1, \tilde{\sigma}_t) \), summarize the beliefs about the evolution of future demand at time \( t \). How firms update these beliefs as they get new information is described in section 3.1.

Note that even though the evolution of the underlying state \( s_t \) is a Markov process under Markov strategies, the evolution of the moment-based industry state \( \hat{s}_t \) may not be. This is because information is lost in the process of aggregating information through moments. Hence, I approximate the Markov process for the moments using empirical transitions, or average transitions consistent with equilibrium play, following Ifrach & Weintraub (2016) as described below.

Let \( \mu \) denote the investment strategy and let \( P_{\mu^{'}, \mu} \) denote the transition kernel of the underlying state \( (x_{it}, s_t) \), when firm \( i \) uses strategy \( \mu^{'}, \) and its competitors use strategy \( \mu \). Then, we can define an

---

28 I explore alternative specifications including a case in which the errors in the AR(1) processes follow heavier-tailed t-distributions and a case in which correlation between demand in the Asia-Europe market and demand in the outside market is allowed. Main results are robust to these alternative specifications.

29 To understand this, suppose that there are three firms. Each of these firms keeps track of its own firm-specific state, \( x_{it} \) and the sum of all three firms’ states as the moment-based industry state such that \( \hat{s}_t = \sum_i x_{it} \). The underlying industry state is \( s_t = (x_{it})_i \). In one case, suppose that the underlying state is \( (10, 10, 10) \), while in another case the underlying state \( (30, 0, 0) \). In both cases, the moment-based industry state is \( \hat{s}_t = 30 \). However, starting from these two different underlying states may not yield the same distribution for the moment-based state in the next period \( (\hat{s}_{t+1}) \).
operator $\Phi$ such that $\hat{P}_{\mu',\mu} = \Phi P_{\mu',\mu}$ where a Markov process $\hat{P}_{\mu',\mu}$ approximates the non-Markov process of the moment-based state, $P_{\mu',\mu}$. In practice, the moment-based industry state’s evolution is defined to be the long-run average of observed transitions from the moment-based state in the current period to the moment in the next period consistent with strategy $\mu$ as follows:

$$\hat{P}_{\mu}[\hat{s}'|\hat{s}] = (\Phi P_{\mu})[\hat{s}'|\hat{s}] = \lim_{T \to \infty} \frac{\mathbb{I}\{\hat{s}_t = \hat{s}, \hat{s}_{t+1} = \hat{s}'\}}{\sum_{t=1}^{T} \mathbb{I}\{\hat{s}_t = \hat{s}\}}$$

where $\hat{s}_t = (\sum_i x_{it})$ includes the moments in the moment-based state.

### 3.3 Demand for Container Shipping and Product Market Competition

In each period, firms choose (a) how much capacity to charter ($h_{it}$), and (b) how much capacity to allocate to the Asia-Europe market ($\bar{q}_{it}$) and the outside market ($\tilde{q}_{it}$) given the state they are in. In other words, a firm chooses how much of its total capacity to allocate to the Asia-Europe market or the outside market where the total capacity is determined as the sum of its chartered and owned capacity. The capacity firms allocate to the Asia-Europe market determines the supply in the market, which along with demand determines the market-clearing price and quantity. Demand for each route in the Asia-Europe market is assumed to have constant elasticity as follows:

$$\log Q_{jt} = z_{jt} + \alpha \log P_{jt}$$  \hspace{1cm} (5)

where $z_{jt}$ denotes the demand state, $P_{jt}$ the price, and $Q_{jt}$ the quantity of route $j$ at time $t$.

The marginal cost of providing services on a route is linearly increasing in quantity up to the firms’ capacity constraint as given in equation (6).

$$mc(q_{it}, \bar{q}_{it}) = \begin{cases} a + \frac{b q_{it}}{\bar{q}_{it}} & \text{if } q_{it} \leq \bar{q}_{it} \\ \infty & \text{otherwise.} \end{cases}$$  \hspace{1cm} (6)

This functional form implies that (i) the marginal cost increases as the firm’s quantity gets closer to the firm’s full capacity; and (ii) firms with higher capacity have a lower marginal cost of producing the same level of quantity. This assumption is based on three institutional details. First, it becomes increasingly hard to schedule loading and unloading as the ship reaches its full capacity. Second, firms that deploy larger capacity on the route has a relative cost advantage due to the fact operating expenses such as crew, insurance, and administration offer scales of economies (Stopford (2009)). Lastly, this functional form allows me to indirectly account for the fact that larger firms tend to have larger ships thus higher fuel efficiency without having to include the size of ships as a state.
variable. This functional form is also similar to that of Ryan (2012) where the marginal cost is increasing as firms operate closer to maximum capacity but only after a threshold.

Then, the supply curve for route \( j \) is given as the horizontal sum of all firms’ supply curves as follows:

\[
P_{jt} = a + \frac{bQ_{jt}}{\bar{Q}_t} \quad \text{if} \quad Q_{jt} \leq \bar{Q}_t
\]

where \( \bar{Q}_t = \sum_i \bar{q}_{it} \). The price in the Asia-Europe market is determined by the intersection of the demand curve given in equation (5) and the supply curve given in equation (7). This assumption implies that in the product market firms are taking prices as given (although making positive profits due to the convex costs) and are not withholding capacity strategically given their capacity constraints. This assumption is motivated by the observation in the data that the “effective capacity” (the total capacity of ships firms make available on the market as a share of the total capacity of ships they have) stays relatively stable over time even in the face of huge excess capacity in the post-crisis period (see figure 13 in appendix A.1). The share stays at 92% on average in 2006-2014 with a dip to 87% in 2009-2010.

The period profit is the sum of profits from providing shipping services on the Asia to Europe and the Europe to Asia routes plus the profit from the outside market minus the charter cost and the fixed cost of capital:

\[
\pi(x_{it}, \hat{s}_t) = \max_{\bar{q}_it, h_{it}} \left\{ \left( \sum_{j=\{1, 2\}} P_{jt}q_{ijt} - c(q_{ijt}, \bar{q}_{it}) \right) + R(\bar{q}_{it}, \bar{Q}_t, \hat{s}_t) - CC(h_{it}, x_{it}, \hat{s}_t) - FC \cdot k_{it} \right\}
\]

where \( FC \) is the fixed cost of holding one unit of capital, \( R \) is the profit from the outside market, \( CC \) is the charter cost, and \( \bar{q}_{it} \) is the capacity deployed in the outside market. The fixed cost of holding ships includes all costs that do not vary with the output level (or how full the ships are) such as docking fees, maintenance costs, canal dues, and port charges. I do not explicitly model the chartering market and the product market competition in the outside market but account for them in a reduced-form way. The detailed specification of the reduced-form functions for the charter cost and the outside-market profit is given in section 4.2.

### 3.4 Equilibrium

The value function can be re-written as the perceived value of a firm using moment-based strategy \( \mu' \) in response to all other firms following strategy \( \mu \):

\[
\hat{V}_{\mu'} \eta (x, \hat{s}) = \pi(x, \hat{s}) - t(\kappa(\hat{s}) + \varepsilon^t) + \nu(\delta, x) \left( \phi(\hat{s}) + \varepsilon^\delta \right) + \beta E_{\mu', \mu} \hat{V}_{\mu'} \eta(x', \hat{s}'| x, \hat{s}).
\]

\textsuperscript{30}The correlation coefficient between the capacity on the A-E market and ship size is 0.83.
The definition of an equilibrium is then given as follows.

**Definition** Equilibrium comprises of an investment and scrapping strategy $\mu$ that satisfies the following conditions:

(a) Firm strategies satisfy the optimality condition:

$$\sup_{\mu' \in \mathcal{M}} \hat{V}^\eta_{\mu',\mu}(x,\hat{s}) = \hat{V}^\eta_{\mu}(x,\hat{s}) \quad \forall (x,\hat{s}) \in \mathcal{X} \times \hat{S}. $$

(b) The perceived transition kernel is given by:

$$\hat{P}_\mu = \Phi P_\mu$$

Equilibrium is computed using an algorithm based on value-function iteration. Appendix C describes the algorithm in detail.

### 4 Estimation and Empirical Results

The estimation of the dynamic model of investment with learning proceeds as follows. First, I estimate demand for shipping services to recover the elasticity of demand and demand states. Second, I estimate parameters governing static competition including the marginal cost of production, the charter cost, and the outside market profit, which are used to compute period profits. Third, I estimate the investment cost and the scrap value based on the pricing data of shipbuilding and demolition as well as other model primitives such as the delivery and depreciation processes. Fourth, I discuss the empirical implementation of the learning model. Lastly, I estimate the dynamic model through the method of simulated moments.

#### 4.1 Estimating Demand for Shipping Services

The goal of this section is to estimate the price elasticity of demand and to construct demand states for the Asia-Europe market and the outside market. The empirical analogue of the constant elasticity demand model in equation (5) is:

$$\log Q_{jt} = \alpha_0 + \alpha_1 \log P_{jt} + \alpha_2 W_{jt} + \epsilon_{jt}$$

where $j$ is an indicator for trade routes, $Q_{jt}$ is the amount of container shipping services in terms of TEU, $P_{jt}$ is the average price per TEU, and $W_{jt}$ is a demand shifter. I estimate equation (9) using

---

31 This section follows the demand estimation of Kalouptsidi (2014) closely.
instrumental variables regression in order to correct for the endogeneity of prices. The price is instrumented with the average size and age of ships and the fraction of ships that are over 20 years old. The size of ships is one of the key determinants of cost efficiency as larger ships require less fuel per TEU on average. The age of ships matters as well, since older ships tend to require higher maintenance costs. Log GDP for the destination area is used as a demand shifter.

The estimation uses data from six major trade routes from 2001:Q2 to 2014:Q4. The demand parameters are identified by the time-series variation as well as the cross-sectional variation across six different routes in the data along with the constant elasticity functional form assumption. In particular, since ships have to go back and forth the two routes in each market they serve, two routes in the same market (e.g. Asia to Europe and Europe to Asia) have the same level of supply while facing different demand shocks, which helps the identification of the demand parameters.

