Detection and Impact of Industrial Subsidies: The Case of World Shipbuilding

Myrto Kalouptsidi[*†]

Department of Economics, Princeton University and NBER

PRELIMINARY

March 2014

Abstract

This paper provides a model-based empirical strategy to, first, detect the presence and magnitue of government subsidies and, second, quantify their impact on industry prices, production reallocation across countries and profits. We apply this strategy to world shipbuilding, an industry long tought to be affected by such policies. We construct a model for the market of new ships, where both demand and supply are dynamic. We find strong evidence that China intervened in its shipbuilding industry and reduced shipyard costs by 15-20%. These production subsidies led to massive reallocation of ship production across the world, with Japan losing significant market share.

Keywords: industry dynamics, government subsidies, China, shipping, shipbuilding

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

[†]I am very thankful to John Asker, Jan De Loecker, Gene Grossman, Tom Holmes, Bo Honore, Eduardo Morales, Richard Rogerson, Junichi Suzuki.

1 Introduction

Government subsidies to industries have been prevalent throughout economic history and in several countries have steered industrialization and growth. Understanding, therefore, their role as determinants of industry prices, market shares, and production reallocation across countries is an important question of interest. In addition, whether this impact varies by different types of subsidies (e.g. subsidies to fixed or variable costs) remains unexplored. A significant challenge in this task is that government transfers to industries are notoriously difficult to detect. Indeed, partly due to WTO agreements that prohibit direct and in-kind subsidies, other than infrastructure,[1] the existence and magnitude of such subsidies is often unknown.

In this paper, we provide a model-based empirical strategy to first, detect the presence and magnitude of government subsidies, and second, quantify the impact of subsidies on industrial evolution. Our strategy draws from the standard IO insights of estimating a cost function from demand variation and applies them to a framework of dynamic demand and supply. Once the cost function is recovered, we use suspected dates of initiation of government subsidization plans and compare costs before and after, as well as across countries.

We apply our strategy to the world shipbuilding industry which is a prototypical example of an industry affected by such policies. Indeed, shipbuilding has historically been a key pillar of countries' industrialization phase (1850's Britain, 1950's Japan, 1970's South Korea, today's China) and is concentrated in a handful of countries: Europe, Japan, South Korea and now China. Several disputes regarding subsidies have occurred.[2] In 2006, China launched its "Long and Medium Term Plan for the Shipbuilding Industry (2006-2015)", which doubled world shipbuilding docks and massively increased China's market share. What is known, observed and admissible by the WTO is the Chinese government's subsidization of the construction of new and the capital expansion of existing shipbuilding plants (shown in the first panel of Figure 1). What is not known, unobserved

[1] In its Agreement on Subsidies and Countervailing Measures, the WTO defines a subsidy as an unrequited financial contribution by a government to enterprises in the form of: (i) direct transfer of funds, (ii) foregone revenue that is otherwise due, (iii) provision of goods or services, except infrastructure, (iv) payments to a funding mechanism to carry out one or more of the type of functions illustrated in (i) to (iii).

[2] A recent example was Europe's accusation for Korean subsidies in 2001, which was not accepted by the WTO: "No progress was achieved, as the Korean Government claimed that it had no influence on the shipyards or on the financial institutions supporting them, and further said that it was convinced business was conducted along free market principles". (EU Commission)

and prohibited by the WTO is the subsidization of shipbuilding production. The second panel of Figure 1 depicts the rapid increase of China's market share; yet this is not necessarily evidence of subsidies: differentiated products, demand variation and inherent cost differences can also explain the observed market share evolution.

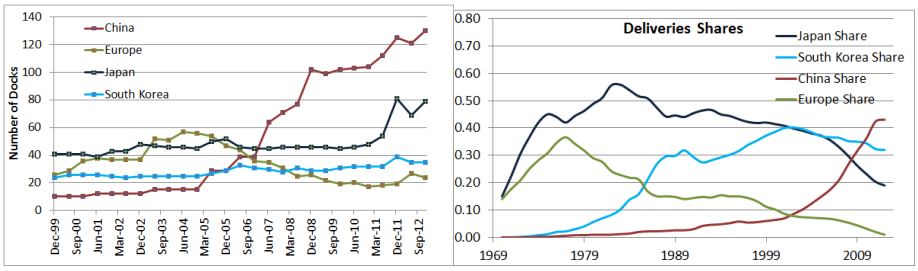


Figure 1: Docks and delivery shares per country.

We construct and estimate a dynamic model of the shipping and shipbuilding industries to disentangle the role of each of the above factors and detect the presence of government subsidies. Our model links the downstream shipping and the upstream shipbuilding industries, providing one of the first empirical analysis in industrial organization looking at dynamic agents on both the demand and the supply side. A shipping firm is a ship. Ships are long-lived and every period they compete in the world market for cargo under convex operating costs that vary with the ship's age and country of built. Demand for freight is uncertain and volatile. Every period a large number of identical potential shipowners decide to enter the market by buying a new ship from world shipyards. Ship prices are bid up to the ship expected discounted lifetime profitability, as in a free entry condition. A large number of shipyards offer differentiated ships and compete by choosing their production level every period, under convex production costs. Due to time to build, inherent in shipbuilding, ships remain in shipyards' backlogs for several years, making production a dynamic choice. A high backlog can either raise production costs because of capacity constraints, or decrease costs because of economies of scale or the accumulation of expertise. We let our estimation dictate which effect dominates. In addition, the backlog affects the shipyard's demand, as it increases its offered time to build. We do not model shipyard entry, but rather assume it is driven by government intervention.

Our model primitive of interest is the cost function of potentially subsidized firms. Our estimation strategy first uses new and used ship prices to estimate the willingness to pay for a new ship and then inserts it into the dynamic optimization problem of shipbuilders. At an intuitive level, we follow the standard techniques of the empirical IO literature (e.g. Berry, Levinshon and Pakes (1995)) where production costs are identified via demand

estimation. Methodologically, however, we take a different approach, since we face a dynamic setup of profit-maximizing firms in both the demand and the supply side of the market. To estimate demand for new ships, we extend Kalouptsidi (2013) where ship value functions are estimated via used ship transaction prices, by adding new ship prices. To estimate costs, we adopt a hybrid approach that is inspired by the recent literature on the estimation of dynamic setups (e.g. Bajari, Benkard and Levin (2007)). Finally, our estimation treats China's 2006 government plan as an unexpected and permanent change from the point of view of industry participants: shipowner expectations and value functions are estimated separately before and after 2006 capturing the change in demand for ships caused by China's intervention.

Once model estimates are obtained, we test whether the cost function of Chinese shipyards is different before and after 2006. We find a strongly significant decline in costs, implying subsidies equal to about 15-20% of costs, which corresponds to about 5 billion US dollars at the observed production levels. We control for various shipyard characteristics, as well as functions of time to alleviate concerns of other time-varying factors leading to cost declines. In addition, we compare our detection method to the price-gap approach used in WTO cases; the latter recovers subsidies equal to 4-7%, less than a third of our retrieved magnitude. Even though our empirical approach is implemented in shipbuilding, it can be also be used in other suspect industries (e.g. steel).

Next, we use our estimated model to quantify the impact of China's capital infrastructure and production subsidization program on ship prices, production reallocation across countries, as well as profits and industry costs. We find that China led to a substantial reallocation of production across the world, significantly reducing Japan's market share and profits. Interestingly, we find that it is production subsidies, rather than capital subsidies, that seem to contribute most to this reallocation.

This paper contributes to the long theoretical (e.g. Jovanovich (1982), Hopenhayn (1992), Ericson and Pakes (1995)) and recent empirical (e.g. Aguirregabiria and Mira (2007), Jofre-Bonet and Pesendorfer (2003), Benkard (2004), Ryan (2011), Collard-Wexler (2008), Xu (2008), Sweeting (2013)) literature on industry dynamics. Methodologically, we lie closest to Rust (1987), Hotz and Miller (1993), Bajari, Benkard and Levin (2007) and Pakes, Ostrovsky and Berry (2007), yet this literature considers either single agent dynamics or dynamic firms and static consumers. To tackle the difficulty of having dynamic consumers (shipowners) and dynamic producers (shipbuilders), we resort to second-hand sale transactions, extending Kalouptsidi (2013). Such transaction prices may be helpful in other markets of durable goods which are characterized by dynamics in both demand

and supply, yet most of the literature has not been able to allow for (with the exception of Chen, Esteban and Shum (2013)).

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 and the estimation results. Section 6 provides the counterfactual experiments and Section 7 concludes.

2 Industry Description

Commercial ships are the largest factory produced product. Based on Stopford (2009), a 30,000 DWT bulk carrier might contain 5,000 tons of steel and 2,500 tons of other components (including the main engine and numerous minor components such as cabling, pipes, furniture and fittings). Materials account for about half the cost of the ship (the steel is around 13%, the main engine around 16%) and labor about 17% of total cost. A shipyard constructs the steel hull and conducts the outfitting of the hull with machinery, equipment services and furnishings (Stopford (2009)); many of these operations are conducted simultaneously, with individual tasks not requiring highly technical skills.

As Figure 2 shows, a small number of countries have always dominated shipbuilding production, often as a result of governmental policies. In the 1850's, Britain was the world leading shipbuilder, until it was overtaken by Japan in the 1950's, which in turn lost its leading position to Korea, in the 1970's. Japan used its shipbuilding industry to rebuild its industrial capability motivated by its strong maritime tradition, while Korea saw shipbuilding as a strategic core for its economic development and was oriented towards exporting (OECD (2008)). Until recently, China was a small player and was regarded as a risky place to build new ships, while Chinese built vessels commanded a significant discount in both the new-building and second-hand market (Stopford (2009)). China's recent shipbuilding expansion, unlike that of Japan's and South Korea's whose slow growth took a couple of decades to complete, consisted of a 500% increase in deliveries between 2006 and 2010 (Stopford (2009)).

Shipbuilding is often seen as a "strategic industry" as it generates employment, accelerates regional development, increases industrial and defence capacity and can have important spill-overs to the iron and steel, electronic, and machinery manufacturing industries (OECD (2008)). Indeed, several of today's leading economies, as mentioned above, developed their production technologies and human capital through a phase of heavy industrialization, in which shipbuilding was one of key pillars, along with steel

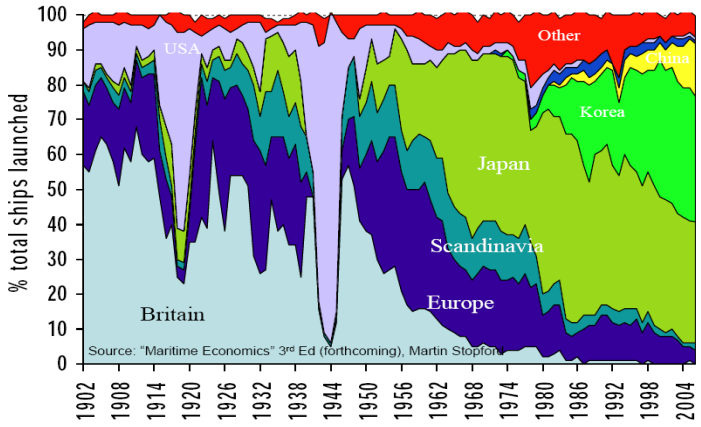


Figure 2: Shipbuilding History.

and petrochemicals. China's "Long and Medium Term Plan" of 2006 urges the country's shipbuilders "to propel the country to world No. 1 status" and "implements plans to strengthen and upgrade the overall shipbuilding industrial capability through the construction of shipyards, while also upgrading existing shipbuilding facilities" (Collins and Grubb (2008)).[3] Nevertheless, it is claimed that the government is not involved in general business operations of individual companies (OECD (2008)), while even state-owned "shipyards largely function as independent corporate entities and handle day-to-day operations and contract bids" (Collins and Grubb (2008)). Finally, China's shipbuilding is mostly geared towards export sales which comprised about 80% of its orderbook in 2006 (Collins and Grubb (2008)). Figure 1 shows China's expansion in both capital infrastructure as measured by shipbuilding docks, as well as market share.[4]

Shipbuilding demand is determined by entry in the shipping industry. In this paper we focus on cargo transportation (as opposed to cruise ships for example) and in particular, bulk shipping, which concerns vessels designed to carry a homogeneous unpacked dry or liquid cargo, for individual shippers on non-scheduled routes. The entire cargo usually

[3]The Chinese government further supports the shipbuilding industry by exempting it from import tariffs of key components necessary for the production of some kinds of high-tech ships, and by providing incentives for investment in R&D and innovation (both not relevant for the bulk carriers that we study here).

[4]Over the years, numerous disputes have arisen whereby countries claim their domestic industries are hurt because of subsidies in foreign countries. In recent years, China has been a target of such complaints (See Haley and Haley (2013) for an overview) that have trickled down to the press. A great example, is "Perverse Advantage", published in The Economist, April 2013: "China is the workshop to the world. It is the global economy's most formidable exporter and its largest manufacturer. The explanations for its success range from a seemingly endless supply of cheap labour to an artificially undervalued currency. (...) another reason for China's industrial dominance: subsidies".

belongs to one shipper (owner of the cargo). 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 can 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 small shipowning firms.

Demand for shipping services is driven by world seaborne trade and is thus subject to world economy fluctuations. In recent years, the growth and infrastructure building at several countries led to increased imports of raw materials, significantly boosting demand for bulk transport. In the short run, the supply of shipping services is determined by the number of voyages carried out by shipowners, who can adjust the ton-miles they offer by adjusting their speed of sail. Even so, short run supply is rather inelastic as voyage costs are convex in speed. In the long run, the supply of cargo transportation adjusts via the building and scrapping of ships. Exit in the industry occurs when shipowners scrap their ships by selling them to scrapyards where they are dismantled and their steel hull is recycled. Entry in the industry occurs when shipowners buy new ships from world shipyards.

3 Model

In this section, we present a dynamic model of the world bulk shipping and shipbuilding industries, which lies within the general class of dynamic models studied in Ericson and Pakes (1995) and Hopenhayn (1992). Time is discrete and the horizon is infinite. Shipowners create demand for shipbuilders, who respond by supplying new ships. We begin by describing shipowner behavior, then turn to the shipbuilders. We also discuss how government subsidies enter.

3.1 Demand for New Ships (Shipowners)

There is a finite number of incumbent shipowners (the fleet) and a large number of identical potential entrant shipowners. We assume constant returns to scale, so that a firm is a ship. Ships are long-lived. The state variable of ship i at time t, s_{it}, includes its:

1. age, $a_{it} \in \{0, 1, ..., A\}$
2. country of built, $c_i \in \mathcal{C}$

while the industry aggregate state s_t includes:

1. the distribution of characteristics in s_{it} over the fleet, $S_t \in \mathbb{R}^{A \times ||\mathcal{C}||}$
2. the backlog $b_t \in \mathbb{R}^{J \times \overline{T}}$, whose $(j,k)^{th}$ element is the number of ships scheduled to be delivered at period $t+k$ by shipyard j
3. the aggregate demand for shipping services, $d_t \in \mathbb{R}^+$, capturing shifts in the inverse demand curve for freight transport
4. the price of steel, $l_t \in \mathbb{R}^+$.[5]

In period t, each shipowner i chooses how much transportation (i.e. ton-miles travelled) to offer, q_{it}. Shipowners face the inverse demand curve:

$$P_t = P\left(d_t, Q_t\right) \tag{1}$$

where P_t is the price per ton-mile, d_t defined above includes demand shifters, such as world industrial production and commodity prices and Q_t denotes the total ton-miles offered, so that $Q_t = \sum_i q_{it}$. Ton-miles are a homogeneous good, but shipowners face heterogeneous convex costs of freight, $c^F\left(q_{it}, s_{it}\right)$. Ship operating costs increase with the ship's age and may differ based on country of built because of varying quality.

We assume that shipowners act as price-takers in the market for freight. Their resulting per period payoffs are $\pi\left(s_{it}, S_t, d_t\right)$.[6]

A ship lives a maximum of A periods. At the same time, a ship can be hit by an exit shock each period. In particular, we assume that a ship at state (s_{it}, s_t) exits with probability $\delta\left(s_{it}, s_t\right)$ and receives a deterministic scrap value $\phi\left(s_{it}, s_t\right)$. Note that $\delta\left(\left[a_{it}, c_i\right], s_t\right) = 1$, for $a_{it} \geq A$ and $\phi\left(\left[a_{it}, c_i\right], s_t\right) = 0$, for $a_{it} > A$ and for all c_i, s_t.[7]

The only dynamic control of shipowners is entry in the industry: each period, a large number of identical potential entrants simultaneously make entry decisions. There is time to build, in other words, a shipowner begins its operation a number of periods after its entry decision. To enter, shipowners purchase new vessels from world shipyards. Shipyard j in period t can build a new ship at price P_{jt}^{NB} and time to build T_{jt}. The assumption of a large number of homogeneous potential shipowners implies that shipyard prices are

[5] The steel price is a state variable for two reasons: first, it determines the ship's scrap value; second and most important, it is a key determinant of shipyard production costs and thus determines future ship entry and competition in the shipping industry.

[6] Note that profits are not equal to zero because of the convex operating costs.

[7] Generalizing to endogenous exit is straightforward (see Kalouptsidi (2013)).

bid up to the ships' values and shipyards can extract all surplus. One can also think of this as a free entry condition in the shipping industry where the entry cost is equal to the shipyard price. Therefore, the following equilibrium condition holds:

$$P_{jt}^{NB} = E\left[\beta^{T_{jt}} V\left(s_{it+T_{jt},s_{t+T_{jt}}}\right)|s_{it}, s_t\right] \tag{2}$$

where s_{it} in this case involves $a_{it} = 0$ and the country of yard j, and the value function $V(s_{it}, s_t)$ satisfies the Bellman equation:

$$V(s_{it}, s_t) = \pi(s_{it}, s_t) + \delta(s_{it}, s_t)\phi(s_{it}, s_t) + (1 - \delta(s_{it}, s_t))\beta E\left[V(s_{it+1}, s_{t+1})|s_{it}, s_t\right] \tag{3}$$

In words, the value function of a ship at state (s_{it}, s_t) equals the profits from cargo transport plus the scrap value which is received with probability $\delta(s_{it}, s_t)$ and the continuation value $E\left[V(s_{it+1}, s_{t+1})|s_{it}, s_t\right]$, which is received with probability $1 - \delta(s_{it}, s_t)$.

In practice, shipowners can also buy a used ship. In this model, ships are indistinguishable from their owners and therefore, transactions in the second-hand market do not affect entry or profits in the industry. In addition, since there is a large number of identical shipowners who share the value of a ship, the price of a ship in the second hand market, P_{it}^{SH}, equals this value and shipowners are always indifferent between selling their ship and operating it themselves. Therefore, in equilibrium:

$$P_{it}^{SH} = V(s_{it}, s_t) \tag{4}$$

We revisit sales in the empirical part of the paper, where both second-hand and newbuilding prices are treated as observations on the value function.

3.2 Supply of New Ships (Shipyards)

There are J long-lived incumbent shipbuilders. The state variable of shipyard j at time t, y_{jt}, includes its:

1. backlog $b_{jt} \in \mathbb{R}^{\overline{T}}$
2. country $c_j \in \mathcal{C}$
3. other characteristics, such as: age, capital equipment (number of docks and berths), number of employees, the length of its largest dock.

Shipyards also share the aggregate industry state, s_t, as this determines their demand in t.

Each period t, shipyard j draws a private iid (across j and t) production cost shock $\varepsilon_{jt} \sim N(0, \sigma)$ and makes its discrete production decision $N_{jt} \in \{0, 1, ..., \overline{N}\}$. Shipyard j faces production costs, $C(N_{jt}, y_{jt}, s_t, \varepsilon_{jt})$. Even though N_{jt} is an integer we assume that the cost function $C(N_{jt}, \cdot)$ can be defined over $[0, \overline{N}]$ and that as such it is convex in N_{jt}. We also assume that the cost shock ε_{jt} is paid for each produced unit, so that:[8]

$$C(N_{jt}, y_{jt}, s_t, \varepsilon_{jt}) = c(N_{jt}, y_{jt}, s_t) + N_{jt}\varepsilon_{jt} \tag{5}$$

In our model N_{jt} corresponds to the number of ships ordered in period t at shipyard j. These ships enter the shipyard's backlog b_{jt} and are delivered a number of years later.[9] Under demand uncertainty, therefore, undertaking a ship order becomes a dynamic choice. To capture these dynamics we assume that the cost function depends on the shipyard's backlog. As in Jofre-Bonet and Pesendorfer (2003), there are two opposing ways the backlog can impact costs: on one hand, increased backlogs can raise costs because of capacity constraints (less available labor etc.); on the other hand, increased backlogs can lower costs because of economies of scale (e.g. it might be easier to order inputs) or the accumulation of expertise. In addition, the shipyard's backlog affects its demand, as it increases the offered time to build.

As discussed above, shipyard j sells its ships at a price equal to the shipowners' entry value:[10]

$$VE_j(s_t) \equiv E\left[\beta^{T_{jt}} V\left(s_{it+T_{jt}}, s_{t+T_{jt}},\right) | s_{it}, s_t\right] \tag{6}$$

[8] Note that it is important to assume that the shipyard faces a single shock regardless of the chosen production level. Even though having for example logit shocks iid across $N = 0, 1, ...$ would simplify the analysis (in that case, the problem falls into the standard dynamic discrete choice framework), it seems implausible that production shocks are independent across production levels. Having production shocks independent across time periods is also a strong assumption but allowing for serially correlated unobserved state variables is a difficult issue that the literature hasn't tackled yet.

[9] As shown in the empirical exercise, N_{jt} is usually a small number between 0 and 20 (with the majority of observations lying between 0 and 5). The reason why we consider the number of orders as the relevant choice variable (as opposed to the number of deliveries or a smoothed version of orders) is that the observed ship prices are paid at the time of order and may be dramatically different from the prevailing prices at the delivery date. It would be, therefore impossible to compute a shipyard's revenue at any other point in time.

[10] Note that the willingness to pay for a new ship from yard j depends only on its country of origin, not j itself. Even though it is straightforward in the model to allow a ship's value to change with j, the hundreds of shipyards we encounter in the data make this generalization impossible.

where s_{it} has $a_{it}=0$ and the country of yard j. We let time to build be shipyard-specific and in particular, $T_{jt}=T\left(y_{jt},s_t\right)$. Note that $VE_j\left(s_t\right)$ does not explicitly depend on N_{jt}; in other words yards do not face a downward sloping demand curve. Indeed, N_{jt} affects the willingness to pay for the ship by entering into b_t and from there into S_t after T_{jt} periods. Typically, N_{jt} is a small integer, while the total fleet is a large number in the order of thousands. Therefore each shipyard, when making its production decision, N_{jt} can ignore the impact it has on $VE_j\left(s_t\right)$; note however, that aggregates do matter so that as the total fleet increases, shipowners' willingness to pay falls, all else equal.

Shipyard j chooses its production level to solve the Bellman equation:

$$W\left(y_{jt},s_t,\varepsilon_{jt}\right)=\max_{N\in\left\{0,1,...,\overline{N}\right\}}VE_j\left(s_t\right)N-c\left(N,y_{jt},s_t\right)-N\varepsilon_{jt}+\beta E\left[W\left(y_{jt+1},s_{t+1},\varepsilon_{jt+1}\right)|N,y_{jt},s_t\right] \tag{7}$$

To ease notation below, we also define the continuation value:

$$Q\left(y_{jt},s_t,N\right)\equiv E\left[W\left(y_{jt+1},s_{t+1},\varepsilon_{jt+1}\right)|N,y_{jt},s_t\right] \tag{8}$$

The expectation in (7), as well as (2) and (3) is over demand for shipping services, d_t, steel prices, l_t and shipyard production N_{jt}, all j. The demand state variable d_t and steel prices l_t evolve according to a first order autoregressive process with trend (see Section 5.1.2). Period t production, N_{jt}, enters in j's backlog, b_{jt}, at position T_{jt}, while the remaining elements of b_{jt} move one period closer to delivery with its first element being delivered. Note that the evolution of all other states is deterministic (see Section 5.1.2). The trend component in demand and steel prices implies that time t is explicitly part of the state (in other words, our state notation $\{s_{it},y_{jt},s_t\}$ incorporates t). Allowing for time to enter the agents' decision making offers some generality and is important in this application, as our empirical analysis of detecting government subsidies hinges on allowing time-varying factors to affect costs.

The shipyard's optimal production policy is as follows: shipyard j chooses $N_{jt}=0$ if:

$$-c\left(0,y_{jt},s_t\right)+\beta Q\left(y_{jt},s_t,0\right)\geq VE_j\left(s_t\right)n'-c\left(n',y_{jt},s_t\right)-n'\varepsilon_{jt}+\beta Q\left(y_{jt},s_t,n'\right),\quad\text{all } n'\in\left\{1,...,\overline{N}\right\}$$

or

$$\varepsilon_{jt}\geq VE_j\left(s_t\right)+\max_{0<n'\leq\overline{N}}\left\{\frac{c\left(0,y_{jt},s_t\right)-c\left(n',y_{jt},s_t\right)+\beta\left(Q\left(y_{jt},s_t,n'\right)-Q\left(y_{jt},s_t,0\right)\right)}{n'}\right\} \tag{9}$$

Similarly, shipyard j chooses $N_{jt} = n \neq 0, \overline{N}$ if:

$$VE_j(s_t)\, n - c(n, y_{jt}, s_t) - n\varepsilon_{jt} + \beta Q(y_{jt}, s_t, n) \geq VE_j(s_t)\, n' - c(n', y_{jt}, s_t) - n'\varepsilon_{jt} + \beta Q(y_{jt}, s_t, n'), \quad \text{all } n' \neq n$$

or

$$\varepsilon_{jt} \geq VE_j(s_t) + \max_{n<n'\leq\overline{N}} \left\{ \frac{c(n, y_{jt}, s_t) - c(n', y_{jt}, s_t) + \beta\left(Q(y_{jt}, s_t, n') - Q(y_{jt}, s_t, n)\right)}{n' - n} \right\} \tag{10}$$

$$\varepsilon_{jt} \leq VE_j(s_t) + \min_{0\leq n'<n} \left\{ \frac{c(n, y_{jt}, s_t) - c(n', y_{jt}, s_t) + \beta\left(Q(y_{jt}, s_t, n') - Q(y_{jt}, s_t, n)\right)}{n' - n} \right\} \tag{11}$$

Finally, shipyard j chooses $N_{jt} = \overline{N}$, if

$$\varepsilon_{jt} \leq VE_j(s_t) + \min_{0\leq n'<\overline{N}} \left\{ \frac{c(\overline{N}, y_{jt}, s_t) - c(n', y_{jt}, s_t) + \beta\left(Q(y_{jt}, s_t, n') - Q(y_{jt}, s_t, \overline{N})\right)}{n' - \overline{N}} \right\} \tag{12}$$

Therefore, a shipyard's optimal policy is characterized by relationships (9), (10) and (12). The next lemma, whose proof is in the Appendix, states that if the cost function $c(n, y, s)$ is convex in n, it is enough to compare each value of n only to $n+1$ and $n-1$, rather than every possible value of n.

Lemma 1 *If the shipbuilding cost function $C(n, \cdot) : [0, \overline{N}] \to \mathbb{R}$, is convex in n, then:*

$$\max_{n<n'\leq\overline{N}} \left\{ \frac{c(n,\cdot) - c(n',\cdot) + \beta\left(Q(\cdot, n') - Q(\cdot, n)\right)}{n' - n} \right\} = c(n,\cdot) - c(n+1,\cdot) + \beta\left(Q(\cdot, n+1) - Q(\cdot, n)\right)$$

$$\min_{0\leq n'<n} \left\{ \frac{c(n,\cdot) - c(n',\cdot) + \beta\left(Q(\cdot, n') - Q(\cdot, n)\right)}{n' - n} \right\} = c(n-1,\cdot) - c(n,\cdot) + \beta\left(Q(\cdot, n) - Q(\cdot, n-1)\right)$$

Proof. See the Appendix. ■

The above lemma, which is very intuitive, implies that the optimal policy of the shipyard takes the following final form: $N^*(y_{jt}, s_t, \varepsilon_{jt}) =$

$$\begin{cases} 0, \quad \text{if } \varepsilon_{jt} \geq VE_j(s_t) + c(0, y_{jt}, s_t) - c(1, y_{jt}, s_t) + \beta\left(Q(y_{jt}, s_t, 1) - Q(y_{jt}, s_t, 0)\right) \\ n, \ \text{if } \varepsilon_{jt} \in \left[\begin{array}{c} VE_j(s_t) + c(n, y_{jt}, s_t) - c(n+1, y_{jt}, s_t) + \beta\left(Q(y_{jt}, s_t, n+1) - Q(y_{jt}, s_t, n)\right), \\ VE_j(s_t) + c(n-1, y_{jt}, s_t) - c(n, y_{jt}, s_t) + \beta\left(Q(y_{jt}, s_t, n) - Q(y_{jt}, s_t, n-1)\right) \end{array} \right] \\ \overline{N}, \quad \text{if } \varepsilon_{jt} \geq VE_j(s_t) + c\left(\overline{N} - 1, y_{jt}, s_t\right) - c\left(\overline{N}, y_{jt}, s_t\right) + \beta\left(Q\left(y_{jt}, s_t, \overline{N}\right) - Q\left(y_{jt}, s_t, \overline{N} - 1\right)\right) \end{cases} \tag{13}$$

The timing in each period is as follows:

1. Incumbent and potential entrant shipowners observe their state (s_{it}, s_t) and shipbuilders observe their state (y_{jt}, s_t)
2. Shipowners are hit by exit shocks and shipbuilders observe their private production cost shocks
3. Shipyards make production decisions
4. Shipowners receive profits from freight services and shipyards receive profits from new ship production
5. Exiting ships receive their scrap value $\phi(s_i, s)$
6. States are updated

We consider a competitive equilibrium which consists of an optimal production policy function $N^*(y_{jt}, s_t, \varepsilon_{jt})$ that is given by (13), as well as value functions $W(y_{jt}, s_t)$ and $V(s_{it}, s_t)$ that solve (7) and (3) respectively, while all expectations employ $N^*(y_{jt}, s_t, \varepsilon_{jt})$. Existence of equilibrium follows from Hopenhayn (1992), Jovanovich (1982) and Doraszelski and Satterthwaite (2010).

Finally, we discuss how China's entry is modelled. We assume that China's 2006 subsidization program was an unexpected, one-shot, permanent change from the point of view of industry participants. Explicitly modeling expectations with regard to policy interventions is extremely complicated and would rely on strong and perhaps ad hoc assumptions. Within our model, the before and after 2006 worlds differ in the number of shipyards, shipbuilding infrastructure (found in y_{jt}) and China's cost function. We also assume that shipyards do not make entry or capital expansion decisions. On one hand, outside of China there is no such action (see the first panel of Figure 1), while within China, these decisions are determined by government policy.

4 Data

We use data from Clarksons, a leading shipbroking firm based in the UK. We employ five different datasets.

The first dataset reports shipbuilding quarterly production (i.e. orders) between Q1-2001 and Q3-2012. For each shipyard and quarter we observe its production in tons

and number of bulk ships, as well as the yard's backlog, its deliveries and average time to build. The shipyard's country is also reported. There are 192 yards that produce Handysize vessels (the segment on which our empirical analysis will focus). There are 119 Chinese, 41 Japanese, 21 South Korean and 11 European shipyards. The majority of bulk ship production occurs in China and Japan; hence even though we include Europe and South Korea in our estimation and counterfactuals, most comparisons will be made between China and Japan.

The second dataset is a sample of shipbuilding contracts, between August 1998 and August 2012. It reports the order and delivery dates, the shipyard, and price in million US dollars. Table 1 reports statistics on new-building prices, showing that Chinese ships are on average 10% cheaper than Japanese both before and after 2006, while all ship prices increased significantly in the post period. It is key to note that this sample is far

Price new-building (million $)	pre- 2006	post- 2006
All	19.21	30.67
	(6.35)	(6.68)
Chinese	18.43	27.67
	(6.8)	(5.51)
European	n.a.	24
	n.a.	(0)
Japanese	20.7	29.8
	(5.26)	(7.9)
South Korean	19.8	34.85
	(6.2)	(5.77)

Table 1: New-building price summary statistics.

from ideal, as prices are reported for only a fraction of contracts. Indeed, in Figure 3 we plot the average reported new ship price per country and quarter and observe that several quarters, especially in the pre-2006 period involve missing prices. Another issue is that for quarter-shipyard combinations that involve zero production such prices do not exist (yet are necessary to compute shipyard optimal policies). To deal with these issues, we introduce a dataset of second-hand ship sale transactions, between August 1998 and August 2012. The dataset reports the date of the transaction, the name and age of the ship, as well as the price in million US dollars. We have a total of 2434 observations of new-building and second-hand sale contracts, of which 1173 are pre-2006 and 1261 are post-2006. Second-had sales are skewed towards Japanese built ships. In contrast, new ship orders are skewed towards China. The average age of used ships sold is 18.5 years

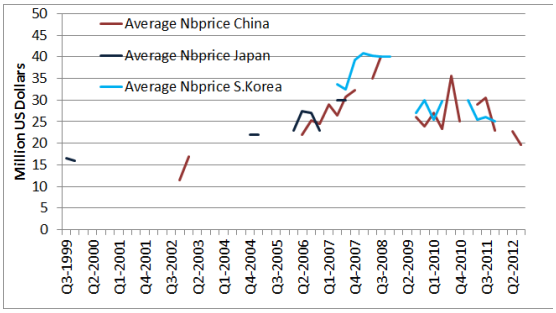


Figure 3: Reported Ship Newbuilding Prices.

old (with a standard deviation of 8).

We match the production data to the fourth dataset employed, which involves shipyard characteristics. First, we observe a snapshot of shipyard characteristics in 2013 which reports: each shipyard's first year of delivery, location, number of dry docks and berths, length of its largest dock, number of employees, total past output and total TEU (i.e. container ships) produced. The first year of delivery is used to compute the shipyard's age.[11] The number of docks and berths are a crude measure of capacity, since production bottlenecks occur during the assembly operations done on the docks/berths. The length of a dock determines the size of the ships built and it is a proxy not only for capacity, but also for overall productivity, since bigger bulk carriers are more complicated versions of smaller ones. Similarly, a shipyard that builds containers is more likely to be overall more efficient (when looking at shipyards that produce Handysize vessels only 5% also produces containers). There are several missing observations in the characteristics dataset and in our empirical analysis we will not be able to incorporate all characteristics simultaneously; we do, however, perform several robustness exercises. Finally, we would like to allow the infrastructure of yards (i.e. docks/berths and length) to be different before and after 2006. To do so, we employ Clarksons's monthly "World Shipyard Monitor" which reports the number of docks, berths and largest dock length for the largest shipyards (about 150 per month) beginning in 2001. We use this information for the before 2006 level and the 2013 snapshot for the after 2006 level. Table 1 reports some summary statistics of shipbuilding capital infrastructure and exhibits China's post 2006 explosion.[12]

[11] Some shipyards took orders before having ever delivered ("greenfields") during the 2007 boom, implying negative shipyard age. We therefore subtract 6 years from every first delivery year of all shipyards, after consulting with Clarksons's analysts.

[12] We do not create a quarterly measure of capital infrastructure (docks, berths, largest dock) for several reasons. First, this changes extremely slowly (and not much outside of China). Second, the information retrieved from the World Shipyard Monitor is rather noisy: there are several missing observations across quarters, the matching between shipyards in the several datasets is sometimes difficult, numbers may

	Average Docks/Berths		**Average Length of Largest Dock**	
	pre-2006	post-2006	pre-2006	post-2006
All	1.065	3.01	92.9	295.6
	(0.183)	(0.3)	(11.41)	(25.51)
Chinese	0.451	3.901	60.47	439.7647
	(0.166)	(.4853)	(15.56)	(39.49)
European	2.375	1.75	91.25	125.875
	(1.224)	(.8814)	(31.66)	(77.5)
Japanese	1.606	2.45	159.5	203.94
	(0.337)	(.45)	(20.42)	(25.68)
Korean	1.25	1.938	59.625	109.9375
	(.536)	(.61)	(26.81)	(40.52)

Table 2: Capacity summary statistics.

Finally, the fifth dataset consists of quarterly time-series for the orders of new ships (i.e. entrants), deliveries, demolitions (i.e. exitors), fleet, the average age of the fleet and total backlog. We also obtain time-series of Japan's steel ship plate commodity price in dollars per ton.

Before turning to the estimation of our model, we ask whether any patterns of the raw data are consistent with the presence of subsidies. One might expect that there should be a short-lived drop in new ship prices in 2006. Unfortunately, as observed in Figure 3, we don't have enough new ship prices to look for such a drop (indeed, there is a gap right on that period). Used ship prices, however, should exhibit the same feature. We, therefore, run a hedonic regression of second hand prices on ship characteristics (i.e. age and country of built) and quarter dummies. Figure 4 shows that indeed there is a short-lived drop in 2006.[13] Of course this finding is not proof of production subsidization; even the announcement of the capital infrastructure subsidization should lead to a temporary drop in prices since shipowners now expect higher competition in the future. Yet if no drop were observed, one may have been concerned about the impact of this policy.

be fluctuating (or even decreasing) out of obvious measurement error. The pre-2006 snapshot we create overcomes these issues.

[13] We have unsuccessfully searched extensively in industry magazines for alternative explanations.

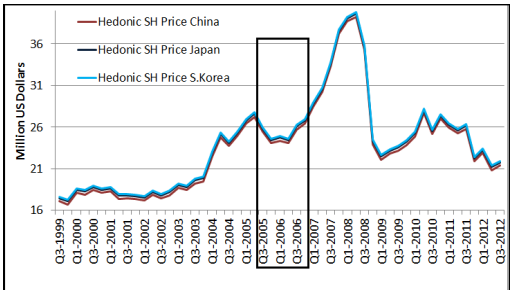


Figure 4: Hedonic regression of second-hand ship prices on ship characteristics (age, country) and quarter dummies.

5 Model Estimation and Detection of Subsidies

To see the main idea of our subsidy detection method, consider a static, perfectly competitive shipbuilding industry, so that $P_{jt}^{NB} = MC_{jt}$ for all j and t. In that case, to detect subsidies we would simply look for a break in 2006 in observed prices which are in fact the marginal cost. There are two complications in our setup: (i) we do not observe enough prices of new ships, and (ii) there are dynamics in the production decision. To address (i), we complement with used ship prices; to address (ii) we use the shipyard's first order condition from its dynamic optimization. The proposed strategy proceeds in two steps.

In the first step, we recover the demand curve that shipbuilders face, which in this case coincides with the value that shipowners place on entering the shipping industry. Retrieving this willingness to pay for a new ship amounts to estimating the value function for a new ship and shipowner expectations. To do so, we treat prices of new and used ships as observations of the expected value function and the value function respectively. In this step we also estimate state transitions.

The second step inserts the estimated willingness to pay for a ship into the optimization problem of shipbuilders to recover their costs. At an intuitive level, this step follows the standard technique of the empirical IO literature (e.g. Berry, Levinshon and Pakes (1995)) where production costs are identified via demand estimation. In this setup, however, the dynamic nature of both demand and supply requires a different approach. To tackle this problem, we combine and extend ideas from the recent literature on the estimation of dynamic games.

5.1 Estimation of the Willingness to Pay for a New Ship

In this step, we estimate ship value functions and state transitions. All ship states are directly observed in the data except for the demand for shipping services, d_t. We construct d_t as in Kalouptsidi (2013) by estimating a demand curve for shipping services and using the intercept. We replicate the analysis in the Appendix, for completeness. Each estimation task is described below and followed by the results. All results presented are for Handysize vessels.

