DETECTION AND IMPACT OF INDUSTRIAL SUBSIDIES:
THE CASE OF WORLD SHIPBUILDING

Myrto Kalouptsidi

Detection and Impact of Industrial Subsidies: The Case of World Shipbuilding
Myrto Kalouptsidi
NBER Working Paper No. 20119
May 2014, Revised October 2016
JEL No. L0

<u>ABSTRACT</u>

This paper provides a model-based empirical strategy to, (i) detect the presence and gauge the magnitude of government subsidies and (ii) quantify their impact on production reallocation across countries, industry prices, costs and consumer surplus. I construct and estimate an industry model that allows for dynamic agents in both demand and supply and apply my strategy to world shipbuilding, a classic target of industrial policy. I find strong evidence consistent with China having intervened and reducing shipyard costs by 13-20%, corresponding to 1:5 to 4:5 billion US dollars, between 2006 and 2012. The subsidies led to substantial reallocation of ship production across the world, with Japan, in particular, losing significant market share. They also misaligned costs and production, while leading to minor surplus gains for shippers.

Myrto Kalouptsidi
Department of Economics
Harvard University
Littauer 124
Cambridge, MA 02138
and NBER
myrto@fas.harvard.edu

1   Introduction

In recent years, Chinese firms have extremely rapidly dominated a number of capital intensive industries, such as steel, auto parts, solar panels, shipbuilding.[1] Government subsidies are often evoked as a possible contributing factor to China's expansion.[2] Yet, even though industrial subsidies have steered industrialization and growth in several countries throughout economic history (e.g. in East Asia), little is known about their quantitative impact on production reallocation across countries, industry prices, costs and surplus. A significant challenge in this task is that government subsidies to industries are notoriously difficult to detect and measure; and in China even more so. Indeed, partly because international trade agreements prohibit direct and in-kind subsidies, "systematic data are non-existent"[3] and thus the presence and magnitude of such subsidies is often unknown.

This paper assesses the consequences of government subsidies on industrial evolution, focusing on the unparalleled and timely Chinese expansion. Since measuring subsidies is a prerequisite for this analysis, I provide a model-based empirical strategy to detect their presence and gauge their magnitude. I apply my empirical strategy to the world shipbuilding industry, a classic target of industrial policy. In 2006, China identified shipbuilding as a "strategic industry" and introduced a plan for its development.[4] Within a short time span, its market share doubled from 25% to 50%, leaving Japan, S. Korea and Europe trailing behind. Many asserted that China's rapid rise was driven by hidden government subsidies that reduced shipyard production costs, as well as by a number of new shipyards constructed through the government plan; here I assess the relative contribution of these interventions.

I develop and estimate a model of the shipbuilding industry that allows for dynamic agents in both demand and supply. A large number of shipyards offer durable, differentiated ships. Their production decisions are subject to a dynamic feedback because of time to build: shipyards accumulate backlogs, which can affect their future production cost, either positively (expertise acquisition) or negatively (capacity constraints). Production cost is also subject to an aggregate stochastic shock, summarized in the price of steel, a key production input. Every period a large number of identical potential shipowners decide to enter the freight market by buying a new ship from world shipyards. Demand for new ships is driven by demand for international sea transport, which is uncertain and volatile. As ships are long-lived investments for shipowners, demand depends on expectations about future demand and fleet development.

---

[1] The share of labor intensive products in Chinese exports fell from 37% to 14% between 2000 and 2010. On a monthly basis, in 2011 the US imported advanced-technology products from China 560% more than it exported to China. By contrast, the monthly US-China trade surplus in scrap (used as raw material) grew by 1187% between 2000 and 2010. (U.S.-China Economic and Security Review Commission (2011)).

[2] "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 [include the] seemingly endless supply of cheap labour (...) another reason for China's industrial dominance: subsidies." ("Perverse Advantage", The Economist, April 2013).

[3] WTO (2006).

[4] Section 2 provides details.

The model primitive of interest is the cost function of potentially subsidized firms. As in many industries, however, costs of production are not observed. My strategy amounts to estimating costs from demand variation in a framework of dynamic demand and supply. In the simplest example of a static, perfectly competitive framework, marginal cost is recovered directly from prices. In that case, the detection strategy amounts to testing for a break in observed ship prices in 2006, when China launched its shipbuilding plan. In my setup, there are two complications: (i) new ship price data are scant, and (ii) the shipbuilding production decision is dynamic as it affects future costs. To address (i), I add used ship prices; to address (ii) I use the shipyard's optimality conditions resulting from its dynamic optimization.

My estimation strategy can be summarized as follows. To estimate shipyard costs, I adopt a novel approach that is inspired by the literature on estimation of dynamic setups, e.g. Hotz and Miller (1993), Bajari, Benkard and Levin (2007) and Pakes, Ostrovsky and Berry (2008), as well as Xu (2008), Ryan (2012), Collard-Wexler (2013) and others. I first invert the shipyards' observed choice probabilities to obtain directly the optimal policy thresholds nonparametrically in an ordered-choice dynamic problem; I then show that the latter lead to a closed-form expression for the shipyards' ex ante optimal per period payoffs, which in turn are sufficient to obtain shipyard value functions. To recover demand for new ships I estimate the willingness to pay using new and used ship prices, as well as shipowner expectations. The estimation allows expectations and value functions to be different before and after 2006, consistent with China's intervention. Finally, in estimating dynamics, I employ sparse approximation techniques (LASSO) that allow for a large state space, as well as significant flexibility. Solving value functions via LASSO, which to my knowledge has not be used before, can be very useful to other applications with dynamic agents and a large state space. Indeed, I provide a detailed Monte Carlo analysis to show that this method works well compared to more standard methods of solving dynamic problems (e.g. state space discretization), which are often intractable in real applications.

I use my framework to detect and measure changes in costs that are consistent with subsidies. I find a strong significant decline in Chinese costs equal to about 13-20% of costs, or 1.5 to 4.5 billion US dollars at the observed production levels. The results hold under flexible cost function specifications, as well as dynamic shipyards whose past production is found to decrease future costs, capturing expertise accumulation. I provide evidence that a number of possible alternative explanations for the recovered cost decline cannot adequately account for the observed patterns. For example, the results are robust to many specifications that control flexibly for time-variation. Moreover, costs did not change in other countries. Convincingly, the results hold when I estimate costs on the subset of shipyards that existed prior to 2001. Hence cost declines are not driven by new shipyards, which may have a different technology (though bulk ship production is not characterized by technological innovations to begin with) or may be optimizing under (a more general form of) learning by doing.

Next, I use my framework to quantify the contribution of government interventions in China's

seizing of the market. My main counterfactual predicts industry evolution (production reallocation across countries, ship prices, industry costs and shipper surplus) in the absence of China's government plan altogether; note that both cost subsidies and investment in new shipbuilding facilities would violate international trade agreements. Moreover, I assess the relative contribution of the new shipbuilding facilities constructed through China's plan in a second counterfactual that removes the new facilities, but maintains the cost subsidies detected above. Here is a brief summary of my main four findings from these counterfactuals.

First, I find that the interventions led to substantial reallocation in production: in the absence of China's government plan, Chinese market share is cut to less than half, while Japan's share increases by 70%. If only new shipyards are removed, China's share falls from 50% to 40%, suggesting that new shipyards played an important, but not the predominant part in China's expansion.

Second, ship prices experience moderate increases in all countries in the absence of China's plan, as the latter shifted supply outward.

Third, in the presence of subsidies, freight rates decrease moderately because of the larger fleet between 2006 and 2012 and more so over time due to time to build. As a result of China's plan, cargo shippers gain about 400 million US dollars in shipper surplus over that time period. Comparing this gain to the 4 billion US dollars of cost subsidies alone implies that the benefits of subsidies for shippers are minimal and perhaps the Chinese government is also aspiring to externalities to different sectors (e.g. steel, defense).

Fourth, subsidies create a wedge in the alignment of market share and production costs: they lead to a large increase in the industry average cost of production (net of subsidies) by shifting production away from low-cost Japanese shipyards towards high-cost Chinese shipyards.

Finally, government subsidies are often related to dynamic considerations, and this is true in this application as well: the long construction lags and the positive production feedback resembling learning by doing are important features of shipbuilding. Indeed, if past production decreases future costs, it is important to take this feedback into account when detecting subsidies, as it may be explaining part of the increase in production. In addition, to evaluate the importance of dynamics in the policy analysis, I consider two extra models: (i) a model with no production feedback so that a firm's backlog does not affect its future production costs; (ii) a model with no construction lags in ship production. In the first model, China experiences a lower increase in its market share of 10%; increased backlog decreases costs and thus market share gains multiply and lead to more reallocation favorable to China. This is consistent with the pro-subsidy arguments of infant industries. In the second model with no time to build, demand for ships is higher in the boom, leading to higher overall production, and higher volatility.

This paper contributes to our understanding of the impact of industrial policy, while recognizing that we most often lack direct information on the measures undertaken by (foreign) governments. The methodology of subsidy detection here relies solely on having a dataset of prices and quantities

(a common dataset in the IO literature) and some knowledge about an implementation date. This allows one to estimate the cost function of potentially subsidized firms from demand variation and to test for a "break" at the corresponding date; the idea being that the government lowers costs via subsidies. In that sense, one could apply this approach to any industry after appropriately tailoring the model (e.g. competition, dynamic/static behavior).[5]

This paper contributes to the recent empirical literature on industry dynamics (e.g. Benkard (2004), Aguirregabiria and Mira (2007), Xu (2008), Ryan (2012), Collard-Wexler (2013)). Methodologically, it lies closest to Hotz and Miller (1993), Bajari, Benkard and Levin (2007) and Pakes, Ostrovsky and Berry (2007). This literature typically considers either single agent dynamics or dynamic firms and static consumers (one exception is Chen, Esteban and Shum (2013)). In contrast, this paper allows for dynamics in both demand and supply; these dynamics are important in many durable goods markets (e.g. autos, TV's), as well as upstream-downstream industries (e.g. airlines-aircraft builders, shipping-shipbuilders). Kalouptsidi (2014) focuses on the *downstream* industry of bulk shipping and studies its cyclical behavior and the role of adjustment costs (time to build as in Kydland and Prescott (1982)) in these fluctuations. That paper is a starting point in modeling the demand of shipbuilders in this paper, yet the focus here is on Chinese government interventions in the *upstream* industry of shipbuilding. Both the questions and the industry under study are different in the two papers.

The paper also naturally contributes to the literature on trade policies. Goldberg (1995) and Berry, Levinsohn and Pakes (1999) consider the impact of voluntary export restraints in the automobile industry. Grossman (1990) provides an excellent survey of the relevant trade literature, while Brander (1995) is a classic reference on the long theoretical literature of strategic trade (see also references within). Not surprisingly, given the constraints in subsidy data availability, there is little empirical work, most of which is in the form of model calibration. Baldwin and Krugman (1987a) and (1987b) explore the impact of trade policies in the wide-bodied jet aircraft and the semiconductor industries, while Baldwin and Flam (1989) in the small commuter aircraft industry. They all discuss the lack of knowledge regarding both the presence and magnitude of subsidies and other policies and compute industrial evolution under different hypothetical scenarios.

The remainder of the paper is organized as follows: Section 2 provides a description of the environment under study including subsidy issues and features of the industry. Section 3 presents the model of shipbuilding subsidies. Section 4 describes the data used and provides some descriptive evidence. Section 5 presents the empirical strategy and the estimation results starting with static shipyards and robustness checks and then proceeding to dynamic shipyards. Section 6 provides the counterfactual experiments. Section 7 performs a robustness analysis in several dimensions to evaluate the paper's results, while Section 8 concludes. Finally, the Appendix provides details on

---

[5]For example, if the industry under consideration involved oligopolists that produce differentiated products, one could estimate a differentiated product demand system (e.g. Berry, Levinsohn and Pakes (1995)) to obtain the cost curve from the firms' first order conditions and test for a break.

several aspects of the analysis, as well as a Monte Carlo for the use of sparse approximations in dynamic programming.

2   Environment

This section discusses the practical difficulties regarding the detection and impact evaluation of subsidies generally and in China in particular. Then, it provides the key features and history of the shipbuilding industry.

Subsidy Disputes   Although government interventions in the form of subsidies are arguably prevalent and subsidy disputes abundant, both domestic and global policy-makers (e.g. WTO) have difficulty in designing and adopting appropriate measures to respond to them.[6] Indeed, deciding on subsidy complaints is difficult for two reasons; first, "systematic data (on industrial subsidies) are non-existent; reliable sources of information are scarce and mostly incomplete [...] because governments do not systematically provide the information" (WTO (2006)). Therefore, detecting and measuring subsidies becomes a difficult task and plaintiffs need to base their allegations mostly on available self-reported data.[7] The second difficulty concerns (dis)proving "injury caused" by the alleged subsidies.[8] Ideally, the question to be answered is "how would have this industry evolved absent the alleged subsidies?"; this is clearly a difficult question, which in practice is answered based on a mostly qualitative analysis of industry indicators on a case by case basis.

China in particular, has had more trade conflicts than any other country in the world, in more industries and with more countries.[9] China provides industry subsidies in the form of free or low-cost loans, as well as subsidies for inputs (including energy), land and technology. Because of institutional and strategic reasons, the information on subsidies that the Chinese government pro-

---

[6]Subsidy disputes are handled both bilaterally (e.g. in the US the relevant agencies are the Department of Commerce and the International Trade Commission (ITC)), as well as internationally through the WTO. "The ITC received a total of 1,606 antidumping and countervailing duty petitions during 1980-2007. These cases involved over $65 billion in imports. Thirty-eight percent of the petitions resulted in affirmative determinations by the Commission and Commerce, culminating in the issuance of an antidumping or countervailing duty order." (ITC (2008))

[7]One possible source of information (other than e.g. national accounts or individual government measures (Sykes (2009))) are subsidy "notifications" required by the WTO under SCM; however "the data contain many gaps and shortcomings. Not all Members fulfill the notification requirements. 29 of the currently 149 WTO Members have so far not submitted any notification. [...] Other Members do not provide quantitative information on subsidy programmes or do not provide this information systematically. In most years, information is available for less than half of the WTO membership." (WTO (2006))

[8]For example, in the US "an interested party may file a petition with ITC and Commerce alleging that an industry in the United States is materially injured or threatened with material injury, or that the establishment of an industry is materially retarded, by reason of imports that are being subsidized" (ITC (2008)). "Material injury" amounts to observing the volume of imports, as well as their impact on the domestic industry's prices and production. WTO's SCM has similar provisions: complaining countries must provide evidence of adverse trade effects in their own or third country markets (as well as evidence on the existence and nature of subsidies as mentioned above).

[9]The rest of the section draws from Haley and Haley (2008).

vides has rampant missing and misreported data.[10] These subsidies are often transferred through financial institutions, most of which are directed by the central and provincial governments.

Shipbuilding and China's Plan[11]   Shipbuilding is one of the major recipients of subsidies globally, along with e.g. the steel, mining and automotive industries. It is often seen as a "strategic industry" as it increases industrial and defense capacity, generates employment and has important spill-overs to other industries (e.g. iron and steel). Indeed, several of today's leading economies developed their production technologies and human capital through a phase of heavy industrialization, in which shipbuilding was one of key pillars. 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 (4.5% of its GDP today).

China's 11th National 5-year Economic Plan 2006-2010 was the first to appoint shipbuilding as a "strategic industry" in need of "special oversight and support"; the central government "unveiled an official shipbuilding blueprint to guide the medium and long-term development of the industry". As part of the national plan, the central government introduced the "Medium and Long-Term Plan for the Shipbuilding Industry" (船舶工业中长期发展规划)[12] which set the goal of China becoming the largest shipbuilding nation within a decade. The plan sets specific output and capacity goals and as such involves investment in shipbuilding infrastructure and financial support for output growth, including increased credit allowance, low input prices, etc.

Consistent with these government programs, Figure 1 shows China's rapid expansion in shipbuilding dry docks (a measure of shipbuilding capacity). It is important to note that the majority of this expansion (82%) was realized through the construction of *new* facilities; so that the industry experienced a massive entry wave in 2005-2006.

In contrast to the capital expansion, which is observed, subsidies that reduce operating costs are not observed. Yet, consistent with such measures, China's production and market share increased dramatically as the programs were announced. Table 1 shows that China's market share more than doubled across all major ship types. Importantly, this market share expansion was driven by both existing shipyards, as well as the new entrant shipyards: the latter accounted for about half (55%) of the output expansion in the case of Handysize bulk carriers (on which the empirical analysis will focus). Figure 2 plots Chinese production of all, new and existing yards over time.

Production in China is also split between public and private firms. A substantial number of

---

[10]In its 2014 report the US-China Economic and Security Commission (a US government body established to monitor and investigate national security implications of the bilateral trade and economic relationship between the United States and China) states that "a full identification of support programmes was not possible for the Secretariat, as they are often the result of internal administrative measures that are not always easy to identify and generally only available in Chinese. In addition, the budget is not a public document hence it is not possible to identify outlays."

[11]This section draws from OECD (2008), Collins and Grubb (2008) and Stopford (2009).

[12]The plan was introduced by the National Development and Reform Commission (NDRC) and the Commission of Science, Technology and Industry for National Defense (COSTIND).
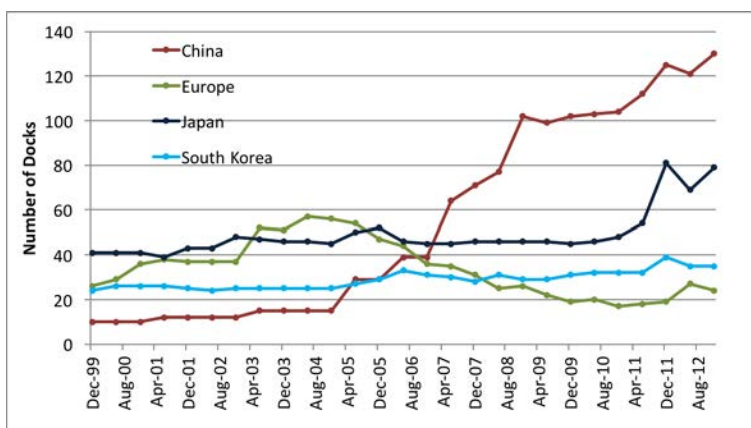
Figure 1: Shipbuilding dry docks.

	Chinese Market Share, Pre-2006	Chinese Market Share, Post-2006
Bulk carriers	0.17	0.57
Tankers	0.15	0.28
Containerships	0.16	0.39
Gas carriers (LNG/LPG)	0.07	0.21

Table 1: China's average quarterly market share before and after 2006.

shipyards operate under the umbrella of two major state-owned enterprises directly administered by the Chinese central government (CSSC and CSIC), or under provincial and local governments. These public firms account for about 40% of annual Chinese ship production on average. It is worth noting that the day-to-day operations and contract bids are handled directly by the individual shipyards, even if the long-term design is carried out at the federal or local level.[13] At the same time, a large number of private firms and foreign joint ventures are active. The private firms accounted for about 40% of the output expansion post 2006. Overall, the global industry is fairly unconcentrated at the firm level with hundreds of firms and low market shares; for instance, the average annual market share in China is 0.8%.[14] In addition, China's shipbuilding is mostly geared towards export sales which comprised about 80% of its production in 2006.

