Time to Build and Fluctuations in Bulk Shipping

Myrto Kalouptsidi[*†]
Department of Economics, Princeton University

September 2011

Abstract

This paper explores the nature of fluctuations in world bulk shipping by quantifying the impact of demand uncertainty and time-to-build on investment and prices. We examine the impact of both construction lags and their lengthening in periods of high investment activity, by constructing a dynamic model of ship entry and exit. A rich dataset of second-hand ship sales allows for a new estimation strategy: resale prices provide direct information on value functions and allow their nonparametric estimation. We find that moving from time-varying to constant to no time-to-build reduces prices, while significantly increasing both the level and volatility of investment.

Keywords: time to build, bulk shipping, industry dynamics, second-hand sales, demand uncertainty, price volatility, investment volatility

[*]Princeton University, Department of Economics, 315 Fisher Hall, Princeton, NJ 08544-1021. email: myrto@princeton.edu

[†]I am extremely thankful to my advisors, Steve Berry, Phil Haile and Lanier Benkard for their numerous useful comments and suggestions. I have also gained a lot from discussions and comments from Eduardo Engel, Costas Arkolakis, Chris Conlon, Rick Mansfield.

1 Introduction

Under uncertainty, time to build can substantially affect investment and prices both in terms of their level and volatility. Time to build was introduced by Kydland and Prescott (1982) into their famous RBC model used to capture aggregate economic fluctuations. Many industries are characterized by demand uncertainty and multi-period construction processes and the world bulk shipping industry is an outstanding example. In recent years, the rapid growth of imports of raw materials in developing countries such as China, was accompanied by a large increase in freight rates, shown in Figure 1. This led to an unprecedented surge in new ship orders, shown in Figure 2. Long construction lags inherent in shipbuilding, as well as bottlenecks due to shipyard capacity constraints that often bind, create time to build: a gap between the order and delivery date of a ship. As a result, when the crisis of 2008 hit and the operating fleet suffered from an inordinately low demand level, another 70% of that fleet was scheduled to be delivered between 2008 and 2012.

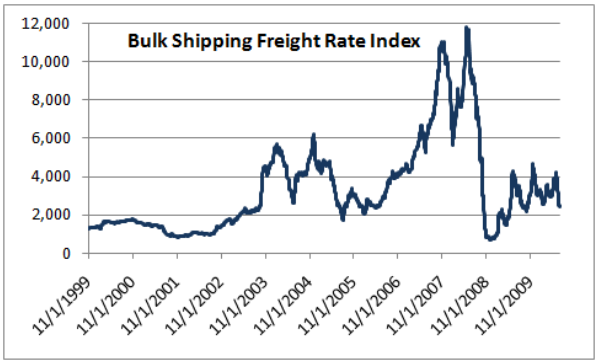

Figure 1: The Baltic Dry Index: daily index based on weighted average of rates on 20 representative bulk routes. Compiled by the Baltic Exchange.11/1/1999 = 1334.

This paper explores the nature of fluctuations in the world bulk shipping industry by quantifying the impact of demand uncertainty and time to build on investment, shipping prices and shipper surplus. We examine the impact of both pure construction lags, as well as the lengthening of delivery lags in periods of high investment activity due to capacity constraints.

We construct a dynamic model of ship entry and exit. We apply a novel estimation strategy using a rich dataset of second-hand ship sales which provides direct information on key dynamic objects of the model and allows for their nonparametric estimation. We find that moving from time-varying to constant to no time to build reduces prices, while

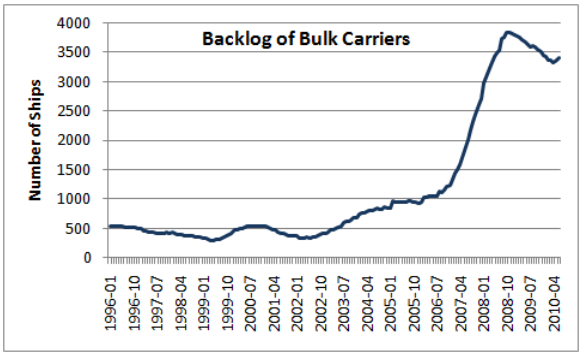

Figure 2: Total number of dry bulk carriers under order.

significantly increasing both the level and volatility of investment.

Our model builds on the two key features: first, demand for sea transport is inherently uncertain and volatile; second, supply adjusts sluggishly due to entry costs, time to build and convex operating costs of ships. In the model, a firm is a ship. Firms enter, age and endogenously exit. A free entry condition every period determines the number of entrants. Due to time to build, operation begins a number of periods after the entry decision. This number of periods varies over time: it increases due to bottlenecks in production in periods of high investment activity. Once a firm is in the market, it ages and decides whether to exit (demolish the ship) or continue to operate in the market.

We use the second-hand price of a ship to estimate its value function nonparametrically: the assumption of a large number of potential shipowners implies that the resale price equals the value of the ship and allows for our new estimation strategy. Our methodology is a "mirror image" of previous work on the estimation of dynamic settings, such as Rust (1987) and Hotz and Miller (1993) on single agent dynamics and Bajari, Benkard and Levin (2008), Pakes, Ostrovsky and Berry (2008), Aguirregabiria and Mira (2007), Jofre-Bonet and Pesendorfer (2003) and others, on dynamic games. In those papers, estimated per period payoffs are used to recover value functions, which in turn reveal exit and entry costs. In this paper, we estimate value functions directly from the data and use them to retrieve nonparametrically entry and exit costs, as well as per period payoffs. This methodology can be of use in other settings with a large number of homogeneous agents.

We test our model and empirical strategy against two time series not used in the estimation. In particular, we compare the estimated profits to an earnings index calculated by shipbrokers, and the estimated entry costs to the average shipbuilding price and find that they match surprisingly well. In addition, we use the estimated profits of the model

to detect that firms act as price takers. The implied operating cost estimates are very close to industry estimates.

We compute the equilibrium of our model in a counterfactual world without shipbuilding capacity constraints, where time to build is a constant number of periods corresponding to pure construction lags, as well as a counterfactual world with no time to build. Governmental subsidies to the shipbuilding industry provide a mechanism that can change the time to build. Time to build becomes a critical constraint at positive demand shocks. Simulating an unexpected positive demand shock we find that moving from time-varying, to constant, to no time to build reduces prices, while increasing the fleet and shipper (consumer) surplus. In particular, following a positive demand shock, in a world with time to build the supply response is limited, as entry is both lower and slower. Indeed, new ships arrive a number of periods later, while in the meantime supply is limited to the existing capacity constrained fleet. In addition, the incentives to enter are dampened since the shock fades due to a mean-reverting demand process and ships that are delivered late will not be able to take advantage of the temporarily increased demand for shipping services. As a result, time to build leads to high profits and prices and low consumer (i.e. shipper) surplus. Long run simulations of our model further shed light on the dynamic evolution of the industry. We find that in the absence of time to build the fleet is about 20% larger. In addition, the fleet is about 50% more volatile under constant time to build and almost three times more volatile under no time to build. In contrast, prices are lower and less volatile as time to build declines. In the absence of time to build, investment is more responsive to demand and as a result, prices are less responsive to demand. Finally, we explore the oscillatory and steady state behavior of investment. We find that time to build leads to investment oscillations that fade extremely slowly.

A large theoretical literature has explored the impact of uncertainty on investment, such as Dixit (1989), Lambson (1991), Pindyck (1993), Caballero and Pindyck (1996) and many others. Empirical work on uncertainty and investment includes Bloom (2009), Collard-Wexler (2008) and Bloom, Bond and Van Reenen (2007). Little empirical work on the impact of time to build can be found in the literature; Rosen, Murphy and Scheinkman (1994) explore time to build and periodicity in breeding cattle. Deaton and Laroque (1992) study the volatility of commodity prices. The contribution of the present paper to the literature is to measure the impact of both uncertainty and time to build in the context of a particular industry, that of oceanic bulk shipping. At the same time, this industry, representing 70% of world seaborne trade (in tons), is important in its own right: both freight rates and inventory costs caused by shipping delivery lags can significantly affect

trade flows, as documented by Alessandria, Kaboski and Midrigan (2009).

The remainder of the paper is organized as follows: Section 2 provides a description of the industry. Section 3 presents the model. Section 4 describes the data used. Section 5 presents the empirical strategy employed and the estimation results. Section 6 provides the counterfactual experiments and Section 7 concludes.

2 Description of the Industry[1]

Bulk shipping concerns vessels designed to carry a homogeneous unpacked dry or liquid cargo, for individual shippers on non-scheduled routes. The entire cargo usually belongs to one shipper (owner of the cargo). Bulk carriers operate like taxi drivers, instead of buses: they carry a specific cargo individually in a specific ship for a certain price. Bulk ships do not operate on scheduled itineraries like container ships, but only via individual contracts. In practice there are a few different types of cargo transport agreements that involve different methods of payment (e.g. price per ton, price per day) and duration. Dry bulk shipping involves mostly raw materials, such as iron ore, steel, coal, bauxite, phosphates, but also grain, sugar and wood chips. There are four different categories of bulk carriers based on size: Handysize (10,000-40,000 DWT), Handymax (40,000-60,000 DWT), Panamax (60,000-100,000 DWT) and Capesize (larger than 100,000 DWT). Vessels in different categories carry different products, take different routes and approach different ports. Practitioners treat them as different markets. Each such market consists of a large number of shipowning firms. For instance, Figure 3 shows the firm size distribution of Handysize bulk carriers (where the empirical analysis will focus) in terms of the number of ships owned by each firm, as well as its market share (captured by its fleet share). Note that the largest share is 3% while there is a long tail of small shipowners (one to two ships per firm). Even though there is some scope for differentiation (based on the age of the ship, the shipyard where it was built at, the reputation of the shipowner, etc.) shipping services are perceived as homogeneous.

Demand for shipping services is driven by world seaborne trade and is thus subject to world economy fluctuations. It is also rather inelastic, due to the lack of alternative shipping methods for bulk commodities. In recent years, the growth and infrastructure building at several countries led to increased imports of raw materials, significantly boosting demand for bulk transport. For instance, Chinese imports grew by 15% every year between 2004 and 2007 (UNCTAD (2009)), while Asian imports represent more than 50%

[1] This section draws from Stopford (2009), as well as Beenstock and Vergottis (1989).

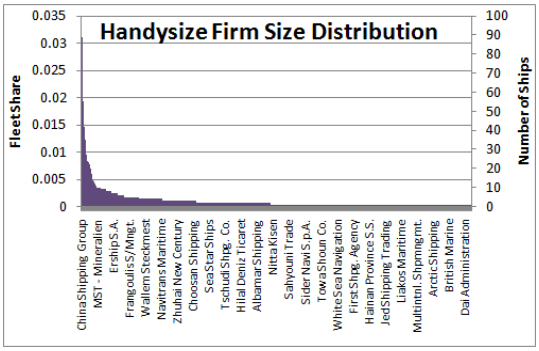

Figure 3: Fleet and fleet shares of shipowning firms.

of world imports; Figure 4 shows China's imports of iron ore and coal.

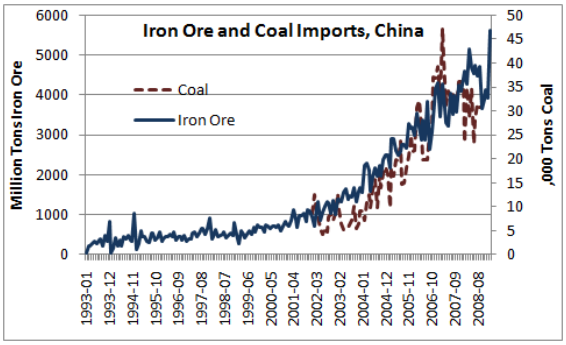

Figure 4: China's iron ore and coal imports.

Supply of shipping services is determined in the short run by the number of voyages carried out by shipowners. The number of voyages is proportional to the speed of the vessel. Voyage costs include fuel, port/canal dues and cargo handling. They are convex in the chosen speed: there are only so many trips that a ship can execute in a given time period. In addition, voyage costs increase with the ship's age, as its fuel efficiency and overall operation deteriorates over time. Ships also face fixed costs that include maintenance and insurance.

In the long run, the supply of cargo transportation is restricted by the size of the world fleet and adjusts via the building and scrapping of ships. Exit in the industry occurs when shipowners scrap their ships. During the scrapping procedure, the interiors of the ship (machinery, furniture, oils, etc.) are removed from the vessel and sold. The ship is then taken to a scrapyard, where it is dismantled and its steel hull is recycled.[2]

[2] A shipowner also has the option of selling the ship in the second-hand market. In this case the ship continues to operate and does not exit the industry.

Entry in the industry occurs when shipowners buy new ships from world shipyards. The building of new ships is characterized by significant construction lags. In addition to the actual construction time, shipyards often face binding capacity constraints due to their limited number of assembly docks. In that case queues of ship orders are formed, that increase the time to build. For instance, the average time to build almost doubled between 2001 and 2008 (documented in Figure 10). Finally, it is worth noting that cancelling a shipyard order is hard, due to the nature of shipbuilding contracts, as well as the fact that bottlenecks occur in the final stages of production, during the assembly of ship parts on the dry dock. In fact even during the 2008-2009 crisis, only about 10% of the entire orderbook (in all ship types) was cancelled (UNCTAD (2009)).

3 Model

In this section, we present a dynamic model of the world bulk shipping industry. Our model lies within the general class of dynamic games studied in Ericson and Pakes (1995)[3] and is closest to the model of entry and exit in Pakes, Ostrovsky and Berry (2007). There are two distinct differences: our model includes firm heterogeneity (due to age) and time to build.