5.1.1 State Transitions

In order to compute the value of entering the shipping industry, defined in (6), we need shipowner expectations over (s_{it}, s_t). The transition of s_{it} is known (age evolves deterministically, while country of built is time invariant). The transition of s_t is computationally complex: on one hand the dimension of the state space is enormous (S_t has dimension $4A$ -where A is a ship's maximum age- in the case of four countries, while b_t has dimension $J\overline{T}$ which in our sample is in the order of several thousand); on the other hand, we have to predict optimal production policies for all shipyards to update b_t. Instead of working with the true transitions (as in Kalouptsidi (2013)) we follow Barwick Jia and Pathak (2012) who take a more flexible approach and we assume that s_t follows a VAR(1) model. This approach is equivalent to the first step of two-step estimation procedures for dynamic games (e.g. Bajari, Benkard and Levin (2007) and Pakes, Ostrovsky and Berry (2007)).

To deal with the state dimension, we make the following simplifying assumptions. First, we replace S_t with two age groups (S_t^1, S_t^2): the number of ships below 20 years old and the number of ships above 20 years old. We do not use the distribution of the fleet over country of built because its evolution is extremely slow due to time to build and it remains practically flat for a big part of our sample. In addition, we replace the matrix b_t with the total backlog $B_t = \sum_{j,l} b_{jl}$.[14]

We have experimented with several variations of the general VAR model:

$$s_t = C_t + R_t s_{t-1} + \xi_t$$

where $\xi_t \sim N(0, \Sigma)$. We allow the VAR parameters (C_t, R_t) to be different before and after 2006: since state transitions are not modeled explicitly, the VAR model embraces equilibrium features of agents' expectations that are likely to change after China's in-

[14] In principle, we would include the distribution of shipyards over y_{jt} as well. Maintaining computational tractability does not allow for such a large state space.

tervention. In particular, since post 2006 shipbuilding capital infrastructure increases, shipowners know that all else equal the supply of ships has permanently increased. This change affects their ship valuations and therefore captures any changes in demand for new ships, brought by China's policies.

We examined several specifications where (C_t, R_t) vary deterministically (e.g. time trend) or randomly with time (random walk model for R_t where it is determined by the Kalman filter), or are time-invariant. Our baseline specification is:

$$\begin{bmatrix} S_t^1 \\ S_t^2 \\ B_t \\ d_t \\ l_t \end{bmatrix} = \begin{bmatrix} c^{S_1} \\ c^{S_2} \\ c^{B} \\ c^{d} \\ c^{l} \end{bmatrix} 1\{t \leq 2006\} + \begin{bmatrix} c^{S_1\prime} \\ c^{S_2\prime} \\ c^{B\prime} \\ c^{d} \\ c^{l} \end{bmatrix} 1\{t > 2006\} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ a^{d} \\ a^{l} \end{bmatrix} t + \begin{bmatrix} \rho^{S_1S_1} & \rho^{S_1S_2} & \rho^{S_1B} & \rho^{S_1d} & \rho^{S_1l} \\ \rho^{S_2S_1} & \rho^{S_2S_2} & \rho^{S_2B} & \rho^{S_2d} & \rho^{S_2l} \\ \rho^{BS_1} & \rho^{BS_2} & \rho^{BB} & \rho^{Bd} & \rho^{Bl} \\ 0 & 0 & 0 & \rho^{d} & 0 \\ 0 & 0 & 0 & 0 & \rho^{l} \end{bmatrix} s_{t-1} + \xi_t \tag{14}$$

and Σ is diagonal. Note that as discussed above, d_t and l_t are exogenous to the model. In contrast, we allow (S_t^1, S_t^2, B_t) to be affected by all variables to account for ship entry and exit. Our baseline specification allows only C to change before and after 2006. Note that we never allow the d_t and l_t processes to change in 2006 since they are exogenously evolving variables. Even though t appears explicitly only in the exogenous variables, it affects (S_t^1, S_t^2, B_t) through their dependence on (d_t, l_t). We estimate the parameters of interest (C, R, Σ) via OLS separately for each variable (note that separate OLS yields identical estimates to Maximum Likelihood estimation) and work with natural logarithms for (S_t, B_t). Table 3 reports the results. All variables are persistent (i.e. diagonal elements of R are positive). Signs are also in general as expected: S_1 is increasing in the backlog and demand and decreasing in steel prices (as steel prices increase, exit increases and the fleet falls); S_2 is decreasing in S_1 as more young ships increase exit and increasing in demand which leads to less exit; the backlog is increasing in demand. Demand and steel prices are strongly persistent. All eigenvalues of R lie inside the unit circle so that the model is stationary conditional on the trend. Finally, the post-2006 world's steady state has significantly higher fleet.

We also experimented heavily with restrictions on C and R both in terms of before and after 2006, allowing Σ to be full, as well as setting some parameters equal to zero (e.g. time to build might imply that $\rho^{S_1 d} = 0$ ignoring ship exit). We have also employed LASSO in a model where all parameters can change in 2006 to choose the relevant terms. Finally, we have allowed d_t to be an $AR(2)$. Our main findings are in general robust to many

c^d	a^d	ρ^d	c^l	a^l	ρ^l
0.4118	0.0077	0.688	0.695	0.0209	0.796
$(0.1656)^*$	(0.0051)	$(0.113)^*$	(0.395)	(0.015)	$(0.091)^*$

	c^{S_1}	c^{S_2}	c^B
pre-2006	-1.47	2.53	-8.49
	(0.97)	$(1.03)^*$	(7.31)
pre-2006	-1.46	2.525	-8.503
	(0.976)	$(1.037)^*$	(7.348)

$\rho^{S_1S_1}$	$\rho^{S_1S_2}$	ρ^{S_1B}	ρ^{S_1d}	ρ^{S_1l}
1.104	0.0802	0.021	0.0046	-0.003
$(0.04)^*$	(0.094)	$(0.0044)^*$	(0.0035)	$(0.00097)^*$
$\rho^{S_2S_1}$	$\rho^{S_2S_2}$	ρ^{S_2B}	ρ^{S_2d}	ρ^{S_2l}
-0.146	0.806	-0.0068	0.0041	0.0022
$(0.043)^*$	$(0.1)^*$	(0.0047)	$(0.0037)^*$	$(0.001)^*$
ρ^{BS_1}	ρ^{BS_2}	ρ^{BB}	ρ^{Bd}	ρ^{Bl}
0.158	1.066	0.87	0.0754	-0.0094
(0.303)	$(0.707)^*$	$(0.033)^*$	(0.026)	(0.0073)

σ^{S_1}	σ^{S_2}	σ^B	σ^b	σ^d
0.0001	0.0001	0.004	0.141	1.225

Table 3: VAR parameter estimates. Stars indicate significance at the 0.05 level.

of these experiments. Our chosen specification combines the following desired properties: it is parsimonious, stationary (conditional on the trend), it includes time explicitly and entering through the exogenous variables and it takes into account the 2006 break.

5.1.2 Ship Value Function

The main object entering the willingness to pay for a new ship in (6), is the ship's value function. In order to estimate it, we treat prices of new and used ships as observations of the value of entry and the value function respectively. In particular, under the assumption of a large number of identical potential entrant shipowners, ship prices are bid up to valuations. The empirical versions of the equilibrium conditions (2) and (4) are:

$$P_{jt}^{NB} = E\left[\beta^{T_{jt}} V\left(s_{it+T_{jt},s_{t+T_{jt}}}\right)|s_{it}, s_t\right] + \zeta^{nb} \tag{15}$$

$$P_{it}^{SH} = V\left(s_{it}, s_t\right) + \zeta^{sh} \tag{16}$$

where ζ^{sh} and ζ^{nb} are measurement error. Kalouptsidi (2013) employes used ship prices alone to nonparametrically estimate ship value functions and provides an extensive discussion on the merits and caveats of this approach, as well as direct and suggestive evidence against worries of sample selection. To this approach we add here the new-building contracts dataset and in order to combine (16) and (15) in a single estimation step we follow a different methodology.[15] In particular, we adopt a parametric technique using a flexible linear sieve approximation for the value function:

$$V\left(s_{it}, s_t\right) = \gamma f\left(s_{it}, s_t\right)$$

where $f\left(\cdot\right)$ is a polynomial function in $\left(s_{it}, s_t\right)$ and γ is a (sparse) vector. We can estimate the parameter vector γ from:

$$P_{jt}^{NB} = \beta^{T_{jt}} \gamma E\left[f\left(s_{it+T_{jt}}, s_{t+T_{jt}}\right) | s_{it}, s_t\right] = \tag{17}$$

$$= \beta^{T_{jt}} \int \gamma f\left(s_{it+T_{jt}}, s_{t+T_{jt}}\right) dP\left(s_{it+T_{jt}}, s_{t+T_{jt}} | s_{it}, s_t\right) \equiv f_{NB}\left(s_{it}, s_t\right)$$

$$P_{it}^{SH} = \gamma f\left(s_{it}, s_t\right) \tag{18}$$

where $P\left(s_{it+1}, s_{t+1} | s_{it}, s_t\right)$ is the state transition and is given by the VAR estimated above. The parameters γ enter (17) and (18) linearly; yet even though (18) can be estimated in a straightforward manner, (17) requires the computation of the right-hand side integrals, which becomes complicated. Indeed, (17) involves the expectation of higher order terms of the following vector:

$$s_{t+T} = R^T s_t + \sum_{k=t+1}^{t+T} R^{t+T-k}(C_k + ak + \xi_k) \tag{19}$$

We derive closed-form expressions for the integrals up to third order terms in s_t in the Appendix.

As the dimensionality of $\left(s_{it}, s_t\right)$ is large, computing high order polynomial terms quickly leads to a very large number of regressors in (17) and (18). We therefore use the LASSO, a method appropriate for sparse regression problems, i.e. for problems with a

[15] An alternative approach would be to work only with estimated expectations and second-hand prices to estimate the value function and then use new ship prices for external validation. We undertook this task (albeit using only an average newbuilding price) in Kalouptsidi (2013) and showed that the three objects are indeed consistent.

large number of potential regressors but where only a small subset of them is important in capturing the regression function accurately. LASSO identifies the relevant regressors by performing a modified OLS procedure which penalizes a large number of nonzero coefficients through regularization by a penalty based on the $\mathcal{L}_1$ norm, so that:

$$\min_{\gamma} \left\{ \sum_{j,t} \left(P_{jt}^{NB} - f_{NB}\left(s_{it}, s_t\right)' \gamma \right)^2 + \sum_{i,t} \left(P_{it}^{SH} - f\left(s_{it}, s_t\right)' \gamma \right)^2 + \lambda \left|\gamma\right|_1 \right\}$$

In our application, the regressors are products of monomials of the states in (s_{it}, s_t) and in particular, third order polynomials in s_t, time t (note t explicitly appears in (14) and is thus part of the state), the ship's age and a dummy variable capturing the country of built, as well as first order interactions between s_{it} and s_t. The discount factor is set to 0.9877 which corresponds to 5% annual interest rate.

The flexible nature of this empirical approach implies that the parameters γ embody equilibrium features which are likely to change in 2006 as agents' expectations and valuations are altered. Therefore, in analogy with the VAR formulation, we allow the value function to change before and after 2006, by adding all monomials multiplied by a post-2006 dummy variable. Differences in ship value functions pre and post 2006 capture changes in ship demand that China's entry may have created. For example, China's capital and/or production subsidies may have led potential shipowners to expect a large increase in the fleet in the coming years, thus reducing the price of ships today. Figure 5 depicts the estimated value function on the observed states for zero year old ships (the relevant value function for the value of entry). Consistent with the raw data, Chinese ships are of lower value, with Japanese and South Korean ships being of higher value.[16]

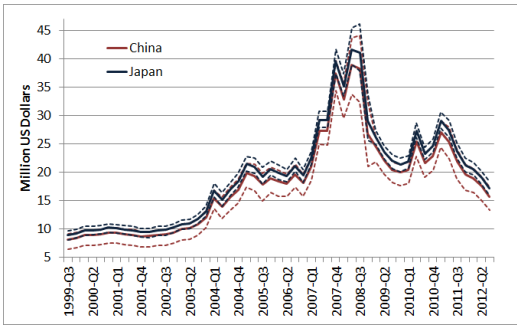


Figure 5: Estimated value function of a 0 years old ship. 0.95 bootstrap confidence intervals.

[16] Pointwise confidence intervals are computed via bootstrap samples, with the resampling done on the error.

5.2 Shipbuilding Production Cost Function

We next turn to our goal of estimating the cost function of shipbuilders and testing whether the cost function of Chinese yards changed after 2006. In particular, we assume that the cost function is parameterized by θ, $c\left(N_{jt}, y_{jt}, l_t; \theta\right)$, and our goal is to estimate θ, as well as the variance σ of the shocks ε_{jt}. We begin by describing our empirical strategy and then present our estimation results, starting with the case of static shipbuilders and then proceeding to dynamic shipbuilders.