The price elasticity of demand is estimated to be -3.89 (see table 8 in Appendix A.2 for detailed results). This implies that a change in price from $1510 per TEU to $1360 per TEU would result in a change in quarterly quantity demanded of approximately 0.92 million TEU on the Asia to Europe route. Stopford (2009) explains that container trade is price elastic because lowering prices encourages the substitution of cheap foreign substitutes for local products. Moreover, other transportation modes are available such as road and rail transportation and air freight. Kalouptsidi (2014) estimates the price elasticity of demand for bulk shipping to be -6.17 under a constant elasticity specification.

Given the elasticity of demand estimates, I construct the demand state for each trade route \((z_{jt})\) as the intercept of the demand curve:

\[
z_{jt} = \hat{\alpha}_0 + \hat{\alpha}_2 W_{jt} + \hat{\epsilon}_{jt}
\] (10)

where \(\{\hat{\alpha}_0, \hat{\alpha}_2\}\) are parameters estimated from the regression and \(\hat{\epsilon}_{jt}\) is the residual. Finally, I construct aggregate demand states for the Asia-Europe market and the outside market from the route-level demand states. For the Asia-Europe market, I take the demand state for the Asia to Europe direction. Since the container trade volume is less than half on the Europe to Asia direction, firms’ investment and capacity deployment decisions in the market are mostly dictated by the trade demand on the Asia to Europe direction. For the outside market, I take the sum of the demand states in the non-Asia-Europe routes. Figure 5 plots the demand states for 1997 to 2014 for the Asia-Europe and the outside markets. There is a large drop in demand in both markets in 2008 to 2009. In the Asia-Europe market, the boom and bust cycles in demand are shorter in length after 2008.

---

32Although the price, quantity, and GDP data available from 1997, the instruments are available starting from 2001:Q2. The included trade routes are Asia to Europe, Europe to Asia, Asia to North America, North America to Asia, Europe to North America, and North America to Europe.
4.2 Estimating the Profit Function

The second step of the estimation is to construct period profits by estimating the marginal cost, charter cost, and outside market profit functions. Firms’ capacity deployment decisions yield a supply curve which along with the demand curve determines the equilibrium prices and quantities for the Asia-Europe market. The marginal cost of providing container shipping services is specified in equation (6), which serves as the basis for the maximum likelihood estimation of the cost parameters \((a, b)\).

The outside market profit and the charter cost functions are specified in a reduced-form way as:

\[
R(\tilde{q}_{it}, x_{it}, \hat{s}_t) = \tilde{q}_{it} \left( r_0 + r_1 \tilde{z}_t + r_2 \tilde{Q}_t \right)
\]

\[
CC(h_{it}, x_{it}, \hat{s}_t) = h_{it} (\gamma_0 + \gamma_1 z_t + \gamma_2 k_{it} + \gamma_3 K_t).
\]

The profit from each unit of capacity deployed in the outside market is allowed to depend on the total deployed capacity in the outside market \((\tilde{Q}_t)\) since higher supply may lead to fiercer price competition and lower profit. The charter cost depends on the firm-level owned capacity \((k_{it})\) since larger firms may get discounts on charter rates. The charter cost is also allowed to depend on the total capacity owned by operator \((K_t)\) as it is likely to affect demand for chartering.

The estimation of the charter cost and outside market profit functions is based on firms’ static profit maximization problem. Given the demand estimates, I estimate these objects via maximum
likelihood based on the first-order conditions with respect to the capacity deployed on Asia-Europe route ($\bar{q}_{ijt}$) and the chartering decisions ($h_{it}$), respectively. The variations in capacity deployment and charter decisions across different firm types and across time along with the first-order conditions and the functional form assumptions provide identification for these parameters.

Table 9 in Appendix A.2 reports the estimates of the profit function parameters. The coefficients on the Asia-Europe market demand state in the outside market profit and charter cost functions ($r_1$ and $\gamma_1$) are positive. This implies that stronger demand leads to higher outside market profits as well as higher charter costs. The estimates also show that when there is more aggregate deployed capacity in the outside market, firms earn less from that market on average, which captures the competitive effects. In addition, larger firms tend to face lower charter costs, and an increase in total industry capacity owned by ship operators lowers charter costs. Finally, the sign on the marginal cost parameter $b$ is positive, capturing the fact that the cost increases as the firm operates closer to its capacity constraint. The estimates also suggest that there is substantial cost heterogeneity across firm size. For example, the marginal cost at $q_{it} = 100,000$ TEU is 38% higher for firms with the maximum capacity of 200,000 TEU than for firms with the constraint of 400,000 TEU. 33

4.3 Estimating Other Model Primitives

This study recovers the investment cost and scrap value directly from the data on shipbuilding prices and scrap prices. Not only are detailed data are available unlike in many other settings, but this approach also helps me in estimating the model of firm beliefs. I use industry-level shipbuilding and demolition price data to estimate the investment cost and the scrap value, respectively, as functions of the industry state variables (industry owned ship and order book capacities, and demand states for the Asia-Europe and outside markets) via least squares. Figure 6 compares investment costs and scrap values observed in the data to predicted values obtained from the regression (see table 10 in Appendix A.2 for the detailed estimates).

The delivery process of newly ordered ships and the depreciation process of existing ships are also estimated separately from the estimation of dynamic parameters. The mean delivery rate is estimated based on a simple regression of delivery on the firm’s order book size with no constant. 34 For the depreciation process, I set an exogenous rate. This is because the data do not differentiate between depreciation and the scrapping of ships that can still be operated physically. Thus, the depreciation rate and the distribution of the private shocks to the scrap value can not be separately

---

33 Note, however, that the marginal cost of operating at $x\%$ of a firm’s capacity constraint is constant regardless of the firm size by construction.

34 The current formulation assumes that the delivery rate depends solely on the firm’s own order book size, since the industry order book size does not have a statistically significant effect on the delivery rate.
Figure 6: Predicted Investment Costs and Scrap Values

Notes: The left panel shows the average shipbuilding price observed in the data and the predicted shipbuilding price from the regression of the shipbuilding price on the industry state variables. The right panel shows the average scrap value and the predicted scrap value.

4.4 Empirical Implementation of the Learning Model

This section discusses the implementation of the learning model described in section 3.1 and presents expectations about demand implied by the model. The truncated nature of the price and quantity data for container trade poses a challenge in implementing the learning model. An agent’s information set in each period includes all observations from the past. However, although firms may have access to observations from the inception of the industry, the researcher may not. This problem arises in most empirical settings when dealing with a learning model. In my particular setting, data on prices and quantities for major trade routes are available starting from 1997, although the first international voyage dates back to 1966. Given this challenge, I explore two alternative methods of empirically implementing an adaptive learning model: the truncation approach and the imputation approach.

35 Although historically the lifespan of container ships was 25 to 30 years, it has fallen in recent years especially for larger ships. Vesselvalues reports that the average age of all sizes of container ships sold for scrap was around 22 years old and the average age at which Post-Panamax container ship was sold for scrap was around 19.5 years.
The truncation approach entails setting the initial period of the information set as the start date of the data. This method is straightforward to implement and is appropriate if firms also do not have access to information beyond the data available to the researcher. However, bias can arise if agents’ information set includes observations going further back than the start date of the data. The bias would be mitigated as agents discount older observations more heavily when forming expectations.

This approach is implemented as follows. I consider the weight parameter \( \lambda_t = \frac{1}{t} \) as well as \( \lambda_t \in (0, 0.04] \) in increments of 0.025.\(^{36}\) If \( \lambda_t = \frac{1}{t} \), equal weights are applied to all past observations. In practice, the estimation procedure under this parameter value amounts to applying least squares to estimate equation (1) for each period separately. The regression at period \( t \) uses demand state data covering from the start date of the data to the current period \( t \), or \( \{\tilde{z}_\tau, \tilde{z}_\tau\}_{\tau=0}^t \) where the steps of recovering these data are provided in section 4.1. If \( \lambda_t \) is a constant, weights on observations geometrically decline with the age. In this case, weighted least squares are applied where the weight on an observation from the period \( \tau \) is given by \( (1 - \lambda_t)^{t-\tau} \).

The imputation approach employs external data that provide information about the missing data. This approach is appealing if agents indeed use longer historical data in forming expectations than observed and the researcher has access to the external data that provide a good approximation to that data. Bias can arise, however, from the imputation process depending on the quality and scope of the external data. For this paper’s setting, one could consider using international trade data to proxy demand for container shipping.

The imputation approach is implemented as follows. I set the start date for firms’ information as the second quarter of 1966, which is the date of the first international container voyage. Then, I employ quarterly data on the value of trade by origin-destination pair from the IMF Direction of Trade Statistics database to impute the missing data on demand states from 1966:Q2-1996:Q4.\(^{37}\) Finally, I estimate the beliefs using the imputed longer time-series data in the same way as the truncation approach.

The truncation approach is adopted in the end because it provides a better data fit. Moreover, it can be more universally applied since the imputation method requires some external data which are not always available. Figure 18 in appendix D.3 compares beliefs under the two approaches.

Figure 7 shows firms’ demand parameter estimates from 2000 to 2014 under adaptive learning with \( \lambda_t = 0.02 \) for the Asia-Europe market (see figure 14 in appendix A.2 for the outside market). The estimates in the shaded area are for 2006 to 2014, which will used in the estimation of the

---

\(^{36}\)Orphanides & Williams (2005) suggest that the constant gain parameter in the range between 0.01 and 0.04 match the data on expectations well.

\(^{37}\)To translate the value of trade to the quantity of container trade, the demand state for the 1997-2014 period was regressed on the de-trended value of trade. Then, the demand states for periods with missing data are constructed as predicted values from the regression. For the 1997-2014 period, actual demand states are used.
Figure 7: Beliefs under Learning for the Asia-Europe Market

Notes: This figure shows firms’ beliefs about demand in the Asia-Europe market for 2000:Q1 to 2014:Q4 under adaptive learning with $\lambda_t = 0.02$. The beliefs are summarized by the three parameters, $\{\sigma_t, \rho^0_t, \rho^1_t\}$, in the AR(1) process as given in equation (1). Beliefs for 2006-2014 in the shaded area are used in the main analysis.

dynamic model. The estimate of the persistent parameter $\rho^1_t$ rises from 2006 to 2007 and shows a general downward trend thereafter. The variance parameter $\sigma_t$ hikes in early 2009 and stays high throughout the end of the sample period.