Finally, commercial ships are the largest factory produced product. Materials account for about half the cost of the ship (steel is about 13%) and labor (mostly low skill tasks) about 17% of total cost. This paper focuses on cargo transportation and in particular, dry bulk shipping, which concerns vessels designed to carry a homogeneous unpacked cargo (mostly raw materials), for individual shippers on non-scheduled routes. The bulk shipping market consists of a large number of small shipowning firms (largest firm has share 3%). See Kalouptsidi (2014) for a detailed

---

[13]A significant portion of CSSC and CSIC financial resources have been devoted to expanding China's shipbuilding infrastructure. As an example, "CSSC had multibillion-dollar projects under way to build new shipyards on Changxing Island in Shanghai and Longxue Island in Guangzhou" (Collins and Grubb (2008)).

[14]The maximum market share ever observed in my sample is about 11%, conditional on positive production.
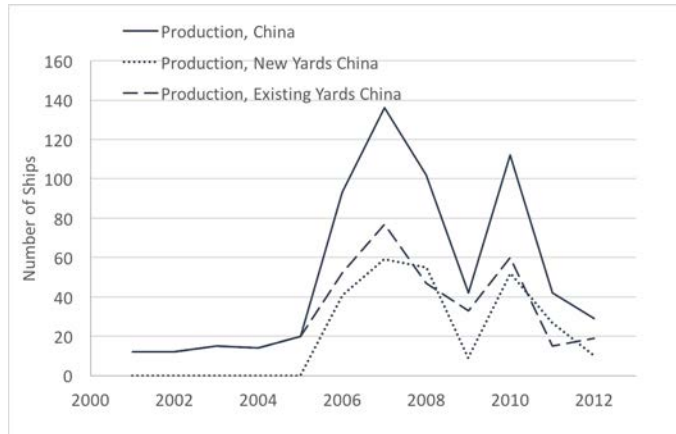
Figure 2: Chinese production expansion for all, new and existing shipyards.

description of the bulk shipping industry.

3 Model

In this section, I present a dynamic model for world shipbuilders, following the tradition of Hopenhayn (1992) and Ericson and Pakes (1995). In Section 3.1, I start by presenting a toy model of shipbuilding subsidies to showcase the main mechanisms at play and to aid the reader in following the main model of Section 3.2. Indeed, in Section 3.2 I substantially enrich this toy model to make it more realistic and amenable to empirical work.

3.1 A Toy Model of Shipbuilding Subsidies

The industry consists of N_c Chinese firms (shipyards) and N_{nc} non-Chinese firms, and both N_c and N_{nc} are large. In each period t, each firm j produces quantity (ships) q_{jt} at cost $c_{jt}(q_{jt})$. The cost function of ship production $c_{jt}(\cdot)$ is quadratic in q_{jt}, and depends on characteristics of firm j, x_{jt}. For instance, consider the cost function:

$$c_{jt}(q_{jt}) = c_1' x_{jt} q_{jt} + c_2 q_{jt}^2$$

where $'$ denotes transpose and c_1, c_2 are parameters.

Shipyards are price-takers in the world market for ships and face price P_t. The ship price fluctuates over time and depends on world market conditions, such as the demand for shipping services and the total fleet in period t, which captures the competition that shipowners are facing. As shipyards build more ships, they reduce the shipowners' willingness to pay, since the latter expect lower profits.[15] Let the aggregate demand for ships be $P_t = \bar{\eta}_{0t} - \eta_1 Y_t$, where $\bar{\eta}_{0t}$ captures

---

[15]Note that each individual shipyard is too small to affect the total fleet (stock of ships) with its period t production.

9

the international trade conditions and Y_t is the total fleet. The fleet evolves according to the rule: $Y_t = (1-\delta)Y_{t-1} + Q_t$ where Q_t is total ship construction at t and δ captures ship exit or depreciation. Letting $\eta_{0t} = \bar{\eta}_{0t} - \eta_1(1-\delta)Y_{t-1}$, the demand curve can be written as $P_t = \eta_{0t} - \eta_1 Q_t$.

Shipyard j solves:

$$\max_q P_t q - c_1' x_{jt} q - c_2 q^2$$

For simplicity, let x_{jt} only indicate whether the shipyard is Chinese or not, so that $c_1' x_{jt} \in \{c_{c1}, c_{nc1}\}$, where subscript c stands for Chinese and nc for non-Chinese. Assume also that $c_{c1} > c_{nc1}$, so that Chinese shipyards are less efficient than non-Chinese shipyards. Shipyard j's optimal production is $q_{jt}^* = (P_t - c_{j1})/2c_2$, for $j \in \{c, nc\}$. Moreover, the equilibrium world ship price is $P_t = (2c_2\eta_{0t} + \eta_1 c_{c1} N_c + \eta_1 c_{nc1} N_{nc})/(2c_2 + \eta_1(N_c + N_{nc}))$.

Now suppose that China provides a subsidy of κ dollars for every ship built, so that the cost function of Chinese firms becomes $(c_{c1}' - \kappa)q_{jt} + c_2 q_{jt}^2$. Then, the optimal production of a Chinese yard becomes $q_{ct}^{**} = (P_t - c_{c1}' + \kappa)/2c_2 > q_{ct}^*$ and the price falls at $P_t = (2c_2\eta_{0t} + \eta_1(c_{c1} - \kappa)N_c + \eta_1 c_{nc1} N_{nc})/(2c_2 + \eta_1(N_c + N_{nc}))$. Essentially, Chinese shipyards become more "efficient" due to their lower cost of production and "steal" market share from firms in other countries whose production declines.

Next suppose that the Chinese government also sponsors new facilities so that the number of Chinese firms grows from N_c to N_c'. In this case, the market price declines, total Chinese production rises and total non-Chinese production falls, as long as the efficiency gap between Chinese and non-Chinese firms, $c_{c1} - c_{1nc}$, is lower than a certain threshold.[16] The higher the per ship subsidy κ, the more likely it is that total Chinese production goes up as well. In other words, conditional on increasing the number of shipyards N_c, the Chinese government might wish to also subsidize their costs to guarantee higher overall production. From a welfare perspective, there are obvious winners and losers. Non-Chinese firms lose market share and profits. Consumers gain because prices go down. Chinese shipyards gain, but at the cost of subsidies.

Figure 3 depicts these equilibria from the firm's and industry's perspective.

Although this toy model illustrates some of the main mechanisms at play, it is too stylized to identify and quantitatively evaluate government subsidies which is the goal of this paper. In the next section I enrich the model in a number of important dimensions. First, I model explicitly the entry behavior of shipowners which essentially constitutes the demand (when a shipowner buys a ship, he effectively enters the shipping industry). Second, building ships may entail dynamic considerations. Ship production does not happen instantly; rather, a ship takes a number of years to be completed (two to five on average) and thus shipyards accumulate backlogs and have ship delivery schedules. This past production (the backlog) may be an important determinant of costs, due to the accumulation of expertise (Thompson (2001)), economies of scale or capacity

---

[16]The trade-off lies in having more firms, each of whom produces less. To be specific, the equilibrium price is decreasing in N_c (i.e. $\partial P/\partial N_c \leq 0$) if and only if $c_{c1} - c_{nc1} \leq 2c_2(\eta_{0t} - c_{c1})/\eta_1 N_{nc}$, ignoring κ. This is less likely to hold if the market is of low value (c_{1c} is close to η_{0t}) or the number of non-Chinese firms is very large.
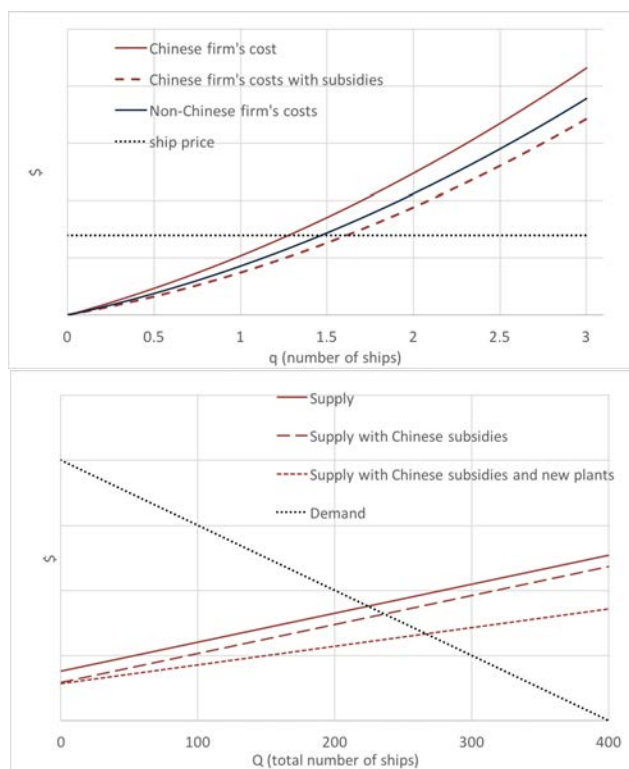
Figure 3: Firm and industry demand and supply for ships.

constraints. This feature is potentially important both for the identification and the evaluation of the impact of subsidies. Therefore, the model of Section 3.2 below involves dynamics in both the demand and supply. Third, I consider discrete production, as each ship is a large project in a firm's production line.

3.2 A Modeling Framework for the Shipbuilding Industry

I now turn to the main model. A key input is demand for new ships, which stems from the decisions of downstream shipowners to purchase new vessels. I therefore begin this section by describing shipowner behavior. Next, I turn to upstream shipbuilder behavior, which is the focus of this paper. I also discuss how government subsidies enter. Variables with superscript "o" refer to shipowners and "y" to yards, while subscripts i and j refer to a particular owner and yard respectively.

3.2.1 Demand for New Ships (Shipowners)

Time is discrete and the horizon is infinite. There is a finite number of incumbent shipowners (the fleet) and a large number of identical potential entrant shipowners. I assume constant returns to scale, so that a shipowner is indistinguishable from his ship. Ships are long-lived.

Shipowner state space   The state variable of ship i at time t, s_{it}^o, includes its:

1. age $\in \{0, 1, ..., A\}$

2. country where built $\in C$

while the industry state, s_t, includes the:

1. distribution of characteristics in s_{it}^o over the fleet, $\overline{s_t^o} \in \mathbb{R}^{A \times ||\mathcal{C}||}$.[17]

2. industry backlog $b_t \in \mathbb{R}^{J \times \overline{T}}$, whose $(j,k)^{th}$ element is the number of ships under construction at time t and scheduled to be delivered at $t+k$ by shipyard $j = 1, ..., J$ and $\overline{T}$ is the maximum possible time to build.[18]

3. aggregate demand for shipping services, $d_t \in \mathbb{R}^+$.

4. price of steel, $l_t \in \mathbb{R}^+$.[19]

Shipowner per period payoffs   In period t, each shipowner i chooses how much transportation (i.e. voyages travelled) to offer, q_{it}^o. Shipowners face the inverse demand curve:

$$P_t^o = P^o (d_t, Q_t^o) \tag{1}$$

where P_t^o is the price per voyage, d_t captures demand shifters, such as world industrial production and commodity prices and Q_t^o denotes the total voyages offered, so that $Q_t^o = \sum_i q_{it}^o$. Voyages are a homogeneous good, but shipowners face heterogeneous convex costs of freight. Ship operating costs increase with the ship's age and may differ based on country of built because of varying quality.

I assume that shipowners act as price-takers in the market for freight. Their resulting per period payoffs are $\pi^o (s_{it}^o, s_t)$.[20]

A ship lives a maximum of A periods. At the same time, a ship can be hit by an exit shock each period.[21] In particular, I assume that a ship at state (s_{it}^o, s_t) exits with probability $\delta (s_{it}^o, s_t)$ and receives a deterministic scrap value $\phi (s_{it}^o, s_t)$. Note that ships exit with certainty at age A.[22]

---

[17]I.e. $\overline{s_t^o}$ is a $A \times ||\mathcal{C}||$ matrix whose (i,j) element is the number of ships that are i years old and built in country j.

[18]For instance, suppose no ship takes more than eight periods to be built; then row j of b_t is $[b_{jt1}, b_{jt2}, ..., b_{jt8}]$ and indicates the scheduled deliveries of yard j for the next 8 periods.

[19]The steel price is part of the state because it: (i) is a key determinant of shipyard production costs; (ii) determines the ship's scrap value.

[20]More accurately, ship profits from freight are $\pi^o (s_{it}^o, \overline{s_t^o}, d_t)$; even though the backlog and steel prices do not affect current payoffs, they affect state transitions and scrap values and are thus part of the state.

[21]Shipowners scrap their ships by selling them to scrapyards where they are dismantled and their steel hull is recycled.

[22]Allowing for endogenous exit is straightforward (see Kalouptsidi (2014)).

Shipowner behavior   The ship's value function, $V^o(s^o_{it}, s_t)$, satisfies the following Bellman equation:

$$V^o(s^o_{it}, s_t) = \pi^o(s^o_{it}, s_t) + \delta(s^o_{it}, s_t)\phi(s^o_{it}, s_t) + (1 - \delta(s^o_{it}, s_t))\beta E\left[V^o\left(s^o_{it+1}, s_{t+1}\right) | s^o_{it}, s_t\right] \quad (2)$$

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

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^N_{jt} 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. Therefore, the following equilibrium condition holds:

$$P^N_{jt} = E\left[\beta^{T_{jt}} V^o\left(s^o_{it+T_{jt}}, s_{t+T_{jt}}\right) | s^o_{it}, s_t\right] \quad (3)$$

where s^o_{it} in this case involves ship age equal to zero and the country of yard j. This is the ship price of the toy model of Section 3.1.

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

$$P^{SH}_{it} = V^o(s^o_{it}, s_t) \quad (4)$$

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

3.2.2   Supply of New Ships (Shipyards)

There are J long-lived incumbent shipbuilders.

Shipyard state space and state transitions   The state variable of shipyard j at time t, s^y_{jt}, includes its:

1. backlog $b_{jt} \in \mathbb{R}^{\overline{T}}$ whose k^{th} element is the number of ships under construction scheduled for

delivery at time $t + k$.[23]

2. country.

3. other characteristics, such as: age, capital equipment (number of docks and berths, length of largest dock), number of employees.

Shipyards also share the industry state, s_t. I allow time itself to be part of the aggregate state as well. Allowing for time to enter the agents' decision-making is important in this application, as my empirical analysis of detecting government subsidies hinges on allowing time-varying factors to affect costs.

The shipyard's backlog changes based on its production choice (more on this below). The demand state variable d_t and steel prices l_t evolve according to exogenous first order autoregressive processes (see Section 5.1.2). The evolution of all other states (for the shipowner, shipyard and aggregate) is deterministic (see Section 5.1.2).

Shipyard actions and per period payoffs In period t, shipyard j draws a private iid (across j and t) production cost shock ε_{jt}, with $\varepsilon_{jt} \sim N(0,1)$, and makes a discrete production decision $q_{jt} \in \{0, 1, ..., \overline{q}\}$. Shipyard j faces production costs, $C\left(q, s_{jt}^y, s_t, \varepsilon_{jt}\right)$. Even though q_{jt} is an integer I assume that the cost function can be defined over the real interval $[0, \overline{q}]$ and that as such it is convex in q, in line with the toy model of Section 3.1. I also assume that the cost shock ε_{jt} is paid for each produced unit, so that the total per period cost is:

$$C\left(q, s_{jt}^y, s_t, \varepsilon_{jt}\right) = c\left(q, s_{jt}^y, s_t\right) + \sigma q \varepsilon_{jt} \tag{5}$$

In this model q_{jt} corresponds to the number of ships ordered in period t at shipyard j.[24] As mentioned previously, producing a ship takes time and thus period t's ship orders enter the shipyard's backlog and are delivered a number of years later (Kydland and Prescott (1982)). In particular, period t production, q_{jt}, enters 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.[25] Under demand uncertainty, therefore, undertaking a ship order becomes a dynamic choice. To capture this dynamic feedback, I 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 backlog can lower costs because of the accumulation of expertise, or economies

---

[23]Note that each b_{jt}, all j, is a row in the total backlog matrix b_t.

[24]I consider the number of orders as the relevant choice variable (as opposed to using the number of deliveries or smoothing orders) because the observed ship prices are paid at the order date and may be dramatically different from the prevailing prices at the delivery date.

[25]Suppose the backlog of yard j at t is $b_{jt} = [b_{jt1}, b_{jt2}, ..., b_{jt\overline{T}}]$ and that the yard received q_{jt} orders to be delivered at $T_{jt} = t + k$ periods with $k \neq \overline{T}$. Then, at $t + 1$, the yard delivered b_{jt1} and $b_{jt+1} = [b_{jt2}, ..., b_{jt+k}, q_{jt+1} + b_{jt+k+1}, b_{jt+k+2}, ..., b_{jt\overline{T}}, 0] \in \mathbb{R}^{\mathbb{T}}$.

of scale (e.g. in ordering inputs); on the other hand, increased backlog can raise costs because of capacity constraints (e.g. less available labor).[26] Allowing for past production to affect current costs is potentially important both for measuring and evaluating the impact of subsidies.

Time to build is shipyard-specific and in particular, $T_{jt} = T\left(s_{jt}^y, s_t\right)$. In words, the time to build that shipyard j can offer at t depends on the shipyard's characteristics (e.g. backlog, docks/berths), as well as the aggregate state (e.g. time, aggregate demand and queuing); one can think of this as a reduced form capturing different delivery times over shipyards and time.

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

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

where $s_{it+T_{jt}}^o$ has zero ship age and the country of yard j.[27] Note that $VE_j^o\left(s_t\right)$ does not explicitly depend on period t's production, q_{jt}; in other words yards do not face a downward sloping demand curve, in line with the toy model of Section 3.1. Indeed, q_{jt} affects the willingness to pay for the ship by entering into the total backlog b_t and from there into the fleet after T_{jt} periods. Typically, q_{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, can ignore the impact it has on $VE_j^o\left(s_t\right)$; note however, that aggregates do matter so that as the total fleet increases, shipowners' willingness to pay falls, all else equal. This latter impact is captured by the aggregate demand of the toy model of Section 3.1.

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

$$V^y\left(s_{jt}^y, s_t, \varepsilon_{jt}\right) = \max_{0 \leq q \leq \overline{q}} VE_j^o\left(s_t\right)q - c\left(q, s_{jt}^y, s_t\right) - \sigma q \varepsilon_{jt} + \beta E\left[V^y\left(s_{jt+1}^y, s_{t+1}, \varepsilon_{jt+1}\right)|s_{jt}^y, s_t, q\right] \tag{7}$$

To ease notation, I also define the continuation value:

$$CV^y\left(s_{jt}^y, s_t, q\right) \equiv E\left[V^y\left(s_{jt+1}^y, s_{t+1}, \varepsilon_{jt+1}\right)|s_{jt}^y, s_t, q\right] \tag{8}$$

The expectation in (7) is over demand for shipping services, d_t, steel prices, l_t, production by other shipyards and j's future cost shocks.