5.2.1 The Empirical Approach

To estimate the cost function parameters, we maximize the following likelihood function, derived in the Appendix:

$$\prod_{j,t:N_{jt}=0} \Pr\left(N_{jt}=0|y_{jt},s_t;\theta\right) \prod_{j,t:N_{jt}=\overline{N}} \Pr\left(N_{jt}=\overline{N}|y_{jt},s_t;\theta\right) \prod_{n} \prod_{j,t:N_{jt}=n} \Pr\left(N_{jt}=n|y_{jt},s_t;\theta\right) \tag{20}$$

where the choice probabilities are computed via the optimal policy (13):

$$\Pr\left(N_{jt}=0|y_{jt},s_t;\theta\right) = 1-\Phi\left(\frac{1}{\sigma}\left[\begin{array}{c} VE_j\left(s_t\right)+c\left(0,y_{jt},s_t;\theta\right)-c\left(1,y_{jt},s_t;\theta\right)+ \\ +\beta\left(Q\left(y_{jt},s_t,1\right)-Q\left(y_{jt},s_t,0\right)\right) \end{array}\right]\right) \tag{21}$$

$$\Pr\left(N_{jt}=\overline{N}|y_{jt},s_t;\theta\right) = \Phi\left(\frac{1}{\sigma}\left[\begin{array}{c} VE_j\left(s_t\right)+c(\overline{N}-1,y_{jt},s_t;\theta)-c\left(\overline{N},y_{jt},s_t;\theta\right)+ \\ +\beta\left(Q\left(y_{jt},s_t,\overline{N}\right)-Q\left(y_{jt},s_t,\overline{N}-1\right)\right) \end{array}\right]\right)$$

$$\Pr\left(N_{jt}=n|y_{jt},s_t;\theta\right) = \Phi\left(\frac{1}{\sigma}\left[\begin{array}{c} VE_j\left(s_t\right)+c\left(n-1,y_{jt},s_t;\theta\right)-c\left(n,y_{jt},s_t;\theta\right)+ \\ +\beta\left(Q\left(y_{jt},s_t,n\right)-Q\left(y_{jt},s_t,n-1\right)\right) \end{array}\right]\right) -$$
$$-\Phi\left(\frac{1}{\sigma}\left[\begin{array}{c} VE_j\left(s_t\right)+c\left(n,y_{jt},s_t;\theta\right)-c\left(n+1,y_{jt},s_t;\theta\right)+ \\ +\beta\left(Q\left(y_{jt},s_t,n+1\right)-Q\left(y_{jt},s_t,n\right)\right) \end{array}\right]\right)$$

Maximizing this likelihood function would be trivial if the continuation value $Q\left(y_{jt},s_t,n\right)$ were known. This is the standard difficulty of estimating dynamic setups and to address it, we adopt a hybrid approach based on the recent literature on estimation of dynamic setups. In particular, we recover the shipyard's optimal policy $N^*\left(y_{jt},s_t,\varepsilon_{jt}\right)$ nonparametrically using choice probabilities, in analogy to the Hotz and Miller (1993) inversion

and the first stage of Bajari, Benkard and Levin (2007). We then use it to obtain ex ante optimal per period payoffs in closed-form and recover a parametric approximation to the shipyard's value function.

Let

$$A\left(y_{jt}, s_t, n\right) \equiv \frac{1}{\sigma}\left[VE_j\left(s_t\right) + \left(c\left(n, y_{jt}, s_t\right) - c\left(n+1, y_{jt}, s_t\right)\right) + \beta\left(Q\left(y_{jt}, s_t, n+1\right) - Q\left(y_{jt}, s_t, n\right)\right)\right]$$

for $n = 0, 1, ..., \overline{N} - 1$. We rewrite the choice probabilities (21) as follows:

$$\Pr\left(N^* = 0 | y_{jt}, s_t\right) \equiv p_0\left(y_{jt}, s_t\right) = \Pr\left(\varepsilon \geq A\left(y_{jt}, s_t, 0\right)\right) \tag{22}$$

$$\Pr\left(N^* = n | y_{jt}, s_t\right) \equiv p_n\left(y_{jt}, s_t\right) = \Pr\left(\varepsilon \leq A\left(y_{jt}, s_t, n-1\right)\right) - \Pr\left(\varepsilon \leq A\left(y_{jt}, s_t, n\right)\right)$$

$$\Pr\left(N^* = \overline{N} | y_{jt}, s_t\right) \equiv p_{\overline{N}}\left(y_{jt}, s_t\right) = \Pr\left(\varepsilon \leq A\left(y_{jt}, s_t, \overline{N} - 1\right)\right)$$

The function $A\left(y_{jt}, s_t, n\right)$ can be recovered in a straightforward manner from the observed choice probabilities $\left\{p_n\left(y_{jt}, s_t\right)\right\}_{n=0}^{\overline{N}}$. Indeed, it is easy to show that[17]

$$A\left(y_{jt}, s_t, n\right) = \Phi^{-1}\left(1 - \sum_{k=0}^{n} p_k\left(y_{jt}, s_t\right)\right), \quad \text{for } n = 0, 1, ..., \overline{N} - 1 \tag{23}$$

Clearly, $A\left(y_{jt}, s_t, n\right)$ is (weakly) decreasing in n. Most important, if $A\left(y_{jt}, s_t, n\right)$ is known, so is the optimal policy: For any $\left(y_j, s, \varepsilon\right)$,

$$N^*\left(y_j, s, \varepsilon\right) = \widehat{n}, \text{ such that } \varepsilon \in \left[A\left(y_j, s, \widehat{n}\right), A\left(y_j, s, \widehat{n} - 1\right)\right]$$

Once the optimal policy is known, we can recover the value function. Indeed, consider shipyard j's Bellman equation (7) which we repeat here for convenience (to ease notation we set $x = \left(y_{jt}, s_t\right)$ and $x' = \left(y_{jt+1}, s_{t+1}\right)$ and suppress (j, t)):

$$W\left(x, \varepsilon\right) = \max_{N \in \left\{0, 1, ..., \overline{N}\right\}} VE\left(x\right) N - c\left(N, x\right) - N\varepsilon + \beta E_{\varepsilon', x'}\left[W\left(x', \varepsilon'\right) | N, x\right]$$

where as a reminder,

$$E_{\varepsilon', x'}\left[W\left(x', \varepsilon'\right) | N, x\right] = Q\left(x, N\right)$$

[17] To show this, begin with $p_0\left(y_{jt}, s_t\right) = 1 - \Phi\left(A\left(y_{jt}, s_t, 0\right)\right)$, so that $A\left(y_{jt}, s_t, 0\right) = \Phi^{-1}\left(1 - p_0\left(y_{jt}, s_t\right)\right)$. Next, $p_1\left(y_{jt}, s_t\right) = \Phi\left(A\left(y_{jt}, s_t, 0\right)\right) - \Phi\left(A\left(y_{jt}, s_t, 1\right)\right) = 1 - p_o\left(y_{jt}, s_t\right) - \Phi\left(A\left(y_{jt}, s_t, 1\right)\right)$, so that $A\left(y_{jt}, s_t, 1\right) = \Phi^{-1}\left(1 - p_0\left(y_{jt}, s_t\right) - p_1\left(y_{jt}, s_t\right)\right)$. The general case follows by induction.

If we use the optimal policy $N^*(x,\varepsilon)$ the value function becomes:

$$W(x,\varepsilon) = VE(x)N^*(x,\varepsilon) - c(N^*(x,\varepsilon),x) - N^*(x,\varepsilon)\varepsilon + \beta E_{\varepsilon',x'}[W(x',\varepsilon')|N^*(x,\varepsilon),x]$$

and the ex ante value function

$$E_\varepsilon W(x,\varepsilon) \equiv W(x) = E_\varepsilon[\pi(x,N^*(x,\varepsilon)) + \beta E_{\varepsilon',x'}[W(x',\varepsilon')|N^*(x,\varepsilon),x]]$$

where

$$\pi(x,N^*(x,\varepsilon)) = E_\varepsilon[VE(x)N^*(x,\varepsilon) - c(N^*(x,\varepsilon),x) - N^*(x,\varepsilon)\varepsilon] \tag{24}$$

is the ex ante per period profit. If $\pi(x,N^*(x,\varepsilon))$ is known then we can solve for the ex ante value function from the following relationship:

$$W(x) = E_\varepsilon\pi(x,N^*(x,\varepsilon)) + \beta E_{\varepsilon,x'}[W(x')|N^*(x,\varepsilon),x] \tag{25}$$

Solving (25) can be done in several ways, such as state space discretization and matrix inversion, or parametric approximation; we opt for the latter. In particular, we approximate the value function by a polynomial function, so that:

$$W(x) = \gamma f(x)$$

then (25) becomes

$$\gamma f(x) = E_\varepsilon\pi(x,N^*(x,\varepsilon)) + \gamma\beta E_{\varepsilon,x'}[f(x')|N^*(x,\varepsilon),x]$$

or

$$(f(x) - \beta E[f(x')|N^*(x,\varepsilon),x])\gamma = E_\varepsilon\pi(x,N^*(x,\varepsilon)) \tag{26}$$

and we can therefore estimate γ via LASSO. We now only need to show how $E_\varepsilon\pi(x,N^*(x,\varepsilon))$ is computed.

We assume that the shipbuilding cost function takes the following form:

$$C(N,x) = c_1(x;\theta)N + c_2(x;\theta)N^2 + \sigma\varepsilon N$$

where $\varepsilon_{jt} \sim N(0,1)$ and $c_2(x;\theta) > 0$. Therefore,

$$\begin{aligned}\pi(x, N^*(x,\varepsilon)) &= E_\varepsilon\left[VE(x) N^*(x,\varepsilon) - c_1(x;\theta) N^*(x,\varepsilon) + c_2(x;\theta) N^*(x,\varepsilon)^2 - \sigma N^*(x,\varepsilon)\varepsilon\right] \\ &= (VE(x) - c_1(x;\theta)) E_\varepsilon N^*(x,\varepsilon) + c_2(x;\theta) E_\varepsilon N^*(x,\varepsilon)^2 - \sigma E_\varepsilon\left[N^*(x,\varepsilon)\varepsilon\right]\end{aligned}$$

We show in the Appendix that

$$E_\varepsilon N^*(x,\varepsilon) = \sum_{n=0}^{\overline{N}-1} \Phi(A(x,n)) \tag{27}$$

$$E_\varepsilon N^*(x,\varepsilon)^2 = 2\sum_{n=1}^{\overline{N}} n\Phi(A(x,n-1)) - \sum_{n=0}^{\overline{N}-1} \Phi(A(x,n)) \tag{28}$$

$$E_\varepsilon\left[N^*(x,\varepsilon)\varepsilon\right] = -\sum_{n=0}^{\overline{N}-1} \phi(A(x,n)) \tag{29}$$

To sum up, our estimation proceeds as follows:

1. We estimate $A(x,n)$ using (23)
2. We compute the statistics of the optimal production in (27), (28) and (29)
3. At each guess of the parameters (θ,σ) in the optimization of the likelihood (20):
   (a) We solve for the approximate value function parameters γ from (26)
   (b) Using γ we compute choice probabilities in the likelihood and update (θ,σ).

We provide further details in the Appendix.

5.2.2 Results

Our baseline specifications involve

$$\begin{aligned}c_1(y_{jt}, s_t;\theta) &= \theta_0^{ch} 1\{\text{China}\} + \theta_0^{ch,post} 1\{t \geq 2006, \text{China}\} \\ &\quad + \theta_0^{EU} 1\{\text{Europe}\} + \theta_0^{J} 1\{\text{Japan}\} + \theta_0^{K} 1\{\text{S.Korea}\} + \theta_1 g(y_{jt}, s_t, t)\end{aligned}$$

and

$$c_2(y_{jt}, s_t;\theta) = c_2$$

where $g(y_{jt}, s_t, t)$ is a polynomial in (y_{jt}, s_t, t). Testing that $\theta_0^{ch,post} \neq 0$ in the above functional form provides evidence of a structural change in China's cost function, for any value of y and N. We report results for the shipbuilding cost function for several specifications in terms of shipyard characteristics included in y_{jt}, the nature of $c_1(y_{jt}, s_t; \theta)$ and $c_2(y_{jt}, s_t; \theta)$, as well as sample cuts. We begin with the case where production is a static choice.

Static Shipbuilders If shipyard j is myopic it solves:

$$\max_{N_{jt} \in \{0,1,\ldots,\overline{N}\}} VE_j(s_t) N_{jt} - \left(c_1(y_{jt}, s_t; \theta) N_{jt} + c_2(y_{jt}, s_t; \theta) N_{jt}^2 + \sigma \varepsilon_{jt} N_{jt}\right)$$

and the continuation value is removed from the decision rule (21). This is essentially an ordered choice problem. We follow Amemiya (1984) and maximize the likelihood over $\left(\frac{1}{\sigma}, \frac{\theta}{\sigma}\right)$ rather than (θ, σ).

Table 7 reports results from some baseline specifications. The number of observations varies depending on the covariates: in specifications that include capital infrastructure (docks plus berths, length of largest dock) we drop observations with missing values.

In all specifications there is a strongly significant decline in China's cost before and after 2006 in the order of 15-20%, as indicated by the significant China-POST dummy. The results suggest that there is significant convexity in costs. Costs are decreasing in capital measures, as expected. The shipyard's age is included to capture learning by doing, which has been documented in military ships (Thompson (2001)). Backlog is negative, even conditional on age, implying cost declines due to economies of scale or expertise. This finding is consistent with industry participants' testimony, who claim shipyards have incentives to produce more ships similar to those they already have under construction. Most specifications imply, not surprisingly, that Europe is the highest cost producer. Either Japan or China post 2006 are the lowest cost producer depending on the specification.

Controlling for time-varying factors is important since one might worry that a decline in costs might be due to an increase in (unobserved) productivity rather than subsidies. To address this concern, we use several functions of time and rely on the identifying assumption that China's policy was a discrete change, while other factors vary continuously. Table 7 presents some of the specifications we have tried with respect to time. Results are robust to several functions of t. We only present a limited number of specifications because of space limitations, but one can essentially add any parametric function of t,

such as polynomial trends. We also use year dummies instead of simple functions of t in which case the estimated subsidies, not surprisingly, are somewhat lower.

