Learning and Investment under Demand Uncertainty in Container Shipping*

Jihye Jeon†

Job Market Paper

November 9, 2016

Click here for latest version.

Abstract

This paper develops a dynamic oligopoly model of investment to investigate how firms invest under demand fluctuations and what drives boom-and-bust cycles of investment. I depart from the standard assumption that firms know the true model of demand and its parameters. Instead, I allow firms to form and revise expectations about demand based on information available at each decision-making moment. I estimate the model using firm-level data from the container shipping industry. Results show that a model with learning successfully predicts the boom-bust investment patterns observed in the data, while a full-information model fails to do so. Counterfactual experiments reveal that (i) strategic incentives play an important role in creating oversupply, as well as increasing the volatility of investment; (ii) scrapping subsidies can reduce excess capacity but cause a loss in consumer surplus; and (iii) under learning higher demand volatility leads to more drastic revisions of beliefs, which amplifies investment boom-bust cycles. I show that the regulator's modeling choice for firms' expectations has important policy implications, namely in merger evaluation.

*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, Michael Dickstein, Myrto Kalouptsidi, Kei Kawai, Peter Newberry, 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, New York University Stern School of Business; jjeon@stern.nyu.edu.

1 Introduction

The container shipping industry, like many other capital-intensive industries, experiences large fluctuations in demand and investment. Since short-run supply is inelastic due to time-to-build and pre-announced schedules, changes in demand conditions often lead to temporary supply-demand imbalances and large changes in shipping rates.[1] In the mid-2000's, when world trade was booming, 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. 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 develops a dynamic oligopoly model of firm investment under demand fluctuations. I depart from the standard full-information assumption of rational expectations that firms know the true model of demand and its parameters. Instead, I allow agents to form and revise expectations about demand based on information available at each decision-making moment. In this model, dynamic strategic interaction among firms and firms' beliefs about future demand determine the timing and quantity of investment and scrapping.

I estimate the model using firm-level data from the container shipping industry to examine whether a model with learning can explain boom and bust cycles in investment better than a full-information model. Based on the model estimates, I conduct counterfactual simulations with respect to competition, the scrap price, and demand volatility. These counterfactuals serve three purposes: first, to understand how strategic incentives, the irreversibility of investment, and demand volatility affect investment cycles and industry outcomes; second, to evaluate the welfare

[1] Shipping firms face time-to-build since there is a lag between the order and delivery of a new ship. Time-to-build in the container shipping industry ranges from 2 to 4 years.

[2] "The hidden opportunity in container shipping", accessed on January 11, 2016. http://www.mckinsey.com/insights/corporate_finance/the_hidden_opportunity_in_container_shipping

[3] "Maersk Line CEO: We Misjudged Container-Shipping Demand", accessed on January 11, 2016. http://www.wsj.com/articles/SB10001424052702303342104579098680549111434.

implications of relevant policy interventions; third, to understand the extent to which the modeling choice for firms' expectations matters in policy evaluation.

Modeling agents' learning. There is a large literature on firm investment under demand uncertainty focusing on factors such as irreversibility (e.g. Bernanke (1983), Pindyck (1988)); time-to-build (e.g. Kydland & Prescott (1982)); and strategic considerations (e.g. Besanko *et al.* (2010)). With rare exceptions, these papers adopt the full-information assumption of rational expectations: firms know the true stochastic process of demand.[4] Under this assumption, the standard empirical approach—sometimes referred to as *rational expectations econometrics*—involves estimating the demand process using as much data as available to the researcher and equating agents' beliefs to the estimated process. The full-information assumption may be too restrictive for settings in which agents have limited information about demand, however. For example, agents may be relatively new to the industry or the environment may be changing due to policy changes or exogenous shocks.

This paper proposes a different methodology. Building on a fast-growing macroeconomics literature (e.g. Cogley & Sargent (2005), Orlik & Veldkamp (2014)), the model allows firms' perceptions about demand to change through learning. Specifically, agent beliefs about demand follow an autoregressive model with parameters that are re-estimated in each period based on available historical data. Instead of imposing a particular learning rule, I consider different models of expectation formation and allow data to determine which model best rationalizes observed investment patterns. I consider Bayesian and adaptive learning models; and within the latter I allow for various relative weights on past observations.[5]

I estimate that an adaptive learning model that places 45% weight on a 10-year-old observation relative to the most current observation provides the best fit to the data.[6] Compared to this, the full-information model with no learning, commonly used in dynamic empirical IO, provides a worse fit and implies different predictions: it predicts that firms withhold investment during demand boom years and suffer less from overcapacity when faced with downturns in demand.[7] Moreover, the full-information model predicts that total investment is lower by 17%, and the volatility of investment lower by 22%, compared to the learning model. In terms of welfare, it implies that

[4]Greenwood & Hanson (2015) provide an exception. They adopt a behavioral approach in studying investment cycles in the bulk shipping industry. They assume that firms overextrapolate exogenous demand shocks.

[5]In particular, I consider cases in which agents put heavier weights on more recent observations, which is often referred to as *constant-gain learning* in the literature. This would be a natural way to model firms' beliefs if firms believed that the demand process changes over time (Evans & Honkapohja (2012)). I estimate the parameter that governs the rate at which firms discount older observations.

[6]This estimate is very close to estimates in 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.* (2014) also find that firms weight recent play disproportionately when forming expectations about competitors' play.

[7]In my setting, a full-information model even predicts a negative relationship between investment and demand. In times of high demand, firms place more orders and there are more backlogs at shipyards, which increases the shipbuilding price and in turn discourages investment. Under full information this negative supply-side effect dominates the positive demand-side effect arising from the fact that high demand is likely to persist.

producer surplus is greater by 85%, and consumer surplus lower by 3% under full information.

Counterfactual experiments. I use my estimated model to perform a series of counterfactual experiments that address various firm-strategy and public-policy issues. The first set of counterfactuals pertain to industry consolidation and highlight the effect of competition and strategic incentives on investment levels and cycles as well as welfare. This is important to understand given theoretical predictions and the industry's trend toward consolidation. A large body of industrial organization research predicts that strategic incentives such as business stealing and preemption can lead to overinvestment (e.g. Mankiw & Whinston (1986), Spence (1977)). Considering that over 40% of the total operator-owned capacity is concentrated in the top three firms, this would seem a relevant issue in the shipping industry. Moreover, the industry is moving towards consolidation in response to fierce competition to fill ships. The top two firms in the industry formed a vessel-sharing agreement (the "2M Alliance") in 2014, and China's two biggest shipping lines also plan to merge. Given the nature of the industry, investment incentives should be an important dimension in merger evaluation.

Motivated by these considerations, I perform a counterfactual experiment whereby the industry becomes monopolized. I predict that investment drops by 34%, and the volatility of investment decreases by 22%. Producer surplus increases by $92 billion both from reduced shipbuilding costs and from higher prices, whereas consumer surplus drops by $42 billion in the Asia-Europe market over the period of 2006 to 2014.[8]

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 lead to a consumer surplus loss but a total welfare gain.[9] From a methodology point of view, I show that a full-information model predicts much smaller changes in investment levels. Considering the efficiency gains provided by investment coordination, I predict that policy that is based on a full-information model is more likely to block welfare-enhancing mergers.

The second counterfactual simulates a ship-scrapping subsidy program. China initiated such a program in 2013, in an effort to support firms struggling with excess capacity. It grants 1500 yuan (approximately $US 240) per gross ton of scrapped old vessels. Scrapping programs have several effects: they encourage firms to scrap when demand conditions become unfavorable, thus alleviating the oversupply problem; however, investment decision become less irreversible, which in turn encourages investment. The net effect is unclear. Counterfactual results show that a subsidy

[8] I also consider an alternative counterfactual of a merger between the top two firms. In this case, I predict investment decreases by 7.5%, producer surplus increases by $14 billion, and consumer surplus decreases by $1 billion.

[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.

program that makes a cash transfer for scrapping leads to more scrapping, especially in the post-crisis period from 2009, but also leads to a slight increase in investment. Overall, this implies a gain in producer surplus but a loss in consumer welfare from reduced supply.

The third counterfactual simulation examines the effect of demand volatility in the presence of learning. I find that demand volatility reduces investment, which is consistent with findings in previous empirical studies. But I also show that introducing learning opens an additional channel through which demand volatility affects investment: 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 and contribution. This paper builds on and contributes to an emerging field, mostly in macroeconomics, 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 Fed's changing beliefs in the monetary policy and Orlik & Veldkamp (2014)) study uncertainty shocks in the Bayesian learning framework. There are many other macroeconomics papers discussing the implications of learning (e.g. Milani (2007)). To my knowledge, this paper is the first to investigate 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.* (2014) examine learning about competitors' play and underlying demand parameters in the context of the UK electricity market. This paper complements this body of the literature by studying learning about aggregate market characteristics outside of a single-agent dynamic model.

This paper is closely related to empirical studies on investment cycles, especially two papers on the bulk shipping industry: Kalouptsidi (2014) and Greenwood & Hanson (2015). Bulk-shipping and container-shipping industries share similar characteristics such as exposure to fluctuations in trade demand, time-to-build, and capital-intensiveness. Compared to the bulk shipping industry, there is a significantly higher concentration of market power in the container shipping industry.[10] Also, container shippers operate according to fixed schedules, while bulk shippers operate on-demand services much like taxis.

The more important difference between my work and the two studies on bulk shipping lies in

[10]Kalouptsidi (2014) assumes that each firm owns on ship only and develops a competitive model of the bulk-shipping industry.

the methodology. Kalouptsidi (2014) employs a fully rational model and uses second-hand ship prices to identify values of owning a ship non-parametrically. The second-hand prices that she uses already reflect sellers' and buyers' beliefs about future demand. As a result, Kalouptsidi indirectly incorporates firms' beliefs about industry demand in the estimation of values of owning ships. By contrast, this study models firms' forecasting process explicitly, and searchers for a model of firm beliefs that is consistent with observed investment patterns. 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.[11] Furthermore, understanding how firms form expectations is interesting in its own right. Greenwood & Hanson (2015) provide an exception to studies based on full-information assumptions. They introduce behavioral biases in persistence in earnings and long-run endogenous supply responses by rivals to explain bulk shippers' investment behavior. In contrast to Greenwood & Hanson, this study does not require biases in firm beliefs. In particular, firms' beliefs about rivals' actions are consistent with the rivals' equilibrium strategies. Lastly, my paper allows strategic interaction among firms through a dynamic oligopoly framework.

Methodologically, this paper makes a contribution to the literature on the structural analysis of industry dynamics: it provides a learning framework that incorporates firms' changing beliefs in a dynamic oligopoly setting (for an overview of the literature, see Ackerberg *et al.* (2007) and Doraszelski & Pakes (2007)). Incorporating learning intensifies the computational burden associated with solving a dynamic model with many firms.[12] Under learning, firms' beliefs change over time as firms receive new data, which means that firms' value function changes over time as well. Therefore, equilibrium needs to be solved separately for each period in time. The industry does not have a large number of independent markets, which precludes the use of the two-step estimator that has become a popular strategy in estimating dynamic oligopoly models. To address this issue, this paper adopts an equilibrium concept that restricts firms' information sets based on the moment-based Markov equilibrium (MME) notion recently proposed by Ifrach & Weintraub (2016) and the experience-based equilibrium (EBE) notion by Fershtman & Pakes (2012).[13] This approach vastly reduces the state space size, while still capturing strategic interaction among firms.

The empirical industrial organization literature analyzing firms' investment problems has typically focused on recovering investment costs, entry costs, or exit values that observed data imply

[11] Adverse selection may arise in the second-hand market if sellers privately observe the quality of the goods. If there is a 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 form second-hand prices will lead to biased estimates.

[12] There are about 17 active firms in my application. Even a specification with a single state variable that can take up to 5 different values would result in over a billion states.

[13] MME 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).

given other primitives (see, for example, Ryan (2012) and Collard-Wexler (2013)).[14] In this paper, I adopt a different estimation strategy: I use detailed data on shipbuilding prices and scrap prices to estimate investment costs and scrap values directly. The use of shipbuilding and scrap prices allows me to focus on identifying the learning model that can best rationalize the observed data.

This paper is also related to the body of empirical studies that quantify the effect of demand uncertainty on investment (e.g. Guiso & Parigi (1999), Bloom (2009), Collard-Wexler (2013), Kalouptsidi (2014), and Kellogg (2014)). The contribution to this line of research is shedding light on the informational channel through which demand fluctuations can affect investment.

The remainder of the paper is organized as follows. Section 2 describes the industry and the data, and provides some preliminary evidence of learning. Section 3 presents the dynamic model of investment with learning about the aggregate demand. Section 4 discusses the empirical implementation of different models of beliefs and diagnoses beliefs that are implied by different models based on GDP forecast data. Section 5 describes the estimation procedure and presents estimation results. Section 6 discusses counterfactual experiments and section 7 concludes.

2 Industry and Data

This section describes key features of the container shipping industry and gives an overview of the data. It then moves to presenting preliminary evidence of learning based on a test of a structural break in firms' investment policy function.

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, 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.[15] Container carriers also rely on chartered vessels, which are leased out by third parties. Chartered vessels account for approximately

[14] The exception is Kalouptsidi (2014), who non-parametrically estimates values of owning a ship using second-hand sales prices.

[15] The construction of new ships happens at shipyards. There are about 300 major shipyards and many smaller ones.

50 percent of the total container ship capacity operated by the largest 20 firms. The majority of charter contracts for container ships time charters which involves 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 move their idle ships into lay-up, container companies must stick to their 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.[16] Hence, the ability to make correct forecasts about future demand and invest accordingly is important in this industry.

Figure 1 shows the industry-level quarterly investment and investment price for 2001 to 2014.[17] Investment is concentrated in the times of high investment costs. Despite the high investment cost from 2006 to 2008, which is on average 42% higher than the average cost for the 2009-2014 period, the quarterly investment is higher by more than 60% in the 2006-2008 period.

2.2 Data

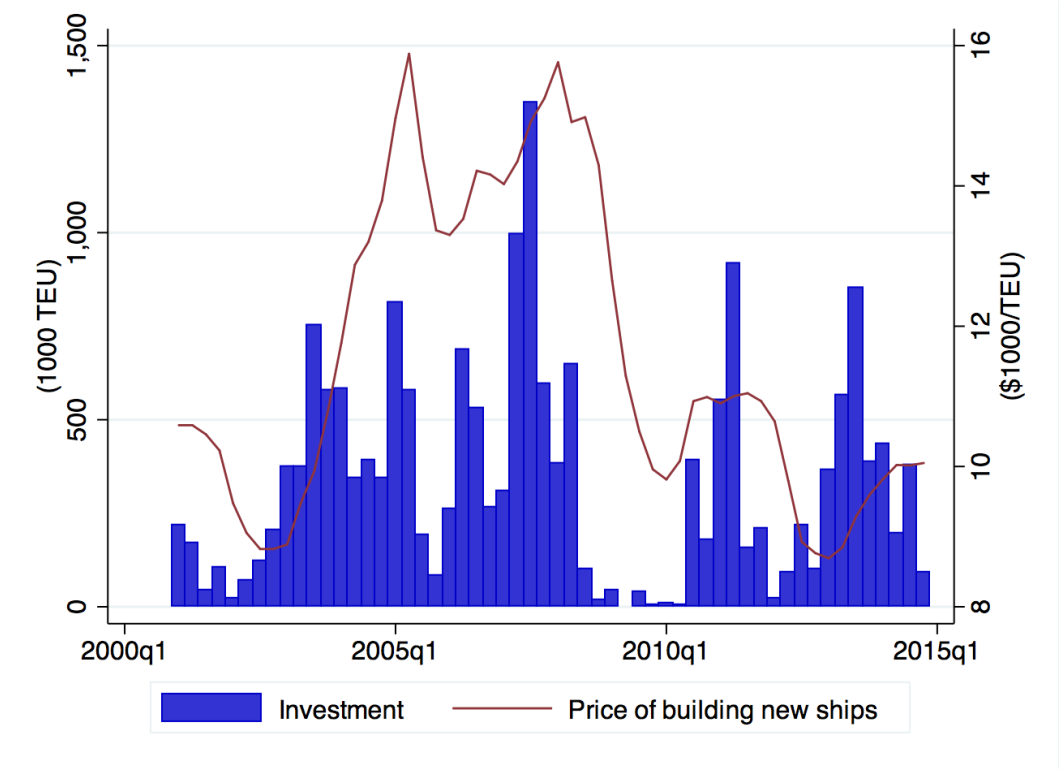
This project uses two main datasets on the container shipping industry. The first dataset combines data collected from two sources: MDS Transmodal, a U.K.-based research company, and Clarksons Research, a U.K.-based ship-broking and research company. This dataset covers quarterly information from 2006 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) charter rates, scrap prices, and shipbuilding prices.

To analyze the data for the period from 2006 to 2014, I need to estimate firms' beliefs for that period, which requires historical price and quantity data that go much 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

[16] 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)).

[17] The average shipbuilding price is constructed based on the prices of building a new ship by size category (2500 TEU, 3700 TEU, 6700 TEU, 8800 TEU, 10000 TEU, 13500 TEU) and the number of ships in the total industry order book by size category. I first obtain a per TEU price for each size category by dividing the ship-level price by the ship size. Then, I construct a weighted average of the per TEU price based on the size distribution. The average scrap value is constructed in a similar way.

