China's Industrial Policy: an Empirical Evaluation

Panle Jia Barwick Myrto Kalouptsidi Nahim Bin Zahur*

Abstract

Industrial policies, broadly defined as policies that shape a country's or region's industry structure by either promoting or limiting certain industries or sectors, have been widely used in developed and developing countries. Despite its importance, few empirical studies directly evaluate the welfare consequences of these policies using micro-level data. This study examines an important industrial policy in China – the policy to develop the country's shipbuilding industry to the largest worldwide in the 2000s. Using comprehensive data on shipyards worldwide, we quantify the magnitude of China's industrial policies in supporting its domestic ship building industry using a dynamic model. Our estimates suggest that the combined policy benefit exceeds 600 billion RMB from 2006 to 2013, boosted China's domestic investment and entry by 300% and 200%, respectively, and enhanced its world market share by 40%. On the other hand, the policy creates sizable distortions and leads to high fragmentation. Entry subsidy is wasteful, though production and investment subsidies can be justified based on market share considerations. Investment subsidy is less distortive than entry or production subsidy, partly because of the dynamic sorting that favors efficient firms. Finally, we find little evidence in supportive of the traditional justifications for industrial policies. Our results point to the importance of non-economic considerations in contributing to the policy design.

Keywords: Industrial Policy, Dynamic, Subsidy, Investment, Welfare

*Barwick: Department of Economics, Cornell University and NBER (email: panle.barwick@cornell.edu); Kalouptsidi: Department of Economics, Harvard University and NBER (email: myrto.kalouptsidi@harvard.edu); Zahur: Department of Economics, Cornell University (email: nz96@cornell.edu). We thank Steve Berry, Julieta Caunedo, JF Houde, and seminar participates at Cornell-Penn State Conference, CRETE, Duke IO-Trade conference, IIOC 2017, IMF, MIT/Harvard, NBER Chinese Economy Working Group, SED, Stanford, and UPenn for helpful comments. Mingyu Chen provided excellent research assistance in the early stage of this project.

1 Introduction

Industrial policies, broadly defined as policies that shape a country's or region's industry structure by either promoting or limiting certain industries or sectors, have been widely used in developed and developing countries. Examples include U.S. and Europe after the World War II, Japan in the 1950s and 1960s (Johnson, 1982; Ito, 1992), South Korea and Taiwan in the 1960s and 1970s (Amsden, 1989), China, India, Brazil, and other developing countries in the last couple of decades (Stiglitz and Lin, 2013). As Rodrik (2010) puts it, "The real question about industrial policy is not whether it should be practiced, but how."

The theoretical literature on industrial policies is divided into two camps. Proponents argue that government intervention via industrial policies can help correct for externalities that are not internalized by individual firms or mitigate negative consequences of inherent wedges in the economy that distort firm decisions (Hirschman, 1958; Wade, 1990; Greenwald and Stiglitz, 2013; Liu, 2018). On the other hand, opponents cautioned the potential disastrous consequences of industrial policies due to the asymmetric information, regulatory capture, and other political economy constraints (Baldwin, 1969; Krueger, 1990).

Despite the contentious debate in the literature and the prevalence of industrial policies in practice, there are remarkably few empirical studies that directly evaluate the costs and benefit of these policies using micro level data. There seems a major disconnect between the vast theoretical literature and the empirical analysis that has largely focused on describing what happens to the benefiting industries (or countries) in terms of output, productivity, growth rate, and has been silent on the welfare consequences (see Harrison and Rodriguez-Clare (2010) for an excellent review).

This study fills this hole in the literature by analyzing an important industrial policy in China – the policy to develop the country's infant shipbuilding industry to be the largest worldwide in the last couple of decades. China's shipbuilding industry was nascent at the turn of this century and accounted for less than 5% of the world production.[1] When former Premier Zhu inspected China State Shipbuilding Corporation (CSSC), one of the two largest shipbuilding conglomerates in China, in Liaoning in May 2002, he pointed out that "China hopes to become the world's largest shipbuilding country (in terms of dead weight tons) ... by 2015." Shipbuilding was dubbed as a strategic industry in the 11th National Five-Year Plan (2006-2010) in need of "special oversight and support". Since then, an unprecedented number of national policies were issued. During the Eleventh (2006-2010) and Twelfth (2011-2015) Five-Year plans, shipbuilding was identified as a pillar or strategic industry by twelve and sixteen provinces, respectively. By 2013, the end of our sample period, China became the largest shipbuilding country in terms of volume, overtaking Japan and South Korea.

[1] Data source: Clarkson Ltd. Ships used for domestic inland navigation are excluded.

Having identified an important historical event, we ask the following questions: how has China's industrial policy shaped its domestic and the world shipbuilding industry? What are the welfare consequences of this policy? How do different policy instruments interact with each other and what are general lessons and guidance that can be applied to other contexts? To answer these questions, we provide the first attempt to collect comprehensive data on all firms in the global shipbuilding industry from 1998 to 2013. Then we quantify the impact of policy support on targeted firms' production, investment, entry and exit, using a dynamic model that incorporates firm learning and accommodates private investment shocks. Finally, we conduct counter-factual analyses to examine the welfare consequences of this policy that spurred the growth of China's shipbuilding industry.

Our analysis delivers three main findings. First, like other policies unleashed by China's central government in the past several decades, the scale of the supportive policy is massive (relative to the size of this industry). Our estimates suggest that the policy support is equivalent to 600 bn RMB during our sample period. It boosted China's domestic investment and entry by 300% and 200%, respectively, and enhanced its world market share by 40%. However, the policy created sizable distortions and generated a net gain of merely 160bn RMB to domestic producers and world wide consumer surplus of 230bn RMB. The policy attracted a large number of inefficient producers, exacerbated the extent of excess capacity, and did not translate into significant higher industry profit over the long run.

Turning to different components of the industry policy, the effectiveness of different policy instruments is mixed. Production and investment subsidies can be justified based on market share considerations, but entry subsidy is extremely wasteful and increases industry fragmentation and idleness. Production subsidy is more effective at achieving output targets, while investment subsidy is less distortive over the long run, due to a more effective targeting at efficient firms (dynamic sorting). Finally, welfare distortions induced by these policies are convex and deteriorate with the magnitude of subsidies. Our results provide limited evidence that the shipbuilding industry generates significant spillover to the rest economy and point to the direction that non-economic arguments might be more important in the design of this policy.

We make three contributions to the existing literature. This is the first empