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

Myrto Kalouptsidi[*†]

Department of Economics, Princeton University and NBER

PRELIMINARY

January 2014

Abstract

Government industrial subsidies, inadmissible by the WTO, are notoriously hard to measure. We provide a strategy for subsidy detection, which is based on estimating costs via an industry model. We apply our strategy to the world shipbuilding industry and in particular China's 2006 Medium and Long-Term Plan for the Shipbuilding Industry. We construct a dynamic model of shipping and shipbuilding and base our estimation on retrieving the willingness to pay for a new ship from new and used ship prices. Both ship demand and ship supply are dynamic. We find strong evidence of production subsidies to Chinese shipyards. We quantify the impact of China's entry into shipbuilding and predict that shipping investment would have been significantly smoother in the absence of China's government programs.

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, 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 industrial evolution has been a classic question of interest to economists. Despite a long theoretical literature, little is known about the quantitative impact of subsidies on prices, market shares, and even production reallocation across countries.

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.[2] As Haley and Haley (2013) point out, "generally researchers have a notoriously hard time measuring subsidies and Chinese subsidies [in particular ... due to] rampant missing and misreported data". The detection method used is the price-gap approach, which 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.[3] In this paper, we first propose a model-based empirical strategy of detecting government subsidies, which draws from the standard IO insights of estimating costs from demand variation and applies them to a framework of dynamic demand and supply. Once costs are recovered, we use suspected dates of initiation of government subsidization plans and compare costs before and after, as well as across countries. We next turn to the question of how subsidies may shape the paths that industries take. To do so, we use the estimated model to explore via counterfactuals the impact of different types of government subsidies on market shares, prices and production allocation across different countries.

Our paper focuses on the world shipbuilding industry, which is a prototypical example of an industry affected by such policies. Indeed, shipbuilding has historically been a

[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] 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.

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

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.[4] In 2006, China launched its "Long and Medium Term Plan for the Shipbuilding Industry (2006-2015)" aimed at expanding the country's capacity "in order to challenge existing mainstream shipbuilders" and "propel the country to world No. 1 status" (OECD (2008)). Since then China has doubled world shipbuilding capacity and massively increased its market share. What is known, observed and admissible by the WTO is the Chinese government's subsidization of the construction of new and the capacity expansion of existing shipbuilding plants (shown in the first panel of Figure 1). What is not known, unobserved 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.

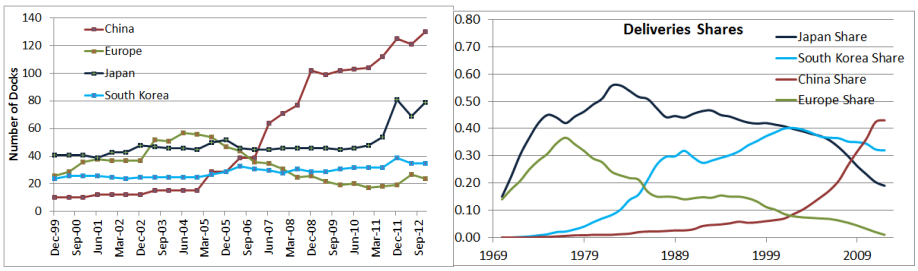


Figure 1: Docks and delivery shares per country.

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. We compare China's recovered cost function before and after 2006 when its Long and Medium Term Plan was announced, as well as with other countries and document a significant decrease in Chinese costs in the order of 15-20%. Even though our empirical approach is implemented in shipbuilding, it can be also be used in other suspect industries (e.g. steel).

Our model links the downstream shipping and the upstream shipbuilding industries, providing one of the first empirical analysis in industrial organization looking at dynamic

[4] 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)

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 extends the recent literature on the estimation of dynamic setups (see below). 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 4 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.

Next, we use our estimated model to quantify the impact of China's capacity and production subsidization program on the shipping and shipbuilding industries. An interesting feature of our sample period is that after 2006, demand for freight services boomed and led to a shipping investment boom. The crisis in 2008 led in turn to a crash. We find that in the absence of China's intervention, this boom and bust in shipping investment would have been significantly smoother. At the same time, China led to a substantial reallocation of production across the world: Japan's market share would have been on average larger by 20% between 2006 and 2012. We also compute the industry's equilibrium under alternative scenarios regarding subsidies. We explore the impact of subsidies on market share allocation, ship prices and investment fluctuations.

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)-henceforth BBL- 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, as in 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)). In addition, we amend the BBL procedure to provide flexible discrete first stage policy functions that depend on unobserved shocks. We also contribute to the mostly theoretical literature on strategic trade (see Brander (1995) for a survey), as well as more recent work on China's governmental programs (e.g. Brandt, Van Biesebroeck, Wang and Zhang (2012)). Finally, our work is also related to papers that use IO methods to understand questions of interest to international trade (e.g. Goldberg (1995), Berry, Levinshon and Pakes (1999)).

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 and offers some descriptive analysis of subsidy detection. 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 (such as 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)).

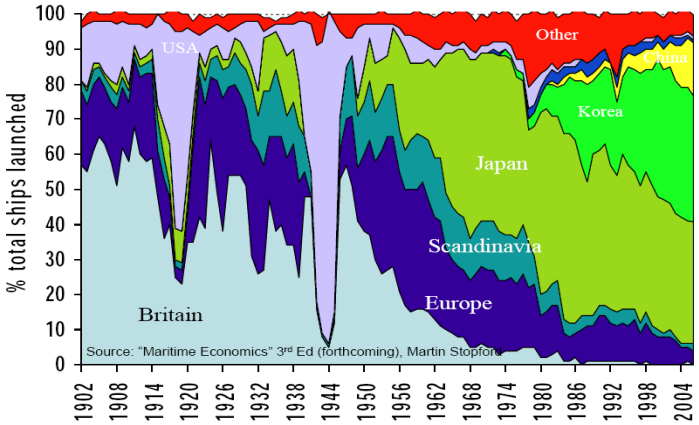


Figure 2: Shipbuilding History.

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 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)).[5] 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 capacity as measured by shipbuilding docks, as well as market share.

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 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

[5]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).

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 The Shipping Industry

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 a ship i at time t consists of an individual state s_{it} and an aggregate state s_t. The individual state s_{it} consists of the ship's:

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

while the industry aggregate state s_t consists of:

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, $l_t \in \mathbb{R}^+$.[6]

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 $\pi\left(s_{it}, S_t, d_t\right)$.[7]

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.[8]

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 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.[9] Therefore, the following equilibrium condition holds:

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

[6] 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.

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

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

[9] In fact, Kalouptsidi (2013) shows that the estimated value of entry in the shipping industry is very close to an average shipbuilding price (taken as exogenous in that work).

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:

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

In words, the value function of a ship at state (s_{it}, s_t) equals to 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 also have the option of selling their ship in the market for used ships. In this model, since all shipowners share the value of a ship, the price of a ship in the second hand market, P_{it}^{SH}, equals this value. Transactions in the second-hand market do not alter the size and characteristics of the fleet (under our assumption that the ship is indistinguishable from its owner). At a sale the ship just changes ownership and shipowners are always indifferent between selling a ship and operating it themselves. Therefore, in equilibrium:

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

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

3.2 The Shipbuilding Industry

There are J long-lived incumbent shipbuilders. The state variable of shipbuilder j at time t consists of an individual state y_{jt} and the aggregate state s_t. Shipyards share the aggregate state of shipowners, s_t, as this determines their demand in t. The individual state y_{jt} consists of the shipbuilder's :

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

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 $N_{jt} \in \left\{0, 1, ..., \overline{N}\right\}$. Shipyard j faces production costs, $C(N_{jt}, y_{jt}, l_t, \varepsilon_{jt})$. Even though N_{jt} is an integer we assume that

the cost function $C(N_{jt}, \cdot)$ can be defined over $\left[0, \overline{N}\right]$ 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:[10]

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

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.[11] It is natural, therefore, to expect that under demand uncertainty, 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:[12]

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

where s_{it} has $a_{it} = 0$ and the country of yard j. We let time to build be shipyard-specific and in particular, $T_{jt} = T(y_{jt}, s_t)$. Note that $VE_j(s_t)$ 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}

[10] 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.

[11] 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.