Figure 1: Total investment and investment cost



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

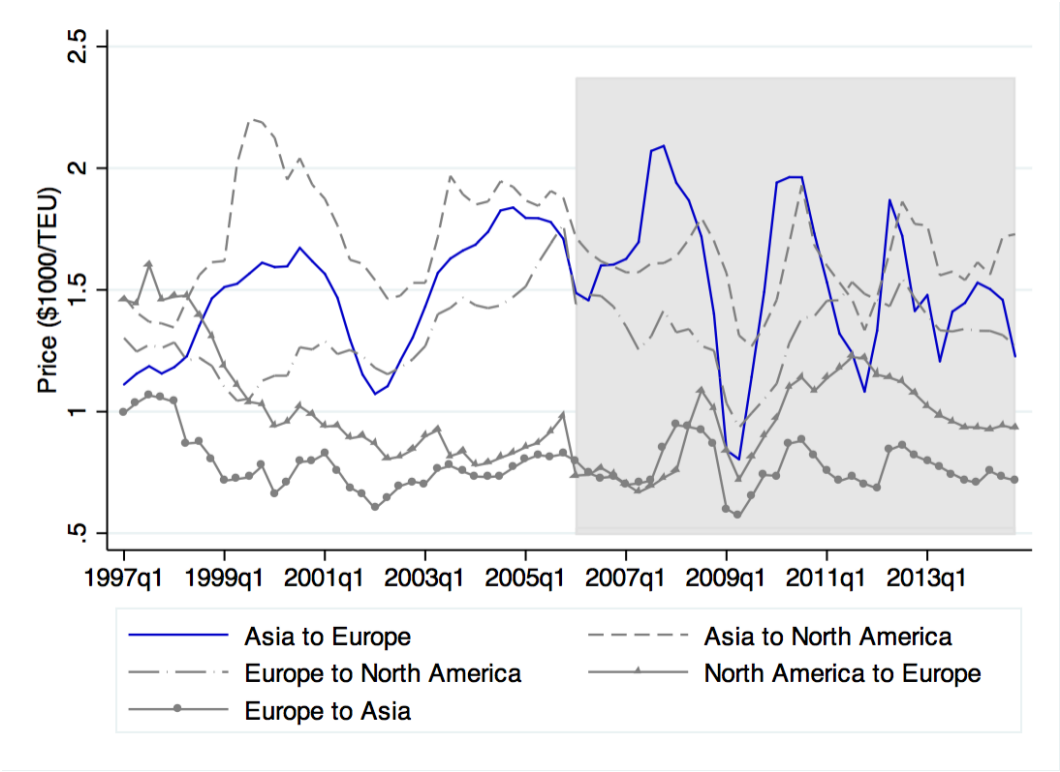
of Maritime Transport published by the United Nations that goes back to 1997.[18] It contains information on the average freight rate and cargo flows on major routes. The volume of trade is available at the yearly level in this dataset, although the price level is available at the quarterly level. The 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 and the assumption that in a given year the quarterly container trade volume is proportional to the value of trade.

The analysis in this paper focuses on the Asia-Europe (A-E) market but also accounts for demand conditions in other markets. In practice, firms have to choose which route to operate on and how much capacity to deploy on each of the routes they operate on. However, it is computationally infeasible to endogenize capacity deployment decisions in every market in this model since there are a large number of firms and the markets are not independent of one another. To make the model tractable, firms are allowed to choose how much capacity to deploy in the Asia-Europe market and the "outside market" which includes all other markets. The demand conditions in

[18] Although this is roughly the start date of the official public data on the aggregate price and quantity of container trade, firms may have data that go further back than 1997 and use them in forming expectations. Subsection 4.1 discusses my empirical strategy in estimating firms' beliefs given the truncated nature of the price and quantity data.

the outside market are estimated from the data from the Asia-North America (A-NA) and Europe-North America (E-NA) markets. The A-NA and E-NA markets along with the A-E market account over 50 % of all interregional trade by volume and 67% by deployed capacity.[19]

Figure 2: Prices on major trade routes



Notes: This figure shows quarterly average prices on major container trade routes from 1997 to 2014. The shaded area covers the period on which the main analysis lies from 2006 to 2014.

The reasons the analysis focuses on the Asia-Europe market are the following. First, the Asia-Europe market is the largest market in the container shipping industry accounting for over 23% of the total interregional container trade by volume and close to 40% in terms of the deployed ship capacity. Second, it was most heavily impacted by the downturns in 2008, so the effect of learning is likely to be more pronounced in this market. Figure 2 shows the fluctuations in the average prices on the five major trade routes by trade volume from 1997 to 2014. The shaded area covers the 2006-2014 period, which is where the main analysis lies. The price fell by over 50 percent from the peak in 2007 to the trough in 2009 on the Asia to Europe route while it fell by less than 30 percent on the Asia to North America route, which is the second largest route. Nonetheless, the model and the empirical analysis account for the aggregate demand state on trade routes outside of

[19]The estimation of the outside market demand uses data from the A-NA and E-NA markets only, since the price and quantity data going back to 1997 are available for the four routes in these two markets. The longer time-series data allow me to estimate firms' beliefs under learning models for the period of interest, which is from 2006 to 2014.

the Asia-Europe market. In addition, the model allows learning with respect to the outside market demand as well as demand in the Asia-Europe market.

Table 1: Descriptive statistics

		Mean	Std. Dev.	Min	Max
Industry-level data (2006-2014)					
Shipbuilding price ($/TEU)		11.62	2.22	8.69	15.76
Scrap price ($/TEU)		2.62	0.55	1.50	3.81
Market-level data (1997-2014)					
Asia to Europe	Quantity (1 mil. TEU)	2.37	1.10	0.70	3.98
	Price ($1000/TEU)	1.51	0.28	0.80	2.09
Europe to Asia	Quantity (1 mil. TEU)	1.08	0.39	0.51	1.76
	Price ($1000/TEU)	0.78	0.10	0.57	1.07
Other routes	Quantity (1 mil. TEU)	5.11	1.41	2.80	7.72
	Price ($1000/TEU)	1.36	0.12	1.06	1.60
Firm-level data (2006-2014)					
Capacity of owned ships (1m TEU)		0.30	0.25	0.04	1.47
Capacity of ships in order book (1m TEU)		0.18	0.13	0.00	0.64
Capacity of chartered ships (1m TEU)		0.31	0.29	0.01	1.55
Capacity of ships deployed in Asia-Europe market (1m TEU)		0.22	0.19	0.04	0.99

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.

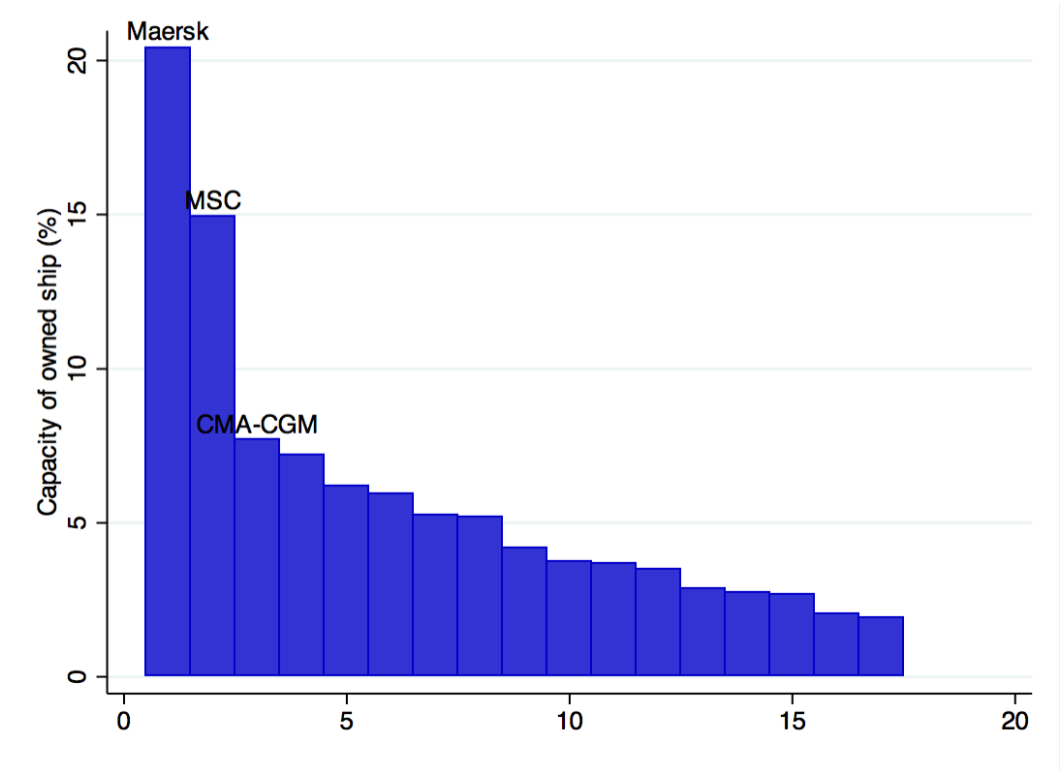
The analysis is further restricted to firms that deployed over 80,000 TEU of ships in the Asia-Europe market on average in the 2006 to 2014 period. These firms account for more than 95 percent of the total capacity of ships deployed in the Asia-Europe market. This results in a quarterly panel of 17 firms from 2006 to 2014. There is no entry into or exit from the Asia-Europe market 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 a book order capacity of 180,000 TEU.

Figure 3 shows the distribution of firm size measured as the share of owned capacity, 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)).[20] While there is considerable size variation among the top two firms, the rest

[20]Kalouptsidi (2014) shows that the largest fleet share is 3% for Handysize bulk carriers.

of the firms are similar in size.[21]

Figure 3: The distribution of firm size



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

2.3 Preliminary Evidence of Learning

It is inherently difficult to test whether firms' beliefs about the demand process are changing since the beliefs are not directly observed. Hence, I search for suggestive evidence of learning by relying on the difference in the predictions that a full-information model and a learning model make. A learning model generally predicts that conditional on being in the same state (where the state summarizes all payoff relevant variables), firms' beliefs will be different before and after experiencing large demand shocks.[22] Therefore, under learning firms' investment policy in a given state may change after firms observe large shocks. For example, after undergoing periods of positive demand shocks, firms may be more likely to invest as their outlook about future demand become more optimistic. 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

[21] The Herfindahl index for the industry is 0.097 when these 17 firms are accounted for.

[22] The 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).

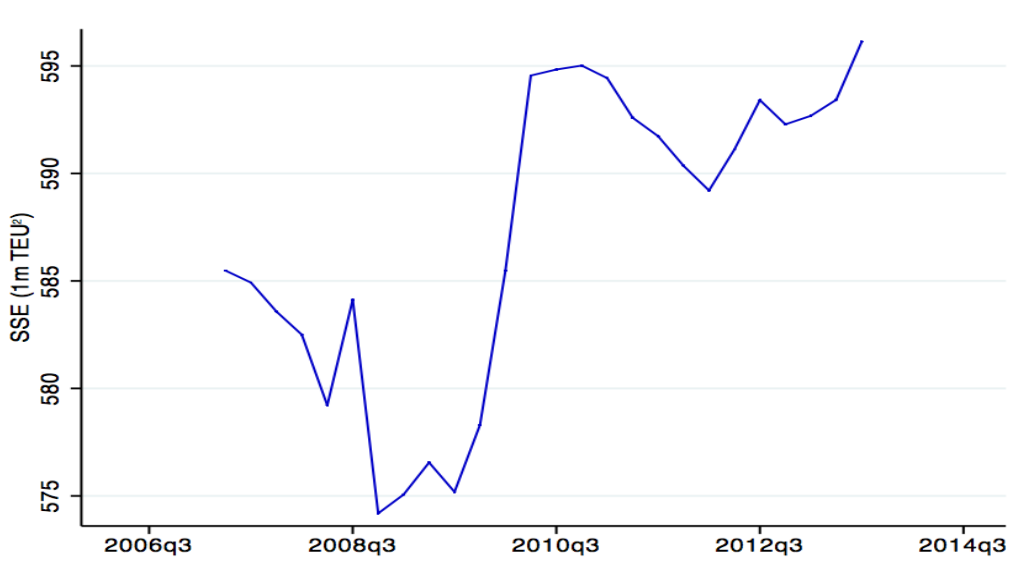
contain any new information. Hence, I examine whether firms' investment behavior change significantly after large demand shocks. In particular, I test for a structural break in the firm's investment policy function with an unknown break date following the approach proposed by Andrews (1993) closely.

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 the capacity of new investment, and x_t includes state variables. The detailed specification is the state space is given in subsection 3.2. 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 order book capacity, the industry state including the aggregate capacity of operator-owned ships and the aggregate order book capacity. The demand states are recovered through demand estimation as given in subsection 5.1.

Figure 4: Estimation of the Break Date in Investment Policy



Notes: This figure plots the sum of squared residuals as defined in equation (1) for each break date from 2007:Q3 to 2013:Q2.

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

$$S(\beta, \bar{t}) = \frac{1}{T} \sum_t (y_t - \beta_1' x_t \mathbb{I}(t < \bar{t}) - \beta_2' x_t \mathbb{I}(t \geq \bar{t}))^2. \tag{1}$$

The break dates from 2007:Q2 to 2013:Q3 are considered as I need sufficient observations before and after the break date to estimate the equation. Figure 4 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 late 2008.

Table 2: Investment Policy Estimation and a Test for 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		
$H_0 : \beta_1 = \beta_2$		
Test statistics	4.38	
p-value	(0.0001)	

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

Table 2 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 test of a structural break in the investment policy function. The test result rejects the null that the investment policy is the same before and after the last quarter of 2008. The results from the regression suggest that in the post-2008 period firms are more responsive to the aggregate industry capacity. 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.

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 information available to them in the current period they are in. In each period a firm decides whether to invest in new ships and whether to scrap existing ships based on its own capital, its order book capacity, and its 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 and how much capacity to station in each market. I start by describing models of firm beliefs in subsection 3.1. Subsection 3.2 presents firms' dynamic problem, and subsection 3.3 demand for container shipping services and product market competition. Subsection 3.4 provides a definition of equilibrium.

3.1 Models of Firms' Beliefs about Demand

This section presents models of how firms form expectations about demand. In a full-information model, agents are assumed to know the true law of motion for demand. Learning models assume that agents form expectations about demand based on information available to them in each period. In other words, agents operate like econometricians who estimate the parameters of the model based on best information at their disposal and make forecasts using their estimates.

This paper considers different models of learning and searches for the model that can best rationalize the data instead of imposing one particular learning rule. This is because the literature provides little guidance on which learning framework should be used. Furthermore, empirically firms in different industries may have different models of demand forecasts depending on, for example, how long they have been in the industry and how fast firms perceive the demand process to be changing. This subsection first presents the full-information model. Then, it moves to present models of adaptive and Bayesian learning. Appendix A.2 also introduces a full-information model in which time-varying volatility is allowed through a GARCH process. The constant volatility model is chosen as the main full-information model for two reasons. First, it serves as a more valid comparison to the learning model which is also based on a constant volatility autoregressive model. Second, most papers in empirical industrial organization have considered models with constant volatility, for example, for example, through autoregressive processes.[23] Therefore, the comparison

[23] Markov processes are widely used as well. In this case, volatility is time-varying but is constant conditional on the state.

between a full-information model with constant volatility and a learning model will highlight the potential improvement that can be achieved by incorporating learning.

Full Information

Agents contemplate a first-order autoregressive model for demand in the Asia to Europe market given as the following:

$$\begin{aligned} z_t &= \rho^0 + \rho^1 z_{t-1} + \omega_t \\ &= \rho' x_t + \omega_t \end{aligned} \tag{2}$$

where z_t is the demand state in the Asia-Europe market, $\omega_t \sim N(0, \sigma_t^2), \rho = [\rho^0, \rho^1]'$, and $x_t = [1, z_{t-1}]'$. Similarly, the model for the demand in the outside market ($\tilde{z}_t$) is given as:

$$\begin{aligned} \tilde{z}_t &= \tilde{\rho}^0 + \tilde{\rho}^1 \tilde{z}_{t-1} + \tilde{\omega}_t \\ &= \tilde{\rho}' \tilde{x}_t + \tilde{\omega}_t \end{aligned} \tag{3}$$

where $\tilde{\omega}_t \sim N(0, \tilde{\sigma}_t^2), \tilde{\rho} = [\tilde{\rho}^0, \tilde{\rho}^1]'$, 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.

Adaptive Learning

Under adaptive learning firms consider the same $AR(1)$ model as given in equations (2) and (3). Unlike in the full-information model, agents revise expectations by re-estimating the parameters of the $AR(1)$ model in each period based on the information set that includes 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 $\eta_t = (\rho_t^0, \rho_t^1, \sigma_t, \tilde{\rho}_t^0, \tilde{\rho}_t^1, \tilde{\sigma}_t)$.

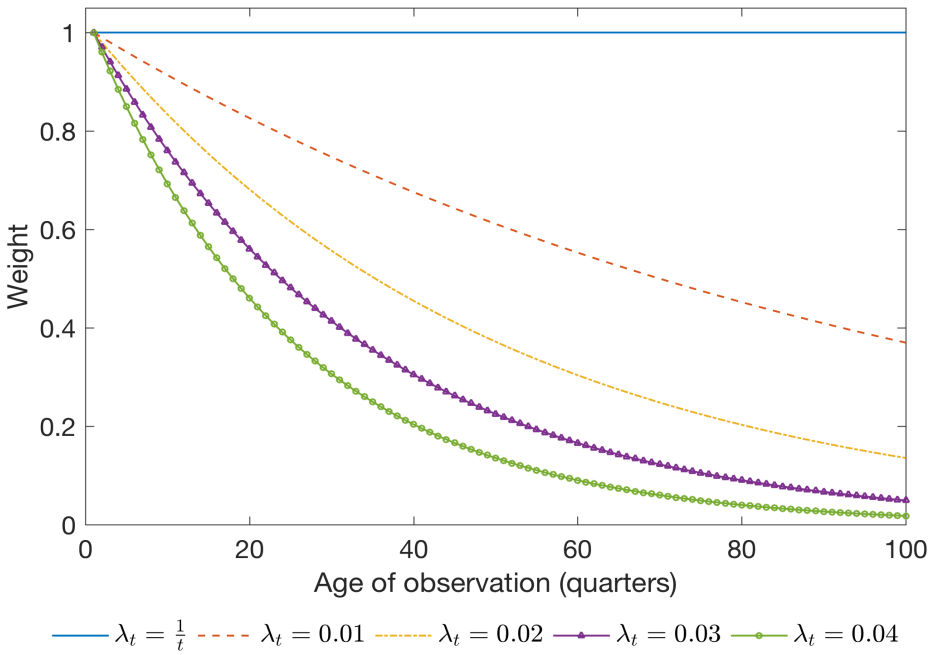
Since the prices and quantities of container trade are public information periodically published in trade journals and other publications, beliefs about the aggregate demand are assumed to be homogenous across all firms. Heterogeneity in firms' beliefs would arise if firms experienced different demand shocks, for example through different customer pools. In container trade, swings in global trade common to all firms are the main source of demand shocks rather than firm-specific demand shocks. At the practical level, there are no publicly available data that provide information on firm-level demand to my knowledge, which are necessary to allow heterogenous firm beliefs. Nevertheless, how heterogeneous firm beliefs affect firm decisions and industry dynamics is an interesting area of study for future work.

The model also assumes that agents do not internalize the possibility of learning in the future; this type of learning is sometimes referred to as *myopic learning* in the literature.[24] The myopic

[24]Suppose information is endogenous to agents' decisions, for example, because agents are making consumption

learning assumption has two behavioral interpretations. The first interpretation is that agents believe current beliefs to be the correct forecasts for future demand. The alternative interpretation is that agents use current beliefs in forecasting as they approximate means of future beliefs.

Figure 5: Weights on observations under adaptive learning



Nots: This figure plots weights that are applied to observations by λ_t in the adaptive learning model where λ_t is the weight parameter that governs how responsive estimate revisions are to new data. (See equation (4)).