Under adaptive learning, the degree to which the parameter estimates react to recent events grows as agents put more weights on recent observations (as shown in figure 15 in appendix A.2). For example, the degree to which $\sigma_t$ jumps around 2009 is the smallest in the case where agents weigh all past demand realizations equally ($\lambda_t = 1/t$). When $\lambda_t$ is a constant, the larger $\lambda_t$, the larger the jump in $\sigma_t$ around 2009. Similarly, the larger the fall in the persistence parameter $\rho^1_t$ in the post-2008 period, the larger $\lambda_t$ becomes. It is this variation in beliefs and the variation in the data in investment and scrapping around demand shocks that identify the model of firm beliefs. The identification of $\lambda_t$ is discussed in more detail in section 4.5.

4.5 Estimating the Dynamic Model of Investment with Learning

The last and most computationally intense step of the estimation entails estimating the model of firm beliefs and the dynamic parameters. The typical empirical strategy of estimating a dynamic game of investment is to recover objects like investment costs, entry costs, and exit values by searching for parameters that minimize the distance between actions observed in the data and the ones that the parameters imply (e.g. Ryan (2012) and Collard-Wexler (2013)). This paper instead employs data on shipbuilding and demolition prices to estimate investment costs and scrap values as described in section 4.3, which opens up the possibility to identify the model of firms beliefs. Although the application is different, the underlying logic of this approach is similar to that of Hortacsu & Puller (2008) in which the authors use marginal cost data to quantify how much firms’
bidding deviates from the optimal bidding benchmark.

I employ the method of simulated moments (MSM) to estimate the dynamic model, which minimizes a distance criterion between key moments from the actual data and the simulated data. Let $\theta$ denote the vector of dynamic and belief parameters such that $\theta = (\sigma^1, \sigma^\delta, FC, \lambda_t)$. I solve for an equilibrium of the dynamic investment model and obtain the optimal investment policy function for each candidate parameter vector.\(^3\) Using equilibrium strategies obtained in the previous step, I simulate the equilibrium path for the 2006 to 2014 period $S = 1000$ times. And from these paths, I obtain the simulated moments as follows:

$$\Gamma(\theta) = \frac{1}{S} \sum_{s=1}^{S} \Gamma_s(\theta).$$

I search for the parameter vector that minimizes the weighted distance between the data and simulated moments given as:

$$f(\theta) = \left( \Gamma^d - \Gamma(\theta) \right)' W \left( \Gamma^d - \Gamma(\theta) \right). \quad (11)$$

where $\Gamma^d$ is the set of data moments.\(^4\)

The moments used in the estimation include the average investment before and after 2008, the volatility of investment, the correlation in demand and investment, and the aggregate capacity of owned and backlogged ships. Table 2 lists these moments and compares the data moments and simulated moments under the parameter estimates.

The results (reported in table 3) indicate that the weighting parameter estimate is $\lambda_t = 0.02$. I will refer to the adaptive learning model with $\lambda_t = 0.02$ as the baseline learning model in the rest of the paper. This implies that agents put approximately 45% weights on a 10-year-old observation compared to the most recent observation. This estimate is very close to the values that previous studies in macroeconomics have estimated based on aggregate survey data such as the Survey of Professional Forecasts or micro data on expectations. For example, Malmendier & Nagel (2016), Milani (2007), and Orphanides & Williams (2005) estimate the constant-gain parameter ($\lambda_t$) to be

\(^3\)Recently, empirical techniques have been proposed to estimate the dynamic industry equilibrium without having to solve for an equilibrium (e.g. Aguirregabiria & Mira (2007), Bajari et al. (2007), Pakes et al. (2007)). The first stage of this approach entails recovering firms’ policy functions by regression observed actions on observed state variables. The second stage involves estimating structure parameters which make these policies optimal. This approach relies on flexible functional forms in the first step, so the data requirement is too high given the global nature of my data set. I use a full solution method instead, which involves solving the model at every guess of the parameter but is more efficient.

\(^4\)The search is done over grids of $(\sigma^1, \sigma^\delta, FC, \lambda_t)$. The grids for $\sigma^1$ and $\sigma^\delta$ are in increments of 0.005 and the grid for $FC$ is in increments of $50/\text{TEU}$. The candidate belief parameter values include $\lambda_t = \frac{1}{t}$ and $\lambda_t = \{0.025, 0.05, 0.075, \ldots, 0.4\}$. I use the inverse of the variance-covariance matrix of the simulated moments as the weighting matrix ($W$).
Table 2: Data and Simulated Moments

<table>
<thead>
<tr>
<th></th>
<th>Data moments</th>
<th>Simulated moments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average investment in 2006-2008 (1 mil. TEU)</td>
<td>0.23</td>
<td>0.23 (0.03)</td>
</tr>
<tr>
<td>Average investment in 2009-2014 (1 mil. TEU)</td>
<td>0.14</td>
<td>0.15 (0.02)</td>
</tr>
<tr>
<td>Total capacity of owned ships (1 mil. TEU)</td>
<td>5.12</td>
<td>5.15 (0.27)</td>
</tr>
<tr>
<td>Total capacity in the order book (1 mil. TEU)</td>
<td>3.01</td>
<td>2.98 (0.14)</td>
</tr>
<tr>
<td>Correlation between demand and investment</td>
<td>0.19</td>
<td>0.22 (0.12)</td>
</tr>
<tr>
<td>Volatility of investment (1 mil. TEU)</td>
<td>0.17</td>
<td>0.17 (0.03)</td>
</tr>
</tbody>
</table>

Notes: This table compares moments observed in the data and moments simulated under the estimated parameters. The simulated moments are computed based on 1000 series of equilibrium paths. Standard deviations are in parentheses.

Table 3: Dynamic Parameter Estimates

<table>
<thead>
<tr>
<th></th>
<th>λ_t</th>
<th>σ^i (1 bil. US dollars)</th>
<th>σ^δ (1 bil. US dollars)</th>
<th>FC (1 bil. US dollars)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.02 (0.005)</td>
<td>0.275 (0.055)</td>
<td>0.43 (0.092)</td>
<td>0.025 (0.0051)</td>
</tr>
</tbody>
</table>

Notes: This table shows estimates of dynamic parameters. λ_t is the weighting parameter in the adaptive learning model which governs how heavily agents discount older observations when forming expectations about demand. σ^i is the standard deviation of the i.i.d. shock around the investment cost of building 100,000 TEU and σ^δ around the scrap value. FC is the fixed cost of holding capacity of 100,000 TEU. Standard errors are in parentheses.

0.0175, 0.0183, and 0.02, respectively, with respect to expectations about macroeconomic conditions and monetary policy. Figure 8 shows that the baseline learning model does well at predicting the investment boom in 2007 and the plunge in investment in 2009.

The fixed cost of holding one unit of capital (100,000 TEU) in one quarter is estimated to be 25 million dollars, which is approximately 36% of the period’s profit from one unit of capital (where the period profit is the sum of profits from the Asia-Europe market and the outside market minus the charter cost and does not include the investment cost and scrap value). This fixed cost includes all costs that owning and operating ships impose regardless of the production level such as maintenance costs, canal dues, and port charges. It also includes the cost of labor needed in the operation of the ships regardless of how full the ships are.

The identification relies on a revealed-preference argument. I have recovered the values of benefits and costs of each of the options that the firm faces—investment, scrapping, and staying for
Figure 8: Model Fits

Notes: The left panel shows the industry evolution simulated under the baseline learning model (adaptive learning with $\lambda_t = 0.02$) and the industry evolution in the data. The right panel shows yearly investment simulated under the baseline learning model and observed in the data, respectively. The simulated moments are based on 1000 equilibrium paths.

each state in the state space as described in section 4.3. As a result, given these values, firms’ choices in various states observed in the data reveal their expectations about future demand.

More concretely, the estimation relies on the variation in firms’ beliefs across different weighting parameter values and the variation in firm behavior across time and firms observed in the data. As firms discount older observations more heavily, their beliefs become more responsive to recent shocks. This will amplify the effect of recent demand shocks on investment, which will increase the correlation between demand and investment. The left panel of 9 illustrates this relationship by plotting the comparative statistics of the correlation of demand and investment for different values of $\lambda_t$ with all other parameters fixed at each estimated value. Similarly, when $\lambda_t$ increases and firms revise their beliefs more dramatically in response to demand shocks, investment becomes more volatile as illustrated in the right panel of figure 9.

In principle, the parameters are identified by both time-series and cross-sectional variations. Nevertheless, the main source of identification is time-series variation in investment and scrapping as well as investment costs and scrap values. And it is essential to observe a boom and a bust in my sample period. The shipping industry provides a great setting in that it is exposed to large exogenous fluctuations in demand coming from cycles in world trade.
5 Alternative Models of Firm Beliefs and Model Diagnosis

In section 5.1, I consider various alternative informational structures in order to understand the economic implications of the uncertainty and learning about the demand process. Comparing results across models of firm beliefs, however, hinges on the assumptions made in various parts of the model and the structural estimation. In section 5.2, I discuss a way of diagnosing these models based on GDP forecast data that relies less heavily on modeling assumptions.

5.1 Alternative Models of Firm Beliefs

Full Information

First, I consider a case in which agents know the parameters of the demand process and the only uncertainty is about which demand will be realized due to the variance in the process. As in the learning model, I assume that agents consider a first-order AR(1) process for the evolution of demand. This specification is expected to perform poorly in matching patterns in the data considering the volatile nature of demand in the sample period. Nevertheless, it provides a direct comparison to the learning model, and thus will help us understand the role of information and agent beliefs arising from learning.

The model governing the evolution of demand is given by equations (1) and (2) as in the adaptive learning model. In the full-information model, however, the parameters in the demand model, \( \{\rho^0, \rho^1, \sigma, \tilde{\rho}^0, \tilde{\rho}^1, \tilde{\sigma}\} \), are known to the agents. Then, estimating beliefs under this model involves estimating the demand process using the full sample of data or as much data as available.
to the researcher. I apply least squares to estimate the $AR(1)$ processes using data from 1997:Q1 to 2014:Q4. Beliefs implied by the full information model are presented in figure 10. The parameter estimates stay constant under full information by construction. Compared to the baseline learning model, the volatility estimate ($\sigma$) is higher in the pre-2008 period and lower afterwards. The persistent parameter ($\rho^1$) by contrast is lower in the pre-2008 period and higher afterwards.