[12] 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.

can ignore the impact it has on $VE_j(s_t)$; note however, that aggregates do matter so that as the total fleet increases, shipowners' willingness to pay falls, all else equal.

Therefore, shipyard j chooses its production level to solve the Bellman equation:

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

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

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

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 deterministic 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 but is also important here, 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:

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

or

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

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

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

or

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

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

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

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

Therefore, a shipyard's optimal policy is characterized by relationships (9), (10) and (12). The next theorem, 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.

Theorem 1 *If the shipbuilding cost function $C(n, \cdot) : \left[0, \overline{N}\right] \to \mathbb{R}$, 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 theorem, which is very intuitive, implies that the optimal policy of the shipyard takes the following final form: $N^*(y_{jt}, s_t, \varepsilon_{jt}) =$

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

The timing of 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 intervention 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 capacities (found in z_j) and China's cost function. We also assume that shipyards do not make entry or capacity 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 six 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 bulker 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, the ship's tonnage 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

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.

sample is far 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

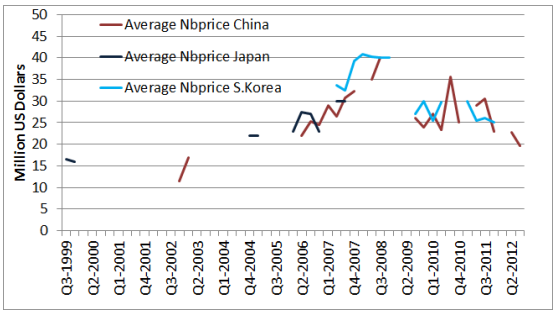


Figure 3: Reported Ship Newbuilding Prices.

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, tonnage 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 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.[13] The number of docks and berths are a 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 in general 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 (especially for number of employees) 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 have a different measure of shipbuilding capacity before 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 capacity and exhibits China's post 2006 explosion.[14]

The fifth dataset consists of shipping voyage contracts and includes the date of transaction, the name and size of the ship and the trip's price per trip. The dataset includes

[13] 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.

[14] We do not create a quarterly measure of capacity (docks, berths, largest dock) for several reasons. First, capacity 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 be fluctuating (or even decreasing) out of obvious measurement error. Third, if capacity were to become a state variable, its (slow and government driven) evolution would be difficult to model.

	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.

a subset of all voyage contracts between January 2001 and August 2012.[15]

Finally, the sixth 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.[16] Of course this finding is not proof of production subsidization; even the announcement of the capacity 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.

[15] See Kalouptsidi (2013) for a more detailed description of this dataset.

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

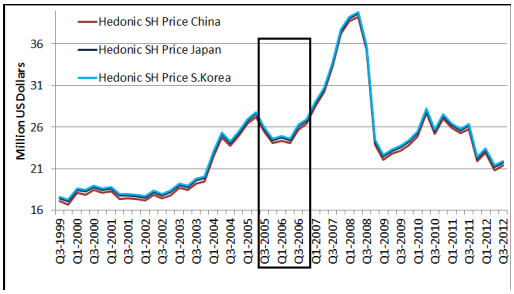


Figure 4: Hedonic regression of second-hand ship prices on ship characteristics (age, country) and quarter dummies.

5 Model Estimation and Detection of Subsidies

The main object of interest is the potentially subsidized cost function of world shipbuilders. To recover it, we estimate our model and test for a structural change in China's shipbuilding cost function in 2006 when its subsidization program began. 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 create the state space (d_t in particular) and estimate its transition.

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 extend and combine ideas from the recent literature on the estimation of dynamic games.

5.1 Estimation of the Willingness to Pay for a New Ship

We begin with the estimation of the inverse demand curve for shipping services. To estimate the expectations of shipowners, we assume that the industry state s_t follows a time-varying VAR and recover its parameters. We then proceed to estimate the value functions via sparse estimation techniques (LASSO). Each estimation task is described

below and followed by the results. All results presented are for Handysize vessels.

5.1.1 Demand for Shipping Services

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

$$P_t = \alpha_0^d + \alpha_1^d X_t^d + \alpha_2^d Q_t + \varepsilon_t^d \tag{14}$$

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 3: Demand IV regression results.