Under convex costs, the shipyard's optimal policy amounts to comparing each production level q to $q+1$ and $q-1$, as stated in the following intuitive lemma:

Lemma 1. *If the shipbuilding cost function $C\left(q, \cdot\right) : [0, \overline{q}] \to \mathbb{R}$, is convex in q, then the shipyard's*

---

[26]Here, the shipyard's backlog also affects its demand, as it increases the time shipowners have to wait for delivery.

[27]Note that the willingness to pay for a new ship from yard j depends only on its country of origin, not j itself. The hundreds of shipyards encountered in the data make this generalization impossible in this empirical application.

optimal policy is given by:

$$q^* \left(s_{jt}^y, s_t, \varepsilon_{jt} \right) = \begin{cases} 0, & if \ \varepsilon_{jt} \geq A \left(s_{jt}^y, s_t, 0 \right) \\[2em] q, & if \ \varepsilon_{jt} \in \left[\ A \left(s_{jt}^y, s_t, q-1 \right), A \left(s_{jt}^y, s_t, q \right) \ \right] \\[2em] \overline{q}, & if \ \varepsilon_{jt} \leq A \left(s_{jt}^y, s_t, \overline{q}-1 \right) \end{cases} \tag{9}$$

where

$$A \left(s_{jt}^y, s_t, q \right) \equiv \frac{1}{\sigma} \left[\begin{array}{l} VE_j^o \left(s_t \right) + \left(c \left(q, s_{jt}^y, s_t \right) - c \left(q+1, s_{jt}^y, s_t \right) \right) + \\[1em] +\beta \left(CV^y \left(s_{jt}^y, s_t, q+1 \right) - CV^y \left(s_{jt}^y, s_t, q \right) \right) \end{array} \right]$$

for $q = 0, 1, ..., \overline{q} - 1$, are the optimal policy thresholds.

Proof. See the Appendix. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ □

Timing and Equilibrium The timing in each period is as follows: incumbent and potential entrant shipowners observe their state (s_{it}^o, s_t), while shipbuilders observe their state (s_{jt}^y, s_t). Shipowners are hit by exit shocks and shipbuilders observe their private production cost shocks. Shipyards make production decisions. Next, shipowners receive profits from freight services and shipyards receive profits from new ship production. Exiting ships receive their scrap value $\phi(s_{it}^o, s_t)$. Finally, states are updated.

I consider a competitive equilibrium which consists of an optimal production policy function $q^* \left(s_{jt}^y, s_t, \varepsilon_{jt} \right)$ that is given by (9), as well as value functions $V^y \left(s_{jt}^y, s_t \right)$ and $V^o \left(s_{it}^o, s_t \right)$ that satisfy (7) and (2) respectively, while all expectations employ $q^* \left(s_{jt}^y, s_t, \varepsilon_{jt} \right)$. Existence of equilibrium follows from Doraszelski and Satterthwaite (2010), Hopenhayn (1992), and Jovanovic (1982).

Finally, I assume that China's plan was an unexpected, one-shot, permanent and immediate change from the point of view of industry participants. Here, there are two equilibria; one before and one after 2006. The post-2006 worlds differs in three dimensions: (i) it has a larger number of Chinese shipyards; (ii) Chinese shipyards obtain subsidies and thus have a lower cost function; (iii) shipyards can have a different number of docks/berths (found in s_{jt}^y).[28]

4 Data and Descriptive Evidence

Data I employ five different datasets, all from Clarksons Research. The first reports shipbuilding quarterly production (i.e. orders) between Q1-2001 and Q3-2012. For each shipyard and quarter I observe its bulk ship production in tons and numbers, as well as the yard's backlog and

---

[28]I am essentially assuming that China is the only country subsidizing. Alternatively, one can think of the recovered subsidies as the differential pre/post China and in comparison to other countries.

time to build. There are 192 yards that produce Handysize vessels (the segment on which my empirical analysis will focus), of which 119 are Chinese, 41 are Japanese, 21 are S. Korean and 11 are European. Of the 119 Chinese shipyards, 57 entered after 2005. The majority of bulk ship production occurs in China and Japan; hence even though I include Europe and S. Korea in the estimation and counterfactuals, most comparisons will be made between China and Japan.

As shown in Table 2, quarterly production at the shipyard level is quite low and often zero. Japan produces on average more than China, while Europe rarely produces over the sample period.

The second dataset is a sample of shipbuilding contracts, between August 1998 and August 2012. It reports the order and delivery dates, the shipyard and price in million US dollars. Unfortunately, prices are reported for only a very small fraction of contracts. Indeed, especially before 2006 most contracts involve missing prices. In addition, for shipyard-quarter combinations that involve zero production, the corresponding price does not exist by default.

To deal with these issues, I introduce a dataset of second-hand ship sale transactions, between August 1998 and August 2012. The dataset reports the date of the transaction, the name and age of the ship, as well as the price in million US dollars. I end up with 418 observations of new ship contracts and 2016 observations of second-hand sale contracts (2434 total), of which 1173 are pre-2006 and 1261 are post-2006.

As shown in Table 2, new ship prices are on average substantially higher than second-hand prices, though this is driven primarily by ship age. Prices are also very volatile, driven mostly by time series variation. Indeed, in Table 3, I regress new and used prices on ship age, country of built and quarter dummies. The R-squared is high, at 0.82. The age coefficient in the used price regressions suggests that a very young ship in the second hand market commands a price roughly equal to that of a new ship. Finally, Chinese ships are slightly cheaper than Japanese and S. Korean.

The fourth dataset employed is a snapshot of shipyard characteristics in 2012, such as shipyards' first year of ship delivery, location, number of dry docks and berths, length of the largest dock, number of employees, total past output. The first year of delivery is used to compute the shipyard's age.[29] The number of docks and berths is a measure of capacity, since production bottlenecks occur during the assembly operations done on the docks/berths. I allow the capital infrastructure of yards (i.e. docks/berths and length) to be different before and after China's plan.[30] The adoption of a pre- and post-2006 capital infrastructure level is consistent with the modeling assumption of two equilibria before and after China's plan. I drop shipyards with missing capital measures (docks/berths) so that the end shipyard sample consists of 4741 shipyard-quarter observations (all results are robust if the full sample is used and docks/berths are omitted as a covariate of costs). As shown in Table 2, a shipyard has on average about 3 docks/berths and there is sizable

---

[29]Because of time to build, I subtract a number of years from every first delivery year of all shipyards, after consulting with Clarksons' analysts. The results reported subtract 3 years (similar findings were obtained when the number of years subtracted varies).

[30]I collect the extra data from a Clarksons publication ("World Shipyard Monitor").

	Mean	St. Dev.	Min	Max
Quarterly firm production (no of ships)	0.35	1.37	0	22
Quarterly firm backlog (no of ships)	3.7	7.2	0	59
Quarterly firm production (no of ships), China	0.28	1.3	0	22
Quarterly firm production (no of ships), Japan	0.48	1.47	0	15
New ship price, million US dollars	28.4	8	6.3	51
Used ship price, million US dollars	9.9	8.6	0.6	56
Firm number of docks/berths, China	3.4	3.4	1	23
Firm number of docks/berths, Japan	2.5	1.96	1	9

Table 2: Sample summary statistics.

	I New Ship Price	II Used Ship Price
Constant	14.8**	17.1**
	(1.5)	(0.69)
Age	-	−0.76**
		(0.01)
Built in Europe	−13**	−0.86**
	(3.6)	(0.42)
Built in Japan	1.7*	0.36
	(0.93)	(0.36)
Built in S. Korea	3.4**	0.53
	(0.51)	(0.51)
Quarter dummies	Yes	Yes
R^2	0.82	0.83
Obs.	418	2016

* $p < 0.1$, ** $p < 0.05$

Table 3: Hedonic regression of new and used ships.

heterogeneity.

Finally, the fifth dataset consists of quarterly time-series between 1998 and 2012 for the fleet, total backlog, orders of new ships, deliveries and demolitions between 1998 and 2012. This dataset is used to create the states and estimate their transition. I also obtain quarterly time-series of Japan's steel ship plate commodity price in dollars per ton during the same time period.

Descriptive Evidence What patterns of the raw data are consistent with the presence of subsidies? One might expect that new ship prices should react around 2006. As discussed above, however, the sparsity of new ship prices (especially in the pre era) makes it impossible to explore this. Used ship prices, however, should also display a reaction. Figure 4 uses the hedonic regression of Table 3 and shows that indeed there is a short-lived drop in late 2005, in a period when ship

prices are trending upward due to increased demand for freight (shifts in demand are also the reason the price decline is not permanent in the raw data).[31] Of course, this finding is not proof of cost subsidies; yet if no drop were observed, one may have been concerned about the impact of this policy.
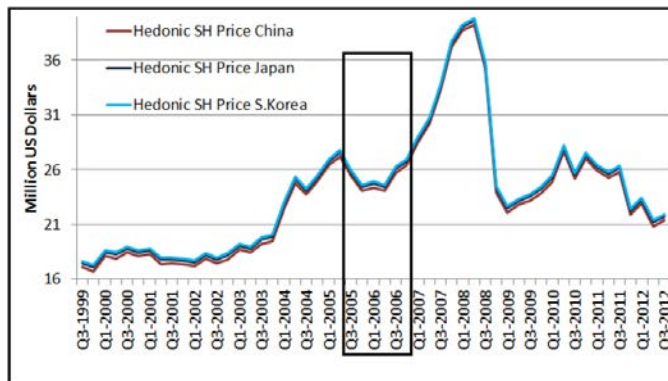
Figure 4: Hedonic regression of used ship prices on ship age, country and quarter dummies.

Despite the importance of a price response, the main insight of this paper, in terms of identifying/measuring subsidies, is that production patterns are equally important. Figure 5 depicts the evolution of China's market share. Between 2005 and 2006, China experiences a large, rapid increase in market share. In this paper, I employ precisely this rapid increase in production (in combination with prices) to identify changes in costs that are consistent with the presence of subsidies. At the same time, the constructed model can control for the different factors leading to this increase and provide a measurement of the alleged subsidies, a fundamental input into subsidy disputes. When I come to the results, I discuss alternative explanations for this pattern (e.g. productivity changes/technological improvements) and argue that they cannot account for the observed industry evolution and the speed of China's upswing.
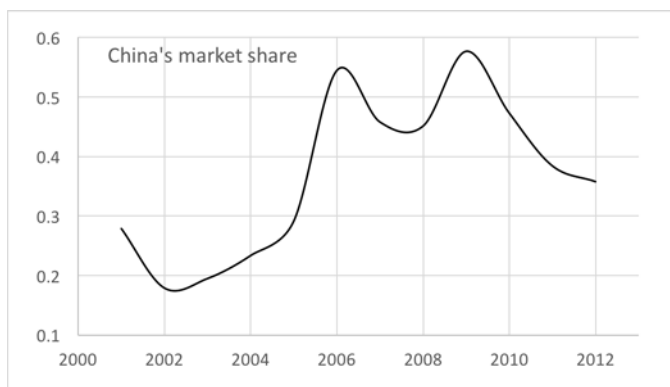
Figure 5: China's market share.

---

[31]I have unsuccessfully searched extensively in industry publications for alternative explanations.

5 Model Estimation and Detection of Subsidies

To see the main idea behind the subsidy detection method, consider the toy model of Section 3.1; i.e. a static, perfectly competitive market, so that $P_{jt}^N = MC_{jt}$ for all j and t. In that case, to detect subsidies one would simply look for a break in observed prices in 2006, since prices are in fact the marginal costs. In my setup, there are two complications: (i) I do not observe enough prices of new ships, and (ii) there are dynamics in the production decision. To address (i), I complement with used ship prices; to address (ii) I use the shipyard's first order condition from its dynamic optimization. Overall the detection methodology relies on a firm level dataset (here, quantities and prices) and some knowledge about the date of the subsidy program implementation. Although the model is tailored to the shipbuilding industry, one can apply the same idea to other setups under suitable adjustments.

The proposed strategy proceeds in two steps. In the first step, I recover the demand curve that shipbuilders face, which 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, as well as shipowner expectations. The second step inserts the estimated willingness to pay for a ship into the optimization problem of shipbuilders to recover their costs.

Each estimation task is described below and followed by the results. A time period is a quarter. I consider Q3-2005 as the first quarter of the post world, consistent with Figures 1 and 4; all results are robust to alternative thresholds around that date. Finally, all results presented are for bulk vessels, in particular, Handysize. There are three good reasons to focus on bulk carriers: China was already an important player before 2006; their production process is not characterized by significant technological advances; product differentiation is limited.

5.1 Estimation of the Willingness to Pay for a New Ship

In this step, I create the state space and then estimate state transitions and ship value functions.

5.1.1 State Space

The ship's state space includes its age and country of built, as well as the industry state, s_t (fleet age distribution, backlog, steel prices and demand for shipping services). All ship states are directly observed in the data except for the demand for shipping services, d_t. I construct d_t following Kalouptsidi (2014) by estimating the demand curve for shipping services (1) and using its intercept. The empirical analog of (1) that I estimate is:

$$P_t^o = \alpha_0^d + \alpha_1^d X_t^d + \alpha_2^d Q_t^o + \varepsilon_t^d$$

where P_t^o is the price per voyage, X_t^d are demand shifters (e.g. world commodity prices and production) and Q_t^o is the number of voyages. This is estimated by IV regression and I set

$d_t \equiv \hat{\alpha}_0^d + \hat{\alpha}_1^d X_t^d + \hat{\varepsilon}_t^d$. Details and estimates are presented in the Appendix.

5.1.2   State Transitions

In order to compute the value of entering the shipping industry, defined in (6), I need shipowner expectations over (s_{it}^o, s_t). The transition of s_{it}^o 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 very large (in the case of four countries, $\overline{s_t^o}$ has dimension $4 \times A$ -where A is a ship's maximum age; while the backlog b_t has dimension $J \times \overline{T}$ which in my sample is in the order of several thousand); on the other hand, updating b_t requires optimal production policies for all shipyards. Instead of working with the true transitions, I follow Barwick and Pathak (2012) who assume that s_t follows a vector autoregressive (VAR) model; this is similar to the first step of two-step estimation procedures for dynamic games (e.g. Bajari, Benkard and Levin (2007) and Pakes, Ostrovsky and Berry (2007)).

To deal with the state dimension, I make the following simplifying assumptions, in the spirit of Krussel and Smith (1996). First, I replace the fleet distribution, $\overline{s_t^o}$, with two age groups $\left(\overline{s_{1t}^o}, \overline{s_{2t}^o}\right)$: the number of ships below 20 years old and the number of ships above 20 years old. I do not use the distribution of the fleet over country of built because its evolution is extremely slow. Second, I replace the backlog, b_t, with the total backlog $B_t = \sum_{j,l} b_{jtl}$. In Section 7, I present a number of robustness exercises regarding these simplifying assumptions.

I consider the vector autoregression (VAR) model:

$$s_t = R_o + R s_{t-1} + \xi_t$$

where $\xi_t \sim N(0, \Sigma_\xi)$. Note that the parameter vector R_0 and the parameter matrix R can 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, shipowners are likely to believe that all else equal the supply of ships is permanently higher post 2006.

As discussed above, demand d_t and steel price l_t are exogenous to the model and are unaffected by the pre/post-2006 regime. Essentially I assume they are autonomous AR(1) processes. In contrast, I allow $\left(\overline{s_{1t}^o}, \overline{s_{2t}^o}, B_t\right)$ to be affected by all variables to account for ship entry and exit. The baseline specification allows only R_0 to change before and after 2006. I estimate the parameters of interest (R_0, R, Σ_ξ) via OLS separately for each variable (Hamilton (1994)).

Table 15 in the Appendix reports the results and provides more details. All variables are persistent. Signs are also in general as expected: $\overline{s_{1t}^o}$ is increasing in the backlog and demand and decreasing in steel prices (as steel prices increase, exit increases); $\overline{s_{2t}^o}$ is decreasing in $\overline{s_{1t}^o}$, as more young ships increase exit, and increasing in demand which leads to less exit; the backlog is increasing in demand. All eigenvalues of R lie inside the unit circle so that the model is stationary.

Finally, the post-2006 world's steady state has significantly higher fleet. In Section 7, I present a number of robustness checks on the transition process for state variables, including a somewhat less sparse model that lies closer to the actual transitions.

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, I 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 (3) and (4) are:

$$P_{jt}^N = E\left[\beta^{T_{jt}} V^o\left(s_{it+T_{jt}}^o, s_{t+T_{jt}}\right) | s_{it}^o, s_t\right] + \zeta_{jt}^N \tag{10}$$

$$P_{it}^{SH} = V^o\left(s_{it}^o, s_t\right) + \zeta_{it}^{SH} \tag{11}$$

where ζ_{it}^{SH} and ζ_{jt}^N are measurement error. Kalouptsidi (2014) employs 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. Here, however, the addition of the new ship prices requires the combination of (10) and (11) in a single estimation step, which is not straightforward. Indeed, even though both V and VE can be written directly as a function of the state, time to build implies that they cannot be written as the *same* function of the state. Indeed, in V the owner already has the ship, while in VE the owner will obtain it in a number of periods. Thus, to combine (10) and (11) in one regression, one needs to model agent expectations. I next describe the approach.

I use a flexible linear sieve approximation for the value function:

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

where $f^o(\cdot)$ is a polynomial function in the ship state (s_{it}^o, s_t) and γ^o is a (sparse) parameter vector. Then:

$$P_{it}^{SH} = \gamma^o f^o\left(s_{it}^o, s_t\right) + \zeta_{it}^{SH} \tag{12}$$

while

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

$$= \gamma^o \beta^{T_{jt}} \int f^o\left(s_{it+T_{jt}}^o, s_{t+T_{jt}}\right) dP\left(s_{it+T_{jt}}^o, s_{t+T_{jt}} | s_{it}^o, s_t\right) + \zeta_{jt}^N \equiv \gamma^o f_N\left(s_{it}^o, s_t\right) + \zeta_{jt}^N \tag{13}$$

where $P\left(s_{it+1}^o, s_{t+1} | s_{it}^o, s_t\right)$ is the state transition probability function and can be determined by the VAR estimated above. The parameter vector γ^o enters (12) and (13) linearly; yet even though

22

(12) can be estimated in a straightforward manner, (13) requires the computation of the right-hand side integrals. Indeed, (13) involves the expectation of higher order terms of the following vector:

$$s_{t+T} = R^T s_t + \sum_{k=t+1}^{t+T} R^{t+T-k}(R_o + \xi_k) \tag{14}$$

In the Appendix, I derive closed-form expressions for the integrals of up to third order terms in the industry state s_t.