**Figure 10: Beliefs under Alternative Models of Beliefs for the Asia-Europe Market**

![Belief Parameter Estimates](image)

Notes: This figure shows firms’ beliefs about demand in the Asia-Europe market for 2000:Q1 to 2014:Q4 under the full-information, Bayesian learning, and baseline learning (adaptive learning with $\lambda_t = 0.02$) models. The beliefs are summarized by the three parameters, $\{\sigma, \rho^0, \rho^1\}$, in the $AR(1)$ process as given in equation (1). Beliefs for 2006-2014 in the shaded area are used in the main analysis.

Figure 11(a) shows the annual investment levels and the second column of table 4 shows the simulated moments under the full information model. Primitives are re-estimated for each alternative model of beliefs. One of the striking features that arises when removing learning and endowing agents with information about the demand process is that the correlation between demand and investment becomes negative such that firms restrain from investing in the high demand periods of 2006-2007 and invest more heavily in the post-2008 period. This happens for the following reason. A demand increase for shipping has two opposing forces on investment. On one hand, high demand raises returns to investment, raising demand for new ships. At the same time, an increased demand for new ships and a higher volume of backlogs drive up the shipbuilding prices, which creates a negative effect. The positive demand-side effect is dominated by the negative supply-side effect in the case of full information resulting in firms investing more under weak demand conditions. By contrast, the positive effect is much stronger under learning, as high demand draws

\[^{40}\text{Results are qualitatively not different when I fix the primitives at the levels recovered under the baseline learning model.}\]
also lead firms to revise their expectations upward, which creates a positive correlation between demand and investment. In addition, investment is less volatile under full information, suggesting that fluctuations in agent beliefs arising from learning amplify cycles of investment.

Table 4: Simulated Moments under Alternative Models of Firm Beliefs

<table>
<thead>
<tr>
<th>Data moments</th>
<th>Full Info</th>
<th>Bayesian</th>
<th>GARCH</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average investment in 2006-2008 (1 mil. TEU)</td>
<td>0.23</td>
<td>0.17</td>
<td>0.20</td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td>(0.009)</td>
<td>(0.01)</td>
</tr>
<tr>
<td>Average investment in 2009-2014 (1 mil. TEU)</td>
<td>0.14</td>
<td>0.22</td>
<td>0.18</td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.009)</td>
<td>(0.005)</td>
</tr>
<tr>
<td>Total capacity of owned ships (1 mil. TEU)</td>
<td>5.12</td>
<td>5.05</td>
<td>5.36</td>
</tr>
<tr>
<td></td>
<td>(0.77)</td>
<td>(1.32)</td>
<td>(0.80)</td>
</tr>
<tr>
<td>Total capacity in the order book (1 mil. TEU)</td>
<td>3.01</td>
<td>3.06</td>
<td>2.96</td>
</tr>
<tr>
<td></td>
<td>(0.23)</td>
<td>(0.01)</td>
<td>(0.19)</td>
</tr>
<tr>
<td>Correlation between demand and investment</td>
<td>0.19</td>
<td>-0.24</td>
<td>0.24</td>
</tr>
<tr>
<td></td>
<td>(0.03)</td>
<td>(0.02)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>Volatility of investment (1 mil. TEU)</td>
<td>0.17</td>
<td>0.14</td>
<td>0.17</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.006)</td>
<td>(0.004)</td>
</tr>
</tbody>
</table>

Notes: The simulated moments are computed based on 1000 series of equilibrium paths. All primitives are re-estimated for each model. Standard deviations are in parentheses.

Figure 11: Predicted Annual Investment under Alternative Models of Firm Beliefs

Bayesian Learning

Under Bayesian learning, each firm starts with prior beliefs about the parameters of the model. Then, based on its information set, \( \{z_\tau, \tilde{z}_\tau\}_{\tau=0}^T \), the firm updates its beliefs about the parameters.
in the demand process, \((\rho_0^0, \rho_1^0, \sigma_t, \rho_0^1, \rho_1^1, \tilde{\sigma}_t)\). The \(AR(1)\) coefficients for the Asia-Europe market, \(\rho = [\rho^0, \rho^1]\), have normal priors given by \(\rho_0 \sim N(\mu_0, \Sigma_0)\). The prior of \(\sigma^2\) follows an inverse Gamma distribution. Then, the posterior distribution \(\rho_t \sim N(\mu_t, \Sigma_t)\) has the mean and the variance given by

\[
\mu_t = \Sigma_t \left( \Sigma_0^{-1} \mu_0 + \sigma^{-2} (X_t'Z_t) \right)
\]

\[
\Sigma_t = \left( \Sigma_0^{-1} + \sigma^{-2} (X_t'X_t) \right)^{-1}.
\]

The beliefs are defined similarly for the outside market.

The first three years of the price and quantity data (1997:Q1-1999Q4) are used in the estimation of the prior beliefs, although I explore alternative estimations of the priors.\(^{41}\) I start from diffuse priors and apply the Gibbs sampling methods (see table 11 in appendix A.3 for the estimates). In the first quarter of 2000, firms start with the prior beliefs about the parameters and revise their beliefs using Bayesian updating in each period based on newly realized data. I apply the Gibbs sampling techniques to estimate the posterior beliefs.

Figure 10 shows beliefs under Bayesian learning. Compared to the baseline model of adaptive learning, the degree to which firms’ beliefs react to new data is smaller under Bayesian learning. This is because there are less weights placed on new data under Bayesian learning as agents place positive weights on their prior beliefs. Consequently, although the timing of the investment boom and bust predicted under Bayesian learning is consistent with the data, the magnitudes of the rise and the fall in investment are smaller than observed in the data as shown in figure 11(b).

**Full Information with Time-Varying Volatility**

Assuming homoskedasticity under full information may be too restrictive given the nature of demand in this industry. Thus, I consider a full-information model with a more flexible specification of the variance of the model: the GARCH model.\(^{42}\) Demand is assumed to follow the same \(AR(1)\) process as in other models. But volatility is assumed to follow a GARCH(1,1) process such that the current period’s variance depends on the last period’s realized error and variance:

\[
\sigma_t^2 = a_0 + a_1 \omega_{t-1}^2 + b_1 \sigma_{t-1}^2
\]

where \(\omega_{t-1}\) is the realized error in period \(t-1\).

\(^{41}\)I explore using the full sample from 1997 to 2014 for estimating the priors with the same updating for posterior beliefs, in which case the posterior beliefs remain almost identical in this case. The choice of the sample period for prior beliefs is thus expected to have little impact on empirical results.

\(^{42}\)As an alternative model of stochastic volatility, I consider a regime-switch mode where the variance is no longer a constant but can take on one of two values \(\sigma_t \in \{\sigma_l, \sigma_h\}\) and state changes are governed by a Markov transition matrix. I omit details of this model as it produces similar results as the GARCH model.
The GARCH model is estimated using the full sample of data (1997:Q1-2014:Q4). The estimates and the inferred conditional variance are presented in table 12 and figure 16 in appendix A.3, respectively. Compared to the volatility estimates in the learning model, the jump in the variance around 2009 is larger. But the increase is more temporary unlike in the learning model where the variance remains high throughout the end of the sample period.

The dynamic model presented in section 3.2 is modified to accommodate time-varying volatility. Since firms need the current period’s variances and errors and the GARCH model parameters as well to predict the next period’s variance, these parameters are included as additional state variables. Allowing time-varying volatility helps generating the correct prediction on the timing of investment, although the magnitude of the investment cycle is still smaller than observed in the data or under the baseline learning model. This finding provides some insights into firms’ beliefs about demand. It shows that the low volatility in demand in the pre-2008 period and the sharp increase in volatility in 2009 help us explain the high level of investment in the pre-2008 period and the subsequent fall in investment. The finding, however, suggests that changes in the level of demand forecasts over time may be also necessary in addition to the changes in the variance to predict firms’ investment behavior fully.

This exercise shows that adding more flexibility to the specification of the demand process can help in correctly predicting firm investment behavior. The downside of this approach, however, is that as it adds more parameters to the specification, an identification issue may arise. Moreover, the additional parameters become state variables, adding computational burden to the already complex dynamic problem. An advantage of the learning model is that it provides a computationally tractable way of modeling more sophisticated information sets.

### 5.2 Diagnosing Models of Firm Beliefs using GDP Data

In this section, I employ an alternative strategy to diagnose different models of firm beliefs that is relatively free of modeling assumptions. In particular, based on the fact that GDP and trade demand are highly correlated, I examine which model of firm beliefs generates beliefs that are most consistent with GDP forecasts. The ECB publishes the Survey of Professional Forecasters (SPF) for the euro area quarterly and reports the mean forecast of one-year-ahead and two-year-ahead GDP growth rates as well as a measure of how uncertain each forecaster is about his or her forecast. For the uncertainty measure, each forecaster is asked to allocate subjective probabilities to ranges of possible outcomes with a width of 0.5 percentage point.\(^{43}\)

I take the forecast for the 2-year ahead GDP growth to construct the mean and the variance

\(^{43}\)For example, forecasters are asked to assign a probability to real GDP rising between 0.0% and 0.4%, 0.5% and 0.9%, and so on.
of the forecasts in each quarter from 2006 to 2014.\footnote{Only two-year ahead forecasts are used in the analysis because there is substantial bunching in the forecasters’ probabilities in end bins for one-year forecasts. The bunching makes it difficult to construct variance estimates.} Then, I construct the mean and the variance of 2-year ahead demand growth that each model of beliefs implies. Finally, I compute correlation between the mean and the variance of GDP forecasts and the mean and the variance of demand forecasts implied by each model of beliefs. The correlation coefficients are reported in table 5.