Let $X_t = [x_0, x_1, ..., x_t]'$ and $R_t = \frac{X_t' X_t}{t}$. The expectations at time t regarding the Asia-Europe market demand can be written recursively as

$$\rho_t = \rho_{t-1} + \lambda_t (R_t)^{-1} x_t \left(z_t - \rho_{t-1}' x_t\right) \tag{4}$$

$$R_t = R_{t-1} + \lambda_t (x_t x_t' - R_{t-1}) \tag{5}$$

where λ_t is the weight parameter that governs how responsive the estimate revisions are to new data. I consider the case with $\lambda_t = \frac{1}{t}$ in which agents put equal weight on all observations in their information set. I also consider cases in which λ_t is some constant between 0 and 1, which is refereed to as *constant-gain learning*. This gives rise to weights that geometrically decline with the age of

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 (investment and scrapping), there is no room for experimentation regardless of assumptions on learning.

observations. In other words, agents assign heavier weights to more recent observations. This would be the natural way to form expectations if agents were concerned about the possibility of structural changes as Evans & Honkapohja (2012) point out. A larger value of λ_t leads to heavier discounting of older observations. The set of values of λ_t considered is $\{\frac{1}{t}, 0.01, 0.02, 0.03, 0.04\}$. Figure 5 plots weights placed on observations from different periods for each value of λ_t. 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.

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, $(\rho_t^0, \rho_t^1, \sigma_t, \tilde{\rho}_t^0, \tilde{\rho}_t^1, \tilde{\sigma}_t)$. 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 priors of σ^2 follow an inverse Gamma distribution. Then, the posterior distribution $\rho_t \sim N(\mu_t, \Sigma_t)$ has the mean and the variance given by

$$\begin{aligned}\mu_t &= \Sigma_t \left(\Sigma_0^{-1}\mu_0 + \sigma^{-2}(X_t' Z_t)\right)\\ \Sigma_t &= \left(\Sigma_0^{-1} + \sigma^{-2}(X_t' X_t)\right)^{-1}.\end{aligned}$$

The beliefs are defined similarly for the outside market.

3.2 Firms' Dynamic Problem

Time is discrete with an infinite horizon and is denoted by $t \in \{0, 1, 2, ...\}$. There are n incumbent firms and the set of incumbent firms is denoted by $N = \{1, 2, ..., n\}$. 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 capacity of firm i's order book.[25] 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 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

[25] The owned capacity space denoted by $\mathcal{K}$ is discretized into 19 points such that $\mathcal{K} = \{k_0, k_1, k_2, ..., k_{18}\}$ and the order book capacity space denoted by $\mathcal{B}$ into 7 points such that $\mathcal{B} = \{b_0, b_1, ..., b_6\}$. $\mathcal{K}$ 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.

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 firms 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.[26] Therefore, I consider an alternative equilibrium concept which can be viewed in the context of the moment-based Markov Equilibrium (MME) introduced by Ifrach & Weintraub (2016), or more broadly the experience-based equilibrium (EBE) developed by 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 to further reduce the size of the state space. Firms' strategies thus depend only 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 A.4, I compute a model that allows firms to keep track of richer information by adding a dominant firm's state into the moment-based industry state and show that model predictions are robust to this change.

Firms make an investment decision ($\iota_{it} \in \{0, 1\}$) and a scrapping decision ($\delta_{it} \in \{0, 1\}$) in order to maximize expected discounted profits. I denote the strategy profile as $\mu_{it} = (\iota_{it}, \delta_{it})$. 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 variables, $\kappa(\hat{s}_t)$, and a privately observed part of the cost, ε^{ι}_{it}. The private cost shock is assumed to be an iid draw from a normal distribution, $\varepsilon^{\iota}_{it} \sim N(0, (\sigma^{\iota})^2)$. If a firm decides to scrap its ships or if there is depreciation, the firm's capital decreases by one unit and the firm receives a scrap value. The scrap value is the sum of the value common to all firms $\phi(\hat{s}_t)$ and an iid private value distributed as $\varepsilon^{\delta}_{it} \sim N(0, (\sigma^{\delta})^2)$. Deprecation occurs with a probability proportional to the firm's current capital amount, given as ζk_{it} for some constant ζ. If a firm scraps its vessels, there is no depreciation in the same period such that the maximum reduction in k_{it} from depreciation or scrapping is one unit.[27] I denote as $\nu(\delta_{it}, x_{it})$ the

[26] 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.

[27] This assumption is made since the data do not provide any observations of a decrease of capital by more than one unit. The interpretation of this assumption can be that when a firm decides to scrap its vessels, it chooses the

expected amount of reduction in capital from depreciation or scrapping before the relization of the depreciation outcome such that $\nu(\delta_{it}, x_{it})$ is one if $\delta_{it} = 1$ 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

$$V^{\eta_t}(x_{it}, \hat{s}_t) = \max_{\iota_{it}, \delta_{it}} \pi(x_{it}, \hat{s}_t) - \iota_{it}(\kappa(\hat{s}_t) + \varepsilon^{\iota}_{it}) + \nu(\delta_{it}, x_{it})(\phi(\hat{s}_t) + \varepsilon^{\delta}_{it}) + \beta E\left[V^{\eta_t}(x_{t+1}, \hat{s}_{t+1} | s_t, \hat{s}_t)\right]$$

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 $d_t = (z_t, \tilde{z}_t)$ will evolve.

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.[28] The model incorporates firm heterogeneity in other areas, however, as it is likely to 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 or may have better access to financing. Second, the marginal cost of production is allowed to vary by the firm's deployed capacity where the deployed capacity is highly correlated with the firm size. The detailed specification of these cost functions are shown in section 3.3.

State Transitions

The transitions of the firm-specific state and the aggregate demand state are Markov processes. 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. When there is scrapping, the firm's owned capacity decreases by one unit in case there is no delivery or stays the same in case there is delivery. There is also a chance of depreciation, in which case the firm's own capacity decreases by one unit. The transition of the firm-specific state is described as:

$$k_{it+1} = k_{it} + \tau_{it} - \min(\delta_{it} + \psi_{it}, 1)$$
$$b_{it+1} = b_{it} + \mu_{it} - \tau_{it}$$

oldest vessels that are about to deprecate on their own. This assumption can be easily relaxed.

[28] It is possible that there exists unobserved persistent heterogeneity, but the data are not rich enough to allow for this type of heterogeneity.

where τ_{it} is delivery and ψ_{it} is depreciation. The probability of delivery is a linear function of the firm's order book capacity. The delivery process is given as

$$\tau_{it} = \begin{cases} 1 & \text{with probability } \kappa b_{it} \\ 0 & \text{with probability } 1 - \kappa b_{it}. \end{cases} \tag{6}$$

The probability of depreciation linearly increases in the stock of capital and the stochastic depreciation process is given as:

$$\psi_{it} = \begin{cases} 1 & \text{with probability } \zeta k_{it} \\ 0 & \text{with probability } 1 - \zeta k_{it}. \end{cases} \tag{7}$$

The perceived evolution of the aggregate demand states for the Asia-Europe market and the outside market follows a first-order autoregressive process and is given as follows:

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

where $\omega_t \sim N(0, \sigma_t^2)$ and $\tilde{\omega}_t \sim N(0, \tilde{\sigma}_t^2)$.[29] The parameters in the $AR(1)$ model, $\eta_t = (\rho_t^0, \rho_t^1, \sigma_t, \tilde{\rho}_t^0, \tilde{\rho}_t^1, \tilde{\sigma}_t)$, summarize the beliefs about the evolution of future demand at time t. The model of firm beliefs governs how firms update these beliefs as they get new information as described in subsection 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. 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 ($\hat{s}_{t+1}$). To address this problem, I approximate the process for the moments using empirical transitions, based on the approaches proposed by Ifrach & Weintraub (2016) and Fershtman & Pakes (2012).

Let μ 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 μ' and its competitors use strategy μ. Then, a Markov process given by $\hat{P}_{\mu',\mu}$ approximates the non-Markov process of the moment-based state $\hat{s}_t$, given

[29] 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.

the strategy (μ', μ). Formally, such an operator is defined as Φ such that

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

In practice, the moment-based industry state's evolution is defined to be the long-run average observed transitions from the moment-based state in the current period to the moment in the next period under strategy μ as follows:

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

where $m_t = (\sum_i x_{it})$ is 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 (lease from a third party charterer), and (b) how much capacity to allocate to the Asia-Europe market and the outside market, given the state they are in. A firm's total capacity is given by the sum of owned capacity and chartered capacity and the firm can choose how much of its total capacity to allocate to the Asia-Europe market or the outside market. 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:

$$\log Q_{jt} = z_{jt} + \alpha_1 \log P_{jt} \tag{8}$$

where z_{jt} denotes the demand state, P_{jt} is the price, and Q_{jt} the quantity of route j at time t. The production cost function is convex in quantity and goes to infinity at the capacity constraint as follows:

$$c(q_{ijt}, \bar{q}_{it}) = \begin{cases} aq_{ijt} + \frac{bq_{ijt}^2}{2\bar{q}_{it}} & \text{if } q_{ijt} \leq \bar{q}_{it} \\ \infty & \text{otherwise} \end{cases}$$

where q_{ijt} is the firm-route-level quantity, and $\bar{q}_{it}$ is the capacity deployed in the Asia-Europe market.[30] This gives rise to the marginal cost of providing services on the Asia-Europe route is

[30] $\bar{q}_{it}$ is not index by route j as firms' deployed capacity in the opposite-direction routes in the same market is the same.

linearly increasing in quantity up to the firms' capacity constraint as follows:[31]

$$mc(q_{ijt}, \bar{q}_{it}) = \begin{cases} a + \frac{bq_{ijt}}{\bar{q}_{it}} & \text{if } q_{ijt} \leq \bar{q}_{it} \\ \infty & \text{otherwise.} \end{cases} \tag{9}$$

Then, the supply curve for route j is given as the horizontal sum of all firms' supply curves as follows:

$$P_{jt} = a + \frac{bQ_{jt}}{\bar{Q}_{jt}} \quad \text{if } Q_{jt} \leq \bar{Q}_t \tag{10}$$

where $\bar{Q}_t = \sum_i \bar{q}_{it}$.

In each period firms decide how much capacity to charter, $h_{it} \geq 0$, and how much capacity to station in the Asia-Europe market, $\bar{q}_{it} \geq 0$, given the state they are in. The capacity in the outside market is then given by

$$\tilde{q}_{it} = k_{it} + h_{it} - \bar{q}_{it}.$$

The price in the Asia-Europe market is determined by the intersection of the demand curve given in equation (8) and the supply curve given in equation (10).

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

$$\pi(x_{it}, \hat{s}_t) = \max_{\bar{q}_{it}, h_{it}} \Big\{ \sum_{j \in \{1,2\}} \big\{ P_{jt} q_{ijt} - c(q_{ijt}, \bar{q}_{it}) \big\} + R(\tilde{q}_{it}, \tilde{Q}_t, \hat{s}_t) - CC(h_{it}, x_{it}, \hat{s}_t) - FC \cdot k_{it} \Big\}. \tag{11}$$

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 function, and $\tilde{q}_{it}$ 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 such as docking fees, maintenance costs, canal dues, and port charges. This study does not explicitly model the chartering market and the product market competition in the outside market but accounts 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 subsection 5.2.

[31] This functional form assumption is based on the fact that it becomes increasingly hard to schedule loading and unloading as the ship reaches its full capacity.

3.4 Equilibrium

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

$$\hat{V}^{\eta}_{\mu',\mu}(x,\hat{s}) = \pi(x,\hat{s}) - \iota(\kappa(\hat{s}) + \varepsilon^{\iota}) + \max(\delta, \zeta k)(\phi(\hat{s}) + \varepsilon^{\delta}) + \beta E_{\mu',\mu}\hat{V}^{\eta}(x',\hat{s}'|x,\hat{s}).$$

The definition of an equilibrium is then given as follows.

Definition Equilibrium comprises of an investment and scrapping strategy μ that satisfies the following conditions:

(a) Firm strategies satisfy the optimality condition:

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

(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 A.1 describes the algorithm in detail.

4 Implementation and Diagnosis of Models of Firm Beliefs

This section begins by estimating demand for container shipping services. Then, I discuss how to implement different models of firm beliefs and presents firms' expectations about demand implied by each model. Once the expectations are recovered, I diagnose the models relying on the fact that trade demand and GDP are highly correlated. Specifically, I examine which of the models of firm beliefs considered in this study generates demand forecasts that are most consistent with the GDP forecast data.

4.1 Estimating Demand for Container 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 based on the estimates. The empirical analogue of the constant elasticity demand model in equation (8) is:

$$\log Q_{jt} = \alpha_0 + \alpha_1 \log P_{jt} + \alpha_2 W_{jt} + \varepsilon_{jt} \tag{12}$$

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 (12) using 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 is an important factor 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 on prices, quantities, and GDP for six major trade routes from 2001:Q2 to 2014:Q4.[32] 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. The demand parameters are identified by the time-series variation as well as the cross-sectional variation across six different routes in data along with the constant elasticity functional form assumption. In particular, since ships have to go around the routes that they serve, opposite-direction 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.

Table 3 presents results from the demand estimation. The first column presents results from the first stage regression, and the second column from the second stage. The price elasticity of demand is estimated to be -3.89, which 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.[33]

Given the elasticity of demand estimates, I construct the demand state variables for each trade route included in the demand estimation. The demand state variable z_{jt} for each j is taken to be the intercept of the demand curve as the following:

$$z_{jt} = \hat{\alpha}_0 + \hat{\alpha}_2 W_{jt} + \hat{\varepsilon}_{jt} \tag{13}$$

where $\{\hat{\alpha}_0, \hat{\alpha}_2\}$ are parameters estimated from the regression and $\hat{\varepsilon}_{jt}$ is the residual. Finally, aggregate demand states for the Asia-Europe market and the outside market are constructed 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

[32] Although the price, quantity, and GDP data available from 1997, the instruments are available starting from 2001:Q2.

[33] Stopford (2009) explains that container trade is price elastic because lowering prices encourages the subsitution of cheap foreign substitues 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 if a constant elasticity specification is chosen and -1.6 if a linear demand specification is chosen.

Table 3: IV Regression Results for Demand for Container Shipping

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

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

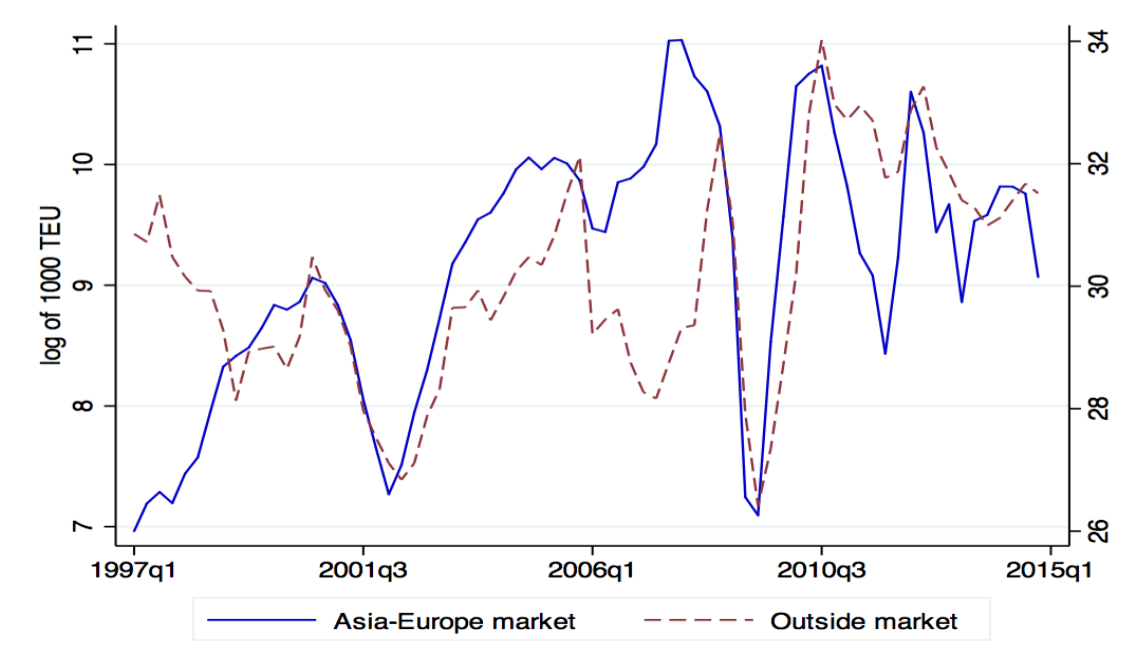
the demand states in the non-Asia-Europe routes. The demand state for the Asia-Europe market is denoted by z_t and the one for the outside market is denoted by $\tilde{z}_t$. Figure 6 plots the demand states for 1997 to 2014 for the Asia-Europe and the outside markets. There is a large drop in demand in the both markets at the end of 2008. In the Asia-Europe market, the boom and bust cycles in demand are shorter as well as larger in magnitude after 2008.

4.2 Empirical Implementation of Different Models of Firm Beliefs

This study considers a full-information model, adaptive learning models with different weighting rules, and a Bayesian learning model. Estimating beliefs under the full-information model involves estimating the demand process using the full sample of data. I apply least squares to estimate the $AR(1)$ processes as given in equations (2) and (3) using data from 1997:Q1 to 2014:Q4.

The truncated nature of the price and quantity data on container trade poses a challenge in implementing an adaptive learning model. An agent's information set in each period includes all observations from the past under adaptive learning. However, unlike firms that may have access to observations from the inception of the industry, the researcher may not have access to observations from early periods. This problem applies to most empirical settings when using an adaptive learning model. In my particular setting, data on prices and quantities for major trade routes are available starting from 1997, although the first container ship voyage dates back to

Figure 6: Demand states



1956 and the first international voyage to 1966. This paper explores two alternative methods of empirically implementing an adaptive learning model given this challenge: the truncation approach and the imputation approach.

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 publicly available data or the data that the researcher has. However, bias can arise if agents' information sets include observations going further back than the start date of the data. The more heavily agents discount older observations when forming expectations, the smaller the bias would be.

This approach is implemented as follows. The set of weight parameters (λ_t) that I consider is $\{\frac{1}{t}, 0.01, 0.02, 0.03, 0.04\}$.[34] If $\lambda_t = \frac{1}{t}$, equal weights are applied to all past observations. In practice, the estimation procedure under this parameter is equivalent to applying least squares to estimate equation (2) for each period separately. The regression at each period t uses data covering from the start date of the data, 1997:Q1 to the current period t, or $\{z_\tau, \tilde{z}_\tau\}_{\tau=0}^{t}$. A constant λ_t gives rise to weights that are geometrically declining in age. And as λ_t increases, agents discount older observations more heavily. In practice, weighted leasts squares is applied on the information set given by $\{z_\tau, \tilde{z}_\tau\}_{\tau=0}^{t}$ at each t where the weight on an observation at τ is given by $(1-\lambda_t)^{t-\tau}$.