Table 3 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.[17]

[17]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 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):

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

5.1.2 Aggregate State Transition

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)) where state transitions and policy functions are estimated "offline" (i.e. directly from data, without imposing equilibrium conditions).

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 $B_t = \sum_{j,l} b_{jl}$.

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 intervention. In particular, since post 2006 shipbuilding capacity increases, shipowners know

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.

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:

$$\begin{bmatrix} S_t^1 \\ S_t^2 \\ B_t \\ d_t \\ l_t \end{bmatrix} = \begin{bmatrix} c^{S_1} \\ c^{S_2} \\ c^{B} \\ c^{d} \\ c^{l} \end{bmatrix} 1\{t \leq 2006\} + \begin{bmatrix} c^{S_1\prime} \\ c^{S_2\prime} \\ c^{B\prime} \\ c^{d} \\ c^{l} \end{bmatrix} 1\{t > 2006\} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ a^{d} \\ a^{l} \end{bmatrix} t + \begin{bmatrix} \rho^{S_1S_1} & \rho^{S_1S_2} & \rho^{S_1B} & \rho^{S_1d} & \rho^{S_1l} \\ \rho^{S_2S_1} & \rho^{S_2S_2} & \rho^{S_2B} & \rho^{S_2d} & \rho^{S_2l} \\ \rho^{BS_1} & \rho^{BS_2} & \rho^{BB} & \rho^{Bd} & \rho^{Bl} \\ 0 & 0 & 0 & \rho^{d} & 0 \\ 0 & 0 & 0 & 0 & \rho^{l} \end{bmatrix} s_{t-1} + \xi_t \tag{15}$$

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 4 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 have 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 $\rho^{S_1d} = 0$ 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 of these experiments. Our chosen specification combines the

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

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.

5.1.3 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:

$$P_{jt}^{NB} = E\left[\beta^{T_{jt}} V\left(s_{it+T_{jt},s_{t+T_{jt}}}\right)|s_{it}, s_t\right] + \zeta^{nb} \tag{16}$$

$$P_{it}^{SH} = V\left(s_{it}, s_t\right) + \zeta^{sh} \tag{17}$$

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 (17) and (16) in a single estimation step we follow a different methodology.[18] 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:

$$P_{jt}^{NB} = \beta^{T_{jt}} \gamma E\left[f\left(s_{it+T_{jt}}, s_{t+T_{jt}}\right) | s_{it}, s_t\right] = \tag{18}$$

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

$$P_{it}^{SH} = \gamma f\left(s_{it}, s_t\right) \tag{19}$$

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 (18) and (19) linearly; yet even though (19) can be estimated in a straightforward manner, (18) requires the computation of the right-hand side integrals, which becomes complicated. Indeed, (18) involves the expectation of higher order terms of the following vector:

$$s_{t+T} = R^T s_t + \sum_{k=t+1}^{t+T} R^{t+T-k}(C_k + ak + \xi_k) \tag{20}$$

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 (18) and (19). We therefore use the LASSO, a method appropriate for sparse regression problems, i.e. for problems with a

[18] 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 (15) 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 capacity 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.

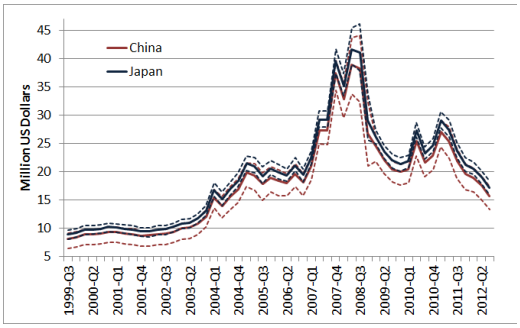


Figure 5: Estimated value function of a 0 years old ship. 0.95 bootstrap con.dence intervals.