As the dimensionality of the state (s_{it}^o, s_t) is large, computing high order polynomial terms quickly leads to a very large number of regressors in (12) and (13). I therefore use the LASSO which is appropriate for sparse regression problems, i.e. problems that involve a large number of potential regressors, only a small subset of which is important in capturing the regression function accurately. Thus γ^o is estimated from:

$$\min_{\gamma^o} \left\{ \sum_{j,t} \left(P_{jt}^N - \gamma^o f_N \left(s_{it}^o, s_t \right) \right)^2 + \sum_{i,t} \left(P_{it}^{SH} - \gamma^o f^o \left(s_{it}^o, s_t \right) \right)^2 + \lambda \left| \gamma^o \right|_1 \right\}$$

where λ penalizes the number of nonzero coefficients in γ^0. In this application, the regressors $f^o \left(s_{it}^o, s_t \right)$ and $f_N \left(s_{it}^o, s_t \right)$ are third order polynomials in s_t and s_{it}^o, as well as interactions of s_{it}^o 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^o embody equilibrium features which are likely to change in 2006 as agents' valuations are altered. Therefore, in analogy with the VAR formulation, I allow the value function to change before and after by a 2006 dummy. Figure 6 depicts the estimated value function on the observed states for zero year old ships. Consistent with the raw data, Chinese ships are of somewhat lower value, yet the differences are small.[32]
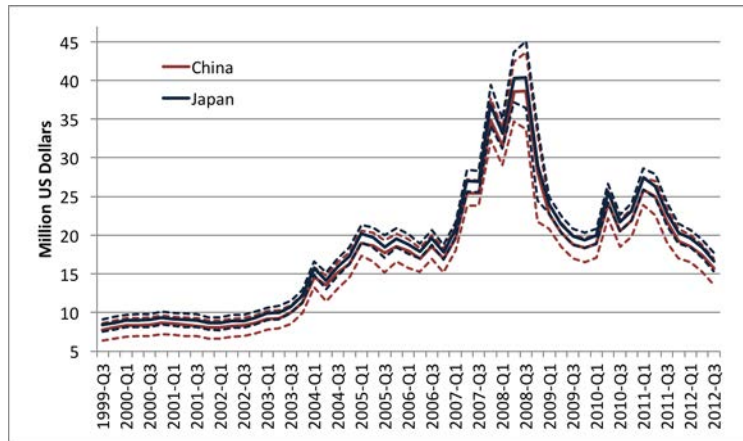
Figure 6: Value function of a ship at age zero. 0.95 bootstrap confidence intervals.

---

[32]Pointwise confidence intervals are computed via 500 bootstrap samples, with the resampling done on the error.

5.2 Shipbuilding Production Cost Function

I next turn to estimating the shipbuilding cost function. I begin with the simple case where shipbuilders are static to illustrate the estimation strategy followed. I present the estimation results along with several robustness exercises. I then proceed to the case of dynamic shipbuilders. Section 7 provides further robustness analysis.

5.2.1 Static Shipbuilders

If shipyard j is myopic it solves:

$$\max_{0 \leq q \leq \bar{q}} VE_j^o(s_t) q - c\left(q, s_{jt}^y, s_t\right) - \sigma \varepsilon_{jt} q$$

This is essentially an ordered choice problem, so to estimate the cost function parameters, I maximize the following likelihood function of observed production decisions over shipyards and time:

$$\prod_{j,t:q_{jt}=0} \Pr\left(q_{jt}=0|s_{jt}^y, s_t\right) \prod_{q=1}^{\bar{q}-1} \prod_{j,t:q_{jt}=q} \Pr\left(q_{jt}=q|s_{jt}^y, s_t\right) \prod_{j,t:q_{jt}=\bar{q}} \Pr\left(q_{jt}=\bar{q}|s_{jt}^y, s_t\right) \tag{15}$$

I consider a cost function of the following flexible form:

$$c\left(q, s_{jt}^y, s_t\right) = c_1' g_1\left(s_{jt}^y, s_t\right) q + c_2\left(q, s_{jt}^y, s_t\right)$$

where $g_1(\cdot)$ is a known vector-valued function of the shipyard characteristics s_{jt}^y and aggregate states s_t (steel prices and time); c_1 is the vector of parameters of the linear component; and $c_2(\cdot)$ is the cost curvature and is convex in the output q. As shown below, I consider several flexible specifications for the latter.

The simple baseline specification assumes that the shipbuilding cost function is quadratic:

$$c\left(q, s_{jt}^y, s_t\right) = c_1' g_1\left(s_{jt}^y, s_t\right) q + c_2' g_2\left(s_{jt}^y, s_t\right) q^2$$

with (c_1', c_2', σ) the parameters of interest. In all specifications, the linear component includes country dummies, as well as a China-post 2006 dummy. Testing whether the latter is different than zero provides evidence of a structural change in China's cost function, for any value of s_{jt}^y, s_t and q.[33]

Table 4 reports the baseline cost function estimates. In all specifications there is a strongly significant decline in China's cost after 2006 in the order of 13-20%, as indicated by the China-post dummy. Multiplying this parameter with China's production, I find that between 2006 and 2012

---

[33]As $VE_j^o(s_t)$ is estimated, to compute standard errors I create 500 bootstrap samples by re-drawing $\left(q_{jt}, s_{jt}^y, s_t\right)$ and combine them with the 500 samples drawn to compute confidence intervals for $VE_j^o(s_t)$. I have also used the block-boostrap where I drew shipyards with replacement, and standard errors are unaffected. Finally, I follow Amemiya (1984) and maximize the likelihood over $\left(\frac{1}{\sigma}, \frac{c_1}{\sigma}, \frac{c_2}{\sigma}\right)$ rather than (c_1, c_2, σ).

Table 4: Baseline static cost estimates

Cost Function $C(q,s) = c_1' g_1(s)q + c_2 q^\gamma$	I	II	III	IV	V	VI	VII	VIII
Linear Term: $c_1' g_1(s)q$								
China	32.00**	32.25**	32.31**	21.48**	32.07**	39.89**	33.66**	30.92**
	(5.24)	(5.05)	(5.10)	(5.30)	(5.46)	(7.71)	(5.20)	(6.51)
China, POST	−6.36**	−6.14**	−7.24**	−2.78**	−5.92**	−8.37**	−6.19**	−6.14**
	(2.19)	(1.93)	(2.98)	(1.56)	(2.28)	(2.93)	(1.96)	(2.10)
Europe	32.16**	32.46**	36.02**	21.58**	32.35**	39.72**	33.78**	31.13**
	(5.68)	(5.40)	(6.85)	(5.56)	(5.65)	(8.28)	(5.46)	(6.50)
Japan	25.91**	26.01**	26.33**	19.54**	25.87**	32.79**	27.57**	24.61**
	(3.25)	(3.25)	(3.37)	(3.66)	(3.32)	(5.37)	(3.37)	(4.29)
Japan, POST					0.24			
					(1.15)			
S.Korea	33.94**	34.41**	38.51**	25.49**	34.40**	39.76**	35.46**	33.20**
	(4.58)	(4.65)	(6.10)	(4.84)	(4.73)	(6.45)	(4.52)	(5.65)
Backlog	−0.63**	−0.61**	−0.63**	−0.29**	−0.60**	−0.84**	−0.59**	−0.62**
	(0.17)	(0.16)	(0.17)	(0.16)	(0.17)	(0.24)	(0.15)	(0.18)
Docks/Berths		−0.14	−0.14		−0.14		−0.15	−0.13
		(0.11)	(0.12)		(0.11)		(0.11)	(0.11)
Max Length		−0.001	−0.001		−0.001		−0.0009	−0.001
		(0.0008)	(0.0009)		(0.0008)		(0.0008)	(0.0008)
Steel Price	0.50**	0.51**	0.52**	0.78**	0.50**	0.50**	0.48**	0.52**
	(0.17)	(0.17)	(0.16)	(0.36)	(0.17)	(0.21)	(0.17)	(0.15)
t	0.14**	0.13**	0.04		0.13**	0.19**	0.13**	0.13**
	(0.05)	(0.05)	(0.06)		(0.06)	(0.07)	(0.05)	(0.05)
China*t			0.13					
			(0.10)					
Japan*t			0.10					
			(0.07)					
Year Dummies	No	No	No	Yes	No	No	No	No
Curvature: $c_2 q^\gamma$								
$c_2 q^2$	1.23**	1.20**	1.24**	0.53*	1.18**			
	(0.35)	(0.32)	(0.34)	(0.33)	(0.36)			
$c_2 * \text{Docks/Berths} * q^2$						0.38**		
						(0.11)		
$c_2 q^3$							0.11**	
							(0.03)	
$c_2 q^\gamma$: c_2								2.47*
								(1.57)
$c_2 q^\gamma$: γ								1.72**
								(0.15)
σ	12.07	11.67	12.13	5.11	11.48	17.31	11.48	11.76
	(3.23)	(3.03)	(3.31)	(3.15)	(3.38)	(4.93)	(2.86)	(3.50)

* $p < 0.1$, ** $p < 0.05$. Standard errors computed from 500 bootstrap samples.

Time t measured in quarters. Countries refer to country dummy variables.

Chinese subsidies cost between 1.5 and 4.5 billion US Dollars. The results suggest some convexity in costs. The coefficient on backlog is negative, implying cost declines due to expertise or economies of scale. This finding is consistent with industry participants' testimony, who claim that shipyards have incentives to produce ships similar to those they already have under construction. In addition, costs are decreasing in capital measures, as expected. Not surprisingly, Europe is the highest cost producer, while either Japan or China post-2006 are the lowest cost producer depending on the specification.

The starting point on the curvature of the cost function is a simple quadratic. Columns I and II are the simplest specifications: they control for the shipyard's backlog, docks/berths, length of the largest dock, as well as for a linear time trend. It is important, however, to reject to the furthest extent possible explanations alternative to subsidies. To do so, in the other columns of Table 4 as well as the following tables, I incorporate rich time variation, perform placebos, adopt very flexible cost curvature, consider only existing shipyards and perform a number of miscellaneous other checks which I discuss next.

Time Variation and Placebos It is important to control for time-varying factors adequately in order to alleviate the concern that the estimated cost declines may be driven by unobserved time variation. The results are robust to any parametric function of time I have tried (e.g. country specific time trends, polynomial trends); as an example, Specification III of Table 4 adds time trends specific to China and Japan. Specification IV moves away from parametric functions of time and adds year dummies; there is still a significant decline in Chinese costs, not surprisingly somewhat lower, at 13%. The most flexible specification in terms of time variation is to estimate China-year dummies. As expected, estimates are more noisy, yet as shown in Figure 7, Chinese costs fall post 2006. One could argue that an arbitrary productivity process or technological improvements can also be consistent with these results. The production process of bulk carriers is old, however, without frequent technological advances. In addition, such a productivity process would have to feature a discontinuity in 2006 in China alone.
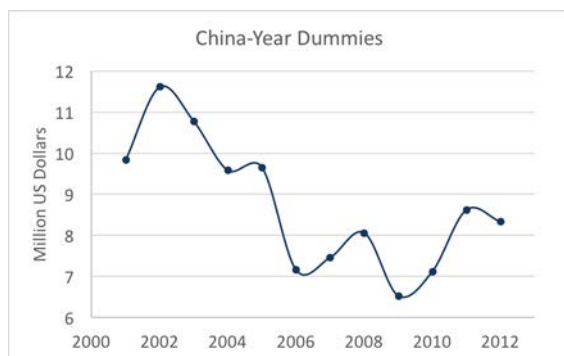

Figure 7: China-year dummies.

Specification V of Table 4 shows that no significant changes occur in 2006 in other countries.

Indeed, I add a Japan-post 2006 dummy, in addition to the China post dummy and find that Japan's costs seem to increase slightly, but the coefficient is not significant (similar findings are obtained if other countries are used).

Cost Curvature Adopting a flexible cost specification is undoubtedly crucial in the subsidy measurement, as misspefication can confound the results. I therefore consider a number of alternative specifications for the cost function and show that the retrieved subsidies are robust.

Columns VI-VIII of Table 4 generalize the quadratic cost function curvature. Column VI makes $c_2' g_2(s_{jt}^y, s_t)$ a linear function of docks/berths, to allow dependence of curvature on capital measures. Results are robust. At the average number of docks/berths the convexity parameter becomes 1.1, close to the simple quadratic estimates. Column VII considers a cubic function of output. Finally, column VIII considers a general power function, so that:

$$c_2\left(q, s_{jt}^y, s_t\right) = c_2 q^\gamma \tag{16}$$

I estimate both c_2 and γ. The estimated power is $\hat{\gamma} = 1.72$ and a LR test cannot reject that it is different from 2.

Next, I turn to a more general cost function that allows for a nonparametric curvature, so that:

$$c\left(q, s_{jt}^y, s_t\right) = c_1' g_1\left(s_{jt}^y, s_t\right) q + c_q \tag{17}$$

I estimate $\{c_0, c_1, ..., c_{\bar{q}}\}$. Note that not all coefficients can be identified; some need to be restricted.[34] Table 5 presents the results for several different restrictions. In Column I, I set $c_0 = 0$ and $c_q = \gamma_1 q^2$ for $q = \bar{q} - 1, \bar{q}$; the parameters to be estimated are $\{c_1, c_2, ..., c_{\bar{q}-2}, \gamma_1\}$. Given the large number of observations with $q = 0$, it might be important to allow for a fixed cost. Thus, in column II of Table 5, I do not restrict c_0 to zero, but rather I set $c_q = \gamma_1 q^2$ for $q = \bar{q} - 2, \bar{q} - 1, \bar{q}$ and estimate $\{c_0, c_1, c_2, ..., c_{\bar{q}-3}, \gamma_1\}$. Next, in column III, I allow c_0 to be country-specific and to change in China pre/post 2006. In other words, this specification allows for subsidies in the fixed cost. Interestingly, the estimated parameters suggest that the cost decline is not manifested in the fixed cost, but rather in positive production levels. The recovered subsidy is robust in all the above specifications.

Overall, the results reveal that the nonparametric cost curvature is rather close to the simple quadratic curvature, suggesting that the latter is a good approximation. One may worry that it is the restriction that the cost for higher quantities follows a quadratic that leads to an overall quadratic-like function. To explore this possibility, I estimate (17) with $c_0 = 0$ and $c_q = \gamma_1 q^{\gamma_2}$, for

---

[34]Roughly, we have $\bar{q}$ equations (the choice probabilities) in the following unknowns: $c_0, ..., c_{10}, c_1', \sigma$. As shown in Kalouptsidi, Scott and Souza-Rodrigues (2016), variation in VE and s_{jt}^y, s_t identifies σ and the corresponding part of c_1'. Thus we are left with $\bar{q} - 1$ equations in $\bar{q} + 2$ unknowns ($c_0, ..., c_{\bar{q}}$ and the constant in c_1'). I either fix the first term (c_0) to zero and force the last terms (e.g. $c_{\bar{q}}, c_{\bar{q}-1}$) to obey a quadratic; or I allow c_0 to be different from zero and force more terms to follow a quadratic. I later generalize the quadratic assumption.

Table 5: Static cost estimates with flexible cost curvature.

Cost Function $C(q,s) = c_1' g_1(s)q + c_q$	I	II	III	IV
Linear Term: $c_1' g_1(s)q$				
China	34.80**	38.48**	35.83**	28.96**
	(5.36)	(7.55)	(5.50)	(6.04)
China, POST	−6.12**	−6.12**	−5.32**	−6.11**
	(1.91)	(2.05)	(1.77)	(2.03)
Europe	35.04**	38.72**	39.36**	29.20**
	(5.63)	(7.73)	(6.80)	(6.37)
Japan	28.49**	32.18**	31.51**	22.66**
	(3.61)	(6.09)	(4.32)	(4.95)
S.Korea	37.09**	40.78**	38.47**	31.27**
	(4.79)	(7.20)	(5.37)	(5.72)
Backlog	−0.62**	−0.62**	−0.62**	−0.62**
	(0.15)	(0.15)	(0.15)	(0.16)
Docks/Berths	−0.13	−0.13	−0.13	−0.13
	(0.11)	(0.11)	(0.11)	(0.12)
Max Length	−0.001	−0.001	−0.0009	−0.001
	(0.0008)	(0.0009)	(0.0008)	(0.0009)
Steel Price	0.53**	0.53**	0.52**	0.53**
	(0.17)	(0.16)	(0.16)	(0.18)
t	0.13**	0.13**	0.13**	0.13**
	(0.04)	(0.05)	(0.05)	(0.05)
Curvature: c_q				
c_0		17.49		4.39
		(19.17)		(4.73)
c_0^{China}			8.96**	
			(3.95)	
$c_0^{China,Post}$			0.88	
			(0.69)	
c_0^{Europe}			12.85**	
			(4.74)	
c_0^{Japan}			11.45**	
			(4.28)	
$c_0^{S.Korea}$			9.12**	
			(4.08)	
c_1	−1.42*	12.38	6.64**	11.63*
	(0.81)	(14.70)	(2.72)	(9.75)
c_2	0.02	10.13	5.58**	22.36*
	(1.47)	(10.38)	(1.80)	(15.28)
c_3	4.96*	11.38*	8.01**	34.92*
	(2.55)	(6.53)	(2.12)	(21.17)
c_4	11.75**	14.47**	12.26**	49.99*
	(4.18)	(4.50)	(3.43)	(27.60)
c_5	21.08**	20.11**	19.05**	66.64*
	(6.52)	(6.73)	(5.44)	(34.45)
c_6	32.01**	27.35**		
	(9.31)	(11.00)		
c_7	44.51**	36.15**		
	(12.55)	(15.88)		
c_8	57.75**			
	(16.01)			
γ_1	0.91**	0.71**	0.76**	3.95
	(0.25)	(0.33)	(0.22)	(3.84)
γ_2				1.58**
				(0.16)
σ	11.77	11.77	11.69	11.75
	(2.92)	(2.96)	(2.87)	(3.25)

* $p < 0.1$, ** $p < 0.05$. Standard errors computed from 500 bootstrap samples.
Time t measured in quarters. Countries refer to country dummy variables.

$q = \bar{q} - 5, ..., \bar{q}$. Results are in Column IV of Table 5. Interestingly, it turns out $\hat{\gamma}_2 = 1.58$.[35]

Existing Shipyards I next reestimate costs using only shipyards that already existed in 2001. Table 6 reports the results, which show that the same cost declines are retrieved when only old shipyards are considered. This finding is consistent with the raw data patterns: existing yards are responsible for a large portion (45%) of the increase in China's production, as discussed in Section 2. Existing yards also produce 56% of total annual Chinese output on average.

Estimating costs using only existing yards speaks to the following two concerns: (i) cost declines are driven by the new facilities built through China's plan, which perhaps are more modern and have entirely different production capabilities (though, to reiterate, bulk shipbuilding technology is not subject to frequent technological innovations); (ii) cost declines are driven by firms' optimizing production under learning by doing (though in the next section I allow for a form of expertise accumulation through the backlog). Indeed, existing shipyards do not change technology and have already gone down their learning curve.

Other Checks It is worth noting that results are also robust to adding several covariates, such as the shipyard's age, total past production (capturing experience), dummy variables for young ages to capture learning by doing (documented in military ships in Thompson (2001)) somewhat more flexibly, administrative region, number of employees (reported only in a subset of yards). Working with tons produced, rather than ships, and thus using a tobit model also does not alter findings.