The imputation approach employs external data that provide information about the missing

[34] Orphanides & Williams (2005) suggest that the constant gain parameter in the range between 0.01 and 0.04 match the data on expectations well.

data. This approach is appealing in that if agents indeed use observations from the beginning of the industry in forming expectations, then it provides the best approximation to the actual weights that agents attach to each observation in their information set. 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. The start date for firms' information is set as the second quarter of 1966, which is the date of the first international container voyage. I use quarterly data on the value of trade by origin-destination pair from the IMF Direction of Trade Statistics database. 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 (1966:Q2-1996:Q4) are constructed as predictions based on the regressions. For the 1997-2014 period, actual demand states are used. Finally, I estimate beliefs that the adaptive learning model implies under different weights using these demand states.

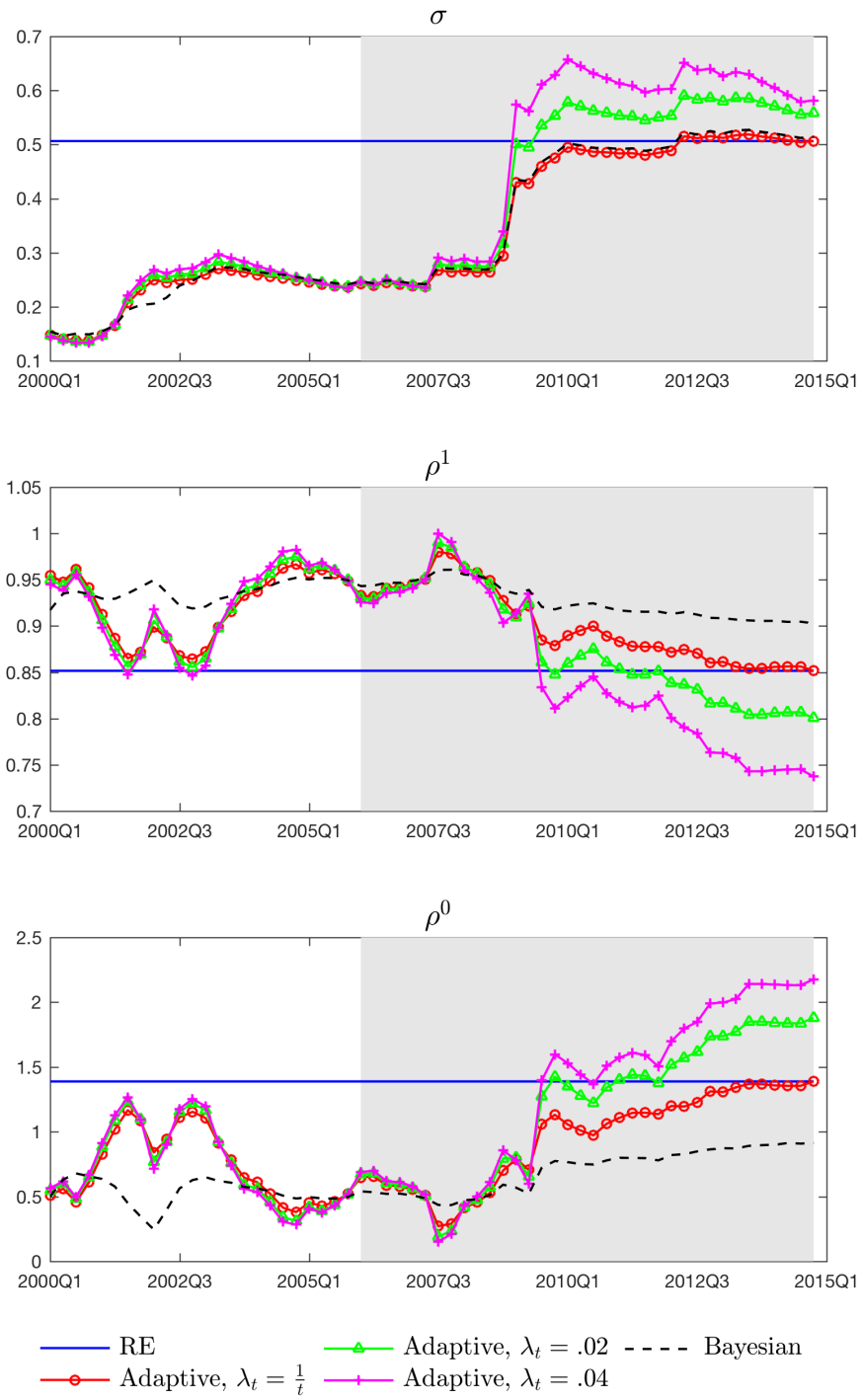
The truncation approach is adopted in the end on the grounds that it provides a better data fit. Moreover, the implementation of the truncation is more straightforward and the truncation approach can be also more universally applied since the external data necessary for the imputation method are not always available. Figures 18 and 19 in appendix A.6 compare beliefs under the two approaches for different values of λ_t and show that as λ_t grows the beliefs under the two approaches become closer to one another. Appendix A.6 also shows that the truncation approach is robust to varying the start date of the sample data used in the estimation of beliefs.

For the Bayesian learning model, the first three years of the historical price and quantity data (1997:Q1-1999Q4) are used in the estimation of prior beliefs.[35] The $AR(1)$ coefficients for the both markets are assumed to have normal priors, and the variance of the error an inverse Gamma prior. I start from diffuse priors and apply the Gibbs sampling methods. Table 4 reports the moments of the prior distributions. In the first quarter of 2000, firms start with prior beliefs about the parameters and update their beliefs using Bayesian updating in each period based on real-time data, $\{z_\tau \ \tilde{z}_\tau\}_{\tau=0}^{t}$ at each t. I apply the Gibbs sampling techniques to estimate posterior beliefs.

Figures 7 and 8 show firms' demand parameter estimates from 2000 to 2014 under different models of beliefs. The estimates in the shaded area are for 2006 to 2014, which will used in the estimation of the dynamic model. For the adaptive learning model, only the cases with $\lambda_t = \{\frac{1}{t}, 0.02, 0.04\}$ are plotted for the Asia-Europe market and the case with $\lambda_t = 0.02$ is plotted for the outside market for visualization purposes. The parameter estimates stay constant for the full-information model with no learning (RE model) by construction. Under learning, the estimate of the persistent parameter ρ_t^1 rises from 2006 to 2007 and shows a general downward trend thereafter for the Asia-Europe market. The variance parameter σ_t hikes in early 2009 and stays high from then.

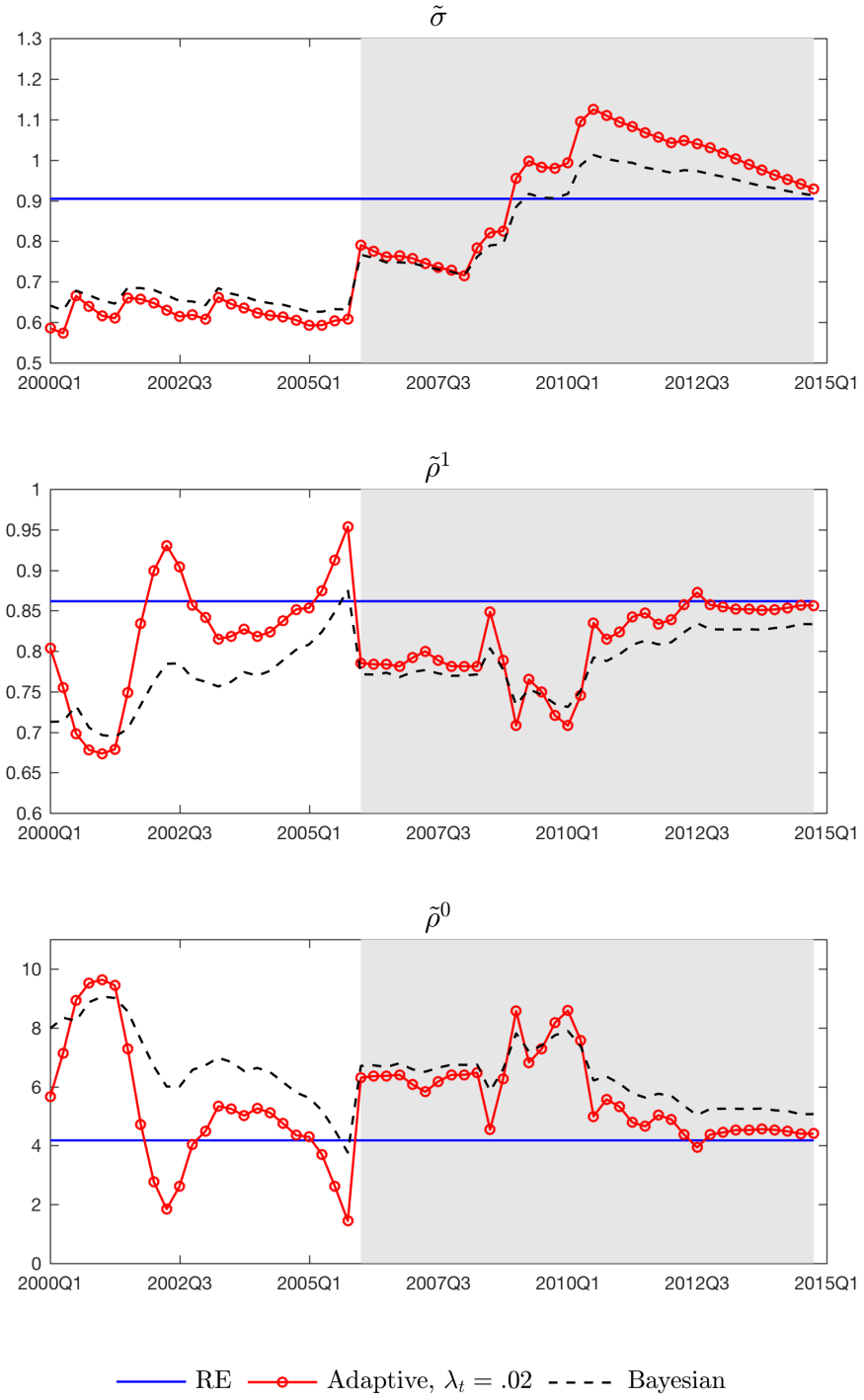
[35] I also explore using the full sample from 1997 to 2014 for estimating priors.

Figure 7: Beliefs under Different Models of Beliefs 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 λ_t, Bayesian learning, and full information (RE). The beliefs are summarized by the three parameters, $\{\sigma_t, \rho_t^0, \rho_t^1\}$, 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 8: Beliefs under Different Models of Beliefs 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 $\lambda_t = 0.02$, Bayesian learning, and full information (RE). The beliefs are summarized by the three parameters, $\{\tilde{\sigma}_t, \tilde{\rho}_t^0, \tilde{\rho}_t^1\}$, in the AR(1) process as given in equation (3). Beliefs for 2006-2014 in the shaded area are used in the main analysis.

Table 4: Moments of the Prior Distributions

Asia-Europe market		
ρ^0	ρ^1	σ
0.51	0.95	0.17
(0.63)	(0.08)	(0.02)
Outside market		
$\tilde{\rho}^0$	$\tilde{\rho}^1$	$\tilde{\sigma}$
8.11	0.72	0.68
(5.19)	(0.17)	(0.27)

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

Under adaptive learning the degree to which the parameter estimates react to recent events is the smaller for the case where agents weigh all past demand realizations equally ($\lambda_t = \frac{1}{t}$) than for the constant-gain cases. Among the constant-gain cases, beliefs respond more to new data as the parameter λ_t becomes larger. For example, as λ_t increases, the magnitude of the increase in the variances of the errors, σ_t, around 2008 and 2009 becomes larger as well. Similarly, the persistence parameter ρ^1 falls more around 2009, as λ_t becomes larger. Compared to adaptive learning models, 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. It is this variation in beliefs around demand shocks and the variation in the data in investment and scrapping before and after observing the shocks that identify the model of firm beliefs.

4.3 Diagnosing Different Models of Firm Beliefs

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 not only the mean forecast of one-year-head and two-year-ahead GDP growth rates but also 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. 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.

I take the forecast for the 2-year ahead GDP growth to construct the mean and the variance of the forecasts in each quarter from 2006 to 2014.[36] I construct the mean and the variance of 2-year

[36] Only two-year head forecasts are used in this analysis, because for one-year forecasts there is substantial bunching

ahead demand growth that each model of beliefs implies as well. Then, I compute correlation coefficients between the mean and the variance of GDP forecasts and the mean and the variance of demand forecasts implied by different models of beliefs. These correlation coefficients are reported in Table 5.

Table 5: Correlation between Demand Forecasts of Different Learning Models and GDP Forecasts

	Learning					Full Information	
	Bayesian	Adaptive				Baseline	GARCH
Correlation Coefficient		$\lambda_t = \frac{1}{t}$	$\lambda_t = .01$	$\lambda_t = .02$	$\lambda_t = .03$		
Mean GDP growth forecast	0.14	0.19	0.20	0.21	0.22	-0.23	-0.23
and mean demand growth	(0.17)	(0.17)	(0.17)	(0.17)	(0.17)	(0.17)	(0.17)
Variance of GDP growth forecast	0.83	0.86	0.86	0.85	0.83	.	0.37
and variance of demand growth	(0.09)	(0.09)	(0.09)	(0.09)	(0.10)		(0.16)

Notes: The mean and variance of GDP growth forecasts are based on two-year ahead forecasts published by the ECB. The mean and variance of two-year ahead demand growth is computed based on expectations about demand estimated under each model of belief.

The correlation coefficient for the mean growth is 0.14 for Bayesian learning and around 0.20 for adaptive learning. The correlation coefficient is negative for the full-information model as the full-information model predicts that the growth rate is high during periods of weak demand. This is because the denominator in calculating the growth rate is small in weak demand periods while firms' expectations about demand are constant over time under full information. Correlation is zero for the variance in demand growth forecasts for the full-information case since variance is constant by construction. On the other hand, the correlation between the variance of the GDP forecast and the variance of demand forecast is around 0.85 for learning models. I compute the correlation for the variance for a full-information GARCH model which allows time-varying volatility. The specification of this model is given in appendix A.2. I find that the full-information GARCH model is able to produce demand forecasts that are much more consistent with the GDP forecasts in terms of the variance but the correlation is still significantly lower than when using learning models.

5 Estimation and Empirical Results

This section presents the estimation of the dynamic model of investment with learning and the estimation results. The goal is to identify parameters in the model that best explain investment patterns in the data. Empirical papers in industrial organization on firm-level investment and industry dynamics typically focus on recovering investment costs, entry costs, or scrap values by searching for parameters that maximize the likelihood of observing the data or similarly minimize

in the forecasters' probabilities in end bins. The bunching makes it difficult to construct variance estimates.

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 applies a different estimation strategy which uses detailed data on shipbuilding prices and demolition prices to estimate investment costs and scrap values separately from the estimation of the dynamic model. Instead, I focus on identifying the model of firms' expectations about demand that can rationalize observed data.

The estimation proceeds as follows. First, given the demand estimates obtained in subsection 4., I estimate static parameters that govern the supply side including the marginal cost of production, the charter cost, and the outside market profit. I use these estimates to compute period profits. Second, I estimate 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 probabilities. Lastly, I estimate the dynamic model. I recover beliefs about future demand in each period that each of the models of firm beliefs implies. Then, I search for the model of firm beliefs and dynamic parameters that yield the best fit to the observed data through the method of simulated moments.

Incorporating learning in the model intensifies the computational burden. Under learning firms' beliefs change over time as firms receive new data. Therefore, an equilibrium needs to be solved separately for each period in the data. 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 regressing observed actions on observed state variables. The second stage involves estimating structure parameters which make these policies optimal. However, due to the global nature of the container shipping industry, I do not have observations from different markets that are independent form one another which are necessary to estimate the reduced-form policy function in the first stage. Therefore, this paper employs the full-solution approach. Although the data requirement is less burdensome under the full-solution approach, there are costs to using the full-solution approach. It is relatively time-consuming since it involves solving for an equilibrium at every guess of the parameter vector. Moreover, it requires a strong assumption of the uniqueness of the equilibrium.

5.1 Estimating Static Parameters

This section demonstrates the estimation of the marginal cost, charter cost, and outside market profit functions. The capacity deployment decisions $\{\bar{q}_{it}\}_i$ 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 on the Asia to Europe route and the Europe to Asia route is specified as (9). The following equation which allows normally distributed

disturbances serves as the basis for the maximum likelihood estimation of the cost parameters (a, b):

$$mc_{ijt} = a + \frac{bq_{ijt}}{\bar{q}_{it}} + \varepsilon_{ijt}\,. \tag{14}$$

The estimation of the charter cost function and the outside market profit function is based on the first-order conditions arising from firms' static profit maximization problem. The charter cost function and the outside market profit function are specified in a reduced-form way as the following:

$$R(\tilde{q}_{it}, x_{it}, \hat{s}_t) = \tilde{q}_{it}\left(r_0 + r_1\tilde{z}_t + r_2\tilde{Q}_t\right)$$
$$CC(h_{it}, x_{it}, \hat{s}_t) = h_{it}(\gamma_0 + \gamma_1 z_t + \gamma_2 k_{it} + \gamma_3 K_t).$$

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 and the outside market demand state since higher supply will likely lead to fiercer price competition and lower profit. The charter cost depends on the firm-level own capacity since larger firms may get discounts on charter rates. The charter cost is also allowed to depend on the industry-level total capacity since demand for charter may be higher when firms own lower capacity themselves.