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, which combines and extends ideas found in the recent literature on the estimation of dynamic setups (Rust (1987), Hotz and Miller (1993), BBL, Pakes, Ostrovsky and Berry (2007), etc.). We 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:

$$\prod_{j,t:N_{jt}=0} \Pr\left(N_{jt}=0|y_{jt},s_t;\theta\right) \prod_{j,t:N_{jt}=\overline{N}} \Pr\left(N_{jt}=\overline{N}|y_{jt},s_t;\theta\right) \prod_{n} \prod_{j,t:N_{jt}=n} \Pr\left(N_{jt}=n|y_{jt},s_t;\theta\right) \tag{21}$$

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

$$\Pr\left(N_{jt}=0|y_{jt},s_t;\theta\right) = 1-\Phi\left(\frac{1}{\sigma}\left[\begin{array}{c} VE_j\left(s_t\right)+c\left(0,y_{jt},s_t;\theta\right)-c\left(1,y_{jt},s_t;\theta\right)+ \\ +\beta\left(Q\left(y_{jt},s_t,1\right)-Q\left(y_{jt},s_t,0\right)\right) \end{array}\right]\right) \tag{22}$$

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

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

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 we adopt a two-step approach to tackle it. First, we recover a first-stage (approximate) optimal

policy function $N^*(y_{jt}, s_t, \varepsilon_{jt})$; second, we use this optimal policy function along with the transitions of the states (y_{jt}, s_t) to create simulated paths of shipyard profits, the average of which is the shipyard value function.

Recovering the Optimal Policy: A complication of our setup compared to other applications is that the optimal policy is a function of the unobserved cost shock ε_{jt}; therefore, we can't simply regress observed production on observed states to retrieve a first stage optimal policy function. We instead recover the policy function nonparametrically using choice probabilities. Indeed, let

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

for $n = 0, 1, ..., \overline{N} - 1$. Then we can rewrite the choice probabilities (22) as follows:

$$\begin{aligned}
\Pr(N^* = 0|y_{jt}, s_t) &\equiv p_0(y_{jt}, s_t) = \Pr(\varepsilon \geq A(y_{jt}, s_t, 0)) \\
\Pr(N^* = n|y_{jt}, s_t) &\equiv p_n(y_{jt}, s_t) = \Pr(\varepsilon \leq A(y_{jt}, s_t, n-1)) - \Pr(\varepsilon \leq A(y_{jt}, s_t, n)) \\
\Pr\left(N^* = \overline{N}|y_{jt}, s_t\right) &\equiv p_{\overline{N}}(y_{jt}, s_t) = \Pr\left(\varepsilon \leq A\left(y_{jt}, s_t, \overline{N} - 1\right)\right)
\end{aligned}$$

The function $A(y_{jt}, s_t, n)$ can be recovered in a straightforward manner from the observed choice probabilities $\{p_n(y_{jt}, s_t)\}_{n=0}^{\overline{N}}$. Indeed, it is easy to show that[19]

$$A(y_{jt}, s_t, n) = \Phi^{-1}(1 - p_0(x_j, s) - p_1(x_j, s) - ... - p_n(x_j, s)), \quad \text{for } n = 0, 1, ..., \overline{N} - 1 \tag{23}$$

Clearly, $A(y_{jt}, s_t, n)$ is (weakly) decreasing in n. Most important, if $A(y_{jt}, s_t, n)$ is known, so is the optimal policy: For any (y_j, s, ε), $N^*(y_j, s, \varepsilon) = \widehat{n}$ such that[20]

$$\varepsilon \in [A(y_j, s, \widehat{n}), A(y_j, s, \widehat{n} - 1)]$$

We recover $A(y_{jt}, s_t, n)$ using (23) on all observed states (y_{jt}, s_t) and for $n = 0, 1, ..., 10$ (99.75% of observations involve $N \leq 10$). Since we don't observe all possible combinations of (y_{jt}, s_t) with n, we smooth the the choice probabilities across states using a normal kernel. This smoothing acts as an interpolation operation that enables us to compute

[19] To show this, begin with $p_0(y_{jt}, s_t) = 1 - \Phi(A(y_{jt}, s_t, 0))$, so that $A(y_{jt}, s_t, 0) = \Phi^{-1}(1 - p_0(y_{jt}, s_t))$. Next, $p_1(y_{jt}, s_t) = \Phi(A(y_{jt}, s_t, 0)) - \Phi(A(y_{jt}, s_t, 1))$, so that $A(y_{jt}, s_t, 1) = \Phi^{-1}(\Phi(A(y_{jt}, s_t, 0)) - p_1(y_{jt}, s_t)) = \Phi^{-1}(1 - p_0(y_{jt}, s_t) - p_1(y_{jt}, s_t))$.