One might also worry that there was a coincident increase in demand for Chinese ships by Chinese shipowners; as discussed in Section 2, however, the overwhelming majority of Chinese built ships are in fact exported during the sample period (about 80%). Another concern is that Chinese shipyards may have switched production towards Handysize vessels and away from other ship types, so that the estimated cost decline reflects a different allocation across products rather than subsidies. As discussed in Section 2, however, the same patterns of increased production are prevalent in virtually all ship types (see Table 1). In addition, the market share of Handysize bulk carriers within China remains practically flat over time during the sample.

As regional governments in China can play an important role (see Section 2), I consider a specification where they implement the subsidy plans at different dates and magnitudes. As no

---

[35]I have also considered the following specification:

$$c(q, s_{jt}^y, s_t) = c_1' g_1 \left(s_{jt}^y, s_t \right) + c_q 1 \{ q \leq \check{q} \} + c_2 q^2 1 \{ q > \check{q} \} \tag{18}$$

which is in the spirit of Ryan (2012). This specification explores the possibility of a kink in the cost function possibly capturing capacity constraints. Since q is discrete, I estimate this for $\check{q} = 1, ..., \bar{q} - 1$; note that when $\check{q} = 1$ we obtain the baseline quadratic model, while when $\check{q} = \bar{q} - 1$ we get (17) of Column II of Table 5. I find that there is no evidence of a kink, which suggests that in fact, we have a fairly smooth convex function (results available upon request). This does not seem unreasonable: a shipyard can always delay production; hire more workers, increase man-hours.

Table 6: Static cost estimates with existing yards.

Cost Function $C(q,s) = c_1' g_1(s)q + c_2 q^2$	I
Linear Term: $c_1' g_1(s)q$	
China	34.79**
	(5.66)
China, POST	−6.03**
	(1.99)
Europe	34.95**
	(6.16)
Japan	27.63**
	(3.80)
S.Korea	36.06**
	(5.22)
Backlog	−0.74**
	(0.19)
Docks/Berths	−0.23
	(0.14)
Max Length	−0.001
	(0.002)
Steel Price	0.57**
	(0.23)
t	0.15**
	(0.05)
Curvature: $c_2 q^2$	1.42**
	(0.38)
σ	13.08
	(3.44)

* $p < 0.1$, ** $p < 0.05$. Standard errors computed from 500 bootstrap samples.
Time t measured in quarters. Countries refer to country dummy variables.

official documentation was found on implementation dates, I consider the first quarter that new shipbuilding docks/berths come online and divide regions into three groups. Results are similar to prior specifications and suggest that the last region to implement, also has the lowest subsidy level.

I also explore whether any firms went through restructuring, mergers or changes in ownership. A total of ten Chinese firms are affected by such changes both within and outside my sample. Removing them from the sample leads to the same results.

One may be concerned that the estimated cost declines are somehow driven by the estimated willingness to pay, $VE_j^o(s_t)$, (for instance due to the different pre and post expectations and value functions). To address this concern I estimate costs without used prices, using an average quarterly price (across shipyards and countries) of a new ship, which is reported by Clarksons. I find that estimated subsidies are significant and of the same magnitude.

Finally, although the vast majority of dock expansion in 2006 results from new firm entry, some shipyards do acquire new docks. Estimating the model using only existing shipyards that did not increase their docks gives the same results; this is consistent with the raw data pattern that these shipyards account for 33% of China's output expansion (recall that new yards account for another 55%).[36]

5.2.2 Dynamic Shipbuilders

I now acknowledge that shipyards may be facing dynamic considerations, as past production can affect their current costs. My empirical approach employs sparse approximation techniques to solve the shipyard's dynamic programming problem.

The shipyard's optimal production now obeys (9). To ease notation, rename the shipyard state $x = \left(s_{jt}^y, s_t\right)$ and $x' = \left(s_{jt+1}^y, s_{t+1}\right)$ and suppress (j, t). Recall the optimal policy thresholds that define the shipyard's optimal production (see Lemma 1):

$$A(x, q) = \frac{1}{\sigma} \left[VE^o(x) + (c(q, x) - c(q+1, x)) + \beta \left(CV^y(x, q+1) - CV^y(x, q) \right) \right] \quad (19)$$

for $q = 0, 1, ..., \bar{q} - 1$. To estimate the cost parameters, I maximize the likelihood (15) with choice probabilities:

$$\Pr(q^* = 0|x) \equiv p_0(x) = \Pr(\varepsilon \geq A(x, 0)) \quad (20)$$

$$\Pr(q^* = q|x) \equiv p_q(x) = \Pr(\varepsilon \leq A(x, q-1)) - \Pr(\varepsilon \leq A(x, q))$$

$$\Pr(q^* = \bar{q}|x) \equiv p_{\bar{q}}(x) = \Pr(\varepsilon \leq A(x, \bar{q} - 1))$$

---

[36]The capacity investment decision is not studied in this paper. This does not bias cost (and thus subsidy) estimates in the static model (the first order condition for optimal production is unaffected by shipyard dock choice). If yards were to choose both production and docks dynamically however, the first order conditions are no longer decoupled, since firms consider the future impact of all their actions.

Maximizing this likelihood function would be trivial if the continuation value $CV^y(x, q)$ were known. This is the standard difficulty of estimating dynamic setups and to address it, I adopt a novel approach that proceeds in two steps (following the recent literature, e.g. Hotz and Miller (1993) and Bajari, Benkard and Levin (2007)). First, I invert observed choice probabilities to directly obtain the optimal policy thresholds nonparametrically. Second, I show that the latter lead to a closed-form expression for ex ante optimal per period payoffs, which in turn are sufficient to obtain the value function. I next describe this approach in detail.

For the first step, note from (20) that clearly, the choice probabilities are a one-to-one function of the optimal policy thresholds $A(x, q)$. Therefore the latter can be recovered from the observed choice probabilities using[37]:

$$A(x, q) = \Phi^{-1}\left(1 - \sum_{k=0}^{q} p_k(x)\right), \quad \text{for } q = 0, 1, ..., \overline{q} - 1 \tag{21}$$

where $\Phi(\cdot)$ is the standard normal distribution. Note that $A(x, q)$ is (weakly) decreasing in q. Most important, if $A(x, q)$ is known, so is the optimal policy: for any (x, ε),

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

For the second step, I show that once the optimal policy is known, the value function can be recovered in a straightforward manner. Indeed, consider shipyard j's Bellman equation (7) which I repeat here for convenience:

$$V^y(x, \varepsilon) = \max_{0 \leq q \leq \overline{q}} \pi^y(x, q, \varepsilon) + \beta E_{\varepsilon', x'}[V^y(x', \varepsilon') | x, q]$$

where $\pi^y(x, q, \varepsilon) \equiv VE^o(x)q - c(q, x) - \sigma q \varepsilon$. The ex ante value function under the recovered optimal policy becomes:

$$V^y(x) \equiv E_{\varepsilon} V^y(x, \varepsilon) = E_{\varepsilon}[\pi^y(x, q^*(x, \varepsilon), \varepsilon) + \beta E_{\varepsilon', x'}[V^y(x', \varepsilon') | x, q^*(x, \varepsilon)]] \tag{22}$$

If the ex ante per period profit, $E_{\varepsilon} \pi^y(x, q^*(x, \varepsilon), \varepsilon)$, were known, then one could solve for the ex ante value function from (22). This can be done in several ways, such as state space discretization and matrix inversion, or parametric approximation; I opt for the latter because of the large dimension of the state space. In particular, I approximate the value function by a polynomial function of the state variables, so that:

$$V^y(x) = \gamma^y f^y(x)$$

---

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

then (22) becomes

$$(f^y(x) - \beta E_{x',\varepsilon}[f^y(x')|x, q^*(x,\varepsilon)])\gamma^y = E_\varepsilon \pi^y(x, q^*(x,\varepsilon), \varepsilon) \tag{23}$$

It is now possible to estimate the approximating parameters γ^y by ordinary regression, provided the ex ante profit can be computed. Note, however, that the large number of states leads to an exploding number of possible terms in $f^y(x)$; in addition, choosing which terms to include in $f^y(x)$ can be an arduous process. Instead, I estimate the sparse vector γ^y via LASSO, which circumvents these issues. Using LASSO to solve for approximate value functions can be useful to the many empirical applications of dynamic setups. Indeed, in the Appendix I present a Monte Carlo exercise that evaluates the performance of this method and shows that it does very well compared to standard approaches.

I now only need to show how $E_\varepsilon \pi^y(x, q^*(x,\varepsilon), \varepsilon)$ is computed. Under the assumption of quadratic costs (i.e. $c_1(x)q + c_2(x)q^2$, where $c_1(x) = c_1' g_1(x)$ and $c_2(x) = c_2' g_2(x)$), ex ante per period payoffs become:

$$E_\varepsilon \pi^y(x, q^*(x,\varepsilon), \varepsilon) =$$

$$= E_\varepsilon\left[VE^o(x)q^*(x,\varepsilon) - c_1(x)q^*(x,\varepsilon) + c_2(x)q^*(x,\varepsilon)^2 - \sigma q^*(x,\varepsilon)\varepsilon\right]$$

$$= (VE^o(x) - c_1(x))E_\varepsilon q^*(x,\varepsilon) + c_2(x)E_\varepsilon q^*(x,\varepsilon)^2 - \sigma E_\varepsilon[q^*(x,\varepsilon)\varepsilon] \tag{24}$$

I show in the Appendix that

$$E_\varepsilon q^*(x,\varepsilon) = \sum_{q=0}^{\bar{q}-1} \Phi(A(x,q)) \tag{25}$$

$$E_\varepsilon[q^*(x,\varepsilon)]^2 = 2\sum_{q=1}^{\bar{q}} q\Phi(A(x,q-1)) - \sum_{q=0}^{\bar{q}-1} \Phi(A(x,q)) \tag{26}$$

$$E_\varepsilon[q^*(x,\varepsilon)\varepsilon] = -\sum_{q=0}^{\bar{q}-1} \phi(A(x,q)) \tag{27}$$

where $\phi(\cdot)$ is the standard normal density.

To sum up, the estimation proceeds as follows (further details are in the Appendix):

1. Estimate the policy thresholds $A(x,q)$ using (21)

2. Compute the statistics of the optimal production in (25), (26) and (27)

3. At each guess of the parameters (c_1', c_2', σ) in the optimization of the likelihood (15):

   (a) Solve for the approximate value function parameters γ^y from (23)

   (b) Using γ^y, compute the choice probabilities and update (c_1', c_2', σ).

Results   Table 7 gives the Maximum Likelihood estimate of the cost function for dynamic shipyards. The implied subsidy is in the order of 18% or 4.3 billion US dollars paid between 2006 and 2012, similarly to the case of static shipyards. Costs are decreasing in the current backlog, consistent with accumulation of expertise or economies of scale. The magnitude of the backlog coefficient suggests that the cost decline accruing from past production is about 11% of total cost at the observed production choices. More docks/berths, as well as longer docks decrease costs. The estimated cost function of dynamic shipyards is somewhat more convex than the one of static shipyards. Since accumulating a backlog decreases future costs and yards take this into account, higher cost parameters are needed to justify the observed low production levels.

A further interesting feature of the dynamic model is that it allows one to compute the expected value of all new Chinese shipyards that are born through China's government plan. I find that this equals 16 billion US dollars. One can think of this amount as a rough estimate of the order of magnitude of the costs of building these shipyards.

The static and dynamic formulations yield similar results in terms of subsidy detection and have similar fit, with the dynamic model having a somewhat higher likelihood. It is not as surprising that retrieved subsidies are similar in the two models- the increase in China's market share is too sharp to be driven by dynamic decision-making. In the next section, I explore the implications of the static and dynamic model, as well as of some extra models in which dynamics are shut down.

Table 7: Dynamic shipyards cost estimates.

Cost Function $C(q,s) = c_1' g_1(s)q + c_2 q^2$	I
Linear Term: $c_1' g_1(s)q$	
China	43.01**
	(9.31)
China, POST	−7.85**
	(2.91)
Europe	43.45**
	(9.81)
Japan	34.49**
	(6.16)
S.Korea	44.34**
	(8.16)
Backlog	−0.64**
	(0.20)
Docks/Berths	−0.17
	(0.12)
Max Length	−0.001
	(0.001)
Steel Price	0.55**
	(0.18)
t	0.003
	(0.04)
Curvature: $c_2 q^2$	1.66**
	(0.55)
σ	15.46
	(5.04)

* $p < 0.1$, ** $p < 0.05$. Standard errors computed from 500 bootstrap samples.
Time t measured in quarters. Countries refer to country dummy variables.

6 Quantifying the Implications of Subsidies

In this section, I quantitatively assess the degree to which industrial subsidies contribute to China's rapid emergence as a world leader in the shipbuilding industry. In particular, I evaluate the impact of government interventions on production reallocation across countries, industry prices, costs and consumer surplus. I use my model to predict the evolution of the industry in the absence of China's government shipbuilding plan, by removing both the cost subsidies retrieved in Section 5, as well as the new shipbuilding facilities. This counterfactual quantifies the adverse trade effects from these two interventions, which are considered actionable by the WTO. Moreover, I assess the relative contribution of the new shipyards by performing a counterfactual that removes the new facilities but maintains the detected cost subsidies.

To implement the main counterfactual of "no interventions", I assume that shipowners maintain their pre-2006 expectations and ship value functions, shipyards keep their pre-2006 costs and docks and new shipyards are removed. To implement the "no entrants" counterfactual, new shipyards are removed and existing shipyards keep their post-2006 cost functions and docks. Also, I assume that shipowners switch to the post-2006 expectations and value functions. In other words, shipowners understand that a change occurred; yet they can't distinguish between different policies. I feed the observed post-2006 values for shipping demand and steel prices into the model and simulate shipyard optimal production and ship prices. Details on the implementation of these counterfactuals can be found in the Appendix.

As shown in Table 8, the industrial subsidies lead to substantial reallocation in production, by increasing China's market share and decreasing Japan's share: if the plan is removed, China's market share falls from 49% to 15%, while Japan's share increases from 46% to 78%. If only the new shipyards are removed, China's share falls from 50% to about 40%, revealing that new facilities played an important but not the predominant part in China's expansion. This is consistent with the raw data patterns of Section 2 that indicated that both new and existing facilities contributed to China's dramatic growth. Figure 8 depicts China's production in the data, as well as the baseline and two counterfactual worlds. In addition to depicting graphically the evolution of the industry in the absence of subsidies, this graph also shows that the model's fit is good since the baseline simulation is not far from the data.

Table 8 also compares ship prices in the baseline and counterfactual worlds and shows that ship prices are higher for all countries in the absence of China's subsidization plans, by about 7%. This is not surprising, given that China's subsidies shifted supply outward.

Next, I turn to profits, shipper surplus and costs, shown in the lower half of Table 8. China's total shipbuilding profits would be substantially lower in the absence of any interventions: from almost 5 billion to less than 1 billion (this is inclusive of the subsidy cost). At the same time, Japan's profits have gone down because of Chinese subsidies by about 10%.

In this model, shipowners neither gain, nor lose from subsidies: because of the free entry

36

	Baseline	No Entrants	No Interventions
Market Share, China	49%	38%	15%
Market Share, Japan	46%	58%	78%
Ship Price, China	20.29	20.7	21.7
Ship Price, Japan	21.4	21.9	22.8
Total Shipyard Profits, China	4807	3186	827
Total Shipyard Profits, Japan	4724	5040	5132
Freight Rate (price per voyage)	1.22	1.23	1.26
Consumer Surplus (shippers)	5968	5826	5569
Industry AVC	1.98	2.3	2.34

Prices, surplus and average variable cost (AVC) measured in million US Dollars.

Profits and surplus refer to the total amount between 2006 and 2012.
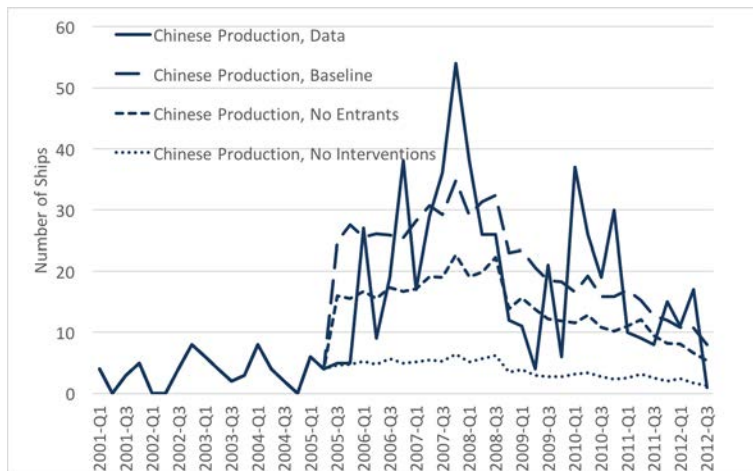
Table 8: Counterfactual results.

Figure 8: China's production in the data, the baseline model and the two counterfactuals.

condition in shipping, they are always indifferent between buying a ship or not (existing shipowners do lose, however, because of the unexpected negative shock to their asset value). Shippers of cargo, however, gain from subsidies as these lead to higher shipbuilding production and thus to a larger fleet. I use the demand curve estimated in Section 5.1.1 and in the Appendix to compute shipping prices and shipper surplus.[38] As shown in Table 8, the freight rate is moderately higher (by 3%) in the absence of Chinese subsidies. The difference in freight rates, however, increases over time between 2006 and 2012: because of time to build it takes time until the different worlds lead to different fleet levels. Indeed, between 2009 and 2012 freight rates are higher by 5%. As a result, cargo shippers benefit from Chinese subsidies; their consumer surplus is higher by 7% because of subsidies and increases over time (between 2009 and 2012 consumer surplus is higher by 9%). Note that between 2006 and 2012 the cost subsidies alone (i.e. ignoring the additional costs of building new yards), are estimated to have cost the government about 4 billion US Dollars and resulted in consumer surplus gains of 400 million US dollars. This calculation implies that a frequent assertion that China developed shipbuilding to benefit from low freight rates for its trade seems to be unsubstantiated; indeed, the benefits of subsidies to shipping are minimal and perhaps the Chinese government is aspiring also to externalities to different sectors (e.g. steel, defense) or, even, national pride (Grossman (1990)).

Finally, I consider the cost implications of subsidies. I compute the average cost of producing a ship at the industry level and find that as expected, subsidies decrease costs of production; see last row of Table 8. If I decompose the average cost to subsidies and market share allocation, however, this picture is entirely different. As implied by the numbers in Table 8, the change in industry average cost brought about by production subsidies is 0.36 million. The role of subsidies in this change in AVC turns out to be -3.8 million, implying a negative reallocation effect equal to $0.36 - 3.8 = -3.44$ million.[39] In other words, the subsidization in costs should have led to a much larger decline in the industry average cost of production; but as subsidies shift production away from the low-cost Japanese shipyards towards the high-cost Chinese shipyards, the industry produces at a much higher average cost net of subsidies.

The Role of Dynamics   Government subsidies are often related to dynamic considerations, and this is true in this application as well: the long construction lags and the positive production feedback resembling learning by doing are important features of shipbuilding.[40] I now explore the importance of these dynamics and their interaction with subsidies. To do so, I consider two extra

---

[38]The cargo shipping demand curve gives the price per voyage as a function of the total number of voyages. I assume that there is a constant fleet utilization rate to map the fleet into voyages.