Based on the profit function given in equation (11), the first-order condition from profit maximization with respect to the capacity deployed on Asia-Europe route ($\bar{q}_{ijt}$) and the first-order condition with respect to the chartering decisions (h_{it}) can be written as

$$\sum_j \left(\frac{\partial P_{jt}}{\partial \bar{q}_{it}} + P_{jt}\frac{\partial q_{ijt}}{\partial \bar{q}_{it}} - \frac{\partial c(q_{ijt}, \bar{q}_{it})}{\partial \bar{q}_{it}} - \frac{\partial c(q_{ijt}, \bar{q}_{it})}{\partial q_{ijt}}\frac{\partial q_{ijt}}{\partial \bar{q}_{it}}\right) + \frac{\partial R(\tilde{q}_{it}, x_{it}, \hat{s}_t)}{\partial \tilde{q}_{it}}\frac{\tilde{q}_{it}}{\bar{q}_{it}} = 0 \qquad \text{and,}$$
$$\frac{\partial R(\tilde{q}_{it}, x_{it}, \hat{s}_t)}{\partial \tilde{q}_{it}}\frac{\partial \tilde{q}_{it}}{\partial h_{it}} - \frac{\partial CC(h_{it}, x_{it}, \hat{s}_t)}{\partial h_{it}} = 0$$

, respectively. Given the demand estimates and the first-order conditions, I then estimate the charter cost and the outside-market profit functions via maximum likelihood. 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 6 shows the estimates of the static parameters and the average values of firm-level marginal costs, outside market profits, and charter costs over the 2006 to 2014 period. The positive sign of the coefficients on the Asia-Europe market demand state in the outside market profit and charter cost functions (r_1 and γ_1) 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. In addition,

Table 6: Estimates of the Static Parameters and Firm-Level Averages

Marginal cost (**MC**)						
Estimates ($ per TEU)				**Mean** (1000 US dollars)		
a	b			Asia to Europe	Europe to Asia	
0.265	1.750			1.518	0.781	
(0.011)	(0.024)			(0.187)	(0.046)	
Outside market profit (**R**)						
Estimates ($ per TEU deployed in the outside market))				**Mean** (1000 US dollars)		
r_0	r_1	r_2				
-1.239	0.089	-0.117		274.481		
(0.177)	(0.006)	(0.008)		(236.508)		
Charter cost (**CC**)						
Estimates ($ per TEU chartered)				**Mean** (1000 US dollars)		
γ_0	γ_1	γ_2	γ_3			
0.206	0.087	-0.084	-0.064	208.016		
(0.096)	(0.007)	(0.021)	(0.009)	(177.946)		

This table reports estimates of the parameters in the marginal cost, outside market profit, and charter cost functions. It also reports average values constructed using these estimates. Averages are taken over firms and periods from 2006:Q1 to 2014:Q4. The aggregate deployed capacity in the outside market ($\bar{Q}_t$) in the function of the outside market profit (R) and the firm-level owned capacity (k_{it}) and the aggregate owned capacity (K_{it}) in the charter cost function (CC) are in the unit of 1 million TEU. Standard errors for the estimates and standard deviations for the means are in parentheses.

larger firms tend to face lower charter costs. The charter cost tends to be lower when operators own more ships in total, which may be due to decreased demand for chartering.

5.2 Estimating Other Model Primitives

This study recovers the investment cost and the scrap price directly from available data using a reduced-form approach. The first reason for doing so is that detailed data on investment costs and scrap values are available, unlike in many other settings. Clarkson Research publishes monthly reports on average shipbuilding prices and scrap values as well as a sample of transaction-level data. More importantly, directly using the investment costs and scrap values from the available data eliminates the need to impose a particular learning model but instead allows me to identify which learning model provides the best fit for the data.

I first examine whether there is firm heterogeneity in the price firms pay for building new ships and the price firms receive for scrapping their ships. If these costs are systematically different across firms, for example because larger firms have greater bargaining power over shipyards, it will be important to incorporate the cost heterogeneity in the model. The analysis in appendix **??**

confirms that the transaction-level prices do not vary significantly across firm size or other firm characteristics. Hence, I use industry-level price data to estimate the investment cost and scrap value as a function of the industry state variables (total capacity of owned ships, total capacity in the order book, demand state for the Asia-Europe market, and demand state for the outside market) via least squares. Table 7 reports the estimates and figure 10 compares investment costs and scrap values in data to predicted values obtained from the regression.

Table 7: Estimates of the Investment Cost and Scrap Value

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

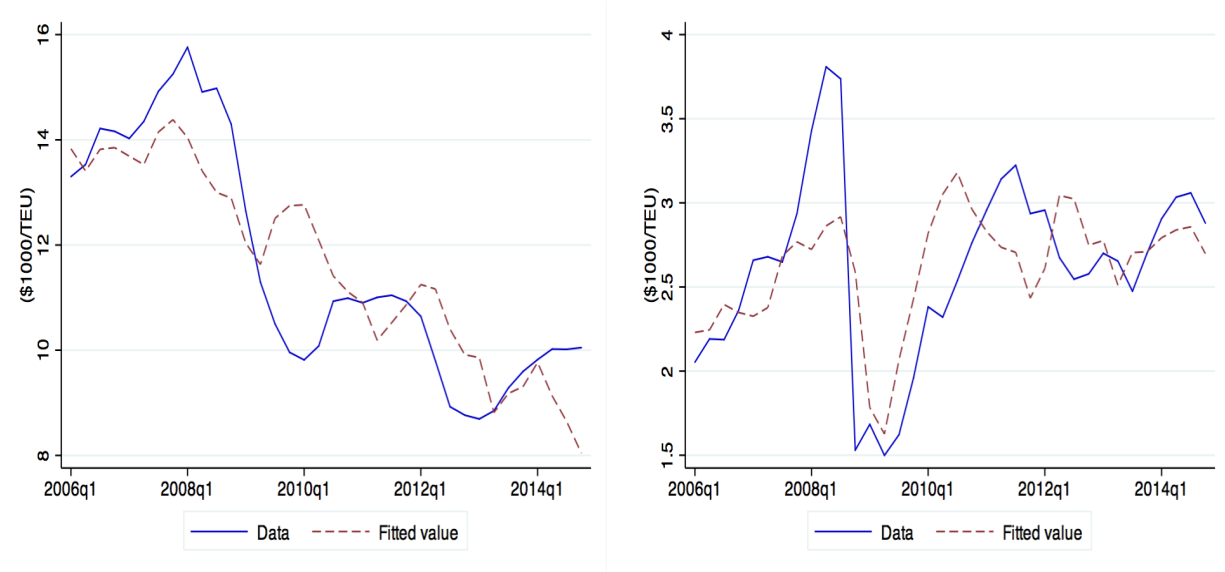
Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.001$.

The delivery process of newly ordered ships and the depreciation process of existing ships are estimated separately from the dynamic model as well. The mean delivery rate, κ, is estimated as a function of the firm's order book capacity as in equation (6).[37] I also set an exogenous depreciation rate instead of estimating it with the dynamic system because the data do not differentiate between the scrapping of ships that have not reached their retirement age from ships and the scrapping of ships that have reached their retirement age. The depreciation rate, ζ, is set such that the average age at which ships naturally depreciate is 20 years based on industry experts' estimates.[38]

[37] The current formulation assumes that the delivery probability depends solely on the current order book capacity and the depreciation probability on the owned ship capacity. However, it may take longer to build ships if shipyards are overbooked due to the high volume of orders that are currently in process. I examine whether the current order book capacity plays a significant role in the delivery rate by including the industry order book capacity in the regression but find that it is not statistically significant.

[38] 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 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.

Figure 9: Predicted Investment Costs and Scrap Values



(a) Investment Costs

(b) Scrap Values

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

5.3 Estimating the Dynamic Parameters and the Model of Firm Beliefs

The last and most computationally intense step of the estimation entails estimating the model of firm beliefs and the dynamic parameters. I employ the method of simulated moments (MSM). Let $\mathcal{H}$ denote the set of models of firm beliefs and Θ the set of dynamic parameter vectors, $\theta = (\sigma^{\iota}, \sigma^{\delta}, FC)$. The models of beliefs considered in this paper are a full-information model, a Bayesian learning model, adaptive learning models with $\lambda_t \in \{\frac{1}{t}, 0.01, 0.02, 0.03, 0.04\}$. Given a model of belief $h \in \mathcal{H}$, I solve for an equilibrium of the dynamic investment model and obtain the optimal investment policy function for each parameter vector, $\theta \in \Theta$. For learning models, I solve the model separately for each period. 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:

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

I search for a pair of a model of beliefs and a parameter vector that minimizes the weighted distance between the data and simulated moments given as:

$$f^h(\theta) = \left(\Gamma^d - \Gamma^h(\theta)\right)' W \left(\Gamma^d - \Gamma^h(\theta)\right). \tag{15}$$

where Γ^d is the set of data moments. The search is done over grids of $(\sigma^\iota, \sigma^\delta, FC)$ for each model of beliefs. The grids for σ^ι and σ^δ are in increments of 0.005 and the grid for FC is in increments of \$50/TEU. I use the inverse of the variance-covariance matrix of the simulated moments as the weighting matrix.

The identification of the dynamic parameters relies on a revealed-preference argument. Each state in the state space implies certain values of benefits and costs for each of the options that the firm faces–investment, scrapping, and staying. As a result, I can back out how firms expect demand to evolve in the future, given the evolution of non-demand states that equilibrium strategies and the observed firm choices imply. In principle, therefore, these dynamic parameters are identified by both time-series and cross-sectional variations. Nevertheless, the main source of identification is time-series variation in investment and investment costs pre- and post-financial crisis. 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 business cycles and demand shocks in world trade. Table 8 lists moments used in the estimation and compares data moments and simulated moments.

Table 8: Data and Simulated Moments

	Data moments	Simulated moments
Average investment in 2006-2008 (1 mil. TEU)	0.23	0.23
		(0.03)
Average investment in 2009-2014 (1 mil. TEU)	0.14	0.15
		(0.02)
Total capacity of owned ships (1 mil. TEU)	5.09	5.15
		(0.27)
Total capacity in the order book (1 mil. TEU)	3.07	2.98
		(0.14)
Correlation between demand and investment	0.19	0.22
		(0.12)
Volatility of investment (1 mil. TEU)	0.17	0.17
		(0.03)

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

My estimates indicate that the adaptive learning model with $\lambda_t = 0.02$ provides the best fit for the observed data moments, which is referred to as the baseline learning model in the rest of the paper. This implies that agents put about 45% weights on a 10-year-old observation compared

Table 9: Dynamic Parameter Estimates

λ_t	0.02 (0.005)
σ^ι (1 bil. US dollars)	0.275 (0.055)
σ^δ (1 bil. US dollars)	0.43 (0.092)
FC (1 bil. US dollars)	0.025 (0.0051)

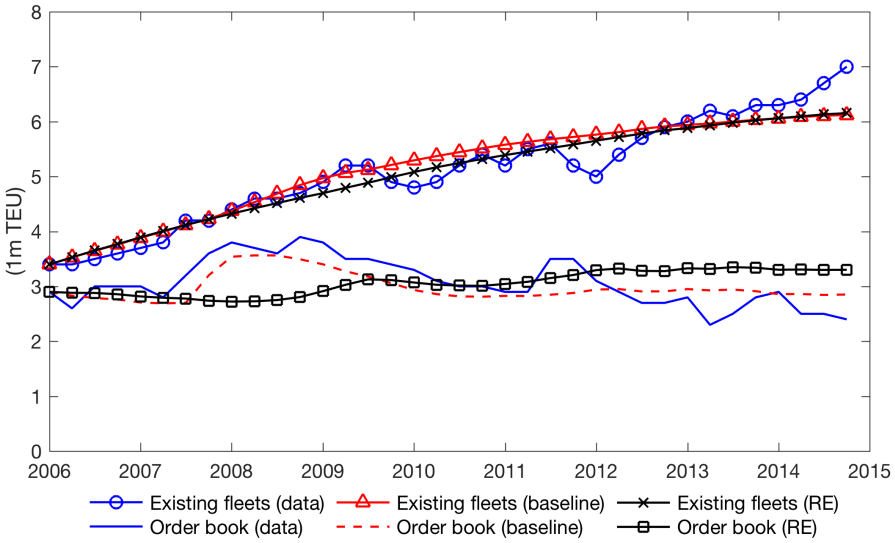
Notes: This table shows estimates of dynamic parameters. λ_t is the weighting parameter in the adaptive learning model which governs the rate at which agents discount older observations when forming expectations about demand. σ^ι is the standard deviation of the i.i.d. shock around the investment cost 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.

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 to be 0.0175, 0.0183, and 0.02, respectively, with respect to expectations about macroeconomic conditions and monetary policy.

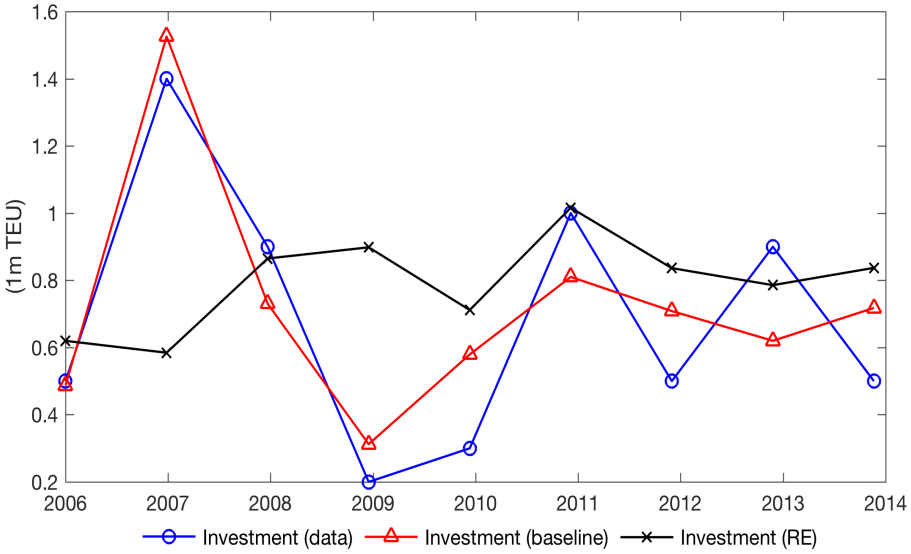
The fixed cost of holding one unit of capital (100,000 TEU) is estimated to be 25 million dollars, which is around 36% of one period's profit from one unit of capital (where the period profit includes the sum of profits from the Asia-Europe market and the outside market minus the charter cost and does not include the investment cost and scrap value). This fixed cost includes all costs that owning and operating ships imposes 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. Table 9 reports the estimates of all dynamic parameters.

Figure 10 shows the simulated industry evolution and investment under the baseline learning model and the full-information model, where the simulation is conducted based on parameters that generate the best fit for each model. The baseline learning model does particularly well at predicting the investment boom in 2006-2007 and the plunge in investment in 2009-2010. On the other hand, the full-information model fails to predict the correct quantity and timing of investment. Specifically, the full-information predicts that firms invest more after demand collapses in 2008 than during high demand periods in 2006-2007. This reallocation of investment across time happens for the following reason. During times of high demand investment is also more costly as demand for new ships increases. At the same time, high demand is likely to persist in the future making investment more profitable. In the case of full information, the negative supply-side effect dominates the positive demand-side effect. Under learning high demand periods make firms more optimistic, which increases the positive demand-side effect and helps predict the positive correlation between demand and investment.

Figure 10: Model Fit of the Baseline Model and the Full-Information Model



(a) Capacity of Owned Ships and Order boook



(b) Yearly Investment

Notes: The top panel shows the industry evolution simulated under the baseline learning model (adaptive learning with $\lambda_t = 0.02$) and the full-information (RE) model, respectively, as well as the industry evolution in the data. The bottom panel shows yearly investment simulated under the baseline learning model and the full-information model as well as yearly investment in the data. The simulated moments are based on 1000 equilibrium paths. For each model of belief, the simulation was conducted for dynamic parameters, $(\sigma^{\iota}, \sigma^{\delta}, FC)$, that yield the best fit.

I simulate the full-information model while holding the dynamic parameters fixed at the estimated values to quantify how far the predictions of investment and welfare would be from the predictions under the baseline model for the 2006-2014 period if a policy maker used the full-

Table 10: Summary of Results and Welfare for Full Information Counterfactuals

	Baseline	RE (%Δ)
Owned capacity (1 mil. TEU)	5.15	4.79 (-6.91)
Order book (1 mil. TEU)	2.98	2.56 (-14.06)
Investment (1 mil. TEU)	0.18	0.15 (-16.66)
Volatility of investment (1 mil. TEU)	0.17	0.13 (-22.17)
Consumer surplus (1 bil. US dollars)	82.63	80.02 (-3.16)
Producer surplus (1 bil. US dollars)	14.89	27.55 (85.06)
Total surplus (1 bil. US dollars)	97.52	107.57 (10.31)

Notes: This table shows results from simulations under the baseline learning model and the full-information (RE) model over the period from 2006:Q1 to 2014:Q4. The owned capacity, order book, and investment are sums across all firms per period. The welfare measures are sums over all periods. The consumer surplus is calculated with respect to the Asia-Europe market only.

information model. As shown in table 10, the total amount of investment would be predicted to be lower by 17%. The producer surplus is predicted to be greater by 85% under full information and total surplus by 10%. The reduced amount of investment combined with the concentration of investment on periods with cheaper costs leads to the gain in producer surplus.

The Bayesian learning model also correctly predicts that investment peaks in 2007 and bottoms in 2009. The model fit is shown in figure 17 in appendix A.3. However, the magnitudes of the rise and the fall in investment around the financial crisis are smaller under Bayesian learning than observed in the data. This is because agents' beliefs are smoother under Bayesian learning as agents put positive weights on prior beliefs and weigh all past observations, which leads to smoother investment over time.

6 Counterfactual Analysis

This section performs counterfactual simulations to quantify the effects of strategic incentives, the irreversibility of investment, and demand volatility on investment and welfare. I also examine whether the modeling choice for firms' expectations is important for policy evaluation. First, I simulate a policy which removes competition and allows coordination among firms. I simulate the industry under a multi-plant monopolist who make decisions to maximize joint profits as well as a case in which the two largest firms merge. These experiments shed light on the role of competition in creating oversupply as well as the potential effects of consolidation. Second, I simulate a scrapping subsidy policy. This policy makes investment less irreversible as it helps firms to scrap ships at a higher rate when demand conditions worsen. This policy may therefore help firms deal with excess capacity. It may also encourage investment, however, as it raises the value of owning ships. Lastly,

I apply the learning framework to address the long-standing question in the literature on the effect of demand volatility on investment.