We experiment by adding several covariates such as: the total TEU that a yard has produced, the yard's total past production (capturing experience), dummy variables for young ages to capture learning by doing somewhat more flexibly, the administrative region, the number of employees (only a subset of observations includes this variable) and the year the shipyard was founded (rather than its age). Results (available upon request) are robust to specifications that include these variables. We also try a "placebo" estimation where we test whether Japan's costs (rather than China's) are different before and after 2006. We overall find that Japan's costs experience a small increase, which is not significant across specifications.

We also try the case where c_2 changes before and after 2006 for Chinese yards (i.e. $c_2 = \theta_0^{ch} 1\{\text{China}\} + \theta_0^{ch,post} 1\{t \geq 2006, \text{China}\} + \theta_0 \{\text{not China}\}$) and find statistically significant decrease in Chinese costs. If both c_2 and c_1 are allowed to change in the above fashion, c_2 declines, while c_1 shows a small increase; yet overall costs fall after 2006 for all y, N. We also make c_2 a function of the backlog to capture further cost convexities; results are robust (it is not straightforward to make c_2 a function of numerous covariates simultaneously, since we need to guarantee $c_2(y_{jt}, l_t; \theta) > 0$ for lemma 1 to hold and the likelihood to be correctly specified).

Next, we turn to some further robustness exercises to test several features of our assumptions and estimation procedure. We begin by estimating costs on the following subsample: we keep only yards that exist in the beginning of our sample; in other words we remove all entrants, most of whom are Chinese and appear post 2006. This robustness check can potentially address the concern that the documented cost declines above are driven by new yards that for some reason are more productive (perhaps because in 2006 China built modern yards very different from the existing ones). In addition, new yards might experience learning by doing, which would lead to initially higher but later lower overall costs. Older shipyards should already be down their learning curve (partly the reason we focus on small bulk carriers is precisely that Chinese shipyards held significant market share before 2006). Table 8 reports the results and shows for two specifications that existing Chinese yards experienced cost declines after 2006. We also repeat estimation assuming that $N_{jt} \geq 0$ so that we have a tobit model, rather than an ordered choice model. Results are both qualitatively and quantitatively robust. Finally, we consider the impact of the yards' administrative regions and allow different regions to implement government plans at different times. To do so we divide regions into three groups based

on when we observe docks coming online regionally. As shown in Table 9, we find that shipyards in the group which implemented in Q3-2005 (Jiangsu region) experience the largest cost declines, shipyards in the group that initiated in Q4-2006 (Fujian, Hainan, Hebei, Hubei, Shandong, Tianjin regions) experience intermediate cost declines, while shipyards in the group that initiated in Q3-2007 (Anhui, Guangdong, Guizhou, Liaoning, Shanghai, Zhejiang regions) experience the lowest cost declines; all declines are significant.

One may be concerned that the estimated cost declines are solely driven by the inherent discontinuity in the estimated $VE_j(s_t)$ due to the different VAR model and LASSO coefficients. To address this concern we estimate costs using the average quarterly price (across shipyards and countries) of a new ship, obtained from Clarksons. We find that estimated subsidies are significant and of the same magnitude.

Note that our model implies that the Chinese government gives the same subsidy amount to all yards. One might worry that subsidies are in fact different for each firm. Suppose instead that the government gives subsidy $x + \xi_{jt}$ to yard j, where $\xi_{jt} \sim N(0, \sigma_\xi)$ across j and t. In that case, our estimated cost parameters θ are still consistent (Wooldridge (2001)) and the estimated subsidy, $\theta_0^{ch,post}$, is the average subsidy across yards (our estimate for σ, however, is no longer consistent). More complicated models where subsidies are targeted for example to the most productive yards are more difficult to handle.

Dynamic Shipbuilders Our preliminary results of dynamic shipbuilders, suggest that Chinese costs experience significant declines post 2006. In analogy with the case of static shipbuilders, the backlog appears to decrease costs, consistent with economies of scale or the accumulation of expertise. More docks/berths, as well as longer docks decrease costs. Retrieved subsidies are again around 20%. We also compute the value function of Chinese shipyards in 2006, which equals 3.2 billion US Dollars. We can think of this amount as a rough estimate of the order of magnitude of the costs of building these shipyards, which may be close to the fixed cost subsidies of China's 2006 plan.

5.2.3 Comparison to the WTO Subsidy Detection Method

Before turning to the impact of subsidies on industrial evolution, we compare our detection approach to the so called price-gap approach, which is followed in WTO subsidy cases. The price-gap approach, compares product end-user prices to reference prices (i.e. prices that would prevail in markets without subsidies); yet the latter can be tough to compute.[18]

[18] See Haley and Haley (2013) for a description of this approach and its caveats.

In the case of ships, where products are close to homogenous and there is one global market with no transportation or other common costs, our understanding of the price-gap approach is that it would essentially compare Chinese to other prices. Table 4 presents results from a hedonic regression of (the few observed) new ship prices. Our interpretation

	parameter	s.e.
constant	15	$(3.5)^{**}$
China	-3.09	$(0.52)^{**}$
Japan	-2.16	$(0.93)^{**}$
delivery lag	-2.08	$(0.61)^{**}$
quarter dummies		

Table 4: Hedonic regression of new ship prices.

of the price-gap approach is that it would detect a 7.3% subsidy (i.e. the discount of Chinese ships), less than half of our magnitude. As Haley and Haley (2013) point out, however, there would be efforts to correct for quality differences. Quality corrections are performed in a case by case basis; one thought would be to explore price differences in the second-hand market where prices may be reflecting quality differences only, rather than cost differences. Table 5 presents results from a hedonic regression of used ship prices and shows that Chinese ships are on average 3.5% cheaper in the second-hand market. Our interpretation is that the price-gap approach would have produced about 4% subsidies, which are dramatically lower from our robustly estimated 15-20%.

	parameter	s.e.
constant	21.95	$(1.36)^{**}$
China	-0.83	$(0.67)^{**}$
Europe	-1.36	$(0.59)^{**}$
Japan	-0.061	$(0.52)^{**}$
age	-0.78	$(0.015)^{**}$
quarter dummies		

Table 5: Hedonic regression of used ship prices.

6 Quantifying the Implications of Subsidies

What is the impact of government subsidies on industry prices, production reallocation across countries, profits and costs? In addition, how do different types of subsidies (e.g.

fixed vs. variable cost) affect the above? We answer these questions in the context of China's intervention in shipbuilding by using our model to predict the evolution of the industry in two counterfactual scenarios: first, no Chinese subsidies of any kind (i.e. no 2006 plan altogether); second, Chinese capital subsidies only (i.e. we remove the (prohibited) production subsidies alone).

To implement the "No China" counterfactual, we assume that shipowners maintain their pre-2006 expectations and ship value functions, while shipyards keep their pre-2006 capital structure (i.e. docks/berths and length) and costs. To implement the "Capital subsidies only" counterfactual, we assume that shipowners switch to the post-2006 expectations and value functions. In other words, we now assume that shipowners understand that a change occurred in 2006; yet they can't distinguish production vs. capital subsidies (which may be reasonable given production subsidies are secret). Shipyards keep their pre-2006 cost functions and their post-2006 capital structures. We feed the observed post-2006 values for shipping demand and steel prices into our model and simulate shipyard optimal production and ship prices. We provide details on the implementation of these counterfactuals in the Appendix.

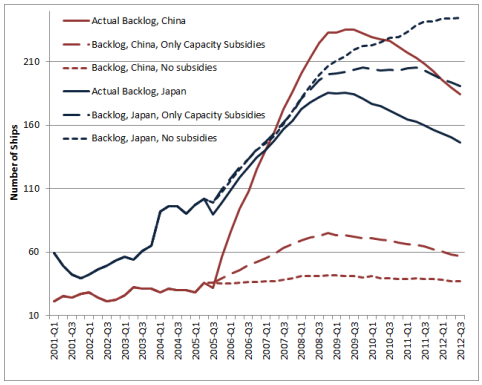


Figure 6: Counterfactual backlogs.

Our results suggest that ship prices are higher for all countries in the absence of China's subsidization plans. This is not surprising, given that China's subsidization shifted supply outward. Moreover, we predict that production and backlog would have been significantly lower had China not subsidized, as shown in Figure 6. Most interestingly, we find that production subsidies play a significant role in increasing China's production and market share. We find that Japan lost about 18 million US dollars from China's production subsidies and about 56 million US dollars from the combination of capital and production subsidies.

A further interesting feature of the post 2006 period is that demand for freight services boomed and led (at least in part) to a shipping investment boom. The crisis in 2008 led in turn to a crash. We find that China's subsidies amplified the boom and bust in shipping investment in the last decade. Indeed, without its massive increase in shipbuilding capital infrastructure, the backlog would not have increased as dramatically and as a consequence would not have crashed as bad in the 2008 crisis.

Tables

	parameter	s.e.
China	-2.62	(28.86)
Europe	-2.87	(28.857)
Japan	-2.74	(28.86)
S. Korea	-2.67	(28.86)
Backlog	0.013	$(0.0036)^*$
t	0.022	$(0.01)^*$
S_t^1	-1.94	(1.36)
S_t^2	2.73	(2.64)
B_t	-0.49	$(0.177)^*$
d_t	0.15	$(0.078)^*$
l_t	0.038	(0.023)

Table 6: Time to build regression results. Time t measured in quarters. Countries refer to country dummy variables. Stars indicate significance at 0.05 level.

	I		II		III		IV	
	Prm	s.e.	Prm	s.e.	Prm	s.e.	Prm	s.e.
China	29.95	$(1.62)^*$	29.72	$(1.6)^*$	28.4	$(1.9)^*$	27	$(1.4)^*$
China,POST	-6.58	$(0.51)^*$	-5.55	$(0.5)^*$	-3.8	$(0.6)^*$	-5.7	$(0.48)^*$
Europe	28.45	$(1.59)^*$	29.8	$(1.8)^*$	27.7	$(1.92)^*$	27.6	$(1.5)^*$
Japan	23.22	$(1.08)^*$	22.31	$(1.01)^*$	25.3	$(1.3)^*$	21.93	$(0.95)^*$
S. Korea	27.99	$(1.26)^*$	29.49	$(1.45)^*$	29	$(1.64)^*$	25.6	$(1.1)^*$
Backlog	-0.54	$(0.04)^*$	-0.51	$(0.04)^*$	-0.3	$(0.05)^*$	-0.52	$(0.04)^*$
Age			-0.017	$(0.002)^*$				
Docks/Berths							-0.12	$(0.01)^*$
Max Length							-0.0003	$\left(2e^{-5}\right)^*$
Steel price	0.316	$(0.03)^*$	0.297	$(0.03)^*$	0.86	$(0.12)^*$	0.35	$(0.03)^*$
t	0.27	$(0.02)^*$	0.21	$(0.02)^*$			0.32	$(0.02)^*$
Year Dummies	NO		NO		YES		NO	
Japan*t			0.089	$(0.02)^*$				
China*t			0.024	$(0.02)^*$				
c_2	1.003	$(0.08)^*$	0.95	$(0.07)^*$	0.55	$(0.11)^*$	1.001	$(0.08)^*$
σ	10.43	(0.83)	9.88	$(0.7)^*$	5.55	$(1.07)^*$	10.82	$(0.84)^*$
No of Obs	8324		8324		8324		4571	

Table 7: Baseline static cost function estimates. Time t measured in quarters. Countries refer to country dummy variables. Stars indicate significance at the 0.05 level.

	I		II	
	Param	s.e.	Param	s.e.
China	29.5	$(1.7)^*$	29.02	$(1.66)^*$
China,POST	-5.88	$(0.53)^*$	-4.5	$(0.46)^*$
Europe	29.2	$(1.76)^*$	30.01	$(1.81)^*$
Japan	23.58	$(1.165)^*$	23.38	$(1.12)^*$
S. Korea	28.5	$(1.36)^*$	27.77	$(1.275)^*$
Backlog	-0.68	$(0.06)^*$	-0.65	$(0.056)^*$
Docks/Berths			-0.184	$(0.015)^*$
Max Length			-0.0037	$(0.0003)^*$
Steel price	0.63	$(0.041)^*$	0.596	$(0.04)^*$
t	0.267	$(0.02)^*$	0.322	$(0.021)^*$
c_2	1.24	$(0.11)^*$	1.23	$(0.11)^*$
σ	11.96	(1.06)	12.43	(1.09)
No of Obs	5029		3350	

Table 8: Static cost function estimates with yards existing prior to 2001. Time t measured in quarters. Countries refer to country dummy variables. Stars indicate significance at the 0.05 level.

	I		II	
	Param	s.e.	Param	s.e.
Region A	26.72	$(1.19)^*$	24.75	(1.14)
Region B	25.03	$(1.15)^*$	24.12	(1.15)
Region C	22.9	$(1.08)^*$	20.76	(1.045)
Region A,POST	-5.7	$(0.41)^*$	-4.92	(0.396)
Region B,POST	-3.78	$(0.27)^*$	-5.03	(0.35)
Region C,POST	-2.8	$(0.22)^*$	-2.12	(0.21)
Europe	25.61	$(1.13)^*$	25.3	(1.16)
Japan	21.4	$(0.82)^*$	20.48	(0.78)
S. Korea	25.78	$(1.013)^*$	23.97	(0.96)
Backlog	-0.44	$(0.026)^*$	-0.45	(0.027)
Docks/Berths			-0.11	(0.006)
Max Length			-0.0003	$\left(2e^{-5}\right)$
Steel price	0.33	$(0.025)^*$	0.36	(0.025)
t	0.25	$(0.017)^*$	0.297	(0.017)
c_2	0.83	$(0.047)^*$	0.86	(0.05)
σ	8.54	$(0.491)^*$	9.25	(0.55)
No of Obs	8324		4571	

Table 9: Static cost function estimates with yards existing prior to 2001. Time t measured in quarters. Countries refer to country dummy variables. Region A includes Jiangsu and Post refers to Q3-2005. Region B includes Hebei, Shandong, Tianjin, Hainan, Fujian and Hubei and Post refers to Q4-2006. Region C includes Liaoning, Shanghai, Zhejiang, Guangdong, Anhui and Guizhou and Post refers to Q3-2007. Stars indicate significance at the 0.05 level.

7 Appendix

7.1 Proof of Lemma 1

We first show the following lemma:

Lemma 2 *If $f(x):[a,b]\rightarrow\mathbb{R}$ is convex in x, then*

$$\min_{\kappa=1,2,...}\left\{\frac{f(x+\kappa)-f(x)}{\kappa}\right\}\leq f(x+1)-f(x)$$

Proof. *It suffices to show that the sequence $\frac{f(x+\kappa)-f(x)}{\kappa}$ is decreasing in $\kappa=1,2,....$ Indeed, the inequality*

$$\frac{f(x+\kappa+1)-f(x)}{\kappa+1}\geq\frac{f(x+\kappa)-f(x)}{\kappa}$$

holds if and only if

$$\kappa f(x+\kappa+1)+f(x)\geq(\kappa+1)f(x+\kappa)\Leftrightarrow\frac{\kappa}{\kappa+1}f(x+\kappa+1)+\frac{1}{\kappa+1}f(x)\geq f(x+\kappa)$$

which holds because of Jensen's inequality, since

$$\frac{\kappa}{\kappa+1}f(x+\kappa+1)+\frac{1}{\kappa+1}f(x)\geq f\left(\frac{\kappa}{\kappa+1}(x+\kappa+1)+\frac{1}{\kappa+1}x\right)=f(x+\kappa)$$

■

Therefore, it suffices to show that the function $c(n,\cdot)-\beta Q(\cdot,n)$ is convex in n. By assumption, $c(n,\cdot)$ is convex. It then follows that $Q(\cdot,n)$ is concave. Indeed, standard dynamic programming arguments yield concavity of the value function if: the per period payoff is continuous, bounded and concave in both the state and the control (which holds in our case by convexity of the cost function), the transition function is continuous, bounded and concave in the control and state (which holds under the backlog transition chosen in our empirical exercise), $\beta\in(0,1)$, and the state spaces is convex.

7.2 Creation of shipping demand state

In this Appendix, we estimate the inverse demand for shipping services via instrumental variables regression, to create the state d_t. The analysis follows Kalouptsidi (2013) and is

reprinted here for completeness. The empirical analogue of the demand curve in (1) that we choose is:

$$P_t = \alpha_0^d + \alpha_1^d X_t^d + \alpha_2^d Q_t + \varepsilon_t^d \tag{30}$$

where P_t is the average price per voyage observed in a quarter, X_t^d includes demand shifters, while Q_t is the total number of voyage contracts realized in a quarter. X_t^d includes the index of food prices, agricultural raw material prices and minerals prices (taken from UNCTAD), the world aluminum (taken from the International Aluminum Institute) and world grain production (taken from the International Grain Council), as well as the Handymax fleet (as a potential substitute). The first stage instruments include the total fleet and its mean age. Both instruments are key determinants of industry supply capacity, as ship operating costs are convex and depend on age. Instrumentation corrects both for endogeneity, as well as measurement error (we only observe the number of trips realized, rather than ton-miles).

	1^{st} stage		2^{nd} stage	
	parameter	s.e.	parameter $*10^5$	s.e.$*10^6$
constant	2066	(2566)	7.0798	(1.5537)
food P	1.63	(2.25)	0.0189	(0.0037)
agr raw mat P	0.73	(1.55)	−0.0087	(0.0027)
mineral P	−0.097	(1.08)	0.0198	(0.0019)
aluminum prod	−0.42	(0.37)	0.0154	(0.0006)**
grain prod	−2.07	(2.9)	0.0227	(0.0045)
subst fleet	0.36	(0.49)	−0.0268	(0.0007)**
fleet	−0.67	(0.73)	−	
mean age fl	99.66	(59.9)	−	
$\widehat{Q_t}$	−	−	−0.0158	(0.0007)**

Table 10: Demand IV regression results.

Table 10 reports the results. In the second stage, the grain and aluminum production positively affect prices, while the number of voyages has the expected negative sign.[19] The impact of all shifters is lumped into the state variable d_t (the residual $\widehat{\varepsilon_t^d}$ is included in d_t as it captures omitted demand shifters):

$$d_t = \widehat{\alpha_1^d} X_t^d + \widehat{\varepsilon_t^d}$$

[19]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.

7.3 Derivation of state expectations used in ship value function

In this Appendix we derive the expressions required for the LASSO estimation of the value functions of Section 5.1.3. Remember that we approximate the value function with a polynomial function, so that:

$$V(x_t) = \gamma f(x_t) = \sum_{i=1}^{d} \gamma_i x_t^{(i)}$$

where $x_t = (s_{it}, s_t)$, and $x_t^{(i)}$ are Kronecker products, so that $x_t^{(2)} = x_t \otimes x_t$, $x_t^{(3)} = x_t^{(2)} \otimes x_t$, etc. Then, note that (17) can be written as:

$$P_{jt}^{NB} = \beta^{T_{jt}} \gamma E\left(f\left(x_{t+T_{jt}}\right)|x_t\right) = \beta^{T_{jt}} \sum_{i=1}^{d} \gamma_i E\left(x_{t+T_{jt}}^{(i)}|x_t\right) \tag{31}$$

The conditional expectation is only necessary for s_t since s_{it} evolves deterministically. We use the general VAR model (6) to get that:

$$s_{t+T} = \phi(t+T, t)\, s_t + \sum_{k=t+1}^{t+T} \phi(t+T, k)\,(C_k + \xi_k)$$

where

$$\phi(t+T, k) = \begin{cases} R_{t+T} R_{t+T-1} ... R_{k+1}, & \text{for } k < t+T \\ I, & \text{for } k = t+T \end{cases}$$

For example, for our empirical exercise where $R_t = R$, all t, we get (19) of the main text.

The above expression takes the form:

$$s_{t+T} = A + v$$

where

$$\begin{aligned} A &= \phi(t+T, t)\, s_t + \sum_{k=t+1}^{t+T} \phi(t+T, k)\, C_k \\ v &= \sum_{k=t+1}^{t+T} \phi(t+T, k)\, \xi_k \end{aligned}$$

Note that conditional on s_t, A is constant. Moreover, v is zero-mean normal with covari-

ance

$$\Sigma_v = Ev'v = \sum_{k=t+1}^{t+T} R^{t+T-k}\Sigma\left(R'\right)^{t+T-k}$$

Therefore, (31) becomes:

$$P_{jt}^{NB} = \beta^{T_{jt}} \sum_{i=1}^{d} \gamma_i E\left(\left(A+v\right)^{(i)}|s_t\right)$$

We next compute the conditional expectations for up to third order terms:

$$\begin{aligned} E\left(A+v|s_t\right) &= A \\ E\left(\left(A+v\right)^{(2)}|s_t\right) &= A^{(2)} + vec\left(\Sigma_v\right) \\ E\left(\left(A+v\right)^{(3)}|s_t\right) &= A^{(3)} + A\otimes vec\left(\Sigma_v\right) + vec\left(\Sigma_v\right)\otimes A + T_{mm^2}A\otimes vec\left(\Sigma_v\right) \end{aligned}$$

where $vec\left(x\right)$ denotes the vector formed by stacking the columns of x one after the other; given a $L\times n$ matrix A, T_{Ln} is an $Ln\times Ln$ matrix defined by $T_{Ln}vec\left(A\right) = vec\left(A'\right)$. The first of the above equations is straightforward. To prove the second, we use:

$$E\left(\left(A+v\right)^{(2)}|s_t\right) = E\left(\left(A+v\right)^{(2)}\right) = A\otimes A + A\otimes E\left(v\right) + E\left(v\right)\otimes A + Ev^{(2)}$$

It is easy to see that $Ev^{(2)} = vec\left(Evv'\right) = vec\left(\Sigma_v\right)$ using the property

$$vec\left(BXC\right) = \left(C'\otimes B\right)vec\left(X\right)$$

Finally, we prove the third order equation. Note that

$$\begin{aligned} E\left(\left(A+v\right)^{(3)}|s_t\right) &= A\otimes E\left(A+v\right)^2 + Ev\otimes\left(A+v\right)^2 \\ &= A\otimes\left(A^{(2)} + vec\left(\Sigma_v\right)\right) + E\left(v\otimes A\otimes v\right) + E\left(v\otimes v\otimes A\right) + Ev^{(3)} \end{aligned}$$

$Ev^{(3)}$ is zero since v is Gaussian. Moreover,

$$E\left(v\otimes A\otimes v\right) = T_{mm^2}\otimes Ev^{(2)}$$

Indeed, if B and C are matrices of dimensions (L,n) and (n,q) respectively, then

$$B\otimes C = T_{pl}\left(C\otimes B\right)T_{nq}$$

The matrix T_{nL} is computed as in Van Loan (1999).

7.4 Derivation of the likelihood function

We proceed as in Rust (1987). Since our data consists of sequences $\{y_{j0}, N_{j0}, y_{j1}, N_{j1}, ..., y_{jT}, N_{jT}\}$ for $j = 1, ..., J$ and T our sample size, the likelihood function is

$$L = \prod_j \Pr\left(\{N_{jt}, y_{jt}, s_t\}_{t=0}^{T}\right)$$

where we use the independence of shipyards to introduce the product over j. We can rewrite this as:

$$\begin{aligned}
L &= \prod_j \Pr\left(N_{jT}, y_{jT}, s_T | \{N_{jt}, y_{jt}, s_t\}_{t=0}^{T-1}\right) \Pr\left(\{N_{jt}, y_{jt}, s_t\}_{t=0}^{T-1}\right) \\
&= \prod_j \Pr\left(N_{jT}, y_{jT}, s_T | N_{jT-1}, y_{jT-1}, s_{T-1}\right) \Pr\left(\{N_{jt}, y_{jt}, s_t\}_{t=0}^{T-1}\right) \\
&= \prod_j \Pr\left(N_{jT} | y_{jT-1}, s_{T-1}\right) \Pr\left(y_{jT} | N_{jT-1}, y_{jT-1}\right) \Pr\left(s_T | s_{T-1}\right) \Pr\left(\{N_{jt}, y_{jt}, s_t\}_{t=0}^{T-1}\right) \\
&= \prod_j \prod_t \Pr\left(N_{jt} | y_{jt-1}, s_{t-1}\right) \Pr\left(y_{jT} | N_{jt-1}, y_{jt-1}\right) \Pr\left(s_t | s_{t-1}\right)
\end{aligned}$$

To move from the first to the second line we use the first-order Markov assumption. Moving from the second to the third line is possible because of the transitions and optimal policies (e.g. only s_t is needed to determine s_{t+1}), while to reach the last line we repeat the previous steps $T-1$ times. Finally, note that the parameters of interest appear only in the terms $\Pr\left(N_{jt} | y_{jt-1}, s_{t-1}\right)$ and we can thus ignore the state transitions when optimizing the likelihood, so that we reach the likelihood function (20).

7.5 Statistics of the Optimal Production

To derive (27) we use (22) to get:

$$\begin{aligned} E_{\varepsilon}N^{*}(x,\varepsilon) &= \sum_{n=1}^{\overline{N}} np_{n}(x) = \sum_{n=1}^{\overline{N}-1} n\left[\Phi\left(A\left(x,n-1\right)\right)-\Phi\left(A\left(x,n\right)\right)\right]+\overline{N}\Phi\left(A\left(x,\overline{N}-1\right)\right) \\ &= \sum_{n=0}^{\overline{N}-2}\left(n+1\right)\Phi\left(A\left(x,n\right)\right)-\sum_{n=1}^{\overline{N}-1} n\Phi\left(A\left(x,n\right)\right)+\overline{N}\Phi\left(A\left(x,\overline{N}-1\right)\right) \\ &= \Phi\left(A\left(x,0\right)\right)+\sum_{n=1}^{\overline{N}-2}\Phi\left(A\left(x,n\right)\right)+\Phi\left(A\left(x,\overline{N}-1\right)\right)=\sum_{n=0}^{\overline{N}-1}\Phi\left(A\left(x,n\right)\right) \end{aligned}$$

Equation (28) follows similarly. Finally, let $\phi(\varepsilon)$ denote the standard normal density. Then,

$$\int_{a}^{b}\varepsilon\phi\left(\varepsilon\right)=-\frac{1}{2\sqrt{\pi}}\int_{a}^{b}de^{-\frac{1}{2}\varepsilon^{2}}=\phi\left(a\right)-\phi\left(b\right)$$

and therefore:

$$\begin{aligned} E_{\varepsilon}\varepsilon N^{*}(x,\varepsilon) &= \int\varepsilon N^{*}\left(x,\varepsilon\right)\phi\left(\varepsilon\right)d\varepsilon= \\ &= \sum_{n=1}^{\overline{N}-1} n\int_{A(x,n)}^{A(x,n-1)}\varepsilon\phi\left(\varepsilon\right)+\overline{N}\int_{-\infty}^{A\left(x,\overline{N}-1\right)}\varepsilon\phi\left(\varepsilon\right) \\ &= \sum_{n=1}^{\overline{N}-1} n\left[\phi(A\left(x,n\right))-\phi\left(A\left(x,n-1\right)\right)\right]-\overline{N}\phi\left(A\left(x,\overline{N}-1\right)\right) \\ &= -\sum_{n=0}^{\overline{N}-1}\phi(A\left(x,n\right)) \end{aligned}$$

7.6 Estimating costs for dynamic shipbuilders: Details

We provide details on each step performed when estimating the cost function of dynamic shipyards.