[20] Even though BBL includes this case for continuous controls, to our knowledge this hasn't been implemented; for example, Ryan (2013) recovers the policy function on observed states and then computes flexible functions of conditional expectations of these shocks that appear in firms' ex ante profits.

$A(\cdot)$ at any state (y, s). This feature is exploited in the forward simulation procedure below.

Recovering the Continuation Value: The objects we want to recover are the continuation values $Q(y_{jt}, s_t, N_{jt})$, $Q(y_{jt}, s_t, N_{jt}+1)$, and $Q(y_{jt}, s_t, N_{jt}-1)$. In words, we need the continuation value at each observed state (y_{jt}, s_t) when the observed N_{jt} is chosen, as well as for two one-shot deviations: $N_{jt}+1$ and $N_{jt}-1$. Indeed, note that by its definition (8) the continuation value $Q(y, s, n)$ implies that the optimal policy $N^*(y, s, \varepsilon)$ is chosen after today's choice n. We recover the continuation value via forward simulation, as in BBL, using the expected discounted sum of profits formulation for the value function:

$$Q(y_{jt}, s_t, n) = E\left[\sum_{\tau=t+1}^{\infty} \beta^t \left(VE_j(s_\tau) N_\tau^* - c(N_\tau^*, y_{j\tau}, l_\tau; \theta) - \sigma\varepsilon_{j\tau} N_\tau^*\right) | y_{jt}, s_t, n\right]$$

where $N_\tau^* = N^*(y_{j\tau}, s_\tau, \varepsilon_{j\tau})$. Note that period t's choice of n affects shipyard j's future states $y_{j\tau}$, $\tau > t$, through the backlog. Therefore, N_τ^* will also differ since the same function $N^*(\cdot)$ is computed on different states.

It is conceptually straightforward to create simulated paths in the above expression. Indeed, using the VAR we can simulate s_t forward. Using the estimated parameters of the ship value function, at each simulated state s_τ we can compute $VE_j(s_\tau)$.[21] Using the estimated policy function above, we can compute N_τ^* at the simulated states $y_{j\tau}, s_\tau$ and draws $\varepsilon_{j\tau}$. Given parameters (θ, σ) we can also compute costs; in our empirical specification $c(\cdot; \theta)$ will be linear in θ, in which case as in BBL the simulated paths need to be computed once, outside of the likelihood maximization. The final element we need to determine is the transition of y_{jt} and the backlog in particular (the remaining elements of y_{jt} are either time invariant, e.g. capacity measures and country, or deterministic, e.g. age). Since keeping track of the entire backlog vector is intractable, we assume that the relevant state is the total backlog (i.e. the sum of all ships under construction). We assume that the backlog has the following transition rule:

$$b_{jt+1} = (1 - \delta_{jt}) b_{jt} + N_{jt}$$

where $\delta_{jt} \sim F_\delta(\cdot | y_{jt})$. In words, $\delta_{jt}\%$ of the backlog is delivered every period and the chosen number of ships enters. We assume that δ_{jt} is drawn from a distribution conditional

[21] Calculating $VE_j(s_\tau)$ also requires j's time to build at $y_{j\tau}$. We, therefore, estimate a time to build function $T(y_{jt}, s_t)$ using the observations with positive production. Table 5 reports our estimated specification.

on the shipyard's individual state.

Once the value function is obtained via simulation at each observed combination (y, s, n), as well as $(y, s, n+1)$ and $(y, s, n-1)$, we insert it in the likelihood function (21) and recover (θ, σ). We next turn to our results.