6.1 Coordination among Firms

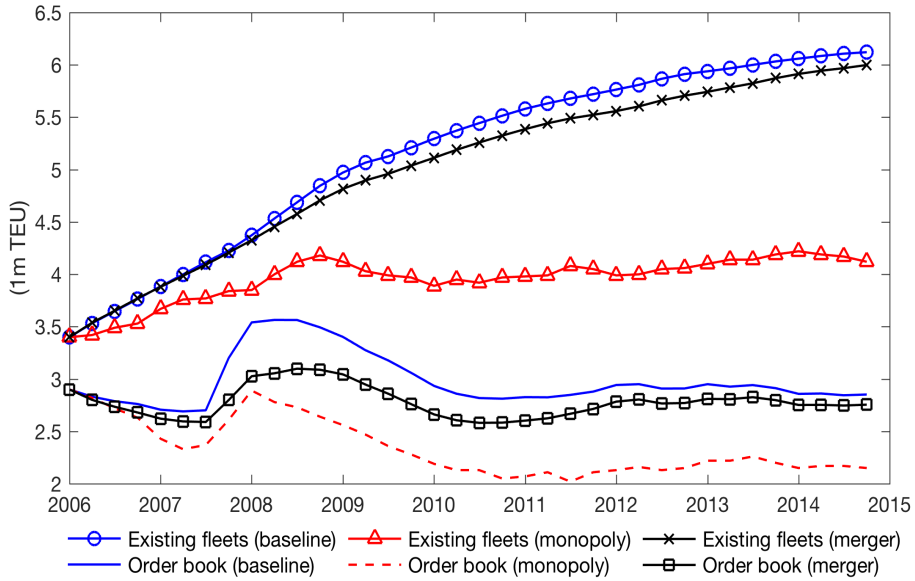
To deal with the excess capacity in the industry, firms have increasingly moved towards cooperation and consolidation. Denmark-based Maersk Line, France-based CMA-CGM and Swiss-based Mediterranean Shipping Co. (MSC)–the world's three biggest container-shipping companies by capacity that jointly control up to 40% of total container capacity–proposed and won approval for their proposed P3 Alliance from the U.S. Federal Maritime Commission as well as European regulators in 2014. Akin to a code-sharing deal between airlines, the alliance was meant for the firms to cut costs by using each other's ships and port facilities as well as reduce competition. However, the Chinese competition authorities rejected the P3 alliance on the grounds that the move would restrict competition on the busiest Asia-Europe container routes. The three P3 carriers subsequently abandoned the plan, with Maersk and MSC later creating a two-carrier alliance dubbed 2M. 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.

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 level of aggregate capacity 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 in order to deter entry or expansion of rivals. Spence (1977) shows that an industry selling a relatively homogenous product may invest in excess capacity to deter entry, and Fudenberg & Tirole (1983) and Reynolds (1987) explore the similar idea.

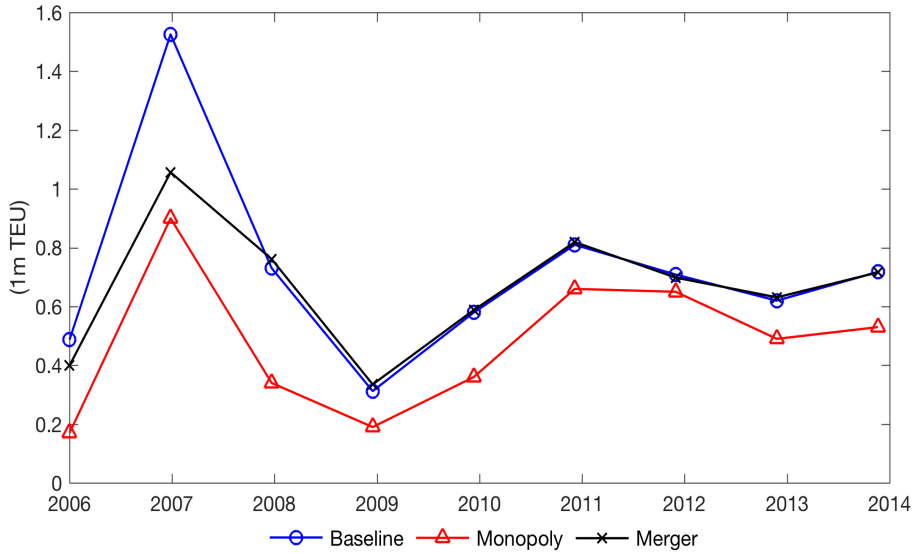
My model incorporates several sources of strategic incentives. First, a firm's deployment of an extra unit of capacity has a negative effect on the market price and the competitors' quantities and profitability. The business stealing effect arises since this negative effect is internalize by all incumbents in the market. When market conditions become adverse, a firm may want to wait till its rivals remove their capacity instead of removing its own. 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 commit to investment before others do when they expect strong demand.

In the monopoly counterfactual, the monopolist operates and makes joint decisions of invest-

Figure 11: Simulations of Monopoly and Merger Cases



(a) Capacity of Owned Ships and Order boook



(b) Yearly Investment

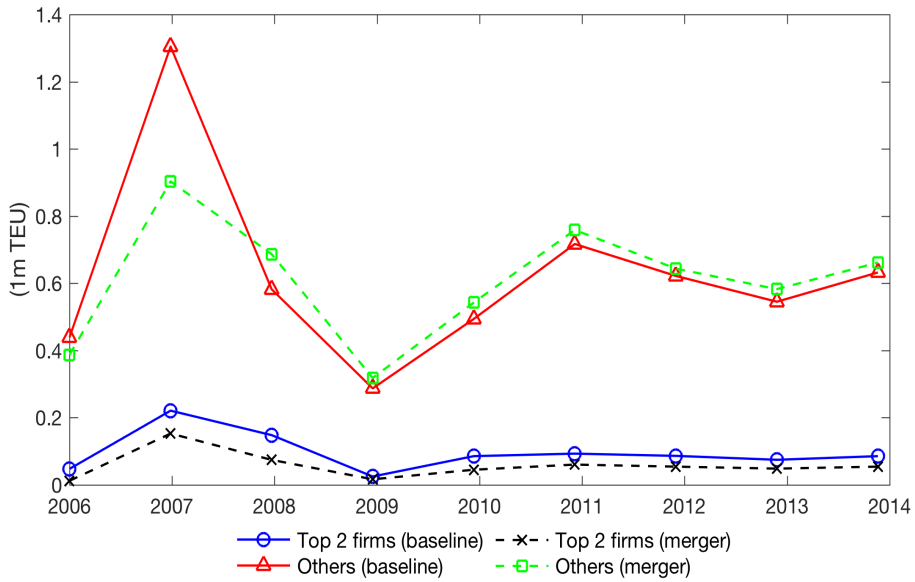
Notes: The top panel shows the industry evolution simulated under the baseline, monopoly, and merger cases, respectively. The bottom panel shows yearly investment. The simulations are based on 1000 equilibrium paths.

ment, scrapping, chartering, and deployment for all firms. I assume that the monopolist maintains the same size distribution as observed in the data. This particular arrangement is adopted instead of the one in which the monopolist operates all ships under one entity, since it helps disentangle the effect of strategic incentives from the effect arising from a change in the firm size distribution (, for example, through cost savings, changes in bargaining power, etc.). The monopolist makes

investment and scrapping decisions and decides how much capacity to charter and to deploy in the Asia-Europe market in order to maximize aggregate profits from all plants. The monopolist is endowed with beliefs in the baseline model.

The merger counterfactual experiment allows a merger between top two firms, Maersk and MSC. Under the 2M Alliance, which is a vessel sharing agreement between these two firms formed in late 2014, the firms aim at cooperation in the deployment of ships over particular routes but do not engage in joint sales or marketing, price fixing, pooling of revenues, or the sharing of profits or losses. A merger would therefore lead to a much higher level of cooperation in which the two firms engage in joint investment and deployment as one entity.

Figure 12: Yearly Investment by Top 2 and Other Firms



Notes: This figure shows yearly investment by top two firms and the rest of the firms simulated under the baseline, monopoly, and merger cases, respectively. The simulations are based on 1000 equilibrium paths.

Removing competition externalities through the monopoly case has a large effect on investment: during the sample period of 2006 to 2014, average investment drops by 34% compared to he baseline case as shown in the second column of table 11. Under the merger case, average 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%. Under monopoly, producer surplus increases by 91 billion dollars but 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%.

Panel B of table 11 shows that the full-information model underestimates changes in investment, thus underestimates welfare changes, especially producer surplus gains, compared to the

Table 11: Monopoly and Merger Counterfactuals

Panel A: Industry Outcomes and Welfare

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

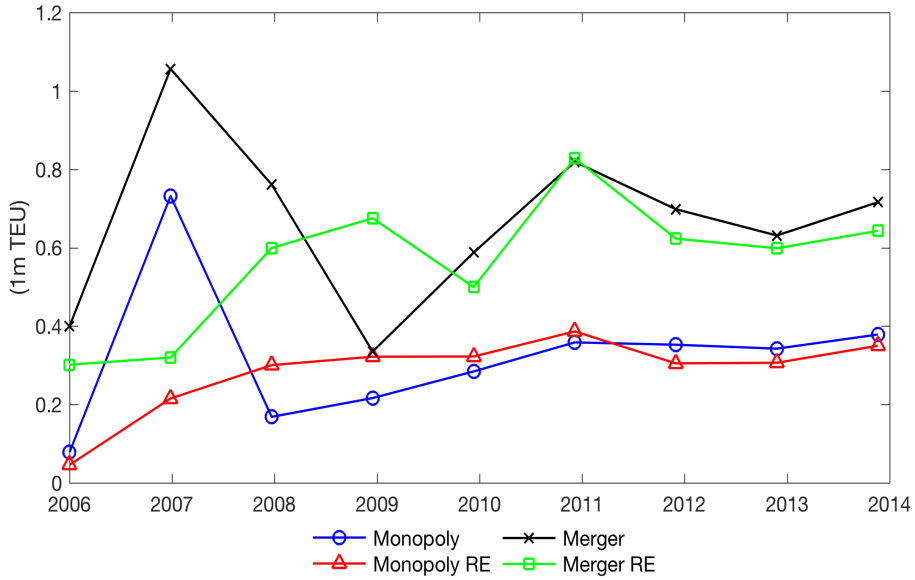
Panel B: Welfare Changes under Baseline and Full-Information Models

	Monopoly			Merger		
	Baseline	RE	RE - Baseline (Col 2 - Col 1)	Baseline	RE	RE - Baseline (Col 5 - Col 4)
Change in investment (1 mil. TEU)	-0.061	-0.039	0.022	-0.014	-0.009	0.005
Change in volatility of investment (1 mil. TEU)	-0.037	-0.020	0.017	-0.025	-0.004	0.021
Change in consumer surplus (1 bil. US dollars)	-41.78	-39.35	2.42	-0.94	-0.46	0.48
Change in producer surplus (1 bil. US dollars)	91.73	83.27	-8.46	13.96	10.04	-3.92
Change in total surplus (1 bil. US dollars)	49.95	43.92	-6.03	13.03	9.58	-3.45

Notes: Panel A shows results form the monopoly and the merger simulations over the 2006:Q1 to 2014:Q4 period. The percent changes from the baseline case are in parentheses. The owned capacity, order book, and investment are sums across all firms per period. The welfare measures are sums over all periods. Panel B compares changes in welfare measures predicted from the learning model and the full-information model. Consumer surplus is calculated with respect to the Asia-Europe market only.

baseline learning model. Therefore, if a regulator used the full-information model, he is likely to underestimate gains from a merger or other forms of consolidation among firms. Under the baseline learning model, firms invest heavily in the 2006 to 2008 period when investment is very costly. Hence, the reduction in investment in this period under the monopoly and merger counterfactuals leads to large savings to the firms. On the other hand, since the full information model predicts that investment is concentrated in the post-crisis-period in which investment is cheaper, cost savings from the reduction in investment are relatively small.

Figure 13: Yearly Investment under Full Information and Learning



Notes: This figure compares monopoly and merger simulations conducted under the baseline learning model with simulations under the full-information model. The simulations are based on 1000 equilibrium paths.

6.2 Scrapping Subsidies

In the container shipping industry, part of the irreversibility in investment stems from the fact that when demand conditions are bad, the scrap price is also low. This is because the scrap market is driven by demand for steel which is highly correlated with demand for trade. Therefore, firms often do not find it profitable to scrap existing ships even though there is excess capacity.

China implemented a subsidy program in 2013 to help Chinese firms that are struggling with overcapacity and also to help its shipyards. The program grants 1500 yuan (around 220 US dollars) per gross ton to replace old ships registered in the country with new vessels. Ships must be within 10 years before their mandatory retirement age to be eligible. Ship owners get half the subsidy when they finish scrapping an old ship and receive the remainder if a new ship is built.

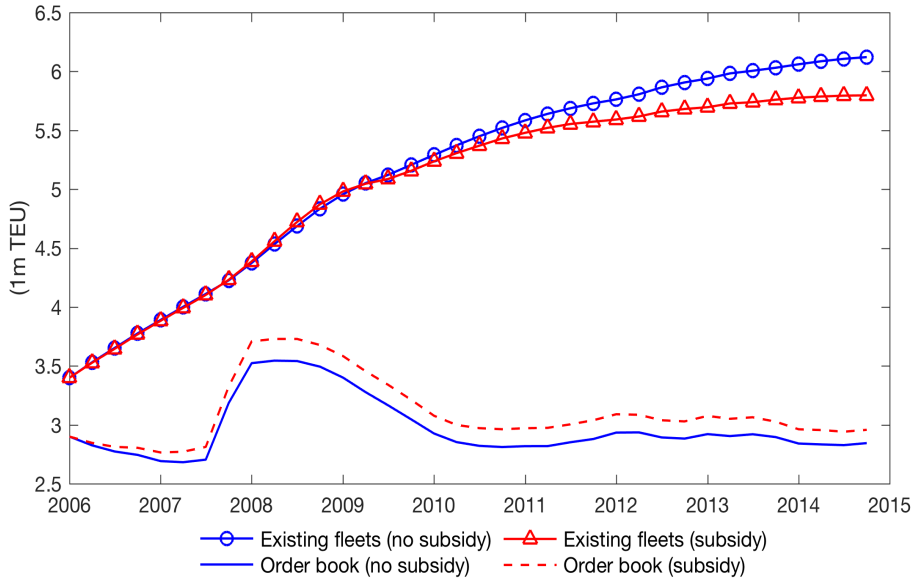
In this counterfactual, I apply a similar subsidy program to all firms which grants 150,000 dollars for scrapping 1000 TEU.[39] This roughly amounts to be 13% of the average new building price or 57% of the average scrap price. The subsidy is transferred to firms at the time of scrapping regardless of whether the firm is replacing ships that are around the retirement age.[40] The program applies to all scrapped ships regardless of their age.

As table 12 shows, scrapped capacity increases by 46% under the subsidy program. The bottom

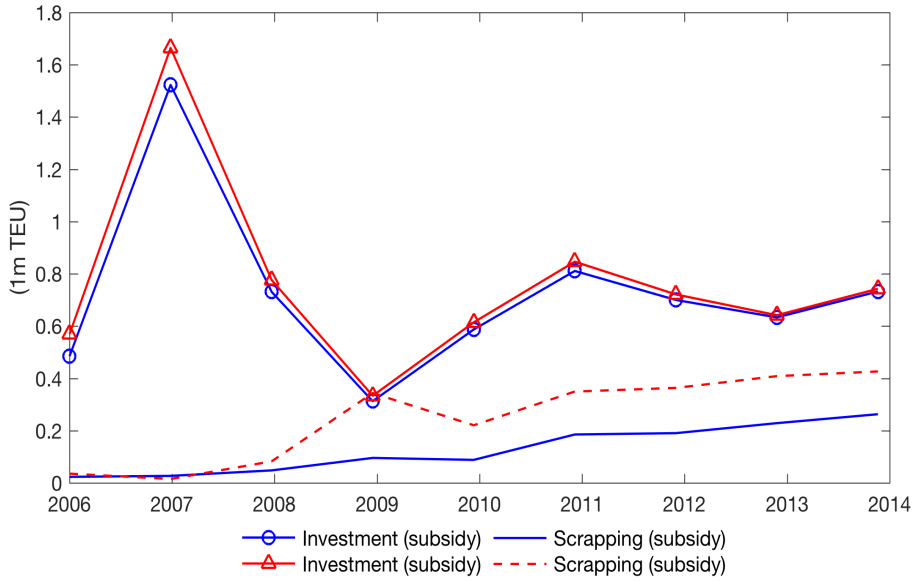
[39]This corresponds to 1500 yuan per gross ton based on the conversion rate of 1 gross ton to 1 dwt suggested by Stopford (2009) and the conventional conversion rate of 1dwt to 14 TEU.

[40]This choice is made since the current model does not allow for efficiency gains from replacing old ships with new ships.

Figure 14: Simulations under Scrapping Subsidies



(a) Capacity of Owned Ships and Order boook



(b) Yearly Investment

Notes: The top panel shows the industry evolution simulated under no subsidy and subsidy, respectively. The bottom panel shows yearly investment. The simulations are based on 1000 equilibrium paths.

panel of figure 14 shows that the scrapping increases especially in the 2009-2014 period. The effect is particularly dramatic in 2009 where the total scrapped capacity under the subsidy program is approximately 3.5 times of the scrapped capacity in the absence of the subsidy. The subsidy also results in a 6.5% increase in investment as the subsidy raises the value of owing ships. In terms of welfare, the policy results in a 5.7 % increase in producer surplus. The decrease in supply from the

Table 12: Scrapping Subsidy Counterfactuals

Panel A: Industry Outcomes and Welfare	
	Subsidy (%Δ)
Owned capacity (1 mil. TEU)	5.04 (-2.11)
Orderbook (1 mil. TEU)	3.10 (4.07)
Investment (1 mil. TEU)	0.19 (6.45)
Scrap (1 mil. TEU)	0.06 (45.87)
Consumer surplus (1 bil. US dollars)	81.78 (-1.02)
Producer surplus (1 bil. US dollars)	15.74 (5.70)
Subsidy (1 bil. US dollars)	3.35 (.)
Total surplus (1 bil. US dollars)	97.52 (0.00)

Panel B: Welfare Changes under Baseline and Full-Information Models			
	Baseline	RE	RE - Baseline (Col 2 - Col 1)
Change in consumer surplus (1 bil. US dollars)	-0.85	-0.98	-0.14
Change in producer surplus (1 bil. US dollars)	0.85	0.87	0.02
Change in subsidy (1bil. US dollars)	3.35	2.61	-0.74
Change in total surplus (1 bil. US dollars)	0.00	-0.11	-0.12

Notes: Panel A shows results form the scrapping subsidy counterfactual over the 2006:Q1 to 2014:Q4 period. The percent changes from the baseline case are in parentheses. The owned capacity, order book, and investment are sums across all firms per period. The welfare measures are sums over all periods. Panel B compares changes in welfare measures predicted from the learning model and the full-information model. Consumer surplus is calculated with respect to the Asia-Europe market only.

subsidy results in a loss in consumer surplus. If the subsidy is factored into total surplus, the net effect of the scrapping subsidy policy is negative.

The findings from this policy experiment confirm the theoretical prediction that the irreversibility of investment reduces investment in an environment with demand uncertainty. From a policy maker's point of view the scrapping subsidy may not be an effective way to address the excess capacity problem since it also encourages firms to invest in new ships. A careful choice of timing might make the policy more effective, for example, through targeting periods in which the policy is less likely to encourage new investment such as 2009 in the sample period. Such targeting is likely to be difficult to implement ex-ante, however. Lastly, this experiment shows that using the right model for firms' expectations about demand is important for policy evaluation. In particular, the full-information model predicts that the subsidy policy will result in a welfare gain as it predicts lower investment and scrapping than the learning model.