3.1 Environment

Time is discrete and the horizon is infinite. A ship is a firm and the terms firm and ship will be used interchangeably. Constant returns to scale with respect to the fleet is a reasonable assumption for this industry, while the firm size distribution shown in Figure 3 suggests that many firms consist indeed of one ship. There are two types of agents: incumbent firms and a large number of identical potential entrants. There is time to build, in other words, a firm begins its operation a number of periods after its entry decision. The state variable of a ship consists of:

1. its age, $j \in \{0, ..., A\}$[4]

[3] A large number of models are built on this framework, such as Benkard (2004) and Besanko and Doraszelski (2004). Our model is also related to industry models such as Hopenhayn (1992).

[4] The maximum age A is determined endogenously by the model. Due to data limitations, in the empirical model below we assume that A is 20 years and that it thus represents the age group of "older than 20 years old". In other words, once the ship reaches age 20 it stays in that age group until it (endogenously) exits. See Section 4 and Appendix A for more details.

2. the age distribution of the fleet, $s_t \in \mathbb{R}^{A+1}$, where $s_t = \left[s_t^0 \; s_t^1 \; ... \; s_t^A\right]$ and s_t^i is the number of ships of age i

3. the backlog $b_t \in \mathbb{R}^{\overline{T}}$, where $b_t = [b_t^1 \; b_t^2 \; ... \; b_t^{\overline{T}}]$ and b_t^i, $i = 1, ..., \overline{T}$, is the number of ships scheduled to be delivered at period $t + i$, while $\overline{T}$ is the maximum possible time to build[5]

4. the aggregate demand for shipping services, $d_t \in \mathbb{R}$, captured by the intercept of the inverse demand curve for freight transport (in logarithmic scale).

Firms face a log-linear inverse demand curve,

$$P_t = D_t Q_t^{\eta} \tag{1}$$

where P_t is the price per voyage, Q_t is the total number of voyages offered in period t and $d_t = \log D_t$. The demand state variable d_t, follows an exogenous first order Markov process. Per period payoffs of a ship of age j are given by $\pi_j(s_t, d_t)$. $\pi_j(s_t, d_t)$ are the profits resulting from the combination of the demand curve (1) and an assumption on the nature of competition (e.g. perfect competition, Cournot). We do not specify the nature of competition, allowing for flexibility. Note that the backlog b_t does not affect per period payoffs.

Every period, incumbents privately observe a scrap value, ϕ drawn from a distribution F_ϕ and decide whether to exit the market and obtain ϕ or continue operating.

There is free entry into the market, subject to an entry cost $\kappa(s_t, b_t, d_t)$. All entrants of period t receive the same time to build, T_t. This number of periods T_t changes with the state of the industry (s_t, b_t, d_t). As the number of orders under construction increases, waiting time is added to the construction time and the time to build increases. Time to build is a deterministic function of the industry state, $T : \mathbb{R}^{A+\overline{T}+2} \to \mathbb{R}$, so that,

$$T_t = T(s_t, b_t, d_t) \tag{2}$$

Specifying time to build as a function of the entire state (rather than only b_t for instance) captures queues due to period t's entry, which is a function of the industry state (i.e. time to build depends on the number of entrants, N_t). Moreover, it may accommodate for changes in shipbuilding capacities that occur in response to the industry state. Ideally time to build and entry costs should also be a function of shipbuilding capacities. In

[5] We assume that $T_t = T(s_t, b_t, d_t) \leq \overline{T}$ (defined in (2)) and ships are delivered in finite time.

that case, shipbuilding capacity would become a state variable in the firms' dynamic optimization problem. It is straightforward to accommodate for this state variable in the theoretical model, as long as its transition is specified. However, introducing capacities in the empirical application is difficult, due to lack of detailed data and the increase of an already large state space. Assuming that the shipbuilding industry is exogenous in the model is not unreasonable, given that governmental support has been crucial for shipyard functioning (Stopford (2009)). Finally, it is worth noting that the empirical dynamic game literature assumes entry costs and scrap values are iid random variables. This is the first empirical paper to our knowledge to allow for non-iid entry costs, a more realistic assumption.

The timing in each period is as follows: incumbents and potential entrants first observe the state of the industry (s_t, b_t, d_t). Incumbents then receive an exogenous shock with some probability, which forces them to sell the ship in the second-hand market. This shock is independent of the value of the ship (a liquidity or retirement shock). In this model, since all shipowners share the value of a ship, the price of a ship in the second hand market equals its value. In addition, as a firm is a ship and the identity of a firm does not matter, sales can be ignored: they are neither entry -the ship was already in operation- nor exit -the ship will continue operating. At a sale the ship just changes ownership. We revisit sales in the empirical part of the paper, where we treat them as observations on the value function. After incumbents observe the state and trade in the second hand market, they privately observe their scrap value and decide whether they will exit the market or continue operating, while potential entrants make their entry decisions simultaneously. Incumbents receive their operating profits, $\pi_j(s_t, d_t)$. Finally, the entry and exit decisions are implemented and the state evolves.

3.2 Firm Behavior

An incumbent firm of age j maximizes its expected discounted stream of profits, deciding only on exit. Its value function before privately observing the scrap value, ϕ, drawn from F_ϕ, is:

$$V_j(s_t, b_t, d_t) = \pi_j(s_t, d_t) + \beta E_\phi \max\{\phi, VC_j(s_t, b_t, d_t)\} \tag{3}$$

where its continuation value is given by

$$VC_j(s_t, b_t, d_t) \equiv E\left[V_{j+1}(s_{t+1}, b_{t+1}, d_{t+1}) \,|s_t, b_t, d_t\right] \tag{4}$$

The expectation in (4) is over the number of firms that will enter, N_t and exit, Z_t^j, as well as the demand for shipping services, d_{t+1}. Randomness in this model results from these variables, as the remainder of the state evolves deterministically conditional on the current state. In particular, the transition of the age fleet distribution, s_t is as follows:

$$\begin{aligned} s_{t+1}^0 &= b_t^1 \\ s_{t+1}^j &= s_t^{j-1} - Z_t^j, \quad j = 1, ..., A \end{aligned}$$

The number of age 0 ships in period $t+1$ equals the first element of the backlog. The remaining elements of s_t age one period and the number of firms that exit at time t, Z_t^j, is subtracted. The transition of the backlog is as follows:

$$\begin{aligned} &b_{t+1}^i = b_t^{i+1}, \ \ i \neq T_t, \overline{T} \\ &b_{t+1}^{T_t} = b_t^{T_t+1} + N_t, \ b_{t+1}^{\overline{T}} = 0, \ \text{ if } T_t < \overline{T} \\ &b_{t+1}^{\overline{T}} = b_t^{\overline{T}} + N_t, \ \text{ if } T_t = \overline{T} \end{aligned}$$

The first element of the backlog enters the market and the remaining elements move one period closer to delivery. Period t entrants, N_t, enter at position $T_t = T(s_t, b_t, d_t)$. If this position is not the maximum time to build $\overline{T}$, zero ships are added at the end of the backlog. Finally, as mentioned above, the demand state variable d_t, follows an exogenous first order Markov process.

An incumbent firm exits if the scrap value drawn, ϕ, is higher than its continuation value, $VC_j(s_t, b_t, d_t)$. This event occurs with probability:

$$\Pr(\phi > VC_j(s_t, b_t, d_t)) = 1 - F_\phi(VC_j(s_t, b_t, d_t)) \tag{5}$$

where F_ϕ is the distribution of the scrap values. The number of firms that exit at age j, Z_t^j, thus follows the Binomial distribution. We approximate the exit rate by a Poisson process. This approximation is strongly supported by the data, as shown in the following sections and formally justified if the number of firms at age j is large, while the exit probability is small.

We next turn to potential entrants. The value of entry for a ship is:

$$VE(s_t, b_t, d_t) \equiv \beta^{T_t} E[V_0(s_{t+T_t}, b_{t+T_t}, d_{t+T_t}) | s_t, b_t, d_t]$$

Note that due to time to build, a firm that decides to enter in period t will begin operating $T_t = T(s_t, b_t, d_t)$ periods later. The firm, thus, needs to compute the time to build

T_t and then forecast the state that will prevail T_t periods in the future. It therefore forms expectations over the number of entrants until delivery, $\{N_t, N_{t+1}, ..., N_{t+T_t-1}\}$, the number of firms that will exit $\left\{Z_t^j, Z_{t+1}^j, ..., Z_{t+T_t-1}^j\right\}$, all j, as well as demand d_{t+T_t}. Finally, the ship is of age 0 upon delivery.

A free entry condition determines the entry rate:

$$VE\left(s_t, b_t, d_t\right) = \kappa\left(s_t, b_t, d_t\right) \tag{6}$$

Potential entrants play a symmetric mixed entry strategy, so that the number of actual entrants, N_t, follows the Poisson distribution with a state-dependent mean, $\lambda\left(s_t, b_t, d_t\right)$.[6]

The equilibrium concept is that of a Markov Perfect Equilibrium. An equilibrium to our model consists of a pair of entry $\lambda\left(s, b, d\right)$ and exit $\mu\left(s, b, d\right)$ that satisfy conditions (3) to (6). In equilibrium, agents' expectations in VC and VE are formed using the equilibrium strategies. Existence of equilibrium follows by the analysis of Doraszelski and Satterthwaite (2010), Weintraub, Benkard and Van Roy (2008) and Ericson and Pakes (1995).

4 Data

The data employed in this paper are taken from the publications of Clarksons, a shipbroking firm based in the UK. Four different datasets are utilized. The first consists of the world second-hand ship sale transactions. It reports the date of the transaction, the name, age and size of the ship sold, the seller and buyer and the price in million US dollars. The dataset includes sales that occurred between August 1998 and June 2010. We use data on Handysize bulk carriers, the size sector with most second hand sales. We are left with 1838 observations for 48 quarters. Table 1 provides summary statistics of this dataset. Prices exhibit high variance unconditional on the state (in contrast to the variance conditional on the state which is very small as discussed below). Moreover, old ships are sold more often.

The second dataset consists of shipping voyage contracts and includes the date of transaction, the name and size of the ship and the trip's price per ton. The dataset includes a subset of all voyage contracts between January 2001 and June 2010. There are two limitations to these data. First, as Handysize vessels represent the smallest type of bulk carriers, they carry out a large variety of routes making the exhaustive tracking

[6] For a proof see Weintraub, Benkard and Van Roy (2008).

	Mean	St. Dev.
Price in million US dollars	9.97	8.54
Age in years	18.7	7.6
Size in DWT	27876	6718
Number of Sales in a Quarter	41.85	15.3
Number of Observations	1838	

Table 1: Second hand sales summary statistics.

of vessel movement impossible. To correct for the fact that the contracts observed are, therefore, a strict subset of contracts realized, we compute and use the fleet utilization rate of Capesize and Panamax bulk carriers (the two largest categories whose routes are limited) where the full sample of contracts is observed. Second, the number of per period contracts only partly captures transportation offered by Handysize vessels, as trips vary in both time and distance; ideally, we would want the ton-miles realized each quarter by Handysize ships.

The third dataset consists of quarterly time-series for the orders of new ships (i.e. entrants), deliveries, demolitions (i.e. exitors), fleet and total backlog. It is used to construct the age fleet distribution, $s_t \in \mathbb{R}^{A+1}$. A limitation of these series is that they date back to 1970, while the second hand sales data begin in 1998. Moreover, the demolitions time series provides no information on the age of ships scrapped. In a small sample of individual scrap contracts, no ships younger than 20 years old are observed. We construct the age fleet distribution up to 20 years of age or older, so that the last element in s_t, s_t^A, represents the absorbing age group of "older than 20 years (80 quarters) old". We also assume that a firm can decide to exit only once it enters the age group of "older than 20 years old". Therefore, in the empirical model $s_t = [s_t^0, s_t^1, ..., s_t^A] \in \mathbb{R}^{A+1}$ where $A = 80$ and its transition is:

$$\begin{aligned} s_{t+1}^0 &= b_t^1 \\ s_{t+1}^j &= s_t^{j-1}, \quad j = 1, ..., A-1 \\ s_{t+1}^A &= s_t^A + s_t^{A-1} - Z_t \end{aligned}$$

As discussed in the following section, we aggregate the age distribution to three age groups: ships younger than 10 years old, ships between 10 and 20 years old and ships older than 20 years old. Using the estimated continuation values in Section 5.2.1, we find that the exit probability of a ship younger than 10 years old (between 10 and 20) is on

average in the order of 10^{-6} (10^{-4}).

Figure 5 shows the evolution of these three age groups. Old ships remain in operation during the boom, while the number of young ships slowly increases. The same figure shows the total Handysize fleet, which increases during and after the boom.

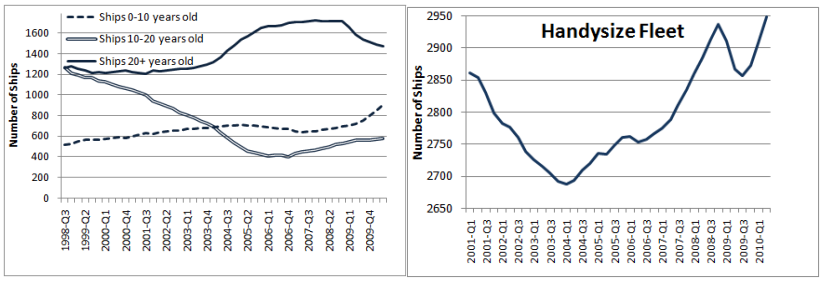

Figure 5: Age distribution across three age groups and total Handysize fleet.

Figure 6 shows the number of ships that enter and exit. The entrant series provides another testament of the large investment swings in bulk shipping: note the spike in new ships orders during 2006-2008. In addition, the variance of entrants seems to follow the mean, as in the Poisson density. Note also the practically zero exit rate between 2005 and 2007 when demand for shipping services was high, followed by a spike in exit at the crisis of 2008. Finally, note that the number of exitors is very small compared to the magnitude of s^A, thus justifying the Poisson approximation to the Binomial random variable Z_t.