1. *We estimate* $A(y,s,n)$ *using* (23). In this step, we first compute the frequencies $\{p_{n}(y_{jt},s_{t})\}_{n=0}^{\overline{N}}$ from observed data. We also include a post-2006 dummy in the state to capture differences in the policy function before and after 2006. As is common in dynamics applications, we don't have numerous observations for all $n=0,1,...,\overline{N}$ at each state (y_{jt},s_{t}). To overcome this sparsity, we first cluster the data finely, using the kmeans algorithm, and compute frequencies on a subset of states. Second,

we smooth the frequency matrix using kernels. In particular, to compute the choice probability $p_n(x)$ at state $x = (y_{jt}, s_t)$ we use the following formula:

$$p_n(x) = \sum_{x'} w(x' - x)\, \widetilde{p}_n(x')$$

where $\widetilde{p}_n(x)$ is the observed frequency count of n at state x and $w(\cdot)$ is a kernel that appropriately weights the distance of x from every other state x'. For numerical states (backlog, docks/berths, length, time, fleet, total backlog, demand, steel price) we use normal kernels with diagonal covariance. For categorical states (country and post-2006 dummy) we use the following kernel:

$$w(x' - x) = \begin{cases} 1 - h, & \text{if } x' = x \\ h/k_x, & \text{if } x' \neq x \end{cases}$$

where k_x is the number of values that x can take (in the case of country it's 4, in the case of the post dummy 2) and h represents the bandwidth of the kernel. As h gets close to 0, this kernel weights states that share the same variable x. We also experimented with parametric specifications for $A(y, s, n)$. In particular, we estimated an ordered probit model using directly the production data, so that:

$$A(x, n) = \beta f(x) + \gamma_n$$

while the observed variables are the production values given by

$$N^*(x, \varepsilon) = \widehat{n}, \text{ such that } \varepsilon \in [A(x, \widehat{n}), A(x, \widehat{n} - 1)]$$

We estimate β and γ_n for $n = 0, ..., \overline{N} - 1$ via Maximum Likelihood. This specification is flexible in terms of n but less so in terms of (y, s).[20] It overall gives similar results to the nonparametric specification above. Finally, we chose $\overline{N} = 10$, since 99.75% of observations involve $N \leq 10$.

2. *We compute the terms* EN^*, EN^{*2}, $E\varepsilon N$ using (27), (28) and (29)

[20] The plot of γ_n with respect to n exhibits small deviation from linearity. This is consistent with the static model where

$$A = \Phi^{-1}\left(\frac{1}{\sigma}\left(VE - c_1 - c_2(2n + 1)\right)\right)$$

This is relevant in case one thought that (in the static case) imposing both a distributional assumption on ε's, as well as a parametric form on $c(n)$ is restrictive.

3. *At each guess of the parameters* (θ, σ) *in the optimization of the likelihood* (20):

   (a) *We solve for the approximate value function parameters* γ *from* (26). Note that the choice probabilities require the continuation value $Q(x,n) = \gamma E[f(x')|n,x]$. To estimate γ from (26) we need

$$E_{\varepsilon,x'}[f(x')|N^*(x,\varepsilon),x] = \gamma \sum_{n=0}^{\overline{N}} p_n(x) E_{x'}[f(x')|n,x] \tag{32}$$

   We use polynomials of third order in all variables (we have also tried fourth order which doesn't alter the results). The aggregate state s evolves by the estimated VAR model described in Section 5.1.1, while the expectations of its polynomial powers are given in Appendix 7.1. We assume that the shipyard's individual backlog, b_{jt}, transitions as follows:

$$b_{jt+1} = (1-\delta) b_{jt} + n$$

   δ% of the backlog is delivered and period t's orders n enter the backlog. We experimented *extensively* with the above transition rule. In particular, we've tried models of time-varying δ (e.g. δ is drawn from a beta distribution estimated from the data whose mean can depend on the shipyard's current backlog, docks/berths or length; alternately, δ is taken as a discrete random variable with probabilities estimated from the data; in other experiments, we used deliveries, instead of δ, described by a binomial random variable whose parameters can again depend on shipyard observables). It was found that the simplest model where δ is taken constant over shipyards and time and equal to the sample mean (which is 10%) performs equally well to more complex models (and even better than several). Given the state transitions it is straightforward to compute (32). We estimate (26) using the LASSO in two ways. First, we call the LASSO within the likelihood maximization with the regularization parameter chosen using Belloni and Chernozhukov (2011). Second, we estimate (26) with LASSO using profits obtained from the static cost estimates. The goal here is to recover which polynomial terms should be kept. We then run OLS within the likelihood with only these terms (and repeat the estimation for many values of the regularization parameter). Results are overall robust to all the above.

(b) *Using the γ's we compute choice probabilities in the likelihood and update (θ, σ).*

7.7 Counterfactual Computation

There are two steps in the implementation of the counterfactual scenarios presented in Section 6. First, we compute the equilibrium of the model in each scenario (if shipyards are static this step is skipped). Second, we simulate the model using the observed paths of demand and steel prices which are exogenous shifters. Note that if we were only interested in the "No China" counterfactual, the first step can be skipped. Indeed, in that case we can simply use the pre-2006 expectations and value functions and simulate the model.

To predict how the industry would evolve under different counterfactual scenarios we need to obtain shipyards' optimal policies and value functions under each scenario. Note that we can no longer use the estimated VAR for state transitions, since this formed an approximation to expectations that hid equilibrium features. We therefore turn to the following accurate state transitions for (S_t^1, S_t^2, B_t), where S_t^1 is the number of ships younger than 20 years old, S_t^2 is the number of ships older than 20 years old and B_t is the total backlog:

$$\begin{aligned} S_{1t+1} &= \delta B_t + (1-\rho_{1t})\, S_{1t} \\ S_{2t+1} &= S_{2t} + \rho_{1t} S_{1t} - \zeta\left(s_t\right) \\ B_{t+1} &= (1-\delta) B_t + \sum_j N_{jt} \end{aligned}$$

where ζ is the number of ships that exit at state s_t, ρ_{1t} is the percentage of ships that transit from 19 years old and 3 quarters to 20 years old and δ is the percentage of backlog that is delivered, consistent with the individual backlog transition used in the estimation and described in Appendix 7.6. In words, the number of young ships S_{t+1}^1 equals last period's young ships plus deliveries from the total backlog, minus exiting ships (as documented in Kalouptsidi (2013) there is virtually no exit in ships younger than 20 years old). The number of old ships S_{t+1}^2 equals last period's old ships plus the aging ships minus exiting ships. Finally, total backlog B_{t+1} equals last period's total backlog minus deliveries, plus total new ship orders. We calibrate ρ_{1t} to 3% which is the sample average. To predict ship exit $\zeta\left(s_t\right)$ we follow Kalouptsidi (2013) where the number of exiting ships is regressed on the aggregate states (in particular, $\log \zeta_t = \beta_\zeta s_t$); note that exit rates are extremely low (even during the 2008 crisis). Demand d_t and steel price l_t retain their original transition processes, since these are exogenous to our model.

To find the equilibrium of the model in any of the counterfactual worlds we use a standard fixed point algorithm with the goal of recovering the shipyard's optimal policy function $p_n^*(x)$, all n and $x = (y, s)$, as well as the shipyard's value function $W^*(x)$. At each iteration l we use the policies $p_n^l(x)$ update to $p_n^{l+1}(x)$ and we keep iterating until $||p_n^{l+1}(x) - p_n^l(x)|| \leq eps$. Each iteration performs the following steps:

1. Update the value function using a parametric approximation and LASSO (we again use third order polynomials in all terms):

$$\left(f(x) - \beta E\left[f(x')|p_n^l(x), x\right]\right)\gamma^{l+1} = \pi\left(x, p_n^l(x)\right)$$

where:

(a) To compute ex ante profits, we use:

$$\pi\left(x, p_n^l(x)\right) = [VE(x) - c_1(x)]\sum_{n=0}^{\overline{N}} np_n^l(x) - c_2\sum_{n=0}^{\overline{N}} n^2 p_n^l(x) + \sigma\sum_{n=0}^{\overline{N}-1}\phi(A^l(x,n))$$

where $A(x,n) = \Phi^{-1}\left(1 - \sum_{k=0}^{n} p_k^l(x)\right)$, $n = 0, 1, ..., \overline{N}$. To derive the above we use (24) and (29).

(b) To compute $E\left[f(x')|p_n^l(x), x\right]$ we use the state transitions above. In particular, we simulate d_t and l_t one period forward since these are the only stochastic states now. We compute next period's (S_t^1, S_t^2, B_t) using the deterministic transitions outlined above. The only tricky part is computing agents' expectations over current total orders that appear in the backlog transition. Due to computational constraints we have assumed throughout that shipyards keep track of the total backlog rather than the distribution of backlogs. Therefore, at this stage shipyards don't have the full information to predict total orders accurately. To circumvent this issue we make the simplifying assumption that shipyards believe they are all at the same state and can thus use the total number of firms to predict total orders.

2. Update the choice probabilities:

$$\begin{aligned}
p_0^{l+1}(x) &= 1-\Phi\left(\frac{1}{\sigma}\left(VE(x)-c_1(x)-c_2+\beta\left(W^{l+1}(1)-W^{l+1}(0)\right)\right)\right) \\
p_n^{l+1}(x) &= \Phi\left(\frac{1}{\sigma}\left(VE(x)-c_1(x)-c_2(2n-1)+\beta\left(W^{l+1}(n)-W^{l+1}(n-1)\right)\right)\right)- \\
&\quad -\Phi\left(\frac{1}{\sigma}\left(VE(x)-c_1(x)-c_2(2n+1)+\beta\left(W^{l+1}(n+1)-W^{l+1}(n)\right)\right)\right) \\
p_{\overline{N}}^{l+1}(x) &= 1-\sum_{n=0}^{\overline{N}-1} p_n^{l+1}(x)
\end{aligned}$$

We solve the above fixed point under three scenarios: the true post 2006 world, a world with no China interventions and a world with only China's capital interventions. These worlds differ in the shipyard cost function, the set of active shipyards and the shipyard capital structure (all shipyards have the post-2006 docks/berths and length levels in the first and third scenarios, and the pre-2006 levels in the second). We perform the fixed point on a set of states chosen by the kmeans algorithm.

Finally, to simulate the model and produce the graphs of Section 6, we employ the retrieved value function approximating parameters γ^* and drawing cost shocks ε we compute $A(x,n)$, all n and obtain optimal production. At each state we visit we still need to compute $E\left[f(x')\left|p_n^*(x),x\right.\right]$ which we do as above, using the retrieved equilibrium choice probabilities.

References

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

[2] Amemiya, Takeshi. 1984. "Tobit Models: A Survey". Journal of Econometrics. 24: 3-61.

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

[4] Barwick Jia, Panle and Parag A. Pathak. 2012. "The Costs of Free Entry: An Empirical Study of Real Estate Agents in Greater Boston". http://economics.mit.edu/faculty/pjia/working

[5] Belloni, Alexandre and Victor Chernozhukov. 2011. "High-Dimensional Sparse Econometric Models, an Introduction". Springer Lecture Notes.

[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] Brander, James A. 1995. "Strategic Trade Policy".

[9] Chen, Jiawei, Susanna Esteban and Matthew Shum. 2013. "When do Secondary Markets Harm Firms?". Forthcoming, American Economic Review.

[10] Collard-Wexler, Allan. 2010. "Demand Fluctuations in the Ready-Mix Concrete Industry". Econometrica

[11] Collins, Gabriel and Michael C. Grubb. 2008. "A Comprehensive Survey of China's Dynamic Shipbuilding Industry". China Maritime Studies Institute. U.S. Naval War College.

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

[13] 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.

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

[15] Goldberg, Penny. (1995). "Product Differentiation and Oligopoly in International Markets". Econometrica. 63: 891-951.

[16] Haley C.V. Usha and George T. Haley. 2013. "Subsidies to Chinese Industry: State Capitalism, Business Strategy and Trade Policy". Oxford University Press.

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

[18] 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.

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

[20] Kalouptsidi, Myrto. 2013. "Time to Build and Fluctuations in Bulk Shipping". American Economic Review. Forthcoming.

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

[22] OECD. 2008. "The Shipbuilding Industry in China".

[23] 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.

[24] Pesendorfer, Martin, and Philipp Schmidt-Dengler. 2008. "Asymptotic Least Squares Estimators for Dynamic Games". Review of Economic Studies 75: 901-928.

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

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

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

[28] Sweeting, Andrew. 2012. "Dynamic Product Positioning in Differentiated Product Markets: The Effect of Fees for Musical Performance Rights on the Commercial Radio Industry". Forthcoming, Econometrica.

[29] Thompson, Peter. 2001. "How Much Did the Liberty Shipbuilders Learn? New Evidence for an Old Case Study". Journal of Political Economy. 109 (1): 103-137.

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

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

[32] Wooldridge, Jeffrey. 2001. "Econometric Analysis of Cross Section and Panel Data". The MIT Press.

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