As table 12 shows, using the full-information model to evaluate the effect of the subsidy policy would lead to an underestimation of the amount of the subsidy as the full-information model

predicts a smaller effect on scrapping. Moreover, the full-information model predicts the effect of the subsidy program on total surplus to be negative as it overestimates the consumer surplus loss, while the baseline model predicts it to be neutral.

6.3 Demand Volatility

Demand volatility can affect investment in several different ways. First, real options theory predicts that 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 volatility in investment costs. Lastly, 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.

Table 13: Demand VolatilityCounterfactuals

	High Volatility	Low Volatility	High Volatility RE	Low Volatility RE
Investment (1 mil. TEU)	0.15	0.16	0.14	0.16
Std. dev. of investment (1 mil. TEU)	0.08	0.04	0.05	0.03
Corr. between demand and investment	0.10	-0.03	-0.05	-0.16
Consumer surplus (1 bil. US dollars)	112.60	85.30	113.27	84.11
Producer surplus (1 bil. US dollars)	24.59	33.28	26.84	35.08
Total surplus (1 bil. US dollars)	137.19	118.58	140.12	119.18

Notes: This table shows results form demand volatility counterfactuals over the 2006:Q1 to 2014:Q4 period. The owned capacity, order book, and investment are sums across all firms per period. The welfare measures are sums over all periods. Consumer surplus is calculated with respect to the Asia-Europe market only.

To quantify the effect of demand volatility, I conduct following counterfactual simulations. For 2006 to 2014, two sets of demand series are simulated–one for a high volatility case and the other for a low volatility case. For both sets of demand series, the $AR(1)$ coefficients are taken from the previous estimation with the full sample of data. But in the high volatility case, the variance in the demand process for the both the Asia-Europe and outside markets, are doubled from the estimates from the full sample of data. In the low volatility case, the variances are halved from the estimates. The 1997-2005 demand realizations are taken from the data in both cases. For each demand series, I solve for an equilibrium and simulate the industry under firms' equilibrium investment and scrapping strategies, under learning and full information, respectively.

Results from the simulations reported in table 13 show that an increase in demand volatility has a negative effect on investment, which is consistent with findings of 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, less volatility in demand raises the expected value of owning a ship. An increase in demand volatility also increases volatility in investment as higher demand volatility leads to more volatile shipbuilding prices. In the learning model, higher demand volatility also leads to larger changes in firms' expectations about future demand, which further increases volatility in investment.

The learning model predicts the correlation between demand and investment to be positive for the high volatility case. When learning is present, higher demand volatility generates agents' beliefs that are more highly correlated with demand. The modeling choice for firms' expectations potentially matters for policy design as the learning model and the full-information model yield different predictions about firms' investment patterns. For the high volatility case, the learning model predicts large investment boom and bust cycles that move in the same direction as the demand cycles, while the full-information model predicts less volatile investment and higher investment in periods with weak demand.

7 Conclusion

This paper evaluates learning as agents' belief-formation process capable of endogenously generating investment boom and bust cycles. Methodologically, the paper contributes to the literature by introducing a dynamic oligopoly model of investment which incorporates learning about the aggregate market demand process. The model departs from the standard practice under full-information assumptions of rational expectations and instead allow agents to form expectations about demand using best information available to them in each period. Agents use their changing forecasts about demand in making their investment decisions. I explore different approaches to modeling firms' expectations and test the models against each other to find the one which can best rationalize observed investment behavior.

I analyze the framework through data from the container shipping industry, which is exposed to sharp demand swings and the problem of excess capacity. The empirical analysis shows that a learning model that places heavier weights on more recent observations successfully predicts investment boom and bust cycles observed in the data. On the other hand, a full-information model, which assumes that firms know the true demand model and its parameters, fails to predict the correct quantity and timing of investment. In particular, such a model underestimates the volatility of investment and predicts that firms withhold investment during high demand periods when investment is more costly.

This paper uses the framework to address various firm-strategy and public-policy issues. The main set of counterfactuals examines the effects of strategic incentives on the level and the volatility of investment. Results reveal that a policy that removes competition among firms would result in a substantial reduction investment and an improvement in overall welfare, but a consumer

surplus loss. This finding thus has implications for antitrust regulations on coordinated investment. In addition, I find that the modeling choice for firms' expectations about demand matters in policy evaluation. This paper also sheds light on the informational channel through which demand fluctuations affect investment. Under learning, high demand volatility leads to more frequent and larger revisions of expectations about future demand, thereby amplifying the magnitude of investment boom-bust cycles.

References

Ackerberg, Daniel, Benkard, C Lanier, Berry, Steven, & Pakes, Ariel. 2007. Econometric tools for analyzing market outcomes. *Handbook of econometrics*, **6**, 4171–4276.

Aguirregabiria, Victor, & Mira, Pedro. 2007. Sequential estimation of dynamic discrete games. *Econometrica*, 1–53.

Andrews, Donald WK. 1993. Tests for parameter instability and structural change with unknown change point. *Econometrica: Journal of the Econometric Society*, 821–856.

Bajari, Patrick, Levin, Jonathan, *et al.* . 2007. Estimating Dynamic Models of Imperfect Competition. *Econometrica*, **75**(5), 1331–1370.

Bernanke, Ben S. 1983. Irreversibility, Uncertainty, and Cyclical Investment. *The Quarterly Journal of Economics*, **98**(1), 85–106.

Besanko, David, Doraszelski, Ulrich, Lu, Lauren Xiaoyuan, & Satterthwaite, Mark. 2010. On the role of demand and strategic uncertainty in capacity investment and disinvestment dynamics. *International Journal of Industrial Organization*, **28**(4), 383–389.

Bloom, Nick. 2009. The Impact of uncertainty shocks. *Econometrica*.

Cogley, Timothy, & Sargent, Thomas J. 2005. The conquest of US inflation: learning and robustness to model uncertainty. *Review of Economic dynamics*, **8**(2), 528–563.

Collard-Wexler, Allan. 2013. Demand Fluctuations in the Ready-Mix Concrete Industry. *Econometrica*, **81**(3), 1003–1037.

Covert, Thomas. 2014. Experiential and social learning in firms: the case of hydraulic fracturing in the Bakken Shale. *Available at SSRN 2481321*.

Dickstein, Michael J. 2011. Efficient provision of experience goods: Evidence from antidepressant choice. *Job Market Paper, Harvard University*.

Doraszelski, Ulrich, & Pakes, Ariel. 2007. A framework for applied dynamic analysis in IO. *Handbook of industrial organization*, **3**, 1887–1966.

Doraszelski, Ulrich, Lewis, Gregory, & Pakes, Ariel. 2014. *Just starting out: Learning and price competition in a new market*. Tech. rept. Citeseer.

Ericson, Richard, & Pakes, Ariel. 1995. Markov-perfect industry dynamics: A framework for empirical work. *The Review of Economic Studies*, **62**(1), 53–82.

Evans, George W, & Honkapohja, Seppo. 2012. *Learning and expectations in macroeconomics*. Princeton University Press.

Fershtman, Chaim, & Pakes, Ariel. 2012. Dynamic games with asymmetric information: A framework for empirical work*. *The Quarterly Journal of Economics*, qjs025.

Fudenberg, Drew, & Tirole, Jean. 1983. Capital as a commitment: Strategic investment to deter mobility. *Journal of Economic Theory*, **31**(2), 227–250.

Greenwood, Robin, & Hanson, Samuel G. 2015. Waves in ship prices and investment. *The Quarterly Journal of Economics*, **55**, 109.

Guiso, Luigi, & Parigi, Giuseppe. 1999. Investment and demand uncertainty. *Quarterly Journal of Economics*, 185–227.

Hansen, Lars Peter. 2007. Beliefs, doubts and learning: Valuing macroeconomic risk. *The American Economic Review*, **97**(2), 1–30.

Ifrach, Bar, & Weintraub, Gabriel Y. 2016. A Framework for Dynamic Oligopoly in Concentrated Industries.

Jovanovic, Boyan. 1982. Selection and the Evolution of Industry. *Econometrica: Journal of the Econometric Society*, 649–670.

Kalouptsidi, Myrto. 2014. Time to build and fluctuations in bulk shipping. *The American Economic Review*, **104**(2), 564–608.

Kellogg, Ryan. 2014. The effect of uncertainty on investment: evidence from Texas Oil Drilling. *The American Economic Review*, **104**(6), 1698–1734.

Kydland, Finn E, & Prescott, Edward C. 1982. Time to build and aggregate fluctuations. *Econometrica: Journal of the Econometric Society*, 1345–1370.

Malmendier, Ulrike, & Nagel, Stefan. 2016. Learning from inflation experiences. *The Quarterly Journal of Economics*, **131**(1), 53–87.

Mankiw, N Gregory, & Whinston, Michael D. 1986. Free entry and social inefficiency. *The RAND Journal of Economics*, 48–58.

Maskin, Eric, & Tirole, Jean. 2001. Markov perfect equilibrium: I. Observable actions. *Journal of Economic Theory*, **100**(2), 191–219.

Milani, Fabio. 2007. Expectations, learning and macroeconomic persistence. *Journal of monetary Economics*, **54**(7), 2065–2082.

O'brien, Patricia A. 1987. Coordinating market forces: the anatomy of investment decisions in the Japanese steel industry, 1945-1975. *Business and Economic History*, 205–209.

Orlik, Anna, & Veldkamp, Laura. 2014. *Understanding uncertainty shocks and the role of black swans*. Tech. rept. National Bureau of Economic Research.

Orphanides, Athanasios, & Williams, John C. 2005. The decline of activist stabilization policy: Natural rate misperceptions, learning, and expectations. *Journal of Economic dynamics and control*, **29**(11), 1927–1950.

Pakes, Ariel, Ostrovsky, Michael, & Berry, Steven. 2007. Simple estimators for the parameters of discrete dynamic games (with entry/exit examples). *The RAND Journal of Economics*, **38**(2), 373–399.

Pindyck, Robert S. 1988. Irreversible Investment, Capacity Choice, and the Value of the Firm. *The American Economic Review*, 969–985.

Reinhardt, Forest L, Casadesus-Masanell, Ramon, & Nellemann, Frederik. 2012. Maersk Line and the Future of Container Shipping.

Reynolds, Stanley S. 1987. Capacity investment, preemption and commitment in an infinite horizon model. *International Economic Review*, 69–88.

Ryan, Stephen P. 2012. The costs of environmental regulation in a concentrated industry. *Econometrica*, **80**(3), 1019–1061.

Spence, A Michael. 1977. Entry, capacity, investment and oligopolistic pricing. *The Bell Journal of Economics*, 534–544.

Stopford, M. 2009. *Maritime Economics, 3rd.*

A Appendix

A.1 Computation

The computational algorithm is analogous to the standard value function iteration algorithm except for an extra simulation step. As discussed in Section 3.1, the transition of the moment-based industry state $\hat{s}$ may not be Markov even if the underlying industry state s is. Therefore, a simulation step is used to generate the Markov approximation of the transition of moment-based industry 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 η. 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, μ^*, that W^{η} implies.

2. Simulate a sample path of $\{\hat{s}_t\}_{t=1}^{T}$ for large T based on μ^*. 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:
$$\tilde{W}^{\eta}(\mu, x, \hat{s}) = \pi(x, \hat{s}) - \iota\kappa(x, \hat{s}) + \nu(\delta, x)\phi(x, \hat{s}) + \beta E_{a,\mu}\left[V^{\eta}(x', \hat{s}'|x, \hat{s})\right]$$
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: $\max_{x,\mu} \sum_{s \in \hat{\mathcal{S}}} |\tilde{W}^{\eta}(\mu, x, \hat{s}) - W^{\eta}(\mu, x, \hat{s})| h(\hat{s})$.

5. If the norm is greater than ε, update the values and the strategy profile with $\tilde{W}$ and $\tilde{\mu}^*$ and repeat steps 2-5.

A.2 A Time-Varying Volatility Model with Full Information

This paper has considered only homoskedastic models so far. However, changes in demand volatility may be able to explain firms' investment behavior under demand fluctuations without the need

to introduce parameter learning in the model. This section explores a full-information model with time-varying volatility with no parameter learning and examines whether such model can successfully predict investment patterns observed in the data.

Demand is assumed to follow an $AR(1)$ process for both the Asia-Europe and outside markets as in the previously studied full-information model. But volatility is assumed to follow a $GARCH(1,1)$ process instead of having a normal distribution with a constant variance such that the current periods' variance depends on the last period's realized error and the last period's variance:

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

where ω_{t-1} is the realized demand shock in period $t-1$. For the outside market the GARCH term $\tilde{b}_1$ is not significantly different from 0 so I estimate an $ARCH(1)$ model instead given as follows:

$$\tilde{\sigma}_t^2 = \tilde{a}_0 + \tilde{a}_1\tilde{\omega}_{t-1}^2.$$

These models are estimated using the full sample of data (1997:Q1-2014:Q4). The estimates are presented in table 14 and inferred conditional variances are plotted in figure 15. The variances for the both markets spike around2009. And the variances are higher on average in the post-2008 period. Compared to volatility estimates under learning models which are shown in figures 9 and 10, the jump in the variance around 2009 is larger and is more temporary.

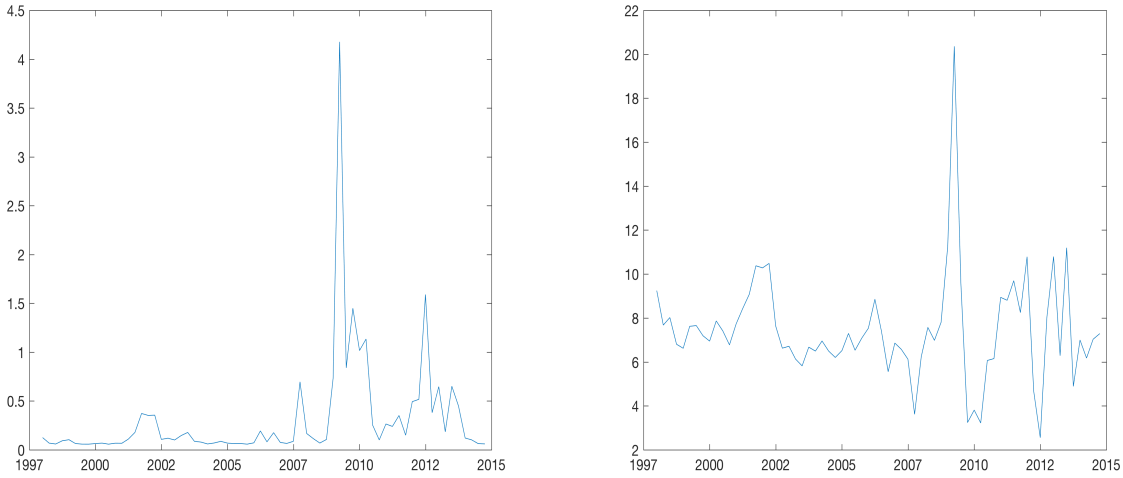
Table 14: Estimates of the Time-Varying Volatility Models

Asia-Europe Market (GARCH)		
a_0	a_1	b_1
0.05	0.83	0.17
(0.01)	(0.28)	(0.16)
Outside Market (ARCH)		
$\tilde{a}_0$	$\tilde{a}_1$	
0.34	0.73	
(0.11)	(0.25)	

Notes: Standard errors are in parentheses.

The dynamic model presented in section 3.1 is modified to accommodate time-varying volatility. Under time-varying volatility firms need the current period's variances for the two markets and the realized error term in the Asia-Europe market in addition to the $AR(1)$ parameters to predict next period's variance so they can predict future demand. Thus, now firms' beliefs at time t is given by $\eta_t = (\rho^0, \rho^1, \sigma_t, \tilde{\rho}^0, \tilde{\rho}^1, \tilde{\sigma}_t, \omega_t)$. Note that there is no time subscript on the constant the slope coefficient of the $AR(1)$ model, since they stay constant over time. The full-information model with

Figure 15: Inferred Conditional Variances under the Time-Varying Volatility Models



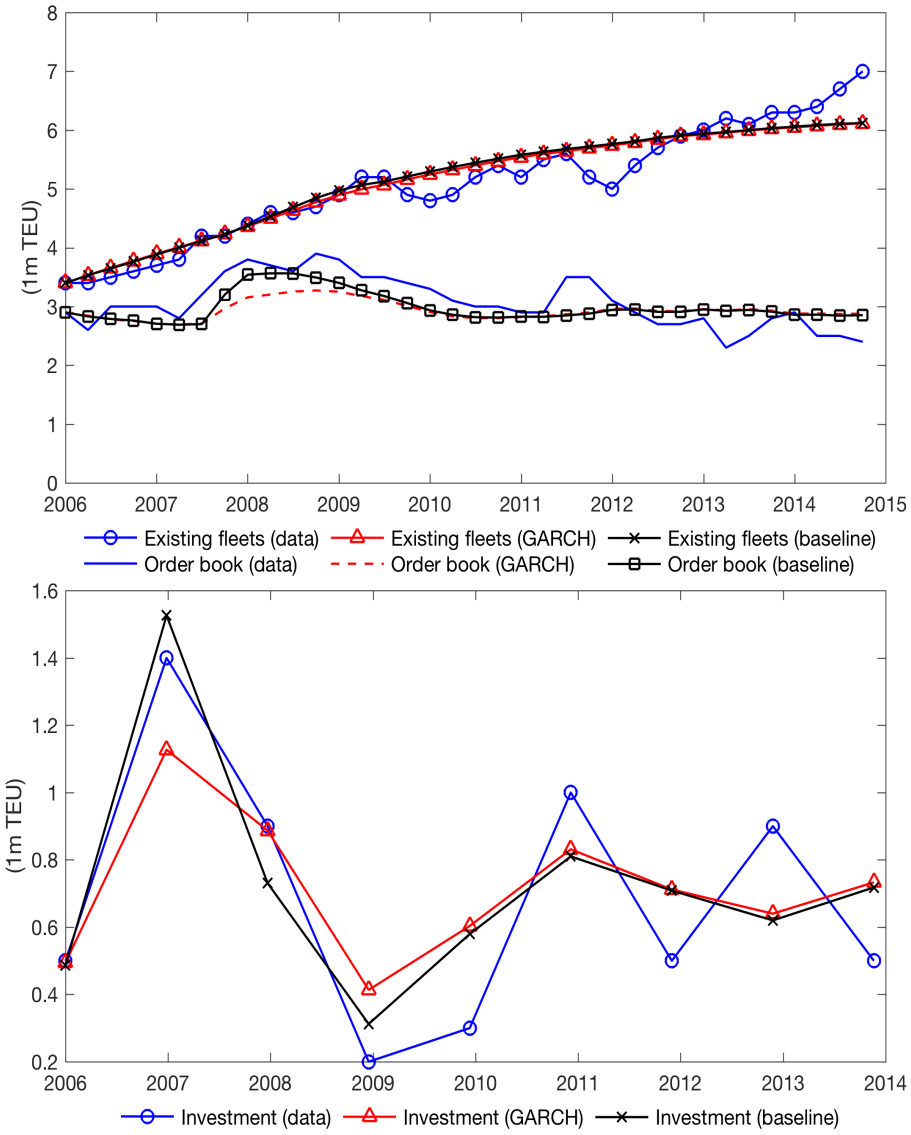
(a) Asia-Europe market

(b) Outside market

time-varying demand volatility is estimated in the same way learning models are estimated. Since the realized error and the conditional variances are different in every period, the model is solved for each period for each candidate set of parameters. Then, I obtain the model fit by searching for the set of parameters that minimizes the distance between simulated moments and data moments as given in equation (15).