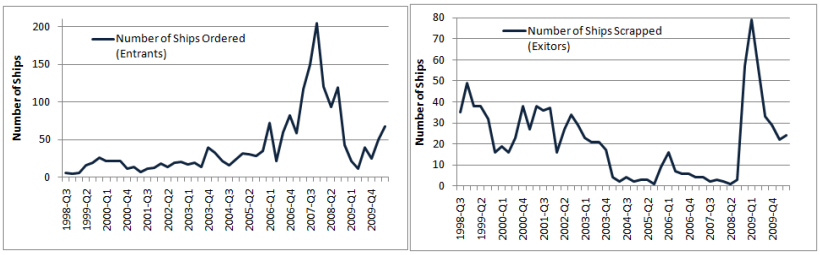

Figure 6: Observed entrants and exitors.

Finally, the fourth dataset is the ship orderbook. It lists all ships under construction each month between 2001 and 2010. The list includes the size and delivery date of each ship under construction. We use this dataset to construct the backlog state variable, b_t and estimate the time to build function (see Appendix A).

5 Empirical Strategy and Estimation Results

The difficulty in the estimation of dynamic games results from the need to compute continuation values. Recently, a number of papers (e.g. Aguirregabiria and Mira (2007), Bajari, Benkard and Levin (2007), Pakes, Ostrovsky and Berry (2007) and Jofre-Bonet and Pesendorfer (2003))[7] have proposed estimation procedures for dynamic games in which per period payoffs are used to compute continuation values, which in turn allow the estimation of entry and exit costs. Following one of these procedures for the bulk shipping industry using our data would be very difficult: estimating per period payoffs requires detailed and reliable information on cargo shipping contracts which we do not have. In this paper, we apply a new estimation strategy using the dataset of second-hand ship sales. We use the second-hand ship sale transactions to recover value functions nonparametrically directly from the data. In particular, since in our model there is a large number of identical potential shipowners, a second-hand transaction price must equal the value of the ship. Value functions, along with estimated state transitions provide the necessary information to recover the primitives of the model: the entry costs, the scrap value distribution and the per period payoffs. These primitives are then used to conduct counterfactual experiments.

The empirical strategy consists of two steps. In the first step we estimate the equilibrium objects necessary to obtain the model primitives. More specifically, we recover the value functions and the transition matrix. In this step we also estimate some exogenous objects that are not determined by the dynamic equilibrium: the demand curve for shipping given in equation (1), the transition of the demand state variable d_t and the time to build function given in (2). These estimates are incorporated in the estimation of the value functions and the transition matrix. Value functions are estimated using the second-hand ship sales. In order to estimate the transition matrix, we combine the exogenous transition rule for d_t and the entry and exit policy functions that we also estimate.

In the second step, we estimate the model primitives, i.e. the profits, the scrap value density and the entry costs. To do so we use the equilibrium objects (value function and transition matrix) obtained in the first step. Profits are calculated from the Bellman equations (3). The scrap value distribution is estimated nonparametrically using the estimated continuation values and observed exit behavior. Finally, the entry costs are obtained from the free entry condition.

[7] Applications of these methodologies have occurred in several contexts, e.g. Snider (2008), Xu (2008), Ryan (2009), Dunne, Klimek, Roberts and Xu (2009).

5.1 First Step: Estimate Exogenous and Equilibrium Objects

We begin with the estimation of the inverse demand curve for shipping services- which we use to construct the state variable d_t. We then proceed to estimate the value functions nonparametrically on a discretized state space. The transition rule for d_t, the time to build function (estimated also "offline") and the estimated entry and exit policy functions are combined to create the transition matrix. Each estimation task is described below and followed by the estimation results.

5.1.1 Demand for Shipping Services

We estimate the inverse demand for shipping services via instrumental variables regression. The empirical analogue of (1) is:

$$\log(P_t) = \eta_0 + \eta_1 \log(H_t) + \eta \log(Q_t) + \varepsilon_t \tag{7}$$

where P_t is the average[8] price per voyage in a quarter, H_t includes demand shifters, while Q_t is the total number of voyage contracts realized in a quarter. H_t includes the index of world industrial production (taken from the OECD), China's steel production, the fleet of Handymax ships (bigger size category of ships), the index of food prices, agricultural raw material prices and minerals prices (taken from UNCTAD). We correct for endogeneity and measurement error in a first stage regression of Q_t on the demand shifters and instruments, which include the following supply shifters: the total fleet, the mean and the standard deviation of the fleet age distribution.

Table 5 in Appendix D reports the results. The index of world industrial production (WIP) and China's steel production positively affects prices, while the number of voyages - corrected for endogeneity and measurement error by instruments- has a negative sign. It is not clear what the appropriate sign of commodity prices is, as these capture shifts in both the demand and the supply of commodities and may affect shipping prices either way. The same is true for the Handymax fleet, which may act as a substitute (as suggested by the negative sign), but it may also capture higher overall demand for shipping services. The impact of all shifters H_t, though, will be lumped into the state variable d_t as described below.

The demand state variable d_t is taken to be the intercept of the inverse demand curve, i.e.

$$d_t = \widehat{\eta_0} + \widehat{\eta_1} \log(H_t) + \widehat{\varepsilon_t}$$

[8] The variance of the price per voyage in a quarter is extremely small.

and is plotted in Figure 7. The residual of the IV regression (7) $\widehat{\varepsilon_t}$ is included in d_t, as it captures omitted demand shifters.

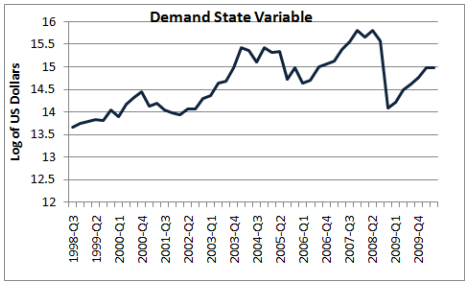

Figure 7: Demand state variable.

5.1.2 Demand State Transition

We specify the evolution of the demand state variable d_t as a first order autoregressive process with non-normal disturbances:

$$d_t = c + \rho d_{t-1} + \varepsilon_t \tag{8}$$

The error ε_t follows a continuous mixture of normals distribution: if its variance is known, ε_t is normally distributed, $\varepsilon_t \sim N(0, \sigma^2)$. The precision parameter (inverse variance, $\frac{1}{\sigma^2}$), however, is not known but is a positive random variable that is assumed to follow the Gamma distribution. The probability density of the resulting model is given in closed form (see Hogg, Craig and McKean (2004)) by:

$$f_\varepsilon(\varepsilon) = \frac{\Gamma\left(\alpha + \frac{1}{2}\right)}{\gamma^\alpha \sqrt{2\pi}\Gamma(\alpha)} \left(\frac{2\gamma}{2 + \gamma\varepsilon^2}\right)^{\alpha + \frac{1}{2}} \tag{9}$$

where $\Gamma(\cdot)$ is the Gamma function and $\alpha > 0$, $\gamma > 0$ are the parameters of the Gamma density. If (α, γ) are restricted to satisfy $4\alpha\gamma = 1$, then f_ε becomes a Student t density with 2α degrees of freedom. The continuous mixture distribution is a generalization of the normal distribution, featuring fat tail. We discuss the validity of this assumption after the estimation results.

We estimate the parameters associated with (8) and the continuous mixture, $(c, \rho, \alpha, \gamma)$ via Maximum Likelihood. Results are shown in Table 2. The demand process exhibits high persistence, with $\rho = 0.954$ at a quarterly level. It is stationary though, as $\rho < 1$.

	c	ρ	a	γ
Param	0.7483	0.9542	1.575	25.356
s.e.	(0.698)	(0.048)	(0.044)	(7.784)

Table 2: Demand transition parameters. Standard errors based on 500 bootstrap samples.

The presence of non-normal innovations significantly increases the likelihood from -8.16 for the case of normal innovations to 35.74, supporting the presence of fat tails in the error distribution. We have experimented with various distributional assumptions for the demand innovations and have found the estimation results to be robust. For instance, two alternative assumptions are the Student t distribution with v degrees of freedom (as in Hamilton (2005)) and the mixture of two normals with zero mean, variances σ_1 and σ_2, and mixture probability p. The likelihood under the Student t error term is 3.071, while the likelihood in the case of the finite mixture is 4.333; both lower than the likelihood under the chosen specification. Moreover, the normal distribution is strongly rejected with a Likelihood Ratio test (the constraint imposed sets the variance of the innovations' random variance equal to zero, i.e. $\alpha\gamma^2 = 10^{-5}$). It is also strongly rejected against the finite mixture model.

The adoption of a distribution with fat tails for the innovations is important because the process allows for larger shocks compared to the normal. We can compute the probability of observing a shock as large as the crisis, i.e. $\Pr\left[\min\left(\widehat{\varepsilon}_t\right)\right] = \Pr\left[\min\left(d_t - \widehat{c} - \widehat{\rho} d_{t-1}\right)\right]$ under each estimated specification. Normally distributed innovations assign zero probability to such shocks, even though these are observed during our sample- in contrast, this probability is significantly raised under the continuous mixture of normals.

Finally, the reported standard errors are based on 500 bootstrap samples (the resampling is done on the error term), to account for the fact that d_t is constructed from the inverse demand curve.

5.1.3 Value Function

Under the assumption of a large number of identical potential shipowners, the second hand sale price of a ship must equal its value. Indeed, since shipowners share the value for the ship a seller is willing to sell at price p^{SH} only if $p^{SH} \geq V_j\left(s, b, d\right)$, while a buyer is willing to buy only if $p^{SH} \leq V_j\left(s, b, d\right)$. As mentioned in the previous section, sales in this model occur due to random shocks, independent to the value of the ship, that hit shipowners. We therefore use the resale dataset to estimate the value function of a ship

as a function of its age j and the remaining state variables (s, b, d). The resale price is thus given by:

$$p^{SH} = V_j(s, b, d) + \varepsilon$$

where ε captures measurement error. Note that this model does not allow for unobservable state variables. Such a state would be an argument of the value function and not in ε. The value function at a state (j, s, b, d) is given by the conditional expectation of the second hand price, p^{SH} on (j, s, b, d):

$$V_j(s, b, d) = E\left[p^{SH}|j, s, b, d\right]$$

We estimate the value function, $V_j(s, b, d)$ nonparametrically on a grid of states, via local linear regression. We first describe how the grid of states is obtained and then proceed to the estimation on this grid, along with the estimation results. Finally, we provide a discussion of the assumptions underlying our estimation procedure.[9]

The state variable consists of the age of the ship, $j \in \{0, 1, ..., A\}$, the age distribution of the fleet s_t of dimension $A+1$, the backlog b_t of size $\overline{T}$ and the scalar demand d_t. The overall state has dimension $A + \overline{T} + 2$. Define:

$$x_t \equiv \left[s_t^0, ..., s_t^A, b_t^1, ..., b_t^{\overline{T}}, d_t\right] \in \mathbb{R}^{A+\overline{T}+2}$$

In our case of quarterly data, $A = 80$ and $\overline{T} = 11$ (the maximum time to build observed), so that the state $[j, x_t]$ has 94 components. Performing local linear regression on 94 coefficients is impossible. In addition to this dimensionality problem, each state $[j, x_t]$ can take on a large number of values, as each of its 94 components can itself take a large number of values: $j \in \{0, 1, ..., A\}$; s_t^i and b_t^i are positive integers, while $d_t \in \mathbb{R}$. The high state dimension and the large number of integer values the state components can take, pose significant computational challenges.

To develop a manageable finite state space we employ aggregation and clustering techniques. Aggregation addresses the first problem of dimensionality, serving to drive down the dimension of the state space. We reduce the 93 dimensions of x_t to 5 by considering three age groups, $S_t = (S_t^1, S_t^2, S_t^3)$, where $S_t^1 = \sum_{i=0}^{A_1} s_t^i$, $S_t^2 = \sum_{i=A_1+1}^{A_2} s_t^i$ and $S_t^3 = \sum_{i=A_2+1}^{A_3} s_t^i$, the total backlog $B_t = \sum_{i=1}^{\overline{T}} b_t^i$ and demand d_t. In particular, the first age group includes the number of ships between 0 and 10 years old, so that $S_t^1 = \sum_{i=0}^{39} s_t^i$, the second group includes the number of ships between 10 and 20 years old

[9] Adland and Koekebakker (2007) use second hand ship sale prices to uncover nonparametrically how a ship's value in the second hand market is affected by its age, its size and freight rates.

so that $S_t^2 = \sum_{i=40}^{79} s_t^i$ and the third group is the number of ships older than 20 years old $S_t^3 = s_t^{80}$. Aggregation linearly compresses x_t into the much smaller vector:

$$X_t = \left[S_t^1, S_t^2, S_t^3, B_t, d_t\right] \in \mathbb{R}^5$$

Clustering techniques are employed to address the second problem of dimensionality and reduce the number of admissible values for these aggregated states X_t, by constructing representative states. The state $[j, X_t]$ can take on a large number of values, as each of its components still takes on a large number of values; S_t^i and B_t are positive integers whose upper bound is considerably larger that the upper bound of their respective summands s_t^i and b_t^i. To reduce the number of possible values that the aggregated state $[S_t, B_t, d_t]$ can take we use vector quantization and the $k-$means algorithm, a method that partitions a large set of vectors into a finite set of groups. Each group is represented by its centroid. The vector quantization algorithm calculates iteratively the group representatives (the codebook values) until the mean square error between the codebook and the provided data is minimized. $[S_t, B_t]$ is treated independently from d_t: vector quantization is applied on $[S_t, B_t]$ which leads to the groups $[S, B]$ and scalar quantization based on the Lloyd algorithm ($k-$means for the case of a scalar variable) is applied on d_t and leads to groups d. We then take all possible combinations of the $[S, B]$ groups and the d groups. We end up with L representative states $X = [S, B, d] \in \mathbb{R}^{L\times 5}$ where L is a few hundred (and combined with the 81 possible ages this leads to about $16,000$ states). The vector quantizer guarantees that the observed data is represented well by the groups, while including unobserved combinations of d with the remaining aggregate state enriches the state space.

We estimate the value function by local linear regression, i.e. we perform linear regression at every grid point $[j, X] = [j, S, B, d]$, including a weighting scheme that down-weighes the contributions of data points away from $[j, X]$. We therefore solve the following local weighted least squares problem, for all grid points $[j, X]$:[10,11]

$$\min_{\beta_0,\beta_j,\beta_s,\beta_B,\beta_d} \sum_i \left\{ \begin{array}{c} p_i^{SH} - \beta_0(j,X) - \beta_j(j,X)(j_i - j) - \beta_s(j,X)(S_i - S) \\ -\beta_B(j,X)(B_i - B) - \beta_d(j,X)(d_i - d) \end{array} \right\}^2 K_h([j_i, X_i] - [j, X])$$

[10] For more information on local linear regression, see Fan and Gijbels, (1996) or Pagan and Ullah (1999).

[11] An alternative is to choose a parametric form for the value function (though it is not clear what is the parametric form of profits corresponding to this value function; high order polynomials in five explanatory variables is demanding). Discretization and parameterization are two different ways of smoothing. The external validation done on entry costs (Section 5.2.3) validates the estimation grid.

where $\left\{p_i^{SH}, j_i, S_i, B_i, d_i\right\}_{i=1}^{1838}$ represent the data of second hand sale transactions (so that $S_i = [S_i^1, S_i^2, S_i^3] = \left[\sum_{l=0}^{39} s_i^l, \sum_{l=40}^{79} s_i^l, s_i^{80}\right]$ and $B_i = \sum_{l=1}^{\overline{T}} b_i^l$, while K_h is a multivariate normal kernel with bandwidth $h \in \mathbb{R}^6$, so that $K_h\left(Y_i - Y\right) = \exp\left(\left(Y_i - Y\right)' H^{-1}\left(Y_i - Y\right)\right) / \sqrt{2\pi} \prod_j h_j$, where $Y_i = [j_i, S_i, B_i, d_i]$ and H is a diagonal with diagonal entries the bandwidths h_j. The unit of observation i corresponds to a quarter in our sample and identifies the common to all ships state $[s_t, b_t, d_t]$. It also designates a specific ship sold identified by its age, j. The bandwidth is different for each state variable based on its observed range. We have experimented with $k-$nearest neighbor bandwidths, cross-validation, as well as non-normal kernels and our results are robust. Under suitable regularity assumptions, the LLR estimator is asymptotically normal, weakly consistent and biased. The biasedness can go away asymptotically under stronger assumptions.

Figure 8 shows the estimated value functions. Each quarter is represented by its corresponding (group) state $X = (S, B, d)$. In other words, we present estimation results on a subset of the state space that is closest to the sample states. Different curves represent different ages and we find that the order is preserved so that older ships have a lower value function for all $X = (S, B, d)$. A subset of all ages (corresponding to 1 year old, 2 years old and up to 20+ years old) is depicted in the figure, for expositional simplicity.

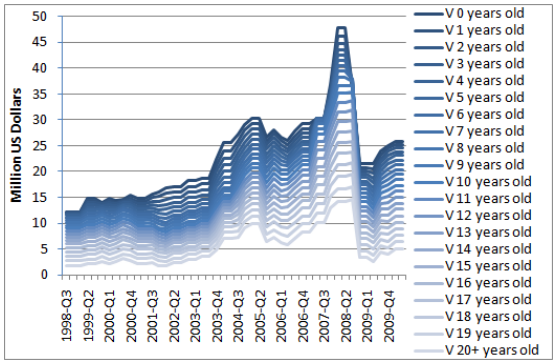

Figure 8: Plots of value functions for different ages over observed state path.

Pointwise confidence intervals are computed via bootstrap samples; we use the standard bootstrap method for local linear regression described in Fan and Gijbels (1996). In particular we resample prices and states without replacement 500 times. Figure 9 shows the value function of a 0 year old ship and that of a 20 year old ship with the computed 95% confidence intervals. As more observations are available for old ships, confidence intervals are tighter.

Estimating value functions using resale prices is central in our empirical strategy. One

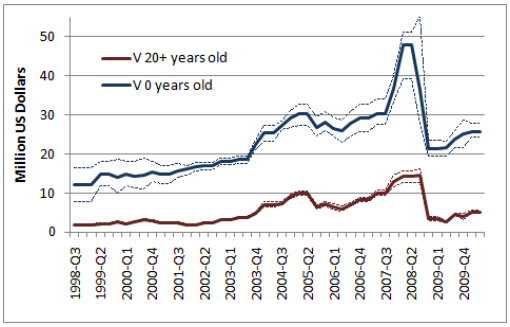

Figure 9: Value Functions and Confidence Intervals.

may worry about both adverse selection, as well as sample selection. Adverse selection would arise in the model if different shipowning firms privately observe the quality of the ship. Practitioners assure that this is not a concern: once a ship enters the resale market its history of maintenance and accidents is public information. In addition, potential buyers thoroughly inspect the ship; quoting Stopford (2009): "Full details of the ship are drawn up, including the specification of the hull, machinery, equipment, class, survey status and general equipment. ... (Inspections) will generally include a physical inspection of the ship, possibly with a dry docking or an underwater inspection by divers ... The buyer will also inspect the classification society records for information about the mechanical and structural history of the ship". Selection on unobservables seems a more worrisome concern. If only low quality ships are traded in the market, even if this is common knowledge, it will lead to a bias in the estimated value functions. The second-hand market for ships is thick: about 10% of the fleet is traded annually. In addition, price variation conditional on a state (j, S, B, d) is extremely small. Unfortunately, we do not observe the distribution of observable characteristics for ships not in our second-hand sample. But our concerns for econometric selection are put to rest when we turn to the external validation of our model and estimation, presented at the end of this section. In particular, Figure 14 compares the estimated entry costs to the average shipbuilding price. The graph is surprisingly good suggesting that selection can not be a severe issue. Finally, let us also discuss the possibility of search costs: based on Stopford (2009), search costs are not a severe concern either; "The shipowner ... offers the vessel through several broking companies. On receipt of the instruction the broker will telephone any client ... (and) call up other brokers in order to market the ship".

In this paper, second-hand prices are the tool we use to estimate our dynamic model of ship entry and exit. The bulk shipping industry provides an ideal setup for our empirical

methodology. A large literature examines resales and their frictions in other industries, such as Gavazza (2010) who examines the case of commercial aircraft.

5.1.4 Time to Build Function

We use the ship orderbook data to construct the average time to build corresponding to each quarter in our sample. Consistent with our state space reduction, we assume that time to build is a function of the group state $X_t = (S_t, B_t, d_t)$. We estimate the following time to build function and Figure 10 presents the actual and predicted time to build.

$$T(X_t) = T(S_t, B_t, d_t) : \log(T_t) = a_0^{TTB} + a_S^{TTB} S_t + a_B^{TTB} B_t + a_d^{TTB} d_t$$

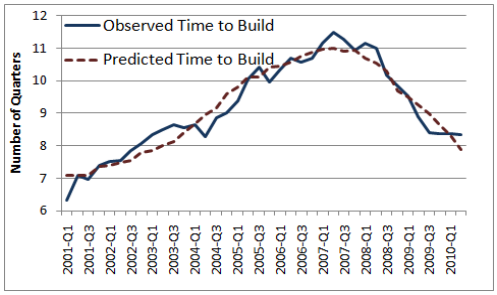

Figure 10: Average observed and estimated time to build.

5.1.5 Transition Matrix

As pointed out above, the state variable consists of the fleet age distribution s_t of dimension $A+1$, the backlog b_t of size $\overline{T}$ and the demand d_t (as well as the ship's age which evolves deterministically and we thus omit from this discussion). The overall state has dimension $A + \overline{T} + 2$ and evolves according to the law of motion described in Section 3. As specified in Section 5.1.2, the evolution of the demand state variable d_t is given by a first order autoregressive process with non-normal disturbances.

Consider the state

$$x_t = \left[s_t^0, ..., s_t^A, b_t^1, ..., b_t^{\overline{T}}, d_t\right]$$

Then,

$$\begin{aligned} x_{t+1} &= \left[s_{t+1}^0, ..., s_{t+1}^A, b_{t+1}^1, ..., b_{t+1}^{\overline{T}}, d_{t+1}\right] \\ &= \left[b_t^1, s_t^0, ..., s_t^{A-2}, s_t^A + s_t^{A-1} - Z_t, b_t^2, .., b_t^{T(X_t)} + N_t, .., b_t^{\overline{T}-1}, 0, c + \rho d_t + \varepsilon_{t+1}\right] \end{aligned}$$

where N_t indicates the number of entrants (ship orders) at time t when the state is x_t and Z_t represents the number of ships destined for scrap when the state is x_t. In words, the first element of the age distribution s_t equals the first element of the backlog (scheduled deliveries for this period), while all remaining elements of s_t age one period (by shifting one position). All elements of the backlog move one period closer to delivery. The time to build function $T(X_t)$ can be used to determine the position of entrants (for state x_t we compute the corresponding group state $X_t = (S_t, B_t, d_t)$ and calculate $T(X_t)$). Finally, exit is subtracted from s^A and entry is added in the last position of the backlog.

The driving stochastic processes are, ε_{t+1}, N_t and Z_t. More precisely, ε_{t+1} follows a continuous mixture of normals and summarizes the uncertainties inherent in demand for shipping services. Entrants N_t and exitors Z_t are distributed Poisson with rates $\lambda(x_t)$ and $\mu(x_t)$ respectively. They form the source of nonlinearity in the above state transitions. Under the assumption that N_t, Z_t and ε_{t+1} are conditionally independent given the state x_t, the transition probability from x_t to x_{t+1} is given by:

$$\Pr(x_{t+1}|x_t) = Q(x_t, x_{t+1}) f_N(N_t = n|x_t) f_Z(Z_t = z|x_t) f_\varepsilon(\varepsilon_{t+1}) \tag{10}$$

where

$$Q(x_t, x_{t+1}) = \delta\left\{s_{t+1}^0 = b_t^1\right\} \prod_{i=1}^{A-1} \delta\left\{s_{t+1}^i = s_t^{i-1}\right\} \prod_{i \neq T(X_t)} \delta\left\{b_{t+1}^i = b_t^{i+1}\right\} \tag{11}$$

and

$$\begin{aligned} n &= b_t^{T(X_t)} - b_t^{T(X_t)+1} \\ z &= s_t^A + s_t^{A-1} - s_{t+1}^A \\ e &= d_{t+1} - c - \rho d_t, \end{aligned}$$

while $\delta\{\cdot\}$ is an indicator function, f_N and f_Z are the Poisson densities with parameter $\lambda(x_t)$ and $\mu(x_t)$ respectively and f_ε is given in (9).

The state transitions described in (10) and (11) pose significant computational chal-

lenges that arise from the high state dimension and the large number of integer values the state components may take. To develop manageable finite state space approximations we employ the grouped and clustered states, $X = [S, B, d] \in \mathbb{R}^{L\times 5}$. Our goal is to create a transition matrix $P \in \mathbb{R}^{L\times L}$ whose (i, j) element P_{ij} specifies the probability of transiting from state X_i to state X_j. Even though the transition from state x_t to x_{t+1} is known (and given in (10) and (11)), the transition from X_i to X_j is more complicated. We construct the transition matrix in two steps that combine the state dynamics of x_t in (10) and (11) with simulation methods. First, we estimate the Poisson processes of entry and exit (results presented below) and the demand transition (8) (results presented in Section 5.1.2). Second, we simulate thousands of states $x_t \in \mathbb{R}^{A+T+2}$ using the estimated Poisson processes for N_t and Z_t and the $AR(1)$ process for d_t. We replace each simulated state by its corresponding group state $X = [S, B, d] \in \mathbb{R}^5$ (based on minimum distance from the L group states). We then compute the frequency transition matrix based on the simulated paths of group states (i.e. count the number of times X_i transits to X_j and divide by the total number of visits to X_i). Finally, we smooth the frequency transition matrix as follows: each element (i, j) of the transition matrix is the inner product of the $j-th$ column and a row of weights reflecting the closeness of the $i-th$ state to the rest of the states. The weights are based on a multivariate normal kernel.[12]

We next turn to the results.

We assume entry is $\lambda(S_t, B_t, d_t) = \exp\left(\gamma_0 + \gamma_{S_1} S_t^1 + \gamma_{S_2} S_t^2 + \gamma_{S3} S_t^3 + \gamma_B B_t + \gamma_{B^2} (B_t)^2 + \gamma_d d_t\right)$ and exit is $\mu(S_t, B_t, d_t) = \exp\left(\delta_0 + \delta_{S_1} S_t^1 + \delta_{S_2} S_t^2 + \delta_{S3} S_t^3 + \delta_B B_t + \delta_{B^2} (B_t)^2 + \delta_d d_t\right)$.[13] We estimate the parameters γ and δ via OLS using the time series of entrants[14] and exitors. This step is similar to the first estimation step in Bajari, Benkard and Levin (2007) where policy functions are estimated offline. Table 7 in Appendix D reports the results. As expected, entry is increasing in demand: as the demand curve for freight shifts outward, more entry is attracted to the market. Similarly, exit is decreasing in demand: as the demand curve for freight shifts outward, exit becomes a less attractive option. Entry is decreasing in the number of competitors, (S^1, S^2, S^3), with the num-

[12] One alternative to our method is the following: choose a number of representative states x within each group state X (for instance by simulating states and then clustering). To compute the probability P_{ij}, sum the probability of moving from each state $x \in X_i$ to each state $x \in X_j$ weighted by the probability of state x within X_i. It is also possible to derive closed form expressions for the probabilities P_{ij} in that case; these expressions however are extremely complicated and computationally intractable.

[13] Since λ and μ are greater or equal than zero, they can be expressed as $\exp(f(S_t, B_t, d_t))$. We can therefore think of the chosen specifications as Taylor approximations to $f(\cdot)$ (first order in S_t^1, S_t^2, S_t^3 and d_t and second order in B_t).

[14] Instead of OLS we could estimate via local linear regression and recover $\lambda = E(N_t|S_t, B_t, d_t)$. If we then perform OLS on the nonparametrically estimated $E(N_t|S_t, B_t, d_t)$, we get the same results.

ber of young competitors S_1 affecting entry more strongly. Exit is increasing in the number of competitors (S^1, S^2, S^3). Note, however, that entry is increasing in the backlog for low backlog levels and then decreasing, while exit is first decreasing and then increasing in the backlog. This problem is due to the "offline" estimation of policy functions found in most two-step estimation procedures for dynamic games, e.g. Bajari, Benkard and Levin (2007), which may suffer from endogeneity of competition variables. To avoid this issue, we choose more parsimonious specifications. We have experimented with various functional forms, as well as several estimation methods (OLS, ML, MME) and concluded to the following forms: $\lambda(S_t, B_t, d_t) = \exp\left(\gamma_0 + \gamma_{S_1} S_t^1 + \gamma_{S_2} S_t^2 + \gamma_d d_t\right)$ and $\mu(S_t, B_t, d_t) = \exp\left(\delta_0 + \delta_{S_1} S_t^1 + \gamma_{S_2} S_t^2 + \gamma_{S_3} S_t^3 + \delta_d d_t\right)$. Table 8 in Appendix D reports the results.

5.2 Second Step: Estimate the Model Primitives

In the second step of our estimation strategy we estimate the model primitives. We use the continuation value of old ships to estimate nonparametrically the scrap value distribution. We estimate the profits directly from the Bellman equations. Entry costs are recovered from the free entry condition.

5.2.1 Scrap Value Distribution

Once we obtain the transition matrix, $P \in \mathbb{R}^{L \times L}$ and the value function $V \in \mathbb{R}^{L \times A+1}$, the continuation value of an age $j < A$ ship is simply

$$VC_j = PV_{j+1} \in \mathbb{R}^L$$

while

$$VC_A = PV_A$$

The estimated continuation value function of old ships VC_A provides information directly on the scrap values. To obtain the density of scrap values, we use the estimated continuation value and the observed exit frequencies combined with an exit equation.

We estimate the scrap value distribution nonparametrically, via local linear regression of the observed exit probabilities Z_t / s_t^A on the continuation value of old ships, $VC_A(x_t)$. Indeed, the probability of exit at continuation value v equals $[1 - F_\phi(v)]$, or

$$\frac{Z}{s^A} = [1 - F_\phi(v)]$$

We run the following weighted regression at all points of interest v,

$$\min_{\beta_0,\beta_{VC}} \sum_t \left\{ \frac{Z_t}{s_t^A} - \beta_0(v) - \beta_{VC}(v)\left(VC_A(X) - v\right) \right\}^2 K_h\left(VC_A(X) - v\right)$$

where K_h is a normal kernel with bandwidth h. We get that $[1 - F_\phi(v)] = \beta_0(v)$. Note also, that the local linear regression estimator implies that $\beta_{VC}(v)$ is equal to the derivative of $\beta_0(v)$ at v. Therefore, $\beta_{VC}(v)$ is the density of the scrap value at v.

The distribution and density of scrap values are shown in Figure 11. Our data provide information directly on the value function of an old ship, which in turn is translated in the quartiles of the scrap value density. The horizontal axis in Figure 11 is essentially derived from the range of continuation values.

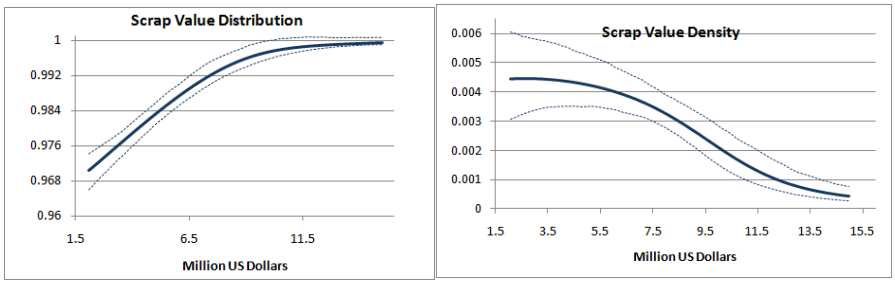

Figure 11: Distribution and density of scrap values. 0.95 bootstrap confidence intervals.

Confidence intervals for the scrap value distribution are constructed as follows. We use the 500 bootstrap samples of the entry and exit regressions, as well as the autoregressive process for the demand state variable to create 500 bootstrap transition matrices. As the errors are independent, we multiply the 500 bootstrap transition matrices with the 500 bootstrap value functions and create 500 bootstrap continuation values for old (20 years old and over) ships. For each bootrstrapped continuation value, we calculate the scrap value distribution and density by following the bootstrap method for local linear regression (i.e. we create a bootstrap sample for Z_t/s_t^A and the corresponding continuation value and perform local linear regression as above).

In the first step of the estimation we used the Poisson exit rate, $\mu(S, B, d)$, towards the creation of the transition matrix. At the same time, the estimated scrap value distribution implies the following exit rate:

$$\widetilde{\mu}(S, B, d) = s^A\left(1 - F_\phi\left(VC_A\right)\right)$$

The presence of two measures for exit is inherent in two-step estimation procedures for dynamic games.[15] Comparing the two obtained measures $\mu(S_t, B_t, d_t)$ and $\widetilde{\mu}(S_t, B_t, d_t)$ we find that they are very close.

5.2.2 Profits

We compute profits $\pi \in \mathbb{R}^{L \times A+1}$ from the Bellman equations (3), which we repeat here for convenience:

$$V_j = \pi_j + \beta E_\phi \max\{\phi, VC_j\}$$

Profits then can be estimated from:

$$\pi_j = V_j - \beta VC_j \tag{12}$$

for $j = 0, 1, ..., A-1$, -i.e. for these ages where there is no exit activity- and $VC_j = PV_{j+1}$, while

$$\pi_A = V_A - \beta E_\phi \max\{\phi, VC_A\}$$

or

$$\pi_A = V_A - \beta\left(1 - p^X\right) VC_A - \beta p^X E\left[\phi | \phi > VC_A\right] \tag{13}$$

where $p^X \in \mathbb{R}^L$ is the exit probability, so that $p^X = \widetilde{\mu}(S, B, d)/s^A$. All V_j as well as the transition matrix, P, were estimated in the first step of the estimation, while the exit rate $\widetilde{\mu}(S, B, d)$ was computed using the estimated scrap value distribution.

Figure 12 presents the estimated profits of ships older than 20 years old, as well as the average profits of ships 10-20 years old. Confidence intervals are constructed by calculating the profits corresponding to each of the 500 bootstrap samples of value functions and continuation values (for the 20+ age category, the bootstrapped scrap value density and exit rate are also used). As we have fewer observations on young ships in our second-hand ship sale dataset, we get noisier estimates of value functions and profits for young ships (in the counterfactuals we will assume firms engage in perfect competition and therefore this will not be a major concern.)

Note that profits fluctuate close to zero between 1998 and 2006, spike during the boom of 2007-2008 and collapse at the crisis. The shipping industry is characterized by the combination of an inelastic and volatile demand curve with an inelastic supply curve.

[15] For instance, Pakes, Ostrovsky and Berry (2007) employ a first step frequency estimator for the exit rate which is used to estimate value functions. In the second step, they estimate a parametric scrap value density which implies a new exit rate (as well as a parametric entry cost distribution which implies a new entry rate).

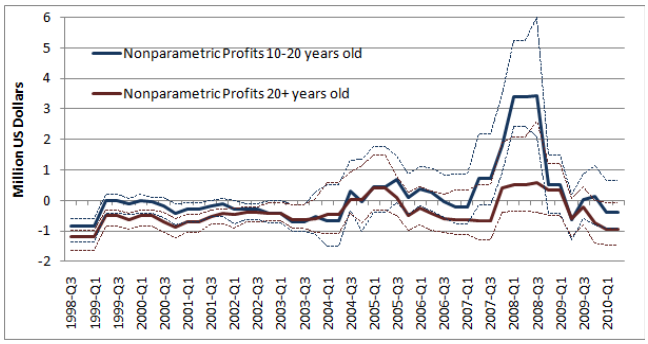

Figure 12: Estimated profits. Average across age 10-20 years and 20+ years old. 0.95 bootstrap confidence intervals.

As a result, shipowners are willing to suffer zero or even negative profits for multiple time periods, waiting for the periods where demand jumps high and the supply response is limited. Figure 12 reports the profits on a subset of the state space. It is worth underlining that we have not imposed the free entry condition to get the profits. As pointed out in Stopford (2009), "Each company faces the challenge of navigating its way through the succession of booms and depressions that characterize the shipping market", while giving a specific example, "For several years the company had accepted this drain on its cashflow, in the hope that the market would improve".

In order to assess the validity of our model and estimation results, we use an earnings index constructed by shipping brokers. This index is not used in the estimation. The scale of this index is uninformative (not only is it an index, but also earnings calculated by brokers do not include fixed costs, opportunity costs, etc.) and we only use it to assess the evolution of profits over time. Our estimated profits track broker calculated profits well, as shown in Figure 13.

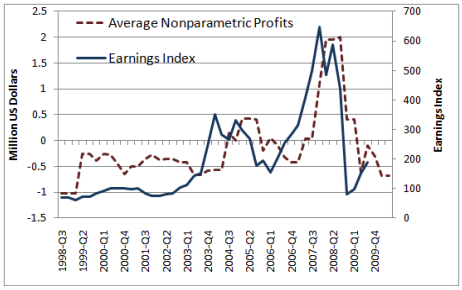

Figure 13: Weighted average (across ages) nonparametric profits and earnings index. Index based on 45000DWT Handymax, based on freight rates and fuel/port costs for representative routes.

5.2.3 Entry Costs

The free entry condition states that at each aggregate state $X = (S, B, d)$, the value of entry equals the entry cost, $\kappa(X)$. The value of entry is the expectation of the value function of age 0, $T(X)$ periods ahead. To construct the value of entry at X we first compute the $T(X)^{th}$ power of the transition matrix. We then multiply the row of that matrix corresponding to state X with $V_0 \in \mathbb{R}^L$ as well as $\beta^{T(X)}$, i.e.,

$$\beta^{T(X)} P_X^{T(X)} V_0 = \kappa(X) \tag{14}$$

where $\kappa \in \mathbb{R}^L$ are the entry costs, $V_0 \in \mathbb{R}^L$ is the value of an age 0 ship and $P_X^{T(X)}$ refers to the row of $P^{T(X)}$ that corresponds to state X. Having computed the transition matrix P, the time to build function $T(X)$ and the value function V_0 we can directly estimate entry costs κ from the free entry condition.[16] Confidence intervals for the entry costs are constructed by computing the entry costs of each bootstrap sample of a 0 year old ship and a bootstrap transition matrix.

In practice, the largest portion of entry costs is the price charged by the shipyard. Additional costs include legal fees, negotiation costs with the shipyard and engineering supervision during construction. We use the average shipbuilding price to evaluate the power of our model and estimation results. The shipbuilding price is not used in the estimation and only acts as an "additional degree of freedom". We use the estimated value function of a young ship, as well as the $T(X)^{th}$ power of our transition matrix to generate entry costs. Figure 14 compares entry costs to the shipbuilding price. Considering that we combine our noisiest estimated value function (due to the shortage of observations on young ships) with the free entry condition implied by our model and the estimated transition matrix, the fit is surprisingly good.

[16] The free entry condition holds with equality only when entry is strictly positive. If mean entry is zero, then the value of entry provides only a lower bound for the entry costs. There is no zero entry at any of our states.

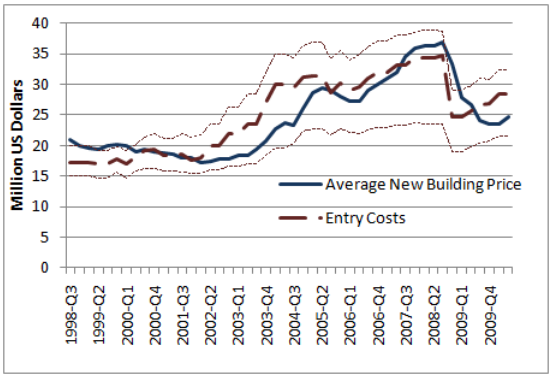

Figure 14: Estimated entry costs and average shipbuilding price. 0.95 bootstrap confidence intervals.

6 Counterfactuals

In this section, we explore the nature of fluctuations in the bulk shipping industry by quantifying the impact of demand uncertainty and time to build on freight rates, investment in new ships and surplus. First, we assume there is perfect competition in the market for freight transport and estimate a parametric form for profits. Next, we compute the equilibrium of our model in two counterfactual worlds of constant and no time to build. We explore the impact of demand impulses and perform long run simulations to study the evolution of the industry.