Table 5: Correlation between GDP Forecasts and Demand Forecasts

<table>
<thead>
<tr>
<th>Correlation coefficient</th>
<th>Full info</th>
<th>Full info GARCH</th>
<th>Bayesian learning $\lambda_t = \frac{1}{3}$</th>
<th>$\lambda_t = .01$</th>
<th>$\lambda_t = .02$</th>
<th>$\lambda_t = .03$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between means of GDP &amp; demand growth</td>
<td>-0.23 (0.17)</td>
<td>-0.23 (0.17)</td>
<td>0.14 (0.17)</td>
<td>0.19 (0.17)</td>
<td>0.20 (0.17)</td>
<td>0.21 (0.17)</td>
</tr>
<tr>
<td>Between variances of GDP &amp; demand growth</td>
<td>0.37 (0.16)</td>
<td>0.37 (0.16)</td>
<td>0.83 (0.09)</td>
<td>0.86 (0.09)</td>
<td>0.86 (0.09)</td>
<td>0.85 (0.09)</td>
</tr>
</tbody>
</table>

Notes: The mean and variance of GDP growth forecasts are based on two-year ahead forecasts published by the ECB.

The results confirm that the adaptive learning model produces beliefs that are most highly correlated with beliefs implied by GDP forecasts (although the exercise does not reveal which value of $\lambda_t$ produces the most favorable outcome). The correlation coefficient for the mean growth is approximately 0.20 for adaptive learning and 0.14 for Bayesian learning. The correlation for the variance ranges from 0.83 to 0.86 for the learning models. By contrast, the correlation for the mean is negative for the full-information models as they predict that the growth rate is higher during periods of weak demand.\footnote{This is because the AR(1) process has the mean-reversion property. So with the constant parameter estimates as in the full-information model, the expected growth rate is larger when current demand is lower.} In addition, under full information with constant volatility the correlation for the variance is zero since the variance is constant by construction. The correlation coefficient is 0.37 under the GARCH model, which is still significantly lower than under learning models.

6 Counterfactual Analysis

In order to understand the mechanisms through which learning affects firm investment behavior, I conduct two sets of counterfactual experiments. First, I conduct counterfactuals with respect to competition and industry consolidation by simulating the industry under a multi-plant monopolist and a merger of top two firms. This exercise helps us understand how competitive forces interact with agent beliefs. In the second set of counterfactuals, I address the long-standing question on the effect of demand volatility on investment. By applying the learning framework I shed light on the informational channel through which demand fluctuations affect investment.
6.1 Coordination among Firms

In this section, I study the effects of strategic incentives and consolidation as well as how these effects interact with agent beliefs. To deal with the recent excess capacity in the industry, container shipping firms have increasingly moved towards consolidation. In July 2014 Maersk Line and MSC—the world’s two biggest container-shipping companies—formed an alliance named 2M, which akin to a code-sharing deal between airlines, is meant to help firms cut costs by using each other’s ships and port facilities and reduce competition. More firms are planning mergers and acquisitions as well. Cosco and CSCL, the sixth and seventh largest carriers by operated fleet capacity, have proposed a merger. CMA-CGM has proposed an acquisition of APL.

On one hand, increased consolidation may hurt consumers through reduced competition. On the other hand, there are potential sources of efficiency gains on the producers’ side, which makes the final direction of the welfare change ambiguous. In particular, consolidation may reduce the business stealing effect and preemption motives that can lead to the capital level that is higher than the socially optimal level. Mankiw & Whinston (1986) show that the business stealing effect can result in socially inefficient levels of entry when there are fixed costs of entry. Also, many theoretical studies predict that strategic incentives can lead to excess capacity, since firms may use investment as a commitment to deter entry or expansion of rivals (e.g. Spence (1977)).

My model incorporates several sources of strategic incentives. First, there is a business-stealing effect. A firm’s deployment of an extra unit of capacity has a negative effect on the market price and the competitors’ profitability. The business-stealing effect arises because this negative effect of increasing one’s own capital is internalize by all incumbents in the market. Second, as the volume of the industry order book grows and shipyards get closer to their full capacity, the price of building a new ship increases. This generates dynamic incentives for firms to preemptively commit to investment before others do when they expect strong demand.

I first consider a monopolist who operates and makes joint decisions of investment, scrapping, chartering, and deployment for all firms in order to maximize the aggregate profits. One could assume that the monopolist operates all ships under one plant. However, this would not only get rid of strategic interaction, but would also result in changes in costs, bargaining power with the charterer, etc., due to the larger firm size. Therefore, to disentangle the effect of strategic incentives from the effect arising from a change in the firm size distribution, I assume that the monopolist operates multiple plants instead of assuming that the monopolist operates all ships under one plant. The size distribution of the plants is assumed to be consistent with the average size distribution of the firms observed in the data. The monopolist is endowed with beliefs from the baseline learning model. Similarly, in a merger counterfactual, I allow the joint profit maximization of the top two firms.
Table 6: Monopoly and Merger Counterfactuals

Panel A: Industry Outcomes and Welfare

<table>
<thead>
<tr>
<th></th>
<th>Monopoly (%Δ)</th>
<th>Merger(%Δ)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Owned capacity (1 mil. TEU)</td>
<td>3.95 (-23.18)</td>
<td>5.02 (-2.53)</td>
</tr>
<tr>
<td>Orderbook (1 mil. TEU)</td>
<td>2.35 (-21.33)</td>
<td>2.78 (-6.80)</td>
</tr>
<tr>
<td>Investment (1 mil. TEU)</td>
<td>0.12 (-33.92)</td>
<td>0.17 (-7.50)</td>
</tr>
<tr>
<td>Volatility of investment (1 mil. TEU)</td>
<td>0.13 (-21.51)</td>
<td>0.15 (-14.66)</td>
</tr>
<tr>
<td>Correlation between investment and demand</td>
<td>0.14 (-35.79)</td>
<td>0.15 (-30.46)</td>
</tr>
<tr>
<td>Consumer surplus (1 bil. US dollars)</td>
<td>40.85 (-50.56)</td>
<td>81.69 (-1.13)</td>
</tr>
<tr>
<td>Producer surplus (1 bil. US dollars)</td>
<td>106.62 (616.14)</td>
<td>28.85 (93.80)</td>
</tr>
<tr>
<td>Total surplus (1 bil. US dollars)</td>
<td>147.47 (51.23)</td>
<td>110.54 (13.36)</td>
</tr>
<tr>
<td>Investment by top two firms (1 mil. TEU)</td>
<td>.</td>
<td>0.01 (-40.12)</td>
</tr>
<tr>
<td>Investment by other firms (1 mil. TEU)</td>
<td>.</td>
<td>0.15 (-2.47)</td>
</tr>
<tr>
<td>Owned capacity of top two firms (1 mil. TEU)</td>
<td>.</td>
<td>1.43 (-5.59)</td>
</tr>
<tr>
<td>Owned capacity of other firms (1 mil. TEU)</td>
<td>.</td>
<td>3.59 (-1.25)</td>
</tr>
<tr>
<td>Producer surplus of top two firms (1 bil. US dollars)</td>
<td>.</td>
<td>25.88 (105.15)</td>
</tr>
<tr>
<td>Producer surplus of other firms (1 bil. US dollars)</td>
<td>.</td>
<td>2.97 (20.85)</td>
</tr>
</tbody>
</table>

Panel B: Welfare Changes under Learning and Full-Information Models

<table>
<thead>
<tr>
<th></th>
<th>Monopoly</th>
<th>Merger</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Learning</td>
<td>RE</td>
</tr>
<tr>
<td>Δ in investment (1 mil. TEU)</td>
<td>-0.061</td>
<td>-0.039</td>
</tr>
<tr>
<td>Δ in investment volatility (1 mil. TEU)</td>
<td>-0.037</td>
<td>-0.020</td>
</tr>
<tr>
<td>Δ in consumer surplus (1 bil. US dollars)</td>
<td>-41.78</td>
<td>-39.35</td>
</tr>
<tr>
<td>Δ in producer surplus (1 bil. US dollars)</td>
<td>91.73</td>
<td>83.27</td>
</tr>
<tr>
<td>Δ in total surplus (1 bil. US dollars)</td>
<td>49.95</td>
<td>43.92</td>
</tr>
</tbody>
</table>

Notes: Panel A shows results from the monopoly and the merger simulations over the sample period (2006:Q1-2014Q4) with the percent changes from the case of no monopolization or merger in parentheses. The owned capacity, order book, and investment are reported as the average over time, and the welfare measures as the sum over the entire period. Panel B compares changes predicted by the learning model and the full-information model. Consumer surplus is calculated with respect to the Asia-Europe market only.

As shown in Table 6 Panel A, removing competition externalities through the monopolization has a substantial effect on investment: during the period of 2006 to 2014, investment drops by 34%. Under the merger case, investment drops by 7.5%. Investment falls heavily for the merging firms by 40%, but also falls for non-top-two firms by 2.5%. In addition to the level of investment, the timing of investment changes as well. In particular, the volatility of investment and the correlation between investment and demand fall under both the monopoly and merger cases. Compared to the baseline case investment is relatively less concentrated in times of strong demand when the price of investment is also very high as figure 12. These results suggest that strategic interaction among firms not only is responsible for raising investment as many theoretical studies suggest, but also
raises the volatility of investment.

![Graph](image.png)

**Figure 12: Yearly Investment under the Monopoly and Merger Cases**

**Notes:** The simulations are based on 1000 equilibrium paths.

In terms of welfare, this means that producer surplus increases by 91 billion dollars under monopoly, while consumer surplus in the Asia-Europe market falls by 42 billion dollars. This amounts to a 51% increase in total surplus, which is computed as the sum of consumer surplus in the Asia-Europe market and producer surplus. In the merger case, producer surplus almost doubles while consumer surplus drops only slightly by 1%.

How do these strategic effects interact with agent learning then? To answer this question, I compare results from the monopolization and merger counterfactuals under full information and learning in panel B of table 6. The results show that, compared to the learning model, the full-information model underestimates changes in investment resulting from the monopolization or the merger, thus underestimates welfare changes, especially the producer surplus gain. For example, the predicted decrease in investment resulting from the merger is approximately 14,000 TEU under the learning model, compared to 9,000 TEU under the full-information model. The underestimation of the changes in investment volatility is also dramatic (25,000 TEU compared to 4,000 TEU). The fact that the effect of strategic incentives is greater under learning sheds light on the relationship between the competitive forces and firm beliefs. Strong demand for shipping raises firms’ strategic incentives, for example, to preemptively commit to investment and to still business from others. But under learning strong demand also makes agents more optimistic, which amplifies
these strategic incentives. These results also suggest that if a regulator used the full-information model in evaluating this merger, he is likely to underestimate the welfare gains.