Figure 16 shows the model fit. A full-information model with time-varying demand volatility predicts the investment boom and the bust better than a full-information model with constant volatility. In particular, the time-varying volatility model predicts that investment peaks in 2007 and bottoms in 2009 as observed in the data. Nevertheless, adaptive learning models still provides a better fit to the data. In particular, the magnitude of the investment boom and the bust is muted under the time-varying-volatility model with full information. This finding provides some insights about firms' beliefs about demand. It shows that low volatility in demand in the pre-2008 period and a sharp increase in volatility around 2008-2009 that the time-varying volatility model implies help predict the high level of investment in the pre-2008 period and the fall in investment in 2009. The finding also suggests that changing means along in addition to changing variances in agents' beliefs may be necessary to correctly predict firms' investment behavior.

Figure 16: Model Fit of the Time-Varying Volatility Model with Full Information

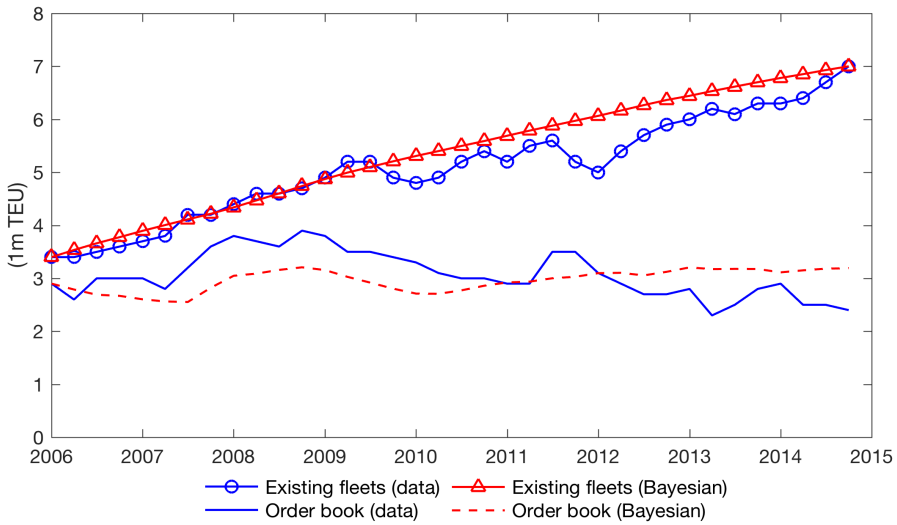


Notes: The top panel compares the industry evolution simulated under the time-varying-volatility model with full information to the industry evolution observed in the data. The bottom panel compares yearly investment simulated under the time-varying volatility model with full information to yearly investment observed in the data. The simulations are based on 1000 equilibrium paths.

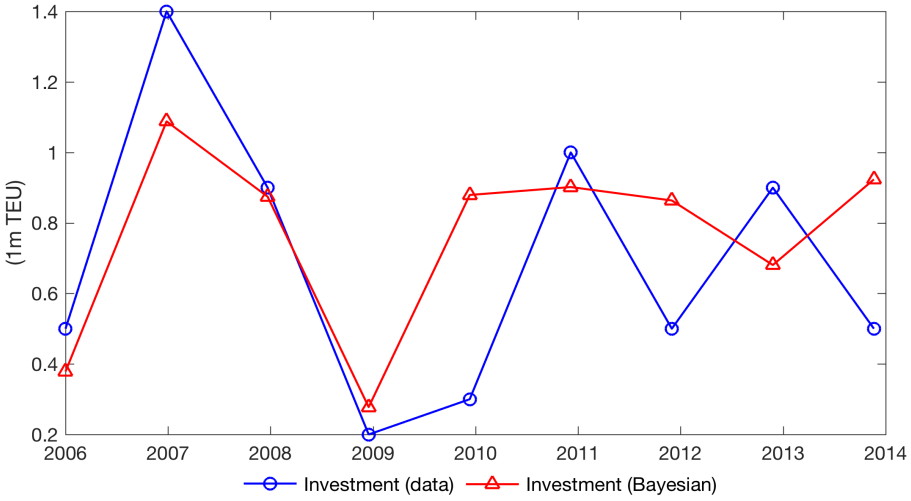
A.3 The Fit of the Bayesian Learning Model

I solve the dynamic model of investment for all candidate parameter vectors based on firms' beliefs under the Bayesian learning model as described in subsection 5.4. Then, I search for the set of parameters that minimizes the distance between the data and simulated moments given in equation (15). Figure 17 shows the fit of the Bayesian learning model.

Figure 17: Model Fit of the Bayesian Learning Model



(a) Capacity of Owned Ships and Order book



(b) Yearly Investment

Notes: The top panel compares the industry evolution simulated under the Bayesian learning model to the industry evolution observed in the data. The bottom panel compares yearly investment simulated under the Bayesian learning model to yearly investment observed in the data. The simulations are based on 1000 equilibrium paths.

A.4 Adding a Dominant Firm's State in the Moment-Based State

The moment-based Markov equilibrium as originally proposed in Ifrach & Weintraub (2016) allows firms to keep track of the detailed state of dominant 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 the sum of states of all firms in the industry. However, MME strategies may not be optimal, if moments do not summarize all payoff relevant information. To address this issue, 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 consider their strategy on the firm-specific state of the largest firm (referred to as 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 μ' 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:

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

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 15 shows that there are no significant difference in model predictions and that the average difference in values is not significantly different from zero for both the model with the dominant firm's capital state and the one with the order book state.

A.5 Credit Market Conditions

In the sample period that this study focuses from 2006 to 2014, there were sharp swings in credit market conditions along with swings in trade demand. This section examines whether credit market conditions played a significant role in determining firms' investment behavior using data from Compustat on company financials information, in particular the firm's debts and liabilities. Only 281 company-quarter-level observations on company financials are available out of 612 observations used in the main analysis.[41] 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.

Using these observations, I regress investment levels on state variables (including the firm's

[41] I am currently in the process of obtaining more data through the Amadeus database.

Table 15: Adding a Dominant Firm's State in the Moment-Based State

Panel A: Simulated moments						
	Baseline		Model with dominant firm's capital state		Model with dominant firm's order book state	
Average investment in 2006-2008 (1 mil. TEU)	0.23	(0.03)	0.23	(0.03)	0.23	(0.03)
Average investment in 2009-2014 (1 mil. TEU)	0.15	(0.02)	0.15	(0.02)	0.15	(0.02)
Total capacity of owned ships (1 mil. TEU)	5.15	(0.27)	5.13	(0.28)	5.14	(0.28)
Total capacity in the order book (1 mil. TEU)	2.98	(0.14)	2.98	(0.14)	2.99	(0.14)
Correlation between demand and investment	0.22	(0.12)	0.21	(0.12)	0.22	(0.12)
Std. dev. in investment (1 mil. TEU)	0.17	(0.03)	0.17	(0.03)	0.17	(0.03)
Panel B: Average difference in values						
All firms (%)			-0.35	(0.46)	-0.42	(0.54)
Dominant firm (%)			-0.18	(0.22)	-0.21	(0.26)
Fringe firms (%)			-0.36	(0.48)	-0.43	(0.56)

capital stock, order book capacity, aggregate capital stock, aggregate order book capacity, demand states, chartered capacity, and deployed capacity in the Asia-Europe market) and variables relating to the firm's credit constraints including long-term debt. If financial constraints are the main driver of the boom and bust in investment, we expect that firms which hold higher amounts of debts thus facing harsher credit constraints will withhold investment more. The regression results presented in table 16, nonetheless, suggest that debt levels do not have significant effects on firms' investment.

A.6 Robustness Checks for the Adaptive Learning Model

As described in section 3.4, the adaptive learning model was implemented under the truncation approach. This section shows beliefs about demand recovered under the imputation approach as well as the model fit. Figures 18 and 19 show the beliefs for the Asia-Europe market and the outside market, respectively, for the baseline learning case of $\lambda_t = 0.02$ and also the case of $\lambda_t = 0.04$. As λ_t increases the beliefs become closer to one another as there are less weights placed on the imputed data. The model fits under the two approaches are very close to one another, although it is slightly better under the truncation approach as shown in table 17.

I also show that for the truncation approach agents' beliefs about demand under the baseline learning model do not vary widely when changing the start date of the sample data used in the estimation. Figure 20 shows beliefs for the 2006-2014 period estimated with start dates of 1997:Q1 and 1998Q1.

Table 16: Regression of investment on debt-related variables

Dependent variable: Investment (1000 TEU)		
Owned ship capacity (1000 TEU)	-.037	(.027)
Order book capacity (1000 TEU)	-.024	(.017)
Aggregate owned ship capacity (1000 TEU)	.012	(.01)
Aggregate order book capacity (1000 TEU)	-.015**	(.0064)
Demand state (Asia to Europe)	1.1	(2.3)
Demand state (Outside market)	.06	(1.3)
Chartered ship capacity (1000 TEU)	-.025	(.024)
Aggregate chartered ship capacity(1000 TEU)	-.019*	(.011)
Deployment in Asia-Europe market (1000 TEU)	.087**	(.043)
Aggregate deployment in Asia-Europe market (1000 TEU)	.019**	(.0078)
Long-term debt (1 bil. US dollars)	.00079	(.002)
Debt in current liablities (1 bil. US dollars)	-.0019	(.0029)
Constant	-11	(38)
Observations	281	
R^2	0.076	

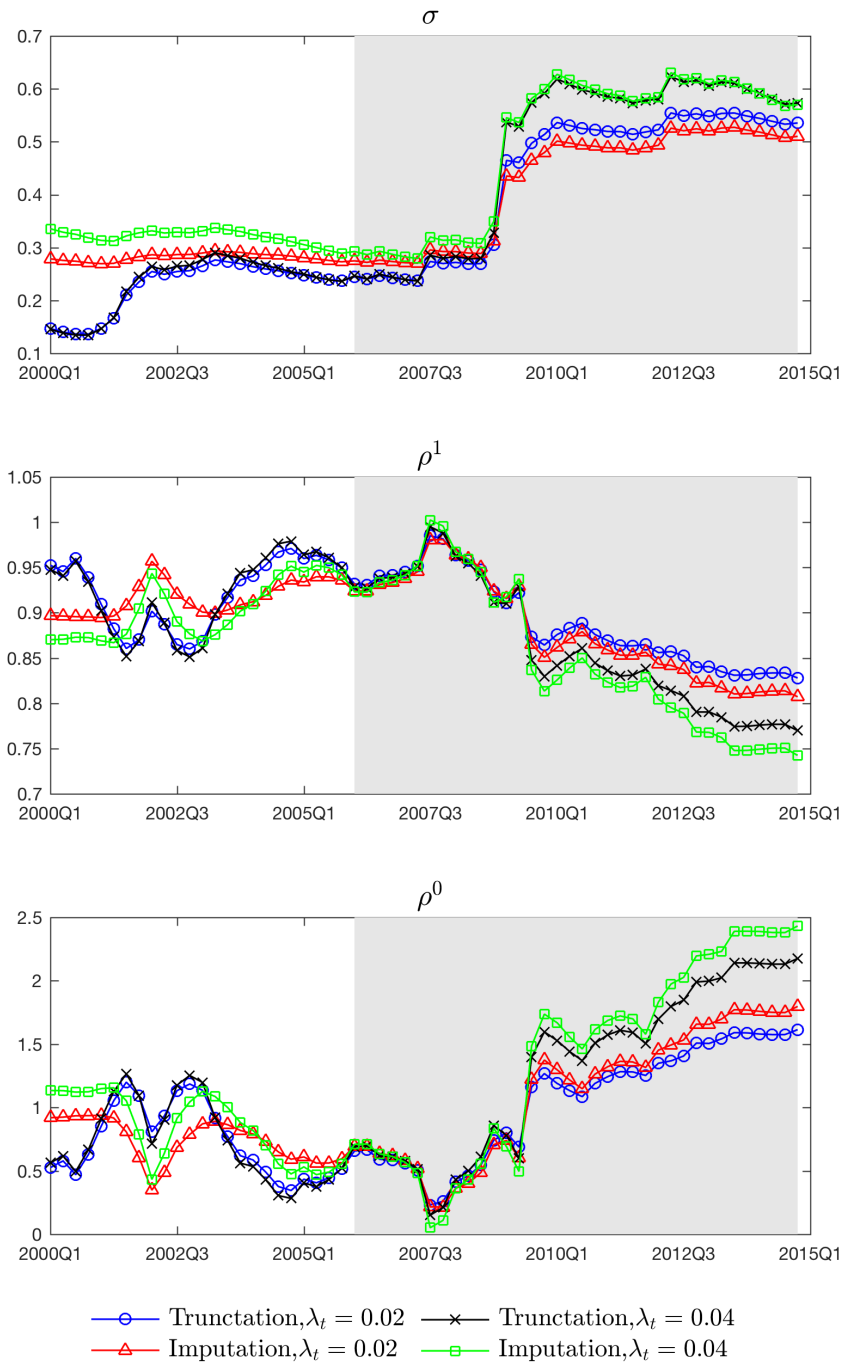
Notes: Standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.001$.

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

	Data	Truncation	Imputation
Average investment in 2006-2008 (1 mil. TEU)	0.23	0.23	0.22
Average investment in 2009-2014 (1 mil. TEU)	0.14	0.15	0.16
Total capacity of owned ships (1 mil. TEU)	5.09	5.15	5.17
Total capacity in the order book (1 mil. TEU)	3.07	2.98	2.98
Correlation between demand and investment	0.19	0.22	0.26
Std. dev. in investment (1 mil. TEU)	0.17	0.17	0.18

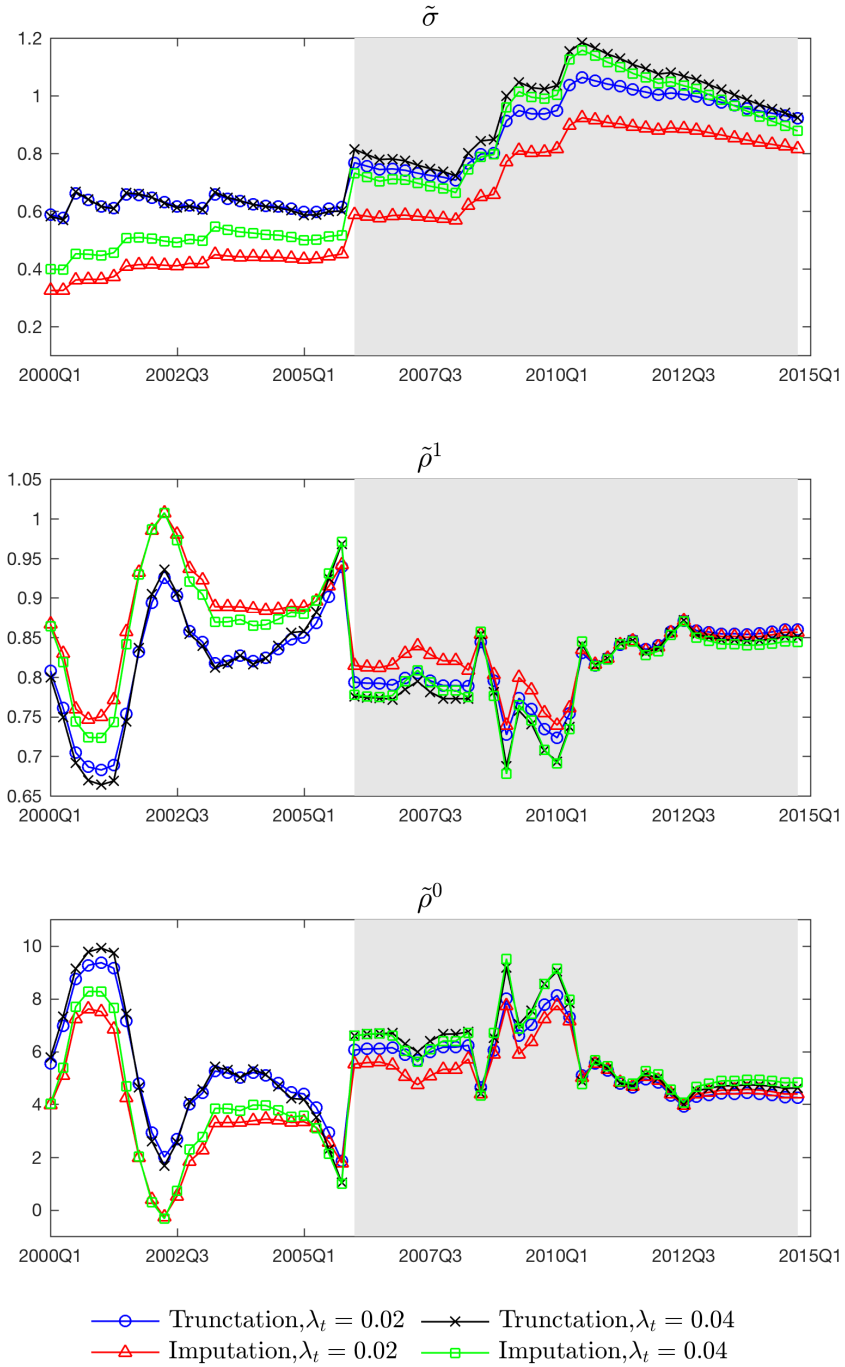
Notes: This table compares moments observed in the data and moments simulated under the truncation and imputation approaches of the baseline learning model. The simulated moments are computed based on 1000 series of equilibrium paths.

Figure 18: Beliefs about the Asia-Europe Market Demand 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 cases of $\lambda_t = 0.02$ and $\lambda_t = 0.04$. Beliefs for 2006-2014 in the shaded area are used in the main analysis.

Figure 19: Beliefs about the Outside Market Demand 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 $\lambda_t = 0.02$ and $\lambda_t = 0.04$. Beliefs for 2006-2014 in the shaded area are used in the main analysis.

Figure 20: Beliefs under the Baseline Learning Model for Different Sample Start Dates