6.1 Perfectly Competitive Profits

In order to analyze the behavior of shipping prices and surplus, we need to take a stand on the nature of per period competition in the shipping industry. Even though counterfactuals concerning the investment in new ships can be performed employing the nonparametric estimates of the model, counterfactuals concerning prices and surplus require structure on the profit function. We therefore assume there is perfect competition in the market for freight transport and use the estimated nonparametric profits to estimate the operating cost parameters. The perfectly competitive behavior ties well with the assumptions of a large number of homogeneous firms. We also experimented with the assumption that firms engage in competition in quantities (Cournot game) but found that the quantities chosen lead to marginal cost pricing.

We assume that firms choose quantities, taking prices as given and subject to convex operating costs that increase in the ship's age. In particular, every period, a ship of age

j chooses the number of voyages q in order to maximize profits:

$$\max_{q} \left\{ Pq - c_j q^3 - F_j \right\} \tag{15}$$

where $P = e^d Q^\eta$ is the inverse demand function (as estimated in step 1) and $(c_j q^3 + F_j)$ are total costs that include fixed costs F_j. Based on Stopford (2009), fuel consumption is proportional to the cube of the vessel's speed. Furthermore, the number of trips is proportional to speed and cost is proportional to fuel cost. Costs change when the ship's age crosses a new decade, so that

$$\begin{aligned} c_j &\in \{c_{0-39}, c_{40-79}, c_{80+}\} \\ F_j &\in \{F_{0-39}, F_{40-79}, F_{80+}\} \end{aligned}$$

Under these assumptions, price can be expressed as a function of the state $[S^1, S^2, S^3, d]$:

$$P\left(S^1, S^2, S^3, d\right) = \left(e^d\right)^{\frac{2}{2-\eta}} \left[\frac{S^1}{3\sqrt{c_{0-39}}} + \frac{S^2}{3\sqrt{c_{40-79}}} + \frac{S^3}{3\sqrt{c_{80+}}} \right]$$

Replacing the above in (15) we obtain the profits $\{\pi_{0-39}, \pi_{40-79}, \pi_{80+}\}$ as a function of the state $[S^1, S^2, S^3, d]$ and the parameters (c_j, F_j).

We use the estimated profits from equations (12) and (13), as well as observed prices and recover the cost parameters via Nonlinear Least Squares. In order to aid the extrapolation to younger ships which are not used in the estimation, we impose a linear dependence of the cost parameters on age, so that $c_{0-39} = \gamma_0 + \gamma_1$, $c_{40-79} = \gamma_0 + 10\gamma_1$, $c_{80+} = \gamma_0 + 20\gamma_1$ and similarly $F_j = \delta_0 + \delta_1$, $F_{40-79} = \delta_0 + 10\delta_1$, $F_{80+} = \delta_0 + 20\delta_1$. We recover $\theta = \{\gamma_0, \gamma_1, \delta_0, \delta_1\}$:

$$\min_{\theta} \left\{ ||\widehat{\pi_{0-39}} - \pi_{40-79}(\theta)||^2 + ||\widehat{\pi_{80+}} - \pi_{80+}(\theta)||^2 + ||P - P(\theta)||^2 \right\}$$

where $\widehat{\pi_{0-39}}$ and $\widehat{\pi_{80+}}$ are the nonparametrically estimated profits from Section 5.2.2, $\pi_{40-79}(\theta)$, $\pi_{80+}(\theta)$, $P(\theta)$ are the profits and price resulting from the perfectly competitive behavior and P the observed prices. The estimated parameters are given in Table 3. Standard errors are calculated using the bootstrapped profits.

We compare our estimates to cost figures found in Stopford (2009) and find that our model and estimation perform surprisingly well. Based on Stopford (2009), fixed costs F include operating costs such as manning, stores, lubricants, repairs, insurance, administration, as well as periodic maintenance. The fixed costs of a 10 year old Handymax

	c_{0-39}	c_{40-79}	c_{80+}	F_{0-39}	F_{40-79}	F_{80+}
param	$491,980$	$459,890$	$424,234$	$377,935$	$648,347$	$948,806$
s.e.	$(225,302)$	$(223,079)$	$(261,356)$	$(276,160)$	$(250,489)$	$(237,577)$

Table 3: Profit Parameters. NLLS results. Standard errors based on 500 bootstrap samples.

vessel, based on Stopford (2009) are $\mathbf{575,000}$\$ quarterly, which is close to our estimate of F_{40-79} at $\mathbf{648,347}$\$. Voyage costs include fuel and diesel costs, as well as port costs, and based on Stopford (2009) are about $\mathbf{775,000}$\$ quarterly for a 10 year old Handymax. Our estimate of $3c_{40-79}q^2$ at the mean level of q is $\mathbf{730,424}$\$. Finally, Figure 15 shows the fit of the perfectly competitive market in terms of the prices and profits, as well as the final perfectly competitive profits for the three age groups.

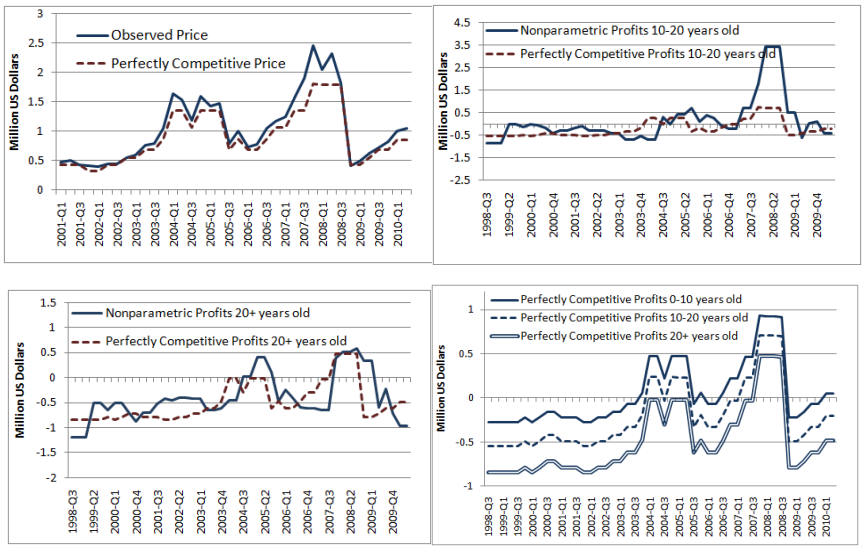

Figure 15: Results of perfectly competitive model.

6.2 Counterfactual Results

We now quantify the impact of demand uncertainty and time to build on freight rates, investment in new ships and surplus. We compute the equilibrium of the game under two counterfactual scenarios: constant time to build and no time to build. We compare these two worlds to the true world of endogenous time to build, i.e. $T_t = T(S_t, B_t, d_t)$. In the extreme case of no time to build, firms that decide to enter at period t, begin operating at

period $t+1$. The case of constant time to build captures a situation without shipbuilding capacity constraints, allowing for only a time invariant period of construction. Regardless of how many ship orders are placed, construction begins immediately and they are all delivered a fixed number of periods later. We choose this constant time to build to equal seven quarters, the minimum time to build we observe. To construct the counterfactual worlds, we compute the equilibrium of the model, i.e. we solve for the entry policy $\lambda \in \mathbb{R}^L$, the exit policy $\mu \in \mathbb{R}^L$ and value function $V \in \mathbb{R}^{L \times A+1}$, using the estimated primitives, i.e. the profits $\pi \in \mathbb{R}^{L \times A+1}$, the scrap value distribution F_ϕ and the entry costs $\kappa \in \mathbb{R}^L$.[17]

6.2.1 Impulse Response

We first explore the impact of a positive shock, as this is the case where time to build becomes a crucial constraint. Following a positive demand shock, supply is restricted because existing ships face capacity constraints and new ships require time to build. This leads to high profits and prices and low shipper surplus, as during the boom of this decade. If time to build were reduced or removed, supply would be more elastic and the size of fleet would increase in order to satisfy the increase in demand, reducing prices and profits. In the case of a negative demand shock, time to build is not a critical constraint, as the fleet can immediately adjust through ship exit. We simulate the model 500 times: after a few quarters, we hit the economy with the positive shock. We continue the simulations for another 20 quarters.

As shown in Figure 16 when the positive shock hits the industry, entry spikes. In a world with time to build, entry is both lower and slower. Indeed, new ships arrive a number of periods later, while in the meantime supply is limited to the existing capacity constrained fleet. In addition, the incentives to enter are dampened since the shock fades due to the mean-reverting demand process and ships that are delivered late will not be able to take advantage of the increased demand. As a result, the response of entry is significantly larger in the absence of time to build.

Figure 17 demonstrates the response of exit to the positive demand shock. Exit occurs in waves as we move from no to constant to endogenous time to build. In the absence of time to build, old ships exit immediately, as a large number of young, more efficient entrants floods the market. As a result, past a positive demand shock, in a world without time to build the turnover of ships is higher compared to a world with time to build, with young ships quickly replacing old ships. In contrast, time to build induces lower and

[17] See Appendix B for details.

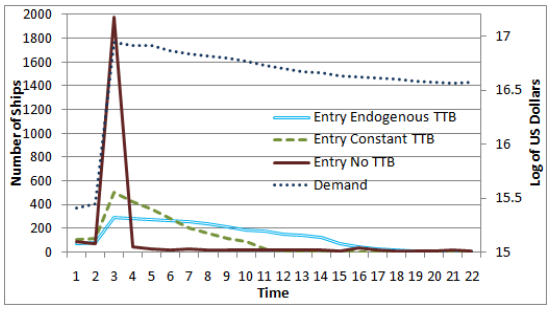

Figure 16: Entry under endogenous, constant and no time to build (TTB). Demand depicted on right axis.

slower exit activity. As it takes time for new ship orders to be delivered, existing old ships prefer to remain in the industry and reap the benefits of increased demand, rather than exit (Figure 5 documents that old ships indeed didn't exit during the boom). As a result, old ships exit only after the new ships are delivered from the shipyards. In addition, the exit spike is lower when there is time to build, as entry was lower.

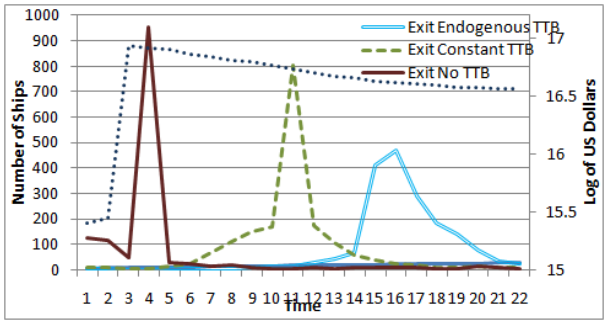

Figure 17: Exit under endogenous, constant and no time to build (TTB). Demand depicted on right axis.

As a result of the entry and exit behavior, freight rates are lower as time to build decreases, as shown in Figure 18. Prices increase when the shock hits, as in all three cases entrants need to wait at least one period before operating. As the shock is temporary and demand is mean reverting, entry, even in the case of no time to build, is not large enough to absorb the entire increase in demand and therefore price increases as well. As a result of the above, consumer surplus (calculated using the model in Section 6.1) is also higher as time to build decreases, as shown in Figure 19. Note that in this case, consumer surplus refers to the surplus of the customers of shipping companies, i.e. the firms that require transportation services (shippers), rather than the consumers of the final goods.

Figure 20 shows total and per ship profits. Total profits are lower as time to build

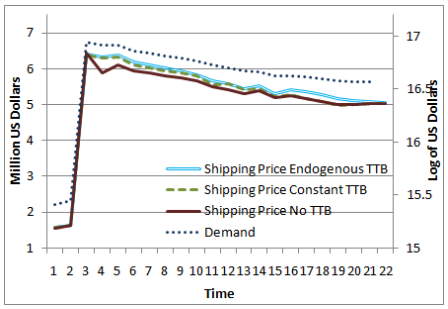

Figure 18: Price under endogenous, constant and no time to build (TTB). Demand depicted on right axis.

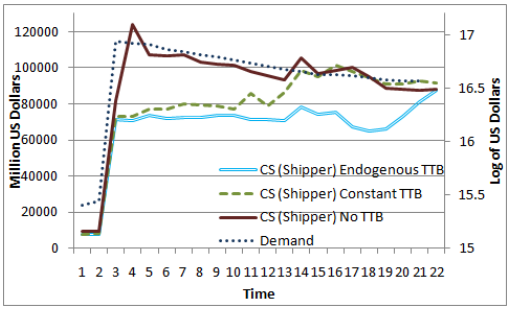

Figure 19: Consumer surplus (i.e. surplus of freight customers) under endogenous, constant and no time to build (TTB). Demand depicted on right axis.

increases. In contrast, profits per ship are higher as time to build increases. Moving from endogenous to constant to no time to build, the fleet increases and total profits increase. In the presence of time to build and a positive demand shock, however, operating firms reap high profits, as new firms can't enter immediately, while fewer firms enter as well. Therefore, moving from no to constant to endogenous time to build, profits per ship increase.

6.2.2 Long Run Simulations

In this section we explore the dynamic evolution of the dry bulk shipping industry by performing long run simulations of the model. Our goal is to understand and quantify the impact of time to build on the fluctuations of investment and prices. We perform 1000 simulations of our model, each 80 years (320 quarters) long. We use these long run simulations to explore the properties of our model both during specific as well as average demand paths.