6.2 Demand Volatility

Demand volatility can affect investment in several ways. First, as real options theory predicts, an increase in demand volatility raises the cost of investment, since once a firm makes an investment it cannot disinvest should market conditions change adversely. Second, an increase in demand volatility may also increase the volatility of investment costs. Finally, the presence of learning opens up an additional channel through which demand fluctuations affect investment, since increased demand volatility makes agents revise their expectations more often and more drastically.

To quantify the effect of demand volatility, I conduct the following counterfactual simulations. I simulate two sets of demand series for 2006 to 2014—one with high volatility and the other with low volatility. In the high volatility case, the variances in the demand processes for the Asia-Europe and outside markets are doubled from the estimates based on the full sample of data. In the low volatility case, the variances are halved from the estimates. The remaining parameters and the demand realizations prior to 2006 are set to the estimated levels.

<table>
<thead>
<tr>
<th>Model</th>
<th>Learning Volatility</th>
<th>RE Volatility</th>
</tr>
</thead>
<tbody>
<tr>
<td>Investment (100,000 TEU)</td>
<td>1.48</td>
<td>1.35</td>
</tr>
<tr>
<td>Volatility of investment (100,000 TEU)</td>
<td>0.75</td>
<td>0.47</td>
</tr>
<tr>
<td>Corr. between demand and investment</td>
<td>0.04</td>
<td>-0.12</td>
</tr>
<tr>
<td>Consumer surplus (1 bil. US dollars)</td>
<td>112.60</td>
<td>113.27</td>
</tr>
<tr>
<td>Producer surplus (1 bil. US dollars)</td>
<td>24.59</td>
<td>26.84</td>
</tr>
<tr>
<td>Total surplus (1 bil. US dollars)</td>
<td>137.19</td>
<td>140.12</td>
</tr>
</tbody>
</table>

Notes: This table shows results from demand volatility counterfactuals. The owned capacity, order book, and investment are reported as the average over time, and the welfare measures as the sum over the entire period. Consumer surplus is calculated with respect to the Asia-Europe market only.

Table 7 shows simulation results for the high and low volatility cases under learning and full information, respectively. An increase in demand volatility has a negative effect on investment, which is consistent with findings in previous studies such as Bloom (2009) and Collard-Wexler (2013). Going from low to high volatility reduces investment by 6% under learning. This suggests that the value function is concave with respect to demand. If the value function is concave, lower volatility in demand raises the expected value of owning a ship, increasing the average level of investment. In addition, an increase in demand volatility also increases the volatility of investment as higher demand volatility leads to more volatile shipbuilding prices.
Note that the learning model and the full-information model yield qualitatively and quantitatively different predictions about investment patterns: under learning higher demand volatility generates larger investment boom and bust cycles that are more highly correlated with demand cycles. First, the increase in the volatility of investment in response to an increase in demand volatility is higher under learning. This is because when learning is present higher demand volatility also leads to larger changes in firms’ expectations about future demand, which further increases the volatility of investment. Second, there is higher a correlation between demand and investment under learning since learning generates agent beliefs that are more correlated with demand. By contrast, under full information the negative effect arising from the fact that higher demand for shipping increases shipbuilding prices dominates the positive effect that an increase in demand has on investment.

7 Conclusion

Information plays a potentially important role when firms invest in long-lived capital while facing large fluctuations and uncertainty in demand. This paper empirically examines the role of information in generating investment boom and bust cycles and overcapacity in the context of the container shipping industry. I develop a dynamic oligopoly model of investment which incorporates an additional layer of uncertainty – uncertainty about the aggregate demand process. In this model, agents form expectations about demand using best information available to them in each period and use their changing forecasts in making their investment and scrapping decisions. This learning framework provides a computationally tractable way to incorporate more sophisticated information sets. This is useful particularly for settings that are highly volatile and complicated with potential structural breaks such that it is hard for economic agents and researchers to observe and estimate the underlying process of interest.

A key empirical strategy of the paper is to adopt the data on shipbuilding and demolition prices, which allows me to identify the model of firm beliefs. I find that the uncertainty about the demand process amplifies investment cycles and raises the correlation between investment and demand, which helps us explain the boom-bust investment patterns. My counterfactual analysis reveals the mechanisms through which learning affects investment cycles. Through simulating policies that reduce competition among firms, I find that learning strengthens firms’ strategic incentives, which amplifies investment cycles. I also find that under learning higher demand volatility leads to more frequent and larger revisions of expectations about demand, thereby amplifying the magnitude of investment cycles.
References


A DETAILED ESTIMATION RESULTS

A.1 ADDITIONAL DESCRIPTIVE STATISTICS

Figure 13: Deployed ship capacity as a share of total (owned and chartered) ship capacity

A.2 DEMAND, PROFITS, AND FIRM BELIEFS

This section presents detailed results from the empirical implementation of the learning model in section 4.4 and the first three steps of the estimation described in sections 4.1 to 4.3.
Figure 14: Beliefs under Learning for the Outside Market

Notes: This figure shows firms’ beliefs about demand in the outside market for 2000:Q1 to 2014:Q4 under adaptive learning with $\lambda_t = 0.02$. The beliefs are summarized by the three parameters, $\{\tilde{\sigma}_t, \tilde{\rho}_0^t, \tilde{\rho}_1^t\}$, in the AR(1) process as given in equation (2). Beliefs for 2006-2014 in the shaded area are used in the main analysis.

Figure 15: Beliefs under Learning with Different Weighting Parameters for the Asia-Europe Market

Notes: This figure shows firms’ beliefs about demand in the Asia-Europe market for 2000:Q1 to 2014:Q4 under adaptive learning for different values of $\lambda_t$. The beliefs are summarized by the three parameters, $\{\sigma_t, \rho_0^t, \rho_1^t\}$, in the AR(1) process as given in equation (1). Beliefs for 2006-2014 in the shaded area are used in the main analysis.
Table 8: IV Regression Results for Demand for Container Shipping

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>First stage</th>
<th>Second stage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size of owned ships (1000 TEU)</td>
<td>-0.13**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.06)</td>
<td></td>
</tr>
<tr>
<td>Age of owned ships (year)</td>
<td>0.03</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.03)</td>
<td></td>
</tr>
<tr>
<td>Fraction of 20+ y.o. ships</td>
<td>-0.02*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td></td>
</tr>
<tr>
<td>Log GDP</td>
<td>0.44***</td>
<td>2.73***</td>
</tr>
<tr>
<td></td>
<td>(0.12)</td>
<td>(0.53)</td>
</tr>
<tr>
<td>Log price</td>
<td>-3.89**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.87)</td>
<td></td>
</tr>
<tr>
<td>Route FE</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Constant</td>
<td>-6.27***</td>
<td>-32.66***</td>
</tr>
<tr>
<td></td>
<td>(1.79)</td>
<td>(7.48)</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.83</td>
<td>0.11</td>
</tr>
</tbody>
</table>

Notes: Standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.001$.

Table 9: Estimates of the Profit Function Coefficient Parameters

<table>
<thead>
<tr>
<th>Marginal cost</th>
<th>$a$ 0.265 (0.011)</th>
<th>$b$ 1.750 (0.024)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Outside market profit (R)</td>
<td>$r_0$ -1.238 (0.177)</td>
<td>$r_1$ 0.089 (0.006)</td>
</tr>
<tr>
<td>Charter Cost (CC)</td>
<td>$\gamma_0$ 0.206 (0.096)</td>
<td>$\gamma_1$ 0.087 (0.007)</td>
</tr>
</tbody>
</table>

Notes: This table reports estimates of the parameters in the marginal cost, outside market profit, and charter cost functions. The unit of the aggregate deployed capacity ($\tilde{Q}_t$) in the outside market profit function; and the firm-level owned capacity $(k_{it})$ and the aggregate owned capacity $(K_{it})$ in the charter cost function is 1 million TEU. Standard errors for the estimates are in parentheses.
Table 10: Estimates of the Investment Cost and Scrap Value

<table>
<thead>
<tr>
<th></th>
<th>Investment cost ($1000/TEU)</th>
<th>Scrap value ($1000/TEU)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total capacity of owned ships (1 mil. TEU)</td>
<td>-1.35***</td>
<td>0.11</td>
</tr>
<tr>
<td></td>
<td>(0.35)</td>
<td>(0.12)</td>
</tr>
<tr>
<td>Total capacity in order book (1 mil. TEU)</td>
<td>1.12**</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>(0.54)</td>
<td>(0.19)</td>
</tr>
<tr>
<td>Demand state: A-E market</td>
<td>0.50</td>
<td>0.25**</td>
</tr>
<tr>
<td></td>
<td>(0.31)</td>
<td>(0.11)</td>
</tr>
<tr>
<td>Demand state: outside market</td>
<td>-0.16</td>
<td>0.08</td>
</tr>
<tr>
<td></td>
<td>(0.17)</td>
<td>(0.06)</td>
</tr>
<tr>
<td>Constant</td>
<td>15.09**</td>
<td>-3.17*</td>
</tr>
<tr>
<td></td>
<td>(4.81)</td>
<td>(1.68)</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.69</td>
<td>0.38</td>
</tr>
</tbody>
</table>

Notes: This table reports coefficient estimates in the investment cost and scrap value functions. Standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.001$. 
A.3 Alternative Models of Firm Beliefs

This section presents detailed estimation results under various alternative models of firm beliefs including the full-information model, Bayesian learning model, and full-information model with time-varying volatility. The specification and implementation of these models are described in section 5.1.

Table 11: Moments of the Prior Distributions under Bayesian Learning

<table>
<thead>
<tr>
<th></th>
<th>Asia-Europe market</th>
<th>Outside market</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho^0$</td>
<td>0.51</td>
<td>8.11</td>
</tr>
<tr>
<td>$\rho^1$</td>
<td>0.95</td>
<td>0.72</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>0.17</td>
<td>0.68</td>
</tr>
<tr>
<td>(0.63)</td>
<td>(0.08)</td>
<td>(0.17)</td>
</tr>
<tr>
<td>(0.02)</td>
<td>(0.27)</td>
<td>(0.27)</td>
</tr>
</tbody>
</table>

Notes: This table shows the estimated means and standard deviations (in parentheses) of the prior distributions of $AR(1)$ parameters. The estimation is based on data from 1997:Q1 to 1999:Q4.