5.2.2 Results

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

$$C\left(N_{jt}, y_{jt}\right) = c_1\left(y_{jt}, l_t; \theta\right) N_{jt} + c_2\left(y_{jt}, l_t; \theta\right) N_{jt}^2 + \sigma \varepsilon_{jt} N_{jt}$$

where $\varepsilon_{jt} \sim N\left(0,1\right)$ and $c_2\left(y_{jt}, l_t; \theta\right) > 0$ and (θ, σ) are our cost parameters of interest. We further assume that $c_1\left(y_{jt}, l_t; \theta\right)$ and $c_2\left(y_{jt}, l_t; \theta\right)$ are linear in θ. In particular, our baseline specifications involve

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

and

$$c_2\left(y_{jt}, l_t; \theta\right) = c_2$$

where $g(y_{jt}, t)$ is a polynomial in (y_{jt}, 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 $c_1\left(y_{jt}, l_t; \theta\right)$ and $c_2\left(y_{jt}, l_t; \theta\right)$, 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:

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

and the continuation value is removed from the decision rule (22). 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 6 reports results from some baseline specifications. The number of observations varies depending on the covariates: in specifications that include capacity measures (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 20%, as indicated by the significant China-POST dummy. The results suggest that there is significant convexity in costs. Costs are decreasing in capacity 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 6 presents some of the specifications we have tried with respect to time. Results are robust to several functions of t and as examples we show different combinations of functions of t and covariates. We only present a limited number of specifications because of space limitations. 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 Theorem 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. Table 7 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 capacity coming online regionally. As shown in Table 8, 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 when shipbuilder dynamics are taken into account suggest that Chinese costs experience significant declines post 2006. It seems that the retrieved subsidies are lower when production dynamics are taken into account. In analogy with the case of static shipbuilders, the backlog appears to decrease costs, consistent with economies of scale or the accumulation of expertise. We plan to perform several robustness exercises as in the static case.

6 Quantifying the Implications of Subsidies

What is the impact of China's subsidization program on the shipping and shipbuilding industries? An interesting feature of this period is that after 2006 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 (see Kalouptsidi (2013) for a detailed analysis). We now ask whether and how much did China's increase in shipbuilding capacity amplify this boom and bust in shipping investment. To answer this question, we use our model to predict the evolution of the industry in the absence of China's entry. Our model and estimation technique allows us to perform this counterfactual in a simple manner, without computing an equilibrium. In particular, since we have estimated the before and after 2006 worlds separately we can feed the observed past-2006 exogenous processes (i.e. demand d_t and steel prices l_t) in the pre-2006 world and compute ship prices and shipbuilding production.

We use the observed post-2006 values for (d_t, l_t) and simulate paths for the fleet and the backlog (S_t, B_t) using the pre-2006 VAR. At all simulated paths, we compute $VE_j(s_t)$ using the pre-2006 LASSO parameters. Figure 6 depicts the evolution of ship prices, had China not subsidized neither capacity, nor production in the shipbuilding industry. Not surprisingly, in the absence of China's government intervention prices remain higher for all countries, since China's subsidization reduced demand for all ships.

We then simulate cost shocks ε_{jt} and use them to compute optimal production for each shipyard. The set of shipyards we consider are those that were actually active in $Q4$-2006 (about 90 yards out of 200 total active in the Handysize segment, with mean age 19 years and median age 9 years old). The cost specification we work with is I of Table 6; since we are working with shipyards active in $Q4$-2006 we did not want to remove shipyards because of missing capacity measures.

Figure 7 presents the observed and predicted total backlog for Japan and China. The

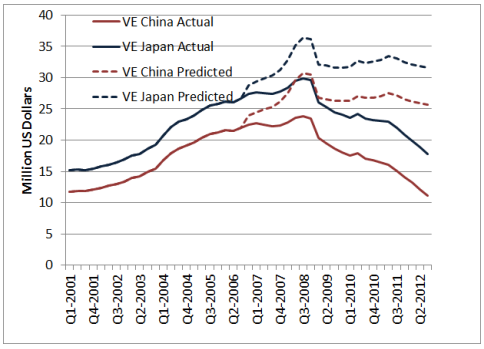


Figure 6: Actual and counterfactual willingness to pay for a new ship.

figure implies that China's subsidies amplified the boom and bust in shipping investment in the last decade. Indeed, without its massive increase in shipbuilding capacity, the backlog would not have increased as dramatically and as a consequence would not have crashed as bad in the 2008 crisis. Moreover, consistent with our estimates that suggest persistently that Japan is a lower cost producer than pre-2006 China, Japan remains market leader in our counterfactual world. In fact, the average difference between Japan's actual and counterfactual market share (based on production) is 22.6% for the above mentioned cost specification.