Table 4 compares the baseline model of endogenous time to build to the two counter-

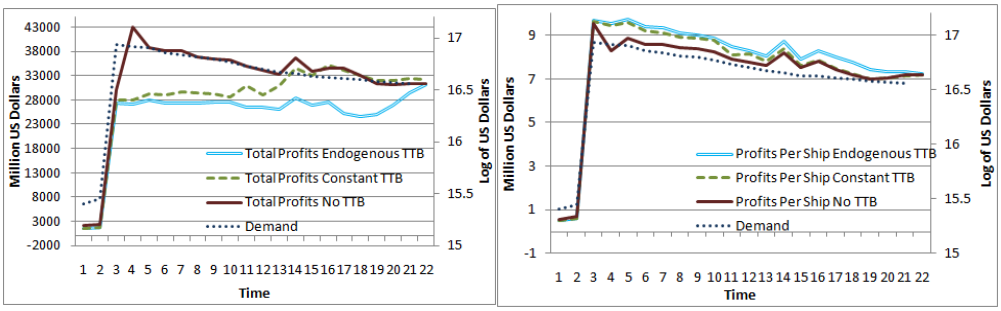

Figure 20: Total (sum of) and per ship profits under endogenous, constant and no time to build (TTB). Demand depicted on right axis.

factual worlds of constant and no time to build. In particular, it compares the ratio of constant to endogenous time to build (CT / ET), as well as the ratio of no to endogenous time to build (NT / ET) in terms of the level and variance of the total fleet and price, as well as their covariance with demand. We find that on average across time and simulations, the fleet is larger by about 20% in the absence of time to build and by about 5% in the case of pure construction lags. As a result, shipping prices are lower in the absence of time to build. The fleet is also younger on average as time to build declines.

	CT / ET	**NT / ET**
Fleet Level	1.055	1.215
Price Level	0.995	0.962
Fleet Variance	1.563	2.830
Price Variance	0.995	0.974
Fleet and Demand Covariance	1.116	2.963
Price and Demand Covariance	0.994	0.942

Table 4: Comparison of ratio of constant to endogenous time to build case (CT/ET) and no to endogenous time to build case (NT/ET). Average across 1000 simulations and 320 quarters.

In the absence of time to build the fleet is more responsive to demand and as a result, prices are less responsive to demand. Indeed, we find that investment volatility is significantly higher as time to build declines: the fleet is about 50% more volatile under constant time to build and almost three times more volatile under no time to build. In contrast, prices are less volatile as time to build declines. Consistent with this finding, the covariance of the fleet and the demand process is larger in the absence of time to build, while the covariance of prices and demand is lower.

As suggested by the analysis of the impulse response function above, in the absence

of time to build entry and exit respond *instantly* to positive demand shocks, while their response is also *larger*: entry is higher as firms can take advantage of the positive demand shock. Even though exit is also higher in the absence of time to build, due to the young firms coming in immediately, the entry effect dominates.[18]

Price volatility persists in a world without time to build mostly because of the presence of entry costs: under a mean reverting demand process, even in the presence of high demand shocks, entry is limited by irreversible costs of entry. Ships are long-lived capital which depreciates over time; even though in the absence of time to build entrants can immediately take advantage of high demand, they still need to compare the stream of profits during the ship's lifetime to the cost of entry. Therefore, if a shock is only temporary, the fleet may not adjust fully and price increases. In addition, as revealed by the simulations, more voyages are travelled, as we move from endogenous to constant to no time to build. The number of voyages per ship, however, is smaller: when there is time to build, the intensive margin takes over and the fewer operating ships offer more transportation. Therefore, in the presence of irreversible entry costs, depreciation, demand mean reversion (as well as some demand elasticity), time to build alone cannot account for the total of price volatility.

Figure 21 shows 60 years from one specific path of the industry, under a specific demand path (this is a single typical simulation of the model). This simulation corroborates the above analysis. The fleet in both the endogenous, as well as the no time to build worlds follows the particular demand path. Note, however, that the fleet response in the endogenous time to build case is much smoother than that in the no time to build case. In the absence of time to build, the investment response is both more immediate as well as larger.

We next examine the properties of investment on average across different demand paths. Figure 22 shows the time variation of the average fleet across the simulations. Note that as expected, demand has settled (after an initial adjustment period) to its long run average level and remains constant there (shocks from the different simulated paths are washed out). This figure is aimed at approximating the steady state for this industry. In the absence of time to build, the total fleet- after a transient period- approaches a constant value under fading small oscillations. In contrast, time to build leads to stronger

[18] In contrast, the industry response to negative demand shocks, depends on the backlog level when the shock hits. If the backlog is low enough, entry is higher in a world with time to build compared to one without, as demand will likely be higher upon delivery of the vessel. At the same time, exit may be lower under time to build due to the option value to waiting for the higher demand, when time to build refrains the new entrants.

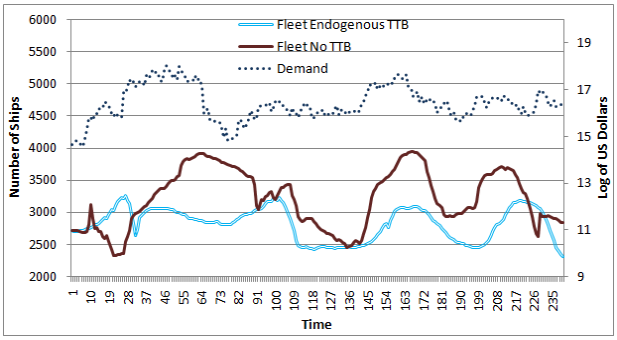

Figure 21: A sample industry path. Demand depicted on right axis.

oscillatory investment behavior. The oscillations seem to fade but extremely slowly and are still strong after 80 years of industry evolution. The fleet cycles in the constant time to build case lead those of the endogenous time to build case, as in this case there are no queues and construction is equal to or smaller than the baseline model.

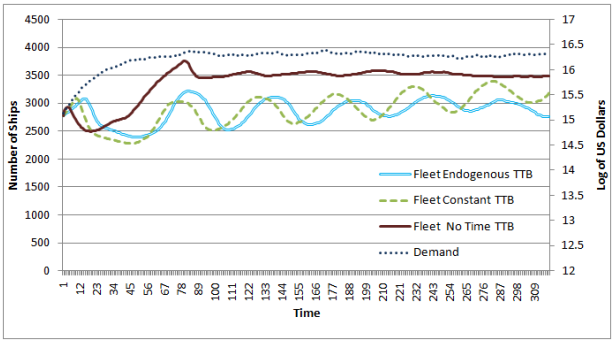

Figure 22: Average of model long run simulations. Demand depicted on right axis.

Appendix C further motivates and explains these results. We analyze the dynamics of the mean and covariance of the state vector. Using linear approximations, we explore the presence of oscillations and steady state behavior. The analysis and findings are reminiscent of Rosen, Murphy and Scheinkman (1994). We find that the matrix governing the dynamics of the mean state has eigenvalues whose norm is always smaller than 1. Several eigenvalues are complex, indicating oscillatory behavior. We thus obtain fleet oscillations that fade away, though slowly because the maximum magnitude eigenvalue is smaller but close to 1.

7 Conclusion

This paper studies the nature of price and investment fluctuations in oceanic bulk shipping. We find that investment volatility is significantly higher as time to build declines: the fleet is about 50% more volatile under constant time to build and almost three times more volatile under no time to build. We also find that the fleet is larger by about 20% in the absence of time to build and by about 5% in the case of pure construction lags. Shipping prices are lower and less volatile as time to build declines.

A dynamic model of entry and exit was constructed to explore the dynamics of the dry bulk shipping industry. The model was estimated using a rich dataset of second-hand ship sales, matched with a new estimation strategy. Our main estimating assumption is that the price of a ship in the second-hand market is equal to its value function. From there, we are able to uncover the primitives that govern the dynamic industry equilibrium. Our model and method are validated with the help of data not included in the estimation; we find that our estimated entry costs match well to the average shipbuilding price, our estimated profits to a broker earnings index and finally, our operating cost estimates to industry estimates. The proposed empirical strategy may be of use in other setups but is applied here on a unique industry characterized by severe demand variability and investment fluctuations.

(ONLINE) APPENDIX A

In this Appendix we describe several data details. The index of world industrial production used in Section 5.1.1, is the weighted average formed by the industrial production index of all OECD countries, as well as Brazil, India, Russia, and South Africa. The indices of individual countries are taken from SourceOECD and the base is $Q1 - 1999$. The weights are based on 2007 annual GDP levels, taken from Penn World Tables. China doesn't report a world industrial production index (only a year-on-year percentage change) and so we add its steel production.

The second-hand sales dataset reports the year of build for each ship sold, as well as its year of built. Since a quarter is the unit of time, age in years must be converted into age in quarters. We construct the distribution of deliveries across quarters of each year in our sample. Then, based on that distribution, we draw an age in quarters for each ship in the second-hand sale sample, from the 4 possible quarters within its year of built.

To construct the backlog vector, we take the following steps. We use the orderbook dataset to identify a quarter's entrants and compute the average time to build offered

to these entrants. We then use the timeseries of entrants and total backlog, as well as the estimated time to build function (using the observed total backlog) and create the backlog vector for each quarter of our sample, $b_t \in \mathbb{R}^{\overline{T}}$.

All prices are deflated using the CPI taken from the US Bureau of Labor Statistics. The basis used is $Q1-2009$ US dollars.

(ONLINE) APPENDIX B

In this Appendix we outline the computation of the equilibrium of our model. For computational reasons we replace the nonparametric scrap value distribution with a normal distribution. Its mean and standard deviation $\left(\mu_\phi, \sigma_\phi\right)$ are estimated via Maximum Likelihood:

$$\max_{\mu_\phi,\sigma_\phi} \sum_t \log\left[F_\phi\left(VC_A\left(x_t\right);\mu_\phi,\sigma_\phi\right)^{s_t^A - Z_t}\left(1-F_\phi\left(VC_A\left(x_t\right);\mu_\phi,\sigma_\phi\right)\right)^{Z_t}\right]$$

where the expression above results from the fact that Z_t follows the Binomial distribution. We find $\mu_\phi = -7.4823$ and $\sigma_\phi = 5.6096$. Figure 23 compares the Poisson exit rate to the exit rate under the assumption of a normal scrap value distribution.

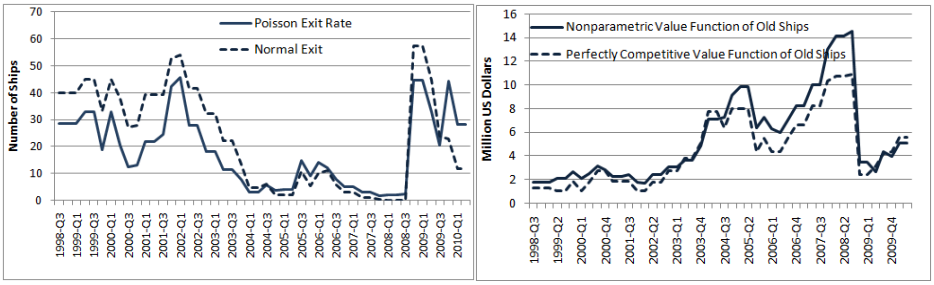

Figure 23: Poisson Exit Rate and Exit Rate under Normal Scrap Value Density. Value function of old ships estimated nonparametrically and computed from perfectly competitive profits.

In addition, to make sure that the primitives are internally consistent, we use the perfectly competitive profits to compute entry costs. In particular, we take the perfectly competitive profits and the normal scrap value distribution as our primitives and compute the corresponding value functions. We start with the Bellman equation for $V_{80} \in \mathbb{R}^L$:

$$V_{80} = \pi_{20+} + \beta\left(1-\widetilde{\mu}\right)PV_{80} + \beta\widetilde{\mu}E\left[\phi|\phi > PV_{80}\right]$$

These equations correspond to a contraction mapping if F_ϕ is log-concave (as shown in Pakes, Ostrovsky and Berry (2007)) and therefore easy to solve in V_{80} via a fixed point algorithm. We compute $E\left[\phi|\phi > P_{\lambda^*,\mu^*}V_{80,n-1}\right]$ by

$$E\left[\phi|\phi > A\right] = \frac{\int_A^\infty \phi f_\phi\left(\phi\right)d\phi}{1 - F_\phi\left(A\right)}$$

Figure 23 compares the nonparametrically estimated value function of old ships to the one that solves the above fixed point and shows that the two are close, as suggested by the good fit of the perfectly competitive behavior.

Once V_{80} is known, we recover all remaining value functions sequentially from:

$$\begin{aligned} V_{79} &= \pi_{20+} + \beta P V_{80} \\ V_{78} &= \pi_{10-20} + \beta P V_{79} \\ &\vdots \\ V_{39} &= \pi_{0-10} + \beta P V_{40} \\ &\vdots \\ V_0 &= \pi_{0-10} + \beta P V_1 \end{aligned}$$

Finally, using the retrieved V_0 the entry costs that correspond to the perfectly competitive environment are given by (14).

We are now ready to solve for the equilibrium of the game. The goal is to recover the entry policy $\lambda \in \mathbb{R}^L$, the exit policy $\mu \in \mathbb{R}^L$ and value function $V \in \mathbb{R}^{L\times A+1}$, using the estimated primitives, i.e. the perfectly competitive profits $\pi \in \mathbb{R}^{L\times A+1}$, the scrap value distribution $N\left(\mu_\phi, \sigma_\phi\right)$ and the entry costs $\kappa \in \mathbb{R}^L$ under different counterfactual scenarios. Below we describe the fixed point algorithm used to compute the equilibrium of the game, which is inspired by Weintraub, Benkard and Van Roy (2009). At each iteration n, we use the policies $\left(\lambda_{n-1}, \mu_{n-1}\right)$ and value functions V_{n-1} from the previous iteration to update to (λ_n, μ_n, V_n). We repeat the iterations until (λ_n, μ_n) don't change in n as follows:

1. Update exit $\mu^* = s^{80}\left(1 - F_\phi\left[P_{\lambda_{n-1},\mu_{n-1}}V_{80,n-1}\right]\right)$
2. Compute the new transition matrix P_{λ_{n-1},μ^*} as described in Section 5.1.5.