Table 12: Estimates of the Time-Varying Volatility Models

<table>
<thead>
<tr>
<th></th>
<th>Asia-Europe Market</th>
<th>Outside Market</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_0$</td>
<td>0.05</td>
<td>0.34</td>
</tr>
<tr>
<td>$a_1$</td>
<td>0.83</td>
<td>0.73</td>
</tr>
<tr>
<td>$b_1$</td>
<td>0.17</td>
<td>(0.17)</td>
</tr>
<tr>
<td>(0.01)</td>
<td>(0.28)</td>
<td>(0.11)</td>
</tr>
<tr>
<td>(0.16)</td>
<td>(0.25)</td>
<td>(0.25)</td>
</tr>
</tbody>
</table>

Notes: Standard errors are in parentheses.
Notes: These figures plot the conditional variance for the Asia-Europe market under the GARCH model as well as the variance under the baseline adaptive learning model.
B Preliminary Evidence of Investment Policy Changes

Learning and full-information models make different predictions about firm behavior. A learning model generally predicts that even after controlling for the state (which includes all payoff-relevant variables), firms’ beliefs, hence firms’ actions will be different before and after experiencing large demand shocks. By contrast, under full information firms’ perceived probabilities of transitioning to different demand states from a given state stay fixed over time as new demand realizations do not contain any new information. Therefore, I examine whether firms’ investment behavior changes significantly after they experience large demand shocks in order to search for indirect evidence of learning. In particular, I test for a structural break in the firm’s investment policy function with an unknown break date following the approach proposed by Andrews (1993) closely.

The structural break equation is given by

\[ y_{it} = \beta'_1 x_{it} \mathbb{I}(t < \bar{t}) + \beta'_2 x_{it} \mathbb{I}(t \geq \bar{t}) + e_{it} \]

where \( \bar{t} \) is the break date, \( y_{it} \) is new investment, and \( x_{it} \) includes payoff-relevant variables. The state variables include the demand states for the Asia-Europe market and the outside market; firm-specific state variables including the owned capacity and the order book capacity; and the industry state including the aggregate capacity of operator-owned ships and the aggregate order book capacity.\(^{46}\)

Instead of imposing an exogenous break date, I first pin down the break date by estimating a structural break equation with different break dates and searching for the one that maximizes the fit of the equation. A break date minimizes the sum of squared residuals function defined as the following:

\[ S(\beta, \bar{t}) = \sum_i \sum_j (y_{it} - \beta'_1 x_{it} \mathbb{I}(t < \bar{t}) - \beta'_2 x_{it} \mathbb{I}(t \geq \bar{t}))^2. \]

The periods from 2007:Q2 to 2013:Q3 are considered as a break date because I need sufficient observations before and after to estimate the equation. Figure 17 plots the sum of squared residuals function for different break dates. The break date that minimizes the SSE is the last quarter 2008, which coincides with the downturns in international trade.

Table 13 reports results for the estimation of the policy function with the last quarter of 2008 as the break point as well as results from the structural break test. Based on the test, I reject the null that the investment policy is the same before and after the last quarter of 2008. The regression results suggest that in the post-2008 period firms’ investment decisions are more responsive to the

\(^{46}\)The demand states are recovered through demand estimation as given in section 4.1.
industry total capacity. That is, firms hold back from investment when there is a greater amount of total fleets available in the industry in the post-2008 periods. On the other hand, the industry capacity does not have a significant effect on investment in the pre-2008 period.

One explanation for the observed policy change is that there is learning about the demand process, that is, firms’ expectations about the demand process are indeed changing over time. Another possibility, however, is that there are important payoff-relevant variables that I failed to control for in this regression that are also causing the change in firms’ investment policy. One most obvious candidate is credit market conditions that are changing over the sample period and might be impacting firms’ investment decisions. Section B.1 investigates whether credit market conditions played an important role in firms’ investment decisions in this period.
Table 13: Investment Policy Estimation and a Test of a Structural Break

**Panel A: Regression**

Dependent variable: New investment (1000 TEU)

$t < 2008Q4$
- Constant: 324** (151)
- Demand state (Asia to Europe): 22** (11)
- Demand state (Outside market): 6.7 (4.1)
- Owned ship capacity (1000 TEU): .037** (.014)
- Order book capacity (1000 TEU): -.029 (.022)
- Aggregate owned ship capacity (1000 TEU): .02 (.015)
- Aggregate order book capacity (1000 TEU): -.055** (.017)

$t \geq 2008Q4$
- Constant: 138** (66)
- Demand state (Asia to Europe): -4.4* (2.4)
- Demand state (Outside market): .77 (1.3)
- Owned ship capacity (1000 TEU): .01 (.0075)
- Order book capacity (1000 TEU): -.0026 (.015)
- Aggregate owned ship capacity (1000 TEU): -.0097** (.0048)
- Aggregate order book capacity (1000 TEU): -.02** (.0069)

Observations: 612
$R^2$: 0.177

**Panel B: Test of a Structural Break**

$H_0: \hat{\beta}_1 = \hat{\beta}_2$
- Test statistics: 4.38
- p-value: (0.0001)

*Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.001$. 
B.1 Credit Market Conditions

In the sample period that this study focuses on (2016-2014), there were sharp swings in credit market conditions along with swings in demand for international shipping. Therefore, one might worry that omitting information about credit market conditions might bias the main results of the paper. This section, therefore, examines whether credit market conditions played an important role in firms’ investment decisions. I use data from Compustat on company financials information, in particular firms’ debts and liabilities.47

Using these data I regress investment levels on state variables and variables relating to the firm’s credit constraints including long-term debt and debt in current liabilities. If financial constraints were the main determinants of investment, we expect that firms that hold a higher amount of debt thus facing harsher credit constraints will withhold investment more. The regression results presented in table 14, nonetheless, suggest that debt levels do not have statistically significant effects on firms’ investment.

Table 14: Regression of investment on debt-related variables

<table>
<thead>
<tr>
<th>Dependent variable: Investment (1000 TEU)</th>
<th>Coefficient</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Owned ship capacity (1000 TEU)</td>
<td>-.037</td>
<td>(.027)</td>
</tr>
<tr>
<td>Order book capacity (1000 TEU)</td>
<td>-.024</td>
<td>(.017)</td>
</tr>
<tr>
<td>Aggregate owned ship capacity (1000 TEU)</td>
<td>.012</td>
<td>(.01)</td>
</tr>
<tr>
<td>Aggregate order book capacity (1000 TEU)</td>
<td>-0.15**</td>
<td>(.0064)</td>
</tr>
<tr>
<td>Demand state (Asia to Europe)</td>
<td>1.1</td>
<td>(2.3)</td>
</tr>
<tr>
<td>Demand state (Outside market)</td>
<td>.06</td>
<td>(1.3)</td>
</tr>
<tr>
<td>Chartered ship capacity (1000 TEU)</td>
<td>-0.25</td>
<td>(.024)</td>
</tr>
<tr>
<td>Aggregate chartered ship capacity(1000 TEU)</td>
<td>-0.19*</td>
<td>(.011)</td>
</tr>
<tr>
<td>Deployment in Asia-Europe market (1000 TEU)</td>
<td>.087**</td>
<td>(.043)</td>
</tr>
<tr>
<td>Aggregate deployment in Asia-Europe market (1000 TEU)</td>
<td>.019**</td>
<td>(.0078)</td>
</tr>
<tr>
<td>Long-term debt (1 bil. US dollars)</td>
<td>.00079</td>
<td>(.002)</td>
</tr>
<tr>
<td>Debt in current liabilities (1 bil. US dollars)</td>
<td>-.0019</td>
<td>(.0029)</td>
</tr>
<tr>
<td>Constant</td>
<td>-11</td>
<td>(38)</td>
</tr>
<tr>
<td>Observations</td>
<td>281</td>
<td></td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.076</td>
<td></td>
</tr>
</tbody>
</table>

Notes: Standard errors are in parentheses. * p < 0.10, ** p < 0.05, *** p < 0.001.

---

47281 company-quarter-level observations on company financials are available out of 612 observations used in the main analysis. There is, however, substantial variation on the magnitude of debts across firms in the data. The average firm-level long-term debt over the sample period varies from 0.06 million dollars for UASC to 4.3 billion dollars for Hyundai.
C Computation

To compute strategies under MME for the model described in 3.2, I adopt a computational algorithm that is analogous to the standard value function iteration algorithm except for an extra simulation step. Because the transition of the moment-based industry state \( \hat{s} \) may not be Markov, a simulation step is used to generate the Markov approximation of the transition of this state. The algorithm starts with a choice-specific value function that maps from the set of state-action pairs to values denoted as \( W^\eta(\mu, x, \hat{s}) \). It contains expected values of different actions prior to drawing random costs of investing and scrapping given beliefs about demand \( \eta \). Then, based on a simulation run in which firms play optimal strategies implied by these choice-specific values, the algorithm constructs the perceived transition kernel \( \hat{P}_\mu[m'|\hat{s}] \). The next step updates the values and strategies using the best response against the current strategy and the perceived transitions kernel. Finally, equilibrium conditions are checked based on the norm of the distance between the values in the memory and the updated values. A more detailed description of the algorithm is provided as follows:

1. Initialize \( W^\eta(\mu, x, \hat{s}) \) for all \((\mu, x, \hat{s})\) \(\in\mathcal{M} \times \mathcal{X} \times \hat{\mathcal{S}}\), and optimal strategies, \( \mu^* \), that \( W^\eta \) implies.