[39]Consider the cost function $c_{1j}q + c_2q^2$, for $j \in$ {China, EU, Japan, S.Korea}, which for China becomes $(c_{1China} - \kappa) q + c_2q^2$ post-2006. The change in the industry average cost because of subsidies is equal to a sum of two terms: the first is $\kappa (q_{China}/ (q_{China} + q_{EU} + q_{Japan} + q_{SK}))$, while the second includes all other terms and is related to the reallocation of production and market share. The first term equals 3.44 million.

[40]Note that the backlog accumulation rule here is identical to a learning and forgetting model (Benkard (2001)), with the forgetting determined by the backlog depreciation parameter.

	I	II	III
	Estimated Model	No Dynamic Feedback	No TTB
%Δ Market Share, China	69%	60%	61%
Average Production Baseline, China	22	20.7	28
Ship Price Baseline, China	20.29	21	21.6
%Δ Ship Price, China	−2%	−5%	−3%
%Δ Total Shipyard Profits, China	83%	79%	79%
%Δ Freight Rate (price per voyage)	−3.3%	−3.2%	−6%

Estimated model refers to the estimated dynamic model.

No Dynamic Feedback sets the parameter of the backlog equal to zero in the cost function.

No TTB assumes that there is no time to build.

%Δ is a % change between the baseline world (Chinese subsidies) and a world with no Chinese interventions.

Table 9: The role of dynamics.

models and contrast them to the estimated one: (i) removes the positive production feedback (i.e. it sets the backlog coefficient equal to zero in the cost function); (ii) removes the time to build. Table 9 reports some results.

Let us first compare the baseline to a model where the backlog has no impact on future costs; see columns I and II. The latter model essentially has no dynamics and in this case, Chinese shipyards experience a lower increase in their market share by 10%. In the baseline world, shipyards are accumulating experience that decreases their future costs and thus market share gains multiply and lead to more reallocation favorable to China. This is consistent with the pro-subsidy arguments of infant industries (though note that here learning is private and provides no externalities, so that there is no real rationale for the government to intervene).

Figure 9 illustrates that the industry is more volatile when shipyards face a dynamic production feedback (column I). When accumulated backlog lowers costs, firms have an incentive to overproduce. In turn, this overproduction during the boom lowers future prices and in the subsequent bust, firms produce less. In contrast, a world that lacks this dynamic feedback (column II) is smoother. Column II of Table 9 also shows that China produces less on average and prices are higher.

Next, I consider a world with virtually no time to build (I allow for a one period construction lag to maintain the model's timing); see column III of Table 9. As argued in Kalouptsidi (2014), time to build tends to smooth out fluctuations. Indeed, in the absence of time to build, firms tend to respond more to shocks: in the case of a positive macro shock, more shipowners enter as ships are delivered instantly and thus shipyards produce more. In contrast, under construction lags (baseline, column I), a positive shock may die out by the time the ship is delivered and thus shipowners demand fewer ships. This is illustrated in Figure 9. As shown in column III of Table 9 as well as Figure 9, under no time to build, average ship production is substantially higher. As the fleet is built instantly, we also see the impact of subsidies on freight rates much faster: the
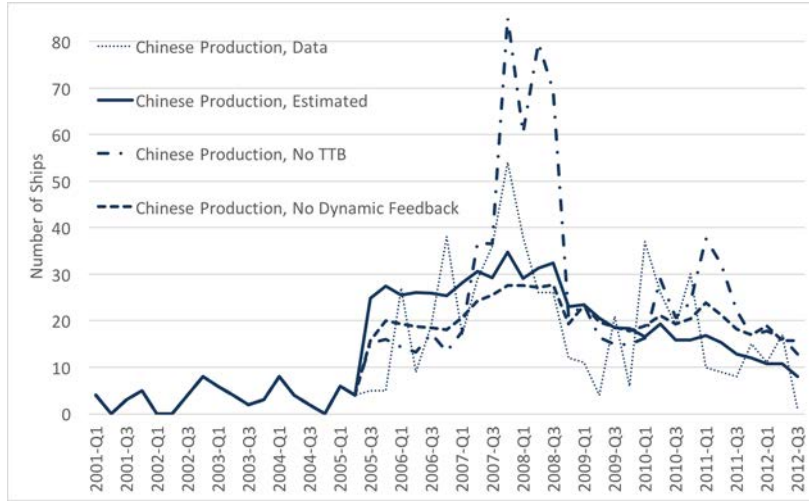
Figure 9: Number of ships produced by China in the data, the estimated model, the no dynamic feedback counterfactual and the no time to build counterfactual.

latter are 6% lower (rather than 3% in column I), consistent with the more long-run impact of the baseline world.

7 Robustness Analysis

In this section I present several robustness exercises to evaluate my empirical strategy. In particular, I explore robustness with respect to: (i) the state space; (ii) the VAR model for state transitions; (iii) the LASSO penalty parameter in the shipyard value function approximation. The latter is complemented by a Monte Carlo (see the Appendix). This section is self-contained and the reader may omit it and proceed to the conclusion if desired.

State Space and Agent Expectations As described in Section 5.1.2, I employ a parsimonious VAR model to capture agents' perceptions of the aggregate state transition process. Here, I explore the robustness of this specification. First, I consider a more flexible VAR model that is different before and after 2006 and includes a time trend. I estimate it via LASSO since the number of parameters is now large. In particular, the model sets

$$s_t = \left(R_0^{pre} + R^{pre}s_{t-1} + R_\tau^{pre}t\right)1\{t < 2006\} + \left(R_0^{post} + R^{post}s_{t-1} + R_\tau^{post}t\right)1\{t \geq 2006\} + \xi_t$$

Note that the process for the demand and steel prices, however, remains unchanged- i.e. the corresponding rows of the VAR are the same as in the baseline. The estimated cost parameters are in Column I of Table 10 and are close to the baseline specifications of Section 5.2.1.

Next, I also alter the state space. As a reminder, the baseline industry state space includes the number of ships younger than 20 years old, the number of ships older than 20 years old, the total

Table 10: Dynamic Cost Estimates - Robustness Specifications

Cost Function $C(q,s) = c_1' g_1(s)q + c_2 q^2$	I	II	III
Linear Term: $c_1' g_1(s)q$			
China	41.78**	38.55**	41.08**
	(7.78)	(5.48)	(9.23)
China, POST	-8.13**	-6.70**	-9.50**
	(2.50)	(1.96)	(3.47)
Europe	41.92**	38.91**	43.59**
	(8.36)	(5.55)	(10.34)
Japan	33.37**	31.43**	32.19**
	(5.20)	(3.58)	(5.84)
S.Korea	42.45**	40.06**	41.64**
	(6.99)	(4.85)	(8.15)
Backlog	-0.61**	-0.56**	-0.81**
	(0.17)	(0.12)	(0.22)
Docks/Berths	-0.16	-0.13*	-0.21
	(0.13)	(0.07)	(0.16)
Max Length	-0.001	-0.001*	-0.003*
	(0.001)	(0.0007)	(0.001)
Steel Price	0.62**	0.48**	0.25
	(0.17)	(0.13)	(0.23)
t	0.005	0.03	0.42**
	(0.03)	(0.02)	(0.09)
Curvature: $c_2 q^2$	1.62**	1.46**	2.12**
	(0.46)	(0.31)	(0.60)
σ	15.04	13.56	19.77
	(4.16)	(2.78)	(5.36)

* $p < 0.1$, ** $p < 0.05$. Standard errors computed from 500 bootstrap samples.

Time t measured in quarters. Countries refer to country dummy variables.

Specification I considers a more flexible VAR model for state transitions.

Specification II considers the total fleet and mean age as state variables.

Specification III considers the entire fleet age distribution as a state variable.

Table 11: Counterfactuals performed using robustness Specification I and III of Table 10

	Robustness Specification I			Robustness Specification III		
	Baseline	No Entrants	No Interventions	Baseline	No Entrants	No Interventions
Market Share, China	50%	39%	14%	51%	40%	17%
Market Share, Japan	43%	56%	79%	42%	55%	77%
Ship Price, China	20.14	20.5	21.8	20	20.6	21.4
Ship Price, Japan	21.37	21.75	23	23	23.7	24.5
Freight Rate (price per voyage)	1.22	1.24	1.26	1.32	1.33	1.35
Consumer Surplus (shippers)	5954	5810	5556	4948	4879	4732
Industry AVC	1.88	2.17	2.3	1.73	2.13	2.07

Prices, surplus and average variable cost (AVC) measured in million US Dollars.

Surplus refers to the total amount between 2006 and 2012.

Specification I considers a more flexible VAR model for state transitions.

Specification III considers the entire fleet age distribution as a state variable.

backlog, demand for shipping services and steel prices. I consider two alternative specifications of the state space. First, I explore different statistics of the age distribution and in particular, the total fleet and the mean age of the fleet. Similar to above, I use a flexible VAR model for this state's transition. Results are presented in Column II of Table 10 and are robust.

Second, I consider a more detailed state space: the entire ship age distribution (i.e. number of ships 0,1,2,...,20+ years old).[41] This is computationally possible due to the sparse approximations which allow one to deal with very large state spaces. Let s_{kt} be the number of ships that are k years old in period t. To form the transition of this state I regress s_{kt} on s_{kt-1} and s_{k-1t-1} for $k = 1, ..., 19$.[42] Similarly, I regress ships of age zero, s_{0t}, on s_{0t-1} and the total backlog. Finally, s_{20+t} depends on optimal ship exit and thus all states; similarly total backlog depends on ship entry and thus all states. Therefore, I assume that the total backlog and s_{20+t} follow a VAR which is also affected by demand and steel prices. Results from this specification are presented in Column III of Table 10 and again are similar to the baseline specifications.

Finally, Table 11 presents a few counterfactual results based on these specifications.[43] Comparing these results to Table 8, we see that counterfactual predictions are overall robust across these different specifications.

LASSO Penalty Parameter   Finally, when performing LASSO, one needs to choose the penalty parameter, λ, that trades off fit and sparsity. In Table 12 I present results of cost estimates across a range of different λ.[44] Results are overall robust, yet some sensitivity is to be

---

[41]The data does not distinguish the age of ships that are older than 20 years old so this last age group is an absorbing age category.

[42]This transition is not deterministic since the data is quarterly rather than annual. However, the recovered coefficients imply a close to uniform age distribution across quarters within each year. Moreover, I use lower degree approximations in the value functions because of the much larger number of states.

[43]I do not perform counterfactuals with robustness specification II, as it is not clear how to construct the "true" transition of the mean age and total fleet from firm decisions.

[44]Twenty penalty values are generated automatically by the LASSO command and are spaced out between the value that holds all regressors and the value that removes all regressors. Here, I remove the first few and last few

Table 12: Dynamic Cost Estimates - Range of λ

Cost Function $C(q,s) = c'_1 g_1(s)q + c_2 q^2$	λ_8	λ_9	λ_{10}	λ_{11}	λ_{12}	λ_{13}	λ_{14}	λ_{15}	λ_{16}	λ_{17}
Linear Term: $c'_1 g_1(s)q$										
China	49.97**	50.34**	48.41**	45.10**	43.01**	41.18**	41.23**	33.65**	31.94**	31.96**
	(13.46)	(13.58)	(12.50)	(10.44)	(9.31)	(8.30)	(8.54)	(5.64)	(4.79)	(4.79)
China, POST	−10.53**	−9.98**	−8.89**	−8.24**	−7.85**	−7.39**	−7.36**	−6.45**	−6.12**	−6.12**
	(4.24)	(4.14)	(3.68)	(3.22)	(2.91)	(2.64)	(2.70)	(2.08)	(1.84)	(1.83)
Europe	50.84**	51.37**	49.35**	46.01**	43.45**	41.60**	41.65**	33.91**	32.16**	32.19**
	(14.39)	(14.57)	(13.29)	(11.17)	(9.81)	(8.76)	(9.00)	(5.99)	(5.12)	(5.13)
Japan	38.89**	39.30**	38.09**	35.70**	34.49**	33.25**	33.33**	26.96**	25.69**	25.71**
	(9.18)	(9.12)	(8.44)	(6.97)	(6.16)	(5.51)	(5.68)	(3.57)	(3.01)	(3.01)
S.Korea	51.71**	51.60**	50.92**	47.46**	44.34**	42.72**	42.76**	35.73**	34.20**	34.21**
	(12.33)	(12.06)	(11.60)	(9.72)	(8.16)	(7.32)	(7.52)	(5.00)	(4.24)	(4.24)
Backlog	−0.77**	−0.82**	−0.75**	−0.70**	−0.64**	−0.61**	−0.61**	−0.52**	−0.50**	−0.50**
	(0.28)	(0.30)	(0.27)	(0.23)	(0.20)	(0.18)	(0.19)	(0.14)	(0.12)	(0.12)
Docks/Berths	−0.16	−0.16	−0.15	−0.15	−0.17	−0.17	−0.16	−0.15	−0.14	−0.14
	(0.17)	(0.16)	(0.17)	(0.15)	(0.12)	(0.11)	(0.14)	(0.12)	(0.11)	(0.11)
Max Length	−0.002	−0.002	−0.002	−0.001	−0.001	−0.001	−0.001	−0.001	−0.001	−0.001
	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.0009)	(0.0009)	(0.0009)
Steel Price	0.72**	0.76**	0.73**	0.70**	0.55**	0.55**	0.55**	0.57**	0.54**	0.54**
	(0.25)	(0.25)	(0.24)	(0.21)	(0.18)	(0.17)	(0.17)	(0.16)	(0.15)	(0.15)
t	−0.02	−0.03	−0.06	−0.02	0.003	0.02	0.01	0.10**	0.13**	0.13**
	(0.06)	(0.06)	(0.06)	(0.05)	(0.04)	(0.04)	(0.04)	(0.05)	(0.05)	(0.05)
Curvature: $c_2 q^2$	2.05**	2.10**	1.96**	1.81**	1.66**	1.56**	1.56**	1.33**	1.24**	1.24**
	(0.78)	(0.81)	(0.73)	(0.63)	(0.55)	(0.49)	(0.50)	(0.37)	(0.32)	(0.32)
σ	19.41	19.83	18.45	17.05	15.46	14.55	14.50	12.32	11.59	11.59
	(7.24)	(7.51)	(6.74)	(5.79)	(5.04)	(4.49)	(4.60)	(3.37)	(2.90)	(2.90)

* $p < 0.1$, ** $p < 0.05$. Standard errors computed from 500 bootstrap samples.
Time t measured in quarters. Countries refer to country dummy variables.

Table 13: Counterfactuals performed using robustness specification λ_8 and λ_{17} of Table 12

	Robustness Specification λ_8			Robustness Specification λ_{17}		
	Baseline	No Entrants	No Interventions	Baseline	No Entrants	No Interventions
Market Share, China	55%	45%	19%	49%	39%	15%
Market Share, Japan	39%	51%	72%	46%	57%	79%
Ship Price, China	19.8	20.1	21	19.9	20.4	21.3
Ship Price, Japan	20.9	21.2	22	21	21.5	22.4
Freight Rate (price per voyage)	1.21	1.22	1.263	1.22	1.23	1.25
Consumer Surplus (shippers)	6047	5984	5817	6019	5916	4731
Industry AVC	2.76	3.38	3.99	2.65	3.2	3.04

Prices, surplus and average variable cost (AVC) measured in million US Dollars.

Surplus refers to the total amount between 2006 and 2012.

expected, since the parameter λ determines the level of smoothing. Finally, Table 13 presents the counterfactual results corresponding to the first and last λ of Table 12. Those too are close to the baseline specifications.

8 Conclusion

The role of industrial policy in China's rapid takeover, especially in industries that are capital intensive, is an open question. To answer, one needs to first know what policies are in place. This paper measures subsidies and quantifies their impact for the striking example of the world shipbuilding industry. I find strong evidence consistent with subsidies that decreased shipyard costs by 13-20%. The interventions led to substantial reallocation of production across countries with no significant surplus gains. Therefore, subsidies may be justified by shipbuilding's important externalities to other sectors, such as the steel industry or the readiness of the military sector. Yet, understanding government objectives is an interesting avenue for future research.

9 Appendix

9.1 Proof of Lemma 1

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

$$-c\left(0, s_{jt}^y, s_t\right) + \beta CV^y\left(s_{jt}^y, s_t, 0\right) \geq VE_j^o\left(s_t\right)q' - c\left(q', s_{jt}^y, s_t\right) - q'\varepsilon_{jt} + \beta CV^y\left(s_{jt}^y, s_t, q'\right)$$

all $q' \in \{1, ..., \bar{q}\}$, or

$$\varepsilon_{jt} \geq VE_j^o\left(s_t\right) + \max_{0 < q' \leq \bar{q}} \left\{ \frac{c\left(0, s_{jt}^y, s_t\right) - c\left(q', s_{jt}^y, s_t\right) + \beta\left(CV^y\left(s_{jt}^y, s_t, q'\right) - CV^y\left(s_{jt}^y, s_t, 0\right)\right)}{q'} \right\}$$

values as both ends tend to be more noisy.

44

Similarly, shipyard j chooses $q_{jt} = q \neq 0, \bar{q}$ if:

$$VE_j^o(s_t)\, q - c\left(q, s_{jt}^y, s_t\right) - q\varepsilon_{jt} + \beta CV^y\left(s_{jt}^y, s_t, q\right) \geq VE_j^o(s_t)\, q' - c\left(q', s_{jt}^y, s_t\right) - q'\varepsilon_{jt} + \beta CV^y\left(s_{jt}^y, s_t, q'\right)$$

all $q' \neq q$, or

$$\varepsilon_{jt} \geq VE_j^o(s_t) + \max_{q < q' \leq \bar{q}} \left\{ \frac{c\left(q, s_{jt}^y, s_t\right) - c\left(q', s_{jt}^y, s_t\right) + \beta\left(CV^y\left(s_{jt}^y, s_t, q'\right) - CV^y\left(s_{jt}^y, s_t, q\right)\right)}{q' - q} \right\}$$

$$\varepsilon_{jt} \leq VE_j^o(s_t) + \min_{0 < q' \leq \bar{q}} \left\{ \frac{c\left(q, s_{jt}^y, s_t\right) - c\left(q', s_{jt}^y, s_t\right) + \beta\left(CV^y\left(s_{jt}^y, s_t, q'\right) - CV^y\left(s_{jt}^y, s_t, q\right)\right)}{q' - q} \right\}$$

Finally, shipyard j chooses $q_{jt} = \bar{q}$, if

$$\varepsilon_{jt} \leq VE_j^o(s_t) + \min_{0 \leq q' < \bar{q}} \left\{ \frac{c\left(\bar{q}, s_{jt}^y, s_t\right) - c\left(q', s_{jt}^y, s_t\right) + \beta\left(CV^y\left(s_{jt}^y, s_t, q'\right) - CV^y\left(s_{jt}^y, s_t, \bar{q}\right)\right)}{q' - \bar{q}} \right\}$$

Proving Lemma 1 amounts to proving that if the shipbuilding cost function $C(q, \cdot) : [0, \bar{q}] \to \mathbb{R}$, is convex in q, then:

$$\max_{q < q' \leq \bar{q}} \left\{ \frac{c(q, \cdot) - c(q', \cdot) + \beta\left(CV^y(\cdot, q') - CV^y(\cdot, q)\right)}{q' - q} \right\} = c(q, \cdot) - c(q+1, \cdot) + $$
$$+ \beta\left(CV^y(\cdot, q+1) - CV^y(\cdot, q)\right)$$

$$\min_{0 \leq q' < \bar{q}} \left\{ \frac{c(q, \cdot) - c(q', \cdot) + \beta\left(CV^y(\cdot, q') - CV^y(\cdot, q)\right)}{q' - q} \right\} = c(q-1, \cdot) - c(q, \cdot) + $$
$$+ \beta\left(CV^y(\cdot, q) - CV^y(\cdot, q-1)\right)$$

I first show the following lemma:

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

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

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

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

$\square$

Therefore, it suffices to show that the function $c\left(q,\cdot\right)-\beta CV^{y}\left(\cdot,q\right)$ is convex in q. By assumption, $c\left(q,\cdot\right)$ is convex. It then follows that $CV^{y}\left(\cdot,q\right)$ 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 this 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 the empirical exercise), $\beta \in (0,1)$, and the state space is convex.

9.2 Creation of shipping demand state

I estimate the inverse demand for shipping services via instrumental variables regression, to create the state d_t. The analysis follows Kalouptsidi (2014). The empirical analogue of the demand curve in (1) chosen is:

$$P_t^o = \alpha_0^d + \alpha_1^d X_t^d + \alpha_2^d Q_t^o + \varepsilon_t^d \tag{28}$$

where P_t^o is the average price per voyage observed in a quarter, X_t^d includes demand shifters, while Q_t^o is the total number of voyages realized (see Kalouptsidi (2014) for a detailed data description). 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 (I observe the number of voyages realized, rather than ton-miles).

Table 14 reports the results. The impact of all shifters is lumped into the state variable d_t (the residual $\hat{\epsilon}_t^d$ is included in d_t as it captures omitted demand shifters):

$$d_t = \hat{\alpha}_0^d + \hat{\alpha}_1^d X_t^d + \hat{\epsilon}_t^d$$

9.3 State Transitions

The baseline specification for the VAR model for state transitions is:

Table 14: IV regression estimates for demand

	I First Stage	II Second Stage[1]
Constant	-2731.3^{**}	-1.403
	(790.28)	(1.26)
Food Price Index	0.61	0.0051
	(0.693)	(0.0038)
Agricultural Raw Material Price	1.35^{**}	0.0022
	(0.48)	(0.0028)
Minerals Prices	-0.43	0.0014
	(0.33)	(0.0018)
Aluminum Production	-0.28^{**}	0.0012^{**}
	(0.11)	(0.0005)
Grain Production	-0.86	0.0047
	(0.9)	(0.0044)
Substitute Fleet	0.38^{**}	-0.0022^{**}
	(0.15)	(0.0005)
Fleet	0.55^{**}	
	(0.22)	
Mean Fleet Age	96.67^{**}	
	(18.5)	
$\hat{Q}_t^o$		-0.0033^{**}
		(0.001)

* $p < 0.1$, ** $p < 0.05$

[1]The parameters reported for the second stage have been divided by 10^6

$$
\begin{bmatrix} \overline{s^o_{1t}} \\ \overline{s^o_{2t}} \\ B_t \\ d_t \\ l_t \end{bmatrix} = \underbrace{\begin{bmatrix} r^{s_1}_o \\ r^{s_2}_o \\ r^B_o \\ r^d_0 \\ r^l_0 \end{bmatrix}}_{R^{pre}_0} 1\{t \le 2006\} + \underbrace{\begin{bmatrix} \tilde{r}^{s_1}_o \\ \tilde{r}^{s_2}_o \\ \tilde{r}^B_o \\ r^d_o \\ r^l_o \end{bmatrix}}_{R^{post}_0} 1\{t > 2006\} + \underbrace{\begin{bmatrix} r^{s_1 s_1} & r^{s_1 s_2} & r^{s_1 B} & r^{s_1 d} & r^{s_1 l} \\ r^{s_2 s_1} & r^{s_2 s_2} & r^{s_2 B} & r^{s_2 d} & r^{s_2 l} \\ r^{B s_1} & r^{B s_2} & r^{BB} & r^{Bd} & r^{Bl} \\ 0 & 0 & 0 & r^d & 0 \\ 0 & 0 & 0 & 0 & r^l \end{bmatrix}}_{R} s_{t-1} + \xi_t \quad (29)
$$

and Σ_ξ is diagonal. I estimate the parameters of interest (R_0, R, Σ_ξ) via OLS separately for each variable (note that separate OLS yields identical estimates to Maximum Likelihood estimation). Table 15 reports the results.

In addition to the robustness analysis performed in Section 7, I have examined several specifications where the VAR parameters can be time-varying, i.e. (R_{ot}, R_t). Indeed, (R_{ot}, R_t) can vary deterministically (e.g. time trend, so that $R_{ot} = R_o + R_\tau t$) or randomly with time (random walk model for R_t estimated by the Kalman filter). The former requires extension of (14) to the time varying case and adaptation of the higher moments formulas; the latter relies on the standard Kalman updates for the time varying parameters which constitute the state of the Kalman filter. End results are robust. I also experimented heavily with restrictions on R_{0t} and R_t both in terms of before and after 2006, allowing Σ_ξ to be full, as well as parameter restrictions (e.g. time to build might imply that $r^{s_1 d} = 0$ ignoring ship exit). Finally, I allowed d_t to be an $AR(2)$. My main findings are robust to these experiments. The chosen specification combines the following desired properties: it is parsimonious, stationary and takes into account the 2006 break.

9.4 Derivation of state expectations in ship value function

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

$$
V^o(x_t) = \gamma^o f^o(x_t) = \sum_{i=1}^{d} \gamma^o_i x^{(i)}_t
$$

where $x_t = (s^o_{it}, s_t)$, and $x^{(i)}_t$ are Kronecker powers, so that $x^{(2)}_t = x_t \otimes x_t$, $x^{(3)}_t = x^{(2)}_t \otimes x_t$, etc. Then, (13) can be written as:

$$
P^N_{jt} = \beta^{T_{jt}} \gamma^o E\left(f^o\left(x_{t+T_{jt}}\right) | x_t\right) = \beta^{T_{jt}} \sum_{i=1}^{d} \gamma^o_i E\left(x^{(i)}_{t+T_{jt}} | x_t\right) \quad (30)
$$

The conditional expectation is only necessary for s_t since s^o_{it} evolves deterministically. Equation

Table 15: VAR parameter estimates

	R_0^{pre}	R_0^{post}	Σ_ξ
$\overline{s_{1t}^o}$	-1.47	-1.46	$7e-05$
	(0.97)	(0.98)	
$\overline{s_{2t}^o}$	2.53^{**}	2.52^{**}	$8e-05$
	(1.03)	(1.04)	
B_t	-8.49	-8.50	0.004
	(7.31)	(7.35)	
d_t		0.38^{**}	0.15
		(0.17)	
l_t		0.74^{*}	1.27
		(0.40)	

	R				
	$\overline{s_{1t}^o}$	$\overline{s_{2t}^o}$	B_t	d_t	l_t
$\overline{s_{1t}^o}$	1.10^{**}	0.08	0.02^{**}	0.005	-0.003^{**}
	(0.04)	(0.09)	(0.004)	(0.003)	(0.001)
$\overline{s_{2t}^o}$	-0.15^{**}	0.81^{**}	-0.007	0.004	0.002^{**}
	(0.04)	(0.10)	(0.005)	(0.004)	(0.001)
B_t	0.16	1.07	0.87^{**}	0.08^{**}	-0.009
	(0.30)	(0.71)	(0.03)	(0.03)	(0.007)
d_t	0	0	0	0.81^{**}	0
				(0.08)	
l_t	0	0	0	0	0.89^{**}
					(0.06)

* $p < 0.1$, ** $p < 0.05$. $\left(\overline{s_{1t}^o}, \overline{s_{2t}^o}, B_t\right)$ are in natural logarithms.

(14) can be written as:

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

where

$$A = R^T s_t + \sum_{k=t+1}^{t+T} R^{t+T-k} R_{0k}$$

$$v = \sum_{k=t+1}^{t+T} R^{t+T-k} \xi_k$$

Note that conditional on s_t, A is constant.[45] Moreover, v is zero-mean normal with covariance:

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

Therefore, (30) becomes:

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

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

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

where $vec(x)$ denotes the vector formed by stacking the columns of x one after the other; the matrix T_{mm^2} is defined as the following linear operator: given a $L \times n$ matrix A, T_{Ln} is an $Ln \times Ln$ matrix specified by the assignment $T_{Ln} vec(A) = vec(A')$, (A' stands for transpose). The first of

---

[45]It is worth noting that this can be generalized to the time-varying VAR model $s_t = R_{0t} + R_t s_{t-1} + \xi_t$, versions of which are employed in the robustness analysis of Section 7. Indeed, in that case we have:

$$s_{t+T} = \phi\left(t+T, t\right) s_t + \sum_{k=t+1}^{t+T} \phi\left(t+T, k\right)\left(R_{0k} + \xi_k\right)$$

where

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

I denotes the identify matrix. Note that indeed in the time invariant case, $R_t = R$, all t, I get (14) of the main text.

the above equations is straightforward. To prove the second, use:

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

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, I prove the third order equation. Note that

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

$Ev^{(3)}$ 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, k) respectively, then

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

9.5   Statistics of the Optimal Production

To derive (25), I use (20) to get:

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

Equation (26) follows similarly. Finally, let $\phi\left(\varepsilon\right)$ denote the standard normal density. Then,

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

and therefore:

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

9.6 Estimating costs for dynamic shipbuilders: Details

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

Step 1: Estimate $A(x, q)$ using (21). In this step, I first compute the choice probabilities $\{p_q (x)\}_{q=0}^{\bar{q}}$ from observed frequencies. I also include a post-2006 dummy in the state to capture differences in the policy function before and after 2006. As is common in dynamic applications, there are not many observations for all $q = 0, 1, ..., \bar{q}$ at each state x. Therefore, I first cluster the data finely, using the kmeans algorithm, and then compute frequencies on this subset of states. Second, I smooth the frequency matrix using kernels. In particular, I compute the choice probability $p_q (x)$ at state x using the following formula:

$$
p_q (x) = \sum_{x'} w (x' - x) \widetilde{p}_q (x')
$$

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

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

where k_x is the number of values that x can take (in the case of country it is 4, in the case of the post dummy 2) and h represents the bandwidth of the kernel. As h gets close to 0, this kernel weights states that share the same variable x.[46] Finally, I chose $\bar{q} = 10$, since 99.75% of observations involve $q \leq 10$.

Step 2: Compute the terms Eq^*, Eq^{*2}, $E\varepsilon q$ using (25), (26) and (27).

---

[46]I also experimented with parametric specifications for $A(x, q)$. In particular, I estimated an ordered probit model using directly the production data. It overall gives similar results to the nonparametric specification above.

Step 3: At each guess of the parameters (c'_1, c'_2, σ) in the optimization of the likelihood (15) the following calculations are performed:

Step 3a: Solve for the approximate value function parameters γ^y from (23). Note that the choice probabilities require the continuation value $CV^y(x, q) = \gamma^y E\left[f^y(x')|x, q\right]$. To estimate γ^y from (23) I need

$$E_{\varepsilon, x'}\left[f^y(x')|x, q^*(x, \varepsilon)\right] = \gamma^y \sum_{q=0}^{\bar{q}} p_q(x) E_{x'}\left[f^y(x')|x, q\right] \tag{31}$$

I use polynomials of third degree in all variables (fourth degree does not alter the results). The industry state s evolves according to the estimated VAR model described in Section 5.1.2, while the expectations of its polynomial powers are given above in this Appendix. I assume that the shipyard's individual backlog, b_{jt}, transitions as follows:

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

i.e. $\delta\%$ of the backlog is delivered and q orders of period t enter the backlog. I experimented extensively with the above transition rule.[47] It was found that the simple model where δ is constant over shipyards and time and equal to the sample mean (which is 10%) performs equally well compared to more complex models. Given the state transitions it is straightforward to compute (31). I estimate (23) using LASSO in two ways. First, call the LASSO within the likelihood maximization with the regularization parameter chosen using Belloni and Chernozhukov (2011). Second, estimate (23) with LASSO using profits obtained from the static cost estimates. The goal here is to recover the relevant nonzero polynomial terms. I then run OLS within the likelihood with only these terms (and repeat the estimation for many values of the regularization parameter). Results are overall robust to all of the above.

Step 3b: Using γ^y, compute choice probabilities in the likelihood and update (c'_1, c'_2, σ).

A concern in two-step approaches to dynamic frameworks is that the first stage policy functions (in this case, the nonparametric $A(x, q)$) may be different from the optimal policy computed using the true parameters and value function, i.e. from (19). To check this, I compute $A(\cdot)$ from (19) and find that it is close to its first stage estimate. I then re-optimize the likelihood using the new $A(\cdot)$. The parameters that I report result from this loop.

9.7 Counterfactual Computation

There are two steps in the implementation of the counterfactual scenarios presented in Section 6. First, I compute the equilibrium of the model in each scenario. Second, I simulate the model

---

[47]In particular: (i) time-varying δ, e.g. drawn from a probability distribution estimated from the data (e.g. beta distribution whose mean depends on the shipyard's backlog, docks/berths or length; discrete random variable with probabilities estimated from the data); (ii) employ deliveries, instead of δ, described by a binomial random variable whose parameters can again depend on shipyard observables.

using the observed paths of demand and steel prices which are exogenous. Note that if one is only interested in the "no interventions" counterfactual, one can simply use the pre-2006 expectations and value functions to simulate the model.

To predict how the industry would evolve under different counterfactual scenarios I need to obtain shipyards' optimal policies and value functions under each scenario. I can no longer use the estimated VAR for state transitions, since this formed an approximation to expectations that hid equilibrium features. I therefore turn to the following state transitions for $\left(\overline{s_{1t}^o}, \overline{s_{2t}^o}, B_t\right)$, where $\overline{s_{1t}^o}$ is the number of ships younger than 20 years old, $\overline{s_{2t}^o}$ is the number of ships older than 20 years old and B_t is the total backlog:

$$
\begin{aligned}
\overline{s_{1t+1}^o} &= \delta B_t + (1 - \rho_{1t}) \overline{s_{1t}^o} && (32) \\
\overline{s_{2t+1}^o} &= \overline{s_{2t}^o} + \rho_{1t} \overline{s_{1t}^o} - \zeta(s_t) \\
B_{t+1} &= (1 - \delta) B_t + \sum_j q_{jt} && (33)
\end{aligned}
$$

where $\zeta(s_t)$ is the number of ships that exit at state s_t, ρ_{1t} is the percentage of ships that transit from 19 years old and 3 quarters to 20 years old and δ is the percentage of the backlog that is delivered, consistent with the individual backlog transition used in the estimation and described above in this Appendix. In words, the number of young ships $\overline{s_{1t+1}^o}$ equals last period's young ships plus deliveries from the total backlog, minus exiting ships (as documented in Kalouptsidi (2014) there is virtually no exit in ships younger than 20 years old). The number of old ships $\overline{s_{2t+1}^o}$ equals last period's old ships plus the aging ships minus exiting ships. Finally, total backlog B_{t+1} equals last period's total backlog minus deliveries, plus total new ship orders. I calibrate ρ_{1t} to 3% which is the sample average. To predict ship exit $\zeta(s_t)$ I follow Kalouptsidi (2014) where the number of exiting ships is regressed on the industry state (in particular, $\log \zeta_t = \beta_\zeta s_t$); note that exit rates are extremely low (even during the 2008 crisis). Demand d_t and steel price l_t retain their original transition processes, since these are exogenous to this model.

To find the equilibrium of the model I use a fixed point algorithm with the goal of recovering the shipyard's optimal policy function $p_q^*(x)$, for all production levels q and states $x = (s^y, s)$, as well as the shipyard's value function $V^y(x)$. At each iteration l the policies $p_q^l(x)$ are used to update to $p_q^{l+1}(x)$ and I keep iterating until $||p_q^{l+1}(x) - p_q^l(x)|| \leq eps$. Each iteration performs the following steps:

Step 1: Update the value function using a sparse parametric approximation and LASSO (third order polynomials are used). The estimation of the value function approximating parameters, $\widetilde{\gamma}^y$, relies on the approximate Bellman equation:

$$
\left(f^y(x) - \beta E \left[f^y(x') | p_q^l(x), x\right]\right) \widetilde{\gamma}^{y,l+1} = E_\varepsilon \left[\pi^y(x, q, \varepsilon); p_q^l(x)\right]
$$

54

where ex ante profits are computed by:

$$E_\varepsilon \left[\pi^y(x, q, \varepsilon); p_q^l(x) \right] = \left[VE(x) - c_1(x) \right] \sum_{q=0}^{\bar{q}} q p_q^l(x) - c_2 \sum_{q=0}^{\bar{q}} q^2 p_q^l(x) + \sigma \sum_{q=0}^{\bar{q}-1} \phi(A^l(x, q))$$

where $A(x, q) = \Phi^{-1} \left(1 - \sum_{k=0}^{q} p_k^l(x) \right)$, $q = 0, 1, ..., \bar{q} - 1$. To derive the above, use (24) and (27). Moreover, to compute $E\left[f(x') | p_q^l(x), x \right]$, I simulate d_t and l_t one period forward since these are the only stochastic states now. State transitions are computed via (32). Due to computational constraints I have assumed throughout that shipyards keep track of the total backlog rather than the distribution of backlogs. Therefore, at this stage shipyards do not have the full information to predict total orders accurately. To circumvent this issue I make the simplifying assumption that shipyards believe they are all at the same state and can predict total orders using the total number of firms.

Step 2: Update the choice probabilities:

$$
\begin{aligned}
p_0^{l+1}(x) &= 1 - \Phi \left(\frac{1}{\sigma} \left(VE^o(x) - c_1(x) - c_2 + \beta \left(CV^{y,l+1}(1) - CV^{y,l+1}(0) \right) \right) \right) \\
p_q^{l+1}(x) &= \Phi \left(\frac{1}{\sigma} \left(VE^o(x) - c_1(x) - c_2(2q - 1) + \beta \left(CV^{y,l+1}(q) - CV^{y,l+1}(q - 1) \right) \right) \right) - \\
&\quad - \Phi \left(\frac{1}{\sigma} \left(VE^o(x) - c_1(x) - c_2(2q + 1) + \beta \left(CV^{y,l+1}(q + 1) - CV^{y,l+1}(q) \right) \right) \right) \\
p_{\bar{q}}^{l+1}(x) &= 1 - \sum_{q=0}^{\bar{q}-1} p_q^{l+1}(x)
\end{aligned}
$$

I solve the above fixed point under three scenarios: the true post-2006 world, a world with no China intervention and a world without the Chinese new shipyards. These worlds differ in the shipyard cost function, the set of active shipyards and the shipyard capital structure. I perform the fixed point on all data.

Finally, to simulate the model, I draw cost shocks ε and obtain the corresponding optimal production using (9) which in turn relies on the value function computed using the parameters $\widetilde{\gamma}^y$ retrieved by the fixed point algorithm. At each state visited, $E\left[f(x') | p_q^*(x), x \right]$ is computed using the retrieved equilibrium choice probabilities.

9.8   Monte Carlo: Sparse Approximations in Dynamic Programming

Here, I present a Monte Carlo to evaluate the performance of the methodology used in the paper based on sparse approximations. To do so, I introduce a simple dynamic shipbuilding model motivated by the paper, but of much smaller dimensionality. Roughly, the model introduces production dynamics in the toy model of Section 3.1. I show that using sparse approximations performs equally well or better than other standard methods (e.g. state space discretization or

Table 16: Monte Carlo parameters.

β	0.95	R_0	$\begin{pmatrix} 200 \\ 1.5 \end{pmatrix}$
δ	0.3		
$\bar{q}$	9		
$c_{1,0}$	35	R	$\begin{pmatrix} 0.9 & 2.5 \\ 0 & 0.93 \end{pmatrix}$
$c_{1,1}$	-1.1		
$c_{1,2}$	0.018		
c_2	2.5	Σ_ξ	$\begin{pmatrix} 3600 & 0 \\ 0 & 0.04 \end{pmatrix}$
σ	8		

parametric approximations) both in terms of solving the model to obtain the optimal policy and value function, as well as in terms of estimating the payoff parameters of interest.

A simplified shipbuilding model The agent (a shipyard) maximizes its expected lifetime payoff from ship production. The shipyard's action is how many units (ships) to produce, $q \in \{0, 1, ..., \bar{q}\}$. Its state comprises of its backlog b and an aggregate state s with two components: the demand for shipping services and the total fleet. A shipyard gets price $P(b, s)$ which is increasing in demand for shipping and decreasing in the total fleet (both variables encoded in s); the price also depends on the yard's backlog b which affects its delivery date.

The backlog transitions deterministically according to the rule: $b' = (1 - \delta)b + q$, as in the Appendix above. In words, the shipyard delivers $\delta\%$ of its backlog every period and production is added to the backlog. It is easy to show that this transition rule confines b in the interval $[0, \bar{q}/\delta]$.

The aggregate state follows a stationary VAR model: $s' = R_0 + Rs + \xi$, where $\xi \sim N(0, \Sigma_\xi)$.

The shipyard's per period payoff is:

$$\pi(q, b, s, \varepsilon) = P(b, s)q - c_1(b)q - c_2 q^2 - \sigma q \varepsilon,$$

where $\epsilon \sim N(0, 1)$. Note that the cost function is quadratic in quantity, while the coefficient c_1 depends on the shipyard's backlog, making the production choice dynamic (I assume that $c_1 = c_{1,0} + c_{1,1}b + c_{1,2}b^2$, where $c_{1,1} < 0$ and $c_{1,2} > 0$).

Table 16 provides the Monte Carlo parameter values for the model.[48]

Solving the model The shipyard's Bellman equation is:

$$V(b, s, \varepsilon) = \max_{q \in \{0,...,\bar{q}\}} \{\pi(q, b, s, \varepsilon) + \beta E_{\varepsilon', s'} V\left((1 - \delta)b + q, s', \varepsilon'\right)\} \tag{34}$$

---

[48] The price is given by:
$$P(b, s) = 4 * demand - 0.015 * fleet - 0.4b.$$

while the ex ante value function is $V(b, s) \equiv E_\varepsilon V(b, s, \varepsilon)$. The shipyard's optimal policy is characterized by the conditional choice probabilities. Let $p_q(b, s)$ be the probability of choosing q at state (b, s). These choice probabilities are determined by optimal thresholds over the range of ε, as shown in Lemma 1 of the paper.

Ideally, one would like to first solve this model "exactly" and then solve it with different "approximate" methodologies (state space discretization, parametric approximations, sparse parametric approximations) to compare their performance. Here, determination of the exact solution is not straightforward as the states are continuous. To address this issue, I discretize the state space very finely and consider this the exact solution (note that all states are bounded). Then, I compare the following three solved models: (i) (coarse) state space discretization; (ii) polynomial approximation of the value function; (iii) sparse polynomial approximation of the value function.

The models are solved via a typical fixed point algorithm that obtains the value function and the optimal policy. The algorithm iterates two steps, similar to those in the paper; one updates the policy function and the other updates the value function. In particular, at iteration l, the algorithm inherits an ex ante value function $V^l(b, s)$ and optimal policy given by the choice probabilities $p_q^l(b, s)$ and then performs the following two steps:

1. Compute the optimal policy $p_q^{l+1}(b, s)$ by calculating the optimal thresholds of Lemma 1 using $V^l(b, s)$.

2. Update the ex ante value function to $V^{l+1}(b, s)$ by first computing the ex ante per period payoff as in the paper and then using the Bellman equation (34). There are two implementations here:

   (a) State space discretization (fine or coarse):

$$V^{l+1}(b, s) = \tilde{\pi}(b, s) + \beta \sum_{q=0}^{\bar{q}} p_q^{l+1}(b, s) \sum_{s'} V^l\left((1-\delta)b + q, s'\right) \Pr(s'|s) \tag{35}$$

   where $\tilde{\pi}(b, s) \equiv E_\varepsilon\left[\pi(q, b, s, \varepsilon); p_q^l(b, s)\right]$ is the ex ante payoff under the optimal policy and $\Pr(s'|s)$ is the transition probability of the aggregate state s in the discretized world (more on this below).

   (b) Parametric approximations: here the value function is of the form $V(b, s) = \gamma f(b, s)$, where $f(b, s)$ denotes a vector of polynomial terms in the state variables and γ a conforming vector of coefficients. To update the value function, use:

$$\gamma^{l+1} f(b, s) = \tilde{\pi}(b, s) + \beta \gamma^{l+1} \sum_{q=0}^{\bar{q}} p_q^{l+1}(b, s) E_{s'} f\left((1-\delta)b + q, s'\right) \tag{36}$$

   where $E_{s'}$ now employs the VAR model for s. γ^{l+1} can be obtained via OLS (standard

parametric approximation) or LASSO (sparse parametric approximation). Then, set
$V^{l+1}(b, s) = \gamma^{l+1} f(b, s)$.

Before proceeding to the results, I provide some further details on the construction of the various models.

- Fine state space discretization (capturing the exact solution): To construct the state space, I use the full range of b, $[0, \bar{q}/\delta]$ and its actual transition. I discretize s using 25 equidistant grid points for each of the two variables. The resulting number of states exceeds 14,000. To construct the transition of this finite aggregate state, I employ the VAR model as follows: I simulate s forward for a long time horizon; then each state of the generated paths is replaced by its closest finite state. The transition probabilities are calculated by the empirical frequencies (this procedure is followed in Kalouptsidi (2014)).

- Coarse state space discretization: The typical difficulty in using state space discretization to solve dynamic programming problems is that often it cannot accommodate large state spaces. To reflect this difficulty, I consider a coarser discretization of the same state space: I now use 5 grid points for each of the two aggregate state variables, but still consider the full range of b. This results in a total of 775 states. I construct the transition matrix following the same procedure as above. It is important to note that since the exact model is itself discretized, this method is likely "positively biased".

- Parametric approximations: I explore polynomial approximations of various degrees (second, third and fourth). Here, I employ the VAR model to compute expectations directly, as in the main body of the paper. The size of the state space used to perform the fixed point, is larger than the one of the coarse discretization (around 5,000 grid points); this is in order to capture the benefit of allowing for continuous states when using approximations.

- Sparse parametric approximations: The dimension of the state space may lead to an explosive number of possible terms to be included in $f(x)$. Additionally, choosing which terms to include in the approximation may prove to be an arduous process. A solution which circumvents both issues is to estimate the sparse vector γ in each iteration via LASSO instead of OLS. This way, the polynomial approximation of the value function can be of higher power, allowing for more flexibility, while term selection takes place at the same time. As in the paper, in most specifications below I use the static model to determine the polynomial terms that are maintained.

Figure 10 depicts the shipyard's value function as a function of the backlog (on average across other states). As shown, the value function obtained from the coarsely discretized model is furthest away from the true. Polynomial and sparse polynomial approximations both do well. Figure 11 plots the evolution of a shipyard's backlog over 35 periods. Again, parametric approximations do better
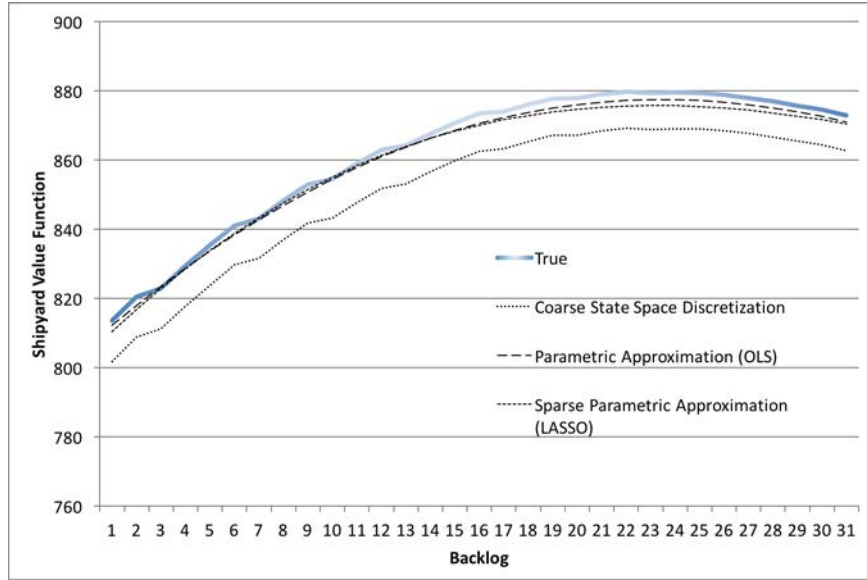
Figure 10: The shipyard's value function as a function of backlog (average across other states).

Table 17: Model solution.

Model	Exact	Coarse	Parametric Approx (OLS)			Sparse Parametric Approx (LASSO)		
State Space Size	14074	775	5766	5766	5766	5766	5766	5766
Polynomial degree	-	-	2	3	4	2	3	4
Mean Value Function	861.495	851.017	855.099	860.496	862.458	857.969	860.333	861.987
Median Value Function	848.347	843.125	839.304	843.454	847.254	845.277	843.753	842.567
St.dev. Value Function	195.598	117.096	201.855	208.150	208.391	201.605	203.648	204.809
Av. Distance Value Function	-	0.119	0.017	0.019	0.019	0.015	0.015	0.016
Av. Distance CPs	-	0.210	0.437	0.399	0.399	0.437	0.429	0.419
Simulated Paths								
Av.Distance Backlog	-	0.034	0.025	0.024	0.022	0.025	0.024	0.024
St.dev. Backlog	-	0.045	0.038	0.036	0.034	0.038	0.036	0.035
Av.Distance Production	-	0.062	0.041	0.040	0.037	0.041	0.040	0.039
St.dev Production	-	0.120	0.109	0.105	0.098	0.109	0.107	0.098

Av. Distance is the average normalized distance from the true model across states or across the points of the simulated path.

than the coarse state space discretization. Table 17 collects a number of statistics pertaining to the different models. In particular, it compares the models in terms of both the obtained value function and choice probabilities, as well as the production and backlog obtained when simulating the model. The (coarsely) discretized model appears to do worse. Both regular and sparse parametric approximations do well.

Estimating the model The second part of the Monte Carlo exercise is to evaluate the use of the different methods as part of an estimation strategy aiming to recover the parameters of the per period payoff function, and in particular those of the cost function, $\{c_{1,0}, c_{1,1}, c_{1,2}, c_2, \sigma\}$.

I create a dataset by simulating the exact model for 50 periods and 200 shipyards.

The estimation procedure closely follows the two-step approach used in the main body of the
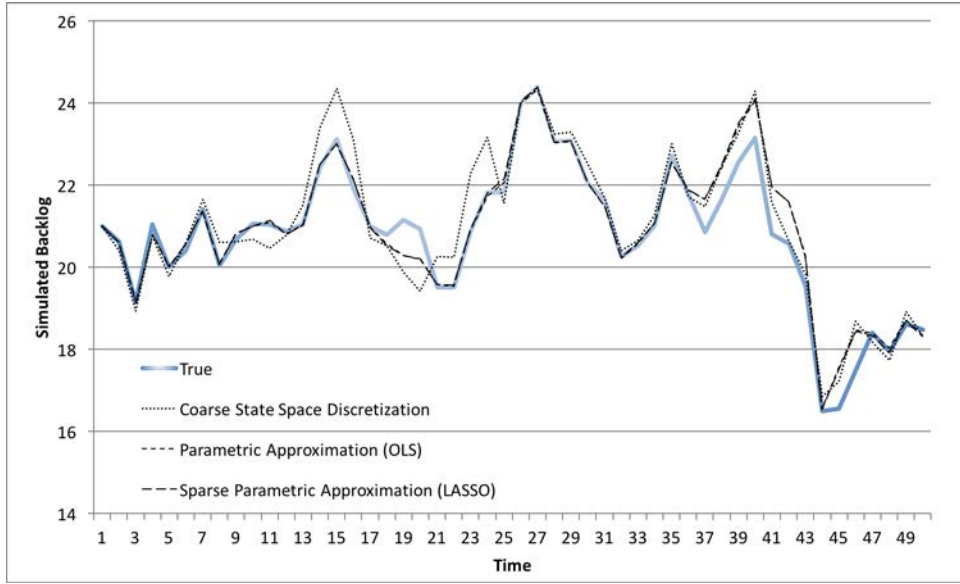
Figure 11: Simulated backlog path.

paper: first, I estimate the optimal policy thresholds using the choice probabilities.[49] In the second step I maximize the likelihood of the paper; different methods differ in how the value function is obtained. When a coarse state space discretization is used, I solve for the value function via a fixed point following (35); when parametric approximations are used, I follow the paper's procedure, summarized in (36).[50]

Table 18 presents the results. I allow for higher order terms in the case of sparse approximations to capture the method's ability to handle larger problems. The recovered parameter estimates via sparse approximations are closer to the true values than regular parametric approximations, though both do well. In contrast, the coarsely discretized model performs worse. Results are also robust over a range of the penalty parameter λ (though some sensitivity is to be expected since λ trades off fit and sparsity), as well as to more structured methods of determining the penalty parameter (cross validation, Belloni and Chernozhukov (2011)). Moreover, similar results are obtained when using Chebyshev polynomials.

---

[49]To simplify the analysis, instead of estimating choice probabilities from the simulated data, I use the choice probabilities from the true model.

[50]The term $E_{s'} f\left((1 - \delta)b + q, s'\right)$ is obtained by simulating the VAR for the aggregate state variables one period forward (1000 simulation draws).

Table 18: Model estimation.

Polynomial degree	True	Coarse	Parametric Approx (OLS)		Sparse Parametric Approx (LASSO)
	-	-	2	3	4
$c_{1,0}$	35	30.26	36.34	35.68	33.21
$c_{1,1}$	-1.1	-1.1	-1.30	-1.26	-1.08
$c_{1,2}$	0.018	0.016	0.0237	0.0224	0.0192
c_2	2.5	3.27	2.57	2.61	2.58
σ	8	10.54	8.38	8.49	8.23
Distance	-	0.18	0.12	0.10	0.04

Distance is the average normalized distance from the true vector of parameters.

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] Baldwin, Richard E. and Harry Flam. 1989. "Strategic Trade Policy in the Market for 30-40 Seat Commuter Aircraft". Weltwritschaftliches Archiv Band 125: 484-500.

[4] Baldwin, Richard E. and Paul R. Krugman. 1987a. "Market Access and International Competition: A Simulation Study of 16K Random Access Memory" in R.C. Feenstra (ed.), Empirical Methods for international Trade, Cambridge, The MIT Press.

[5] Baldwin, Richard E. and Paul R. Krugman. 1987b. "Industrial policy and international competition in wide-bodied aircraft" in R.E. Baldwin (ed.), Trade Policy lssues and Empirical Analysis, Chicago, University of Chicago Press for the National Bureau of Economic Research.

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

[7] Barwick Jia, Panle and Parag A. Pathak. 2012. "The Costs of Free Entry: An Empirical Study of Real Estate Agents in Greater Boston". Working Paper.

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

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

[10] Berry Steven, James Levinsohn and Ariel Pakes. 1995. "Automobile Prices in Market Equilibrium". Econometrica 7 (1): 841-890.

[11] Berry Steven, James Levinsohn and Ariel Pakes. 1999. "Voluntary Export Restraints on Automobiles: Evaluating a Trade Policy". American Economic Review 89(3): 400-430.

[12] Brander James E. 1995. "Strategic Trade Policy". Handbook of International Economics, Chapter 27, Vol. III, Edited by G. Grossman and K. Rogoff.

[13] Chen, Jiawei, Susanna Esteban and Matthew Shum. 2013. "When do Secondary Markets Harm Firms?". American Economic Review (103): 2911-2934.

[14] Collard-Wexler, Allan. 2013. "Demand Fluctuations in the Ready-Mix Concrete Industry". Econometrica 81 (3): 1003–1037.

[15] 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.

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

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

[18] Goldberg Koujianou, Pinelopi. 1995. "Product Differentiation and Oligopoly in International Markets: The Case of the U.S. Automobile Industry". Econometrica 63 (4): 891-951.

[19] Grossman, M. Gene. 1990. "Promoting New Industrial Activities: A Survey of Recent Arguments and Evidence". OECD Economic Studies 14: 87-126.

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

[21] Hamilton, James D. 1994. "Time Series Analysis". Princeton, NJ: Princeton University Press.

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

[23] 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.

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

[25] Kalouptsidi, Myrto. 2014. "Time to Build and Fluctuations in Bulk Shipping". American Economic Review 104(2): 564-608.

[26] Kalouptsidi, Myrto, Paul T. Scott and Eduardo Souza-Rodrigues. 2016. "Identification of Counterfactuals in Dynamic Discrete Choice Models". NBER Working Paper 21527.

[27] Krussel, Per and Anthony A. Smith. 1998. "Income and Wealth Heterogeneity in the Macroeconomy". Journal of Political Economy 106(5): 867-896.

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

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

[30] 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.

[31] Ryan, P. Stephen. 2012. "The Costs of Environmental Regulation in a Concentrated Industry". Econometrica 80(3): 1019–1062.

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

[33] Sykes, Alan O. 2009. "The Questionable Case for Subsidies Regulation: A Comparative Perspective". Stanford University School of Law, Law and Economics Research Paper Series, no. 380.

[34] 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.

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

[36] U.S.-China Economic and Security Review Commission. 2011. Annual Review.

[37] US International Trade Commission. 2008. "Antidumping and Countervailing Duty Handbook". 13th Edition

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

[39] WTO. "Agreement on Subsidies and Countervailing Measures"

[40] WTO. 2006. "World Trade Report"

[41] Xu, Y. Daniel. 2008. "A Structural Empirical Model of R&D, Firm Heterogeneity and Industry Evolution". Working Paper.