3. Update the value functions:

$$\begin{aligned} V_{80,n} &= \pi_{20+} + \beta\left(1-\mu^*\right) P_{\lambda_{n-1},\mu^*} V_{80,n-1} + \beta\mu^* E\left[\phi|\phi > P_{\lambda_{n-1},\mu^*} V_{80,n-1}\right] \\ V_{79,n} &= \pi_{20+} + \beta P_{\lambda_{n-1},\mu^*} V_{80,n} \\ &\vdots \\ V_{0,n} &= \pi_{0-10} + \beta P_{\lambda_{n-1},\mu^*} V_{1,n} \end{aligned}$$

4. Update entry $\lambda^*(X) = \lambda_{n-1}(X) \frac{\beta^{T(X)}\left(P_{\lambda_{n-1},\mu^*,X}\right)^{T(X)} V_{n-1}}{\kappa(X)}$, all X, where $P_{\lambda_{n-1},\mu^*,X}$ is the Xth row of P_{λ_{n-1},μ^*}.

5. Set $\lambda_n = \alpha_n\lambda^* + (1-\alpha_n)\lambda_{n-1}$ and $\mu_n = \gamma_n\mu^* + (1-\gamma_n)\mu_{n-1}$ where α_n and γ_n go to zero as n increases.

Once the optimal policies are recovered, we regress them on the state.

APPENDIX C

In this Appendix, we analyze the dynamics of the mean of the state vector, in order to explore the oscillatory and steady state behavior of the industry. The covariance has also been considered but is omitted due to space limitations.

Consider the state $x_t \equiv \left[s_t^0, ..., s_t^A, b_t^1, ..., b_t^{\overline{T}}, d_t\right] \in \mathbb{R}^{A+\overline{T}+2}$ and its transition as described in Section 5.1.5, which we rewrite here as follows:

$$x_{t+1} = v + Fx_t + G\begin{bmatrix} N_t \\ Z_t \end{bmatrix} + he_{t+1} \tag{16}$$

The precise form of (16) depends on the time to build assumption (endogenous, constant, no). We focus on constant time to build of T periods; the no time to build case results when $T = 1$. The endogenous time to build case is harder to analyze because the matrices F and G vary with the state. Under the assumption of constant time to build, (16) takes the form:

$$F = \begin{pmatrix} A & 0 & B \\ 0 & C & 0 \\ 0 & 0 & A \\ 0 & 0 & \rho \end{pmatrix} \in \mathbb{R}^{(81+T+1)\times(81+T+1)}$$

where

$$A = \begin{pmatrix} 0_{1\times 39} & 0_{39\times 1} \\ I_{39} & 0 \end{pmatrix} \qquad C = \begin{pmatrix} 0_{1\times 39} & 0_{39\times 1} \\ I_{39} & 1 \end{pmatrix}$$

and $B \in \mathbb{R}^{T\times T}$ is a zero matrix whose $(1,T)$ element equals 1, $v = [0\ 0 \cdots 0\ c]' \in \mathbb{R}^{81+T+1}$, $h = e_1 \in \mathbb{R}^{81+T+1}$, $G = [e_2\ \ -e_{81+T+1}] \in \mathbb{R}^{(81+T+1)\times 2}$ and e_i a vector which has all entries zero except for the i-th entry which equals 1. F has multiple eigenvalues at 0, 1, ρ. Let $m_t = E[x_t|x_0]$ for some initial condition x_0. Then, after some manipulations we get that $m(0) = x_0$ and:

$$m_{t+1} = v + Fm_t + G \begin{bmatrix} E[\lambda(x_t)|x_0] \\ E[\mu(x_t)|x_0] \end{bmatrix} \tag{17}$$

The stationary points of (17), m, satisfy (17) for $m_{t+1} = m_t = m$ and are of the form:

$$m = \left[a, a, ..., a, a^{80}, a, ..., a, \frac{c}{1-\rho}\right]$$

where $a = E[\lambda(x_t)|x_0] = E[\mu(x_t)|x_0]$ and α^{80} solves $a = E[\mu(x_t)|x_0]$. For the counterfactual performed in Section 6 with $T = 7$ we find that the steady state fleet which is equal to $80a + a^{80}$ is close to the mean fleet value obtained by the long run simulations.

The behavior of (small) deviations from stationary points is explored by linearizing (17), using the functional forms $\lambda(x) = e^{\gamma_0+\gamma' x}$ and $\mu(x) = e^{\delta_0+\delta' x}$. The matrix that governs the dynamics of the linearized system is

$$F + aGK, \text{ where } K = \begin{bmatrix} \gamma' \\ \delta' \end{bmatrix}$$

The eigenvalues of $F + aGK$ determine the return of the system to the stationary point, the rate of convergence and the oscillatory features. They also influence the behavior of variables that are a function of the state, such as the total fleet and prices. We compute the eigenvalues of $F + aGK$ using the estimated c, ρ and the (γ, δ) found from the constant time to build counterfactual. $F + aGK$ has multiple complex eigenvalues leading to oscillatory behavior. The largest eigenvalue (in terms of its norm) is 0.99, which means that the system returns to steady state slowly, with fluctuations that fade very slowly. This finding is consistent with Figure 22. As the total fleet is a linear function of the state, we are able to isolate the eigenvalues of $F + aGK$ that determine its behavior; these are also complex, with the largest one equal to 0.98. Price is a nonlinear function of the state and the dynamics of its mean and covariance are more complicated, unless

first order approximations are made.

APPENDIX D

This Appendix provides the estimated parameters from the first step of our estimation strategy. Table 5 provides the estimated parameters of the inverse demand curve, found in Section 5.1.1. Table 6 presents the results of the time to build function, found in Section 5.1.4. Tables 7 and 8 present the entry and exit regression results, found in Section 5.1.5.

	const	WIP	agr raw mat P	mineral P	food P	China steel	Handymax	$\widehat{Q_t}$
param	-7.601	9.501	2.969	-1.658	-0.346	1.534	-4.705	-0.162
s.e.	(23.8)	$(4.51)^*$	$(1.32)^*$	$(0.565)^*$	(0.702)	$(0.592)^*$	$(1.324)^*$	(0.597)

Table 5: Inverse demand curve for freight transport: IV regression results. Star indicates significance at the 0.05 level.

	constant	S^1	S^2	S^3	B	d
Param	2.536	-0.00082	-0.00063	0.00011	$1.93e-005$	0.0303
s.e.	(1.266)	(0.00058)	(0.00036)	(0.00036)	$(8.3e-005)$	(0.019)

Table 6: Time to Build regression estimates. Standard errors based on 500 bootstrap samples. Coefficients jointly significant at the 0.01 level.

Entry							
	constant	S^1	S^2	S^3	B	B^2	d
Param	-3.68	-0.00134	-0.00019	-0.00014	0.0069	$-5.37e-006$	0.486
s.e.	(7.99)	(0.0049)	(0.0019)	(0.0018)	(0.0024)	$(2.12e-006)$	(0.22)
Exit							
Param	22.68	0.0040	0.000178	0.00050	-0.0042	$4.141e-006$	-1.585
s.e.	(10.15)	(0.0065)	(0.0023)	(0.0023)	(0.0036)	$(3.09e-006)$	$(0.34)^*$

Table 7: Entry and exit regression estimates. Standard errors based on 500 bootstrap samples. Star indicates significance at the 0.05 level.

Entry					
	constant	S^1	S^2	S^3	d
Param	-8.425	-0.0024	-0.00045		0.934
s.e.	(4.90)	(0.0025)	(0.00075)		$(0.244)^*$
Exit					
Param	22.728	0.0073	0.00093	0.00104	-1.859
s.e.	$(4.89)^*$	$(0.0016)^*$	(0.00092)	(0.0008)	$(0.242)^*$

Table 8: Entry and exit regression estimates. Standard errors based on 500 bootstrap samples. Star indicates significance at the 0.05 level.

References

[1] Adland, Roar, and Steen Koekebakker. 2007. "Ship Valuation Using Cross-Sectional Sales Data: A Multivariate Non-parametric Approach". Maritime Economics and Logistics 9: 105-118.

[2] Aguirregabiria, Victor, and Pedro Mira. 2007. "Sequential Estimation of Dynamic Discrete Games". Econometrica 75 (1): 1-53.

[3] Alessandria, George, Kaboski Joseph, and Virgiliu Midrigan. 2009. "Inventories, Lumpy Trade and Large Devaluations". American Economic Review. Forthcoming.

[4] Bajari, Patrick, Benkard Lanier C., and Jonathan Levin. 2007. "Estimating Dynamic Models of Imperfect Competition". Econometrica 75 (5): 1331-1370.

[5] Beenstock, Michael, and Andreas Vergottis. 1989. "An Econometric Model of the World Market for Dry Cargo Freight and Shipping". Applied Economics 21 (3): 339-356.

[6] Besanko, David, and Ulrich Doraszelski. 2004. "Capacity Dynamics and Endogenous Asymmetries in Firm Size". Rand Journal of Economics 35 (1): 23-49.

[7] Benkard, Lanier C. 2004. "A Dynamic Analysis of the Market for Wide-Bodied Commercial Aircraft". Review of Economic Studies 71: 581-611.

[8] Bloom, Nicholas. 2009. "The Impact of Uncertainty Shocks". Econometrica 77 (3): 623-685.

[9] Bloom, Nicholas, Bond Stephen and John Van Reenen. 2007. "Uncertainty and Investment Dynamics". Review of Economic Studies 74: 391-415.

[10] Caballero, J. Ricardo, and Robert S. Pindyck. 1996. "Uncertainty, Investment and Industry Evolution". International Economic Review 37 (3): 641-662.

[11] Collard-Wexler, Allan. 2010. "Demand Fluctuations in the Ready-Mix Concrete Industry" http://pages.stern.nyu.edu/~acollard/

[12] Deaton, Angus, and Guy Laroque. 1992. "On the Behavior of Commodity Prices", Review of Economic Studies 59: 1-23.

[13] Dixit, Avinash. 1989. "Entry and Exit Decisions Under Uncertainty". The Journal of Political Economy 97 (3): 620-638.

[14] Dixit, Avinash, and Robert S. Pindyck. 1994. Investment Under Uncertainty. Princeton University Press.

[15] Doraszelski, Ulrich, and Mark Satterthwaite. 2009. "Computable Markov-perfect Industry Dynamics". Rand Journal of Economics 41 (2): 215-243.

[16] Dunne, Timothy, Klimmek D. Shawn, Roberts J. Mark, and Daniel Y. Xu. 2007. "The Dynamics of Market Structure and Market Size in Two Health Service Industries". In Producer Dynamics: New Evidence from Micro Data, edited by Dunne, Timothy, Jensen J.Bradford, and Mark J. Roberts.

[17] Ericson, Richard, and Ariel Pakes. 1995. "Markov-Perfect Industry Dynamics: A Framework for Empirical Work". The Review of Economic Studies 62(1): 53-82.

[18] Fan, Jianqing, and Irene Gijbels. 1996. Local Polynomial Modelling and Its Applications. Chapman & Hall/CRC.

[19] Gavazza, Alessandro. 2010. "Leasing and Secondary Markets: Theory and Evidence from Commercial Aircraft". The Journal of Political Economy. Forthcoming.

[20] Hopenhayn, Hugo. 1992. "Entry, Exit and Firm Dynamics in Long Run Equilibrium". Econometrica 60 (5): 1127-1150.

[21] Hotz, V. Joseph, and Robert A. Miller. 1993. "Conditional Choice Probabilities and the Estimation of Dynamic Models". Review of Economic Studies 60 (3): 497-529.

[22] Hamilton, D. James, 2005. "What's Real About the Business Cycle?". Federal Reserve Bank of Saint Lewis Review 87 (4): 435-452.

[23] Hogg, V. Robert, Craig T. Allen, and Joseph W. McKean. 2004. Introduction to Mathematical Statistics. Prentice Hall.

[24] Jofre-Bonet, Mireia and Martin Pesendorfer. 2003. "Estimation of a Dynamic Auction Game". Econometrica 71 (5): 1443-1489.

[25] Kydland, E. Finn, and Edward C. Prescott. 1982. "Time to Build and Aggregate Fluctuations". Econometrica 50(6): 1345-1370.

[26] Lambson, E. Val. 1991. "Industry Evolution with Sunk Costs and Uncertain Market Conditions". International Journal of Industrial Organization 9: 171-196.

[27] Pagan, Adrian, and Aman Ullah. 1999. Nonparametric Econometrics. Cambridge University Press.

[28] Pakes, Ariel, Ostrovsky Michael and Steven T. Berry. 2007. "Simple Estimators for the Parameters of Discrete Dynamics Games". The RAND Journal of Economics 38 (2): 373-399.

[29] Rosen, Sherwin, Murphy M. Kevin and Jose A. Scheinkman. 1994. "Cattle Cycles". The Journal of Political Economy 102 (3): 468-492.

[30] Rust, John. 1987. "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher". Econometrica. 55 (5): 999-1033.

[31] Ryan, P. Stephen. 2009. "The Costs of Environmental Regulation in a Concentrated Industry" http://econ-www.mit.edu/faculty/sryan

[32] Snider, Connan. 2008. "Predatory Incentives and Predation Policy: The American Airlines Case" http://sites.google.com/site/connansnider/

[33] Stopford, Martin. 1999. Maritime Economics. New York: Routledge.

[34] UNCTAD. 2009. "Review of Maritime Transport".

[35] Weintraub, Y. Gabriel, Benkard Lanier C. and Benjamin Van Roy. 2008. "Markov Perfect Industry Dynamics with Many Firms". Econometrica 76 (6): 1375-1411.

[36] Weintraub, Y. Gabriel, Benkard Lanier C. and Benjamin Van Roy. 2009. "Computational Methods for Oblivious Equilibrium" http://www.stanford.edu/~lanierb/.

[37] Xu, Y. Daniel. 2008. "A Structural Empirical Model of R&D, Firm Heterogeneity and Industry Evolution" http://sites.google.com/site/yixusite/