2. Simulate a sample path of \( \{\hat{s}_t\}_{t=1}^T \) for large \( T \) based on \( \mu^* \). Calculate the empirical frequencies of industry state \( h(\hat{s}) = \frac{1}{T} I\{\hat{s}_t = \hat{s}\} \) for all \( \hat{s} \in \hat{\mathcal{S}} \). Calculate the empirical transition kernel as
   \[
   \hat{P}_\mu[m'|\hat{s}] = \frac{\sum_{t=1}^T I\{\hat{s}_t = \hat{s}, m_{t+1} = m'\}}{\sum_{t=1}^T I\{\hat{s}_t = \hat{s}\}}.
   \]

3. Calculate the new values for each state-action pair \((\mu, x, \hat{s})\) as:
   \[
   \tilde{W}^\eta(\mu, x, \hat{s}) = \pi(x, \hat{s}) - \kappa(\hat{s}) + v(\delta, x)\phi(\hat{s}) + \beta E_{a,\mu}[V^\eta(x', \hat{s}'|x, \hat{s})]
   \]
   and obtain the new best response \( \tilde{\mu}^* = \arg\max_\mu W(\mu, x, \hat{s}|\mu, \mu^*) \) for all \((x, \hat{s})\) \(\in\mathcal{X} \times \hat{\mathcal{S}}\).

4. Calculate the following norm: \( \max_{x, \mu} \sum_{\hat{s} \in \hat{\mathcal{S}}} |\tilde{W}^\eta(\mu, x, \hat{s}) - W^\eta(\mu, x, \hat{s})| h(\hat{s}) \).

5. If the norm is greater than \( \epsilon \), update the values and the strategy profile with \( \tilde{W} \) and \( \tilde{\mu}^* \) and repeat steps 2-5.
D Robustness

D.1 Relaxing the Myopic Learning Assumption

The main specification of this paper assumes that agents do not internalize the possibility of learning in the future and use their current beliefs in their forecasts. This implies that although the parameters \( \{ \rho_0^0, \rho_1^1, \sigma_0, \rho_1, \tilde{\rho}_0^0, \tilde{\rho}_1^1, \tilde{\sigma}_t \} \) that summarize agents’ current beliefs are still state variables, they are not ‘active’ ones in the sense that they stay fixed over time since forecasting demand for all future periods requires only the current beliefs. Therefore, I can solve the model while fixing the belief parameters at the levels implied by demand realizations observed in the data and the specified learning model.

In contrast, if I allow agents to forecast using beliefs that change as they receive new draws of demand, the belief parameters now become state variables that evolve stochastically depending on the realizations of demand. Therefore, implementing this requires solving the model for each point on a grid of belief parameter values in addition to values of all other parameters.

Since fully relaxing the myopic learning assumption is computationally infeasible, I consider partially relaxing it by allowing agents to internalize future learning for one period ahead. That is, at time \( t \) agents use their current beliefs in predicting the distribution of demand for \( t + 1 \) and use their updated beliefs based on demand realized at \( t + 1 \) to predict demand from \( t + 2 \) onwards. Although still restrictive, this exercise will be informative, especially because due to discounting the effect should be the strongest for \( t + 1 \) and subside with time. I solve this model for the baseline case where all other parameters are held at their estimated values.

<table>
<thead>
<tr>
<th></th>
<th>Non-myopic learning</th>
<th>Baseline (myopic learning)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average investment in 2006-2008 (1 mil. TEU)</td>
<td>0.24</td>
<td>0.23</td>
</tr>
<tr>
<td></td>
<td>(0.03)</td>
<td>(0.03)</td>
</tr>
<tr>
<td>Average investment in 2009-2014 (1 mil. TEU)</td>
<td>0.15</td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>Total capacity of owned ships (1 mil. TEU)</td>
<td>5.16</td>
<td>5.15</td>
</tr>
<tr>
<td></td>
<td>(0.27)</td>
<td>(0.27)</td>
</tr>
<tr>
<td>Total capacity in the order book (1 mil. TEU)</td>
<td>3.00</td>
<td>2.98</td>
</tr>
<tr>
<td></td>
<td>(0.14)</td>
<td>(0.14)</td>
</tr>
<tr>
<td>Correlation between demand and investment</td>
<td>0.27</td>
<td>0.22</td>
</tr>
<tr>
<td></td>
<td>(0.11)</td>
<td>(0.12)</td>
</tr>
<tr>
<td>Volatility of investment (1 mil. TEU)</td>
<td>0.17</td>
<td>0.17</td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td>(0.03)</td>
</tr>
</tbody>
</table>

Notes: The simulated moments are computed based on 1000 series of equilibrium paths. Standard deviations are in parentheses.
Table 15 shows simulated moments under this new specification referred to as ‘non-myopic learning’ in comparison to the moments under the baseline specification. The results show the correlation between demand and investment is higher under non-myopic learning while other moments are almost identical under the two specifications. That is, when firms internalize the possibility of learning in the future, their recent demand draws and thus their current beliefs have stronger positive effects on investment. For example, when firms draw favorable demand draws, their current beliefs are revised upward, but this also makes them believe that future draws will be more favorable and they will continue to be more optimistic relative to the myopic learning case. Thus, under non-myopic learning the effect of recent demand on investment is further reinforced.

D.2 Adding a Dominant Firm’s State in the Moment-Based State

The moment-based Markov equilibrium as proposed by Ifrach & Weintraub (2016) allows firms to keep track of the detailed state of dominant firms (strategically important firms) as well as moments describing the state of fringe firms as their moment-based industry state. In my application, firms’ industry states are further reduced to the sum of states of all firms However, MME strategies may not be optimal (i.e. there may be a profitable unilateral deviation to a strategy that depends on more detailed information), if moments do not summarize all payoff-relevant information. In order to investigate how robust equilibrium strategies are to changes in the moment-based industry state, I consider a version in which richer information is allowed in the industry state and compare model predictions and values to the baseline case.

In particular, firms condition their strategy on the firm-specific state of the largest firm (the dominant firm) in addition to the states in the baseline case including their own firm-specific state, the sum of all firms’ states, and demand states. In one version, the dominant firm’s capital, denoted as $k_1$ is included in the information set and in the other version, the dominant firm’s order book, $b_1$. Let $\hat{s}'$ denote the new industry state and let $\mu'$ and $\hat{V}$ denote the optimal strategy and the value of the new game based on $\hat{s}'$ as the industry state. The difference in the values of the baseline model and the model that includes the dominant firm’s state for each underlying state $s$ is defined as:

$$\Delta_{\mu'}(x, s) = \frac{V_{\mu', \mu}(x, \hat{s}') - \hat{V}_{\mu}(x, \hat{s})}{\hat{V}_{\mu}(x, \hat{s})}.$$  

The expected value of this deviation is computed as the weighted average through a simulation where the weights come from simulations based on the baseline model, or $\hat{V}$. Table 16 shows that model predictions stay robust when either of the dominant firm states is added. The average difference in the values is not significantly different from zero for both cases.
Table 16: Adding a Dominant Firm’s State in the Moment-Based State

<table>
<thead>
<tr>
<th>Panel A: Simulated moments</th>
<th>Baseline</th>
<th>Model with dominant firm’s capital state</th>
<th>Model with dominant firm’s order book state</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average investment in 2006-2008 (1 mil. TEU)</td>
<td>0.23 (0.03)</td>
<td>0.23 (0.03)</td>
<td>0.23 (0.03)</td>
</tr>
<tr>
<td>Average investment in 2009-2014 (1 mil. TEU)</td>
<td>0.15 (0.02)</td>
<td>0.15 (0.02)</td>
<td>0.15 (0.02)</td>
</tr>
<tr>
<td>Total capacity of owned ships (1 mil. TEU)</td>
<td>5.15 (0.27)</td>
<td>5.13 (0.28)</td>
<td>5.14 (0.28)</td>
</tr>
<tr>
<td>Total capacity in the order book (1 mil. TEU)</td>
<td>2.98 (0.14)</td>
<td>2.98 (0.14)</td>
<td>2.99 (0.14)</td>
</tr>
<tr>
<td>Correlation between demand and investment</td>
<td>0.22 (0.12)</td>
<td>0.21 (0.12)</td>
<td>0.22 (0.12)</td>
</tr>
<tr>
<td>Std. dev. in investment (1 mil. TEU)</td>
<td>0.17 (0.03)</td>
<td>0.17 (0.03)</td>
<td>0.17 (0.03)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B: Average difference in values</th>
<th>All firms (%)</th>
<th>Dominant firm (%)</th>
<th>Fringe firms (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.35 (0.46)</td>
<td>-0.42 (0.54)</td>
<td>-0.36 (0.48)</td>
<td>-0.43 (0.56)</td>
</tr>
</tbody>
</table>

D.3 Robustness Checks for the Adaptive Learning Model

As described in section 4.4, the adaptive learning model was implemented under the truncation approach. This section presents results from the imputation approach in which imputed data from 1966 to 1996 are used in the belief estimation. Figure 18 show the beliefs for the Asia-Europe market for the adaptive learning model with $\lambda = 0.02$. The beliefs under the truncation and the imputation approaches are closer to one another especially for the period where the main analysis lies from 2006 to 2014. The model fits under the two approaches are also close to one another, although they are better under the truncation approach especially for the correlation between demand and investment as shown in table 17.

Table 17: Data Moments and Simulated Moments under the Truncation and Imputation Approaches

<table>
<thead>
<tr>
<th>Data Truncation</th>
<th>Imputation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average investment in 2006-2008 (1 mil. TEU)</td>
<td>0.23</td>
</tr>
<tr>
<td>Average investment in 2009-2014 (1 mil. TEU)</td>
<td>0.14</td>
</tr>
<tr>
<td>Total capacity of owned ships (1 mil. TEU)</td>
<td>5.09</td>
</tr>
<tr>
<td>Total capacity in the order book (1 mil. TEU)</td>
<td>3.07</td>
</tr>
<tr>
<td>Correlation between demand and investment</td>
<td>0.19</td>
</tr>
<tr>
<td>Std. dev. in investment (1 mil. TEU)</td>
<td>0.17</td>
</tr>
</tbody>
</table>

Notes: This table compares moments observed in the data and moments simulated under the truncation and imputation approaches of the baseline learning model.
Figure 18: Beliefs under Adaptive Learning Based on Two Alternative Approaches

Notes: This figure shows firms’ beliefs about future demand under adaptive learning estimated with the truncation approach and the imputation approach, respectively, for the case of $\lambda_t = 0.02$. Beliefs for 2006-2014 in the shaded area are used in the main analysis.