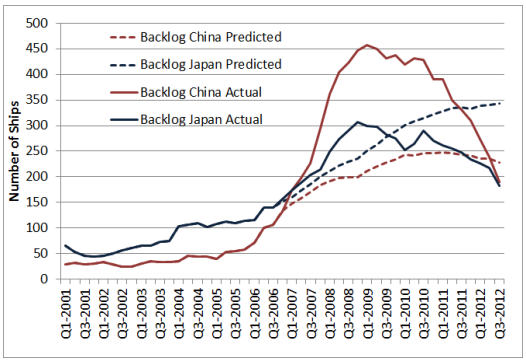


Figure 7: Actual and counterfactual backlog.

In work in progress we quantify the impact of the estimated, prohibited, production subsidies. We compute the equilibrium of our model under different counterfactual scenarios regarding subsidies. We compute the optimal production strategy of shipbuilders under these scenarios and compare the case of only production subsidies, only capacity subsidies, both and neither.

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 5: 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		V	
	Prm	s.e.	Prm	s.e.	Prm	s.e.	Prm	s.e.	Prm	s.e.
China	29.95	$(1.62)^*$	29.72	$(1.6)^*$	28.4	$(1.9)^*$	26.88	$(1.4)^*$	19.26	(2.04)
China,POST	-6.58	$(0.51)^*$	-5.55	$(0.5)^*$	-3.8	$(0.6)^*$	-5.7	$(0.49)^*$	-4.21	$(0.5)^*$
Europe	28.45	$(1.59)^*$	29.8	$(1.8)^*$	27.7	$(1.92)^*$	26.98	$(1.46)^*$	19.3	$(2.2)^*$
Japan	23.22	$(1.08)^*$	22.31	$(1.01)^*$	25.3	$(1.3)^*$	20.81	$(0.89)^*$	15.3	$(1.3)^*$
S. Korea	27.99	$(1.26)^*$	29.49	$(1.45)^*$	29	$(1.64)^*$	25.4	$(1.13)^*$	19.25	$(1.6)^*$
Backlog	-0.54	$(0.04)^*$	-0.51	$(0.04)^*$	-0.3	$(0.05)^*$	-0.52	$(0.04)^*$	-0.37	(0.05)
Age			-0.017	$(0.002)^*$			0.039	$(0.003)^*$	0.027	(0.004
Docks/Berths							-0.16	$(0.01)^*$	-0.11	(0.02)
Max Length							-0.0002	$\left(2e^{-5}\right)^*$	-0.0002	$\left(2e^{-5}\right.$
Steel price	0.316	$(0.03)^*$	0.297	$(0.03)^*$	0.86	$(0.12)^*$	0.36	$(0.03)^*$	0.28	(0.03)
t	0.27	$(0.02)^*$	0.21	$(0.02)^*$			0.31	$(\ 0.018)^*$	0.53	(0.04)
t^2									-0.004	(0.001
Year Dummies	NO		NO		YES		NO		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	$(0.11)^*$	1.02	$(0.08)^*$	0.71	$(0.1)^*$
σ	10.43	(0.83)	9.88	$(0.7)^*$	5.55	$(1.07)^*$	10.97	(0.856)	7.6	(1.13)
No of Obs	8324		8324		8324		4571		4571	

Table 6: 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 7: 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 8: 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.

Appendix A

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 (18) can be written as:

$$P_{jt}^{NB} = \beta^{T_{jt}} \gamma E\left(f\left(x_{t+T_{jt}}\right)|x_t\right) = \beta^{T_{jt}} \sum_{i=1}^{d} \gamma_i E\left(x_{t+T_{jt}}^{(i)}|x_t\right) \tag{24}$$

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 (20) 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, (24) 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).

Appendix B

Derivation of the likelihood function (21): 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 (21).

Proof of Theorem 1. We first show the following lemma:

Lemma 2 *If $f(x) : [a, b] \to \mathbb{R}$ 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.

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

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

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

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

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

[10] 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.

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

[12] 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.

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

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

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

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

[17] 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.

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

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

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

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

[22] 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.

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

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

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

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

[27] 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.

[28] 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.

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

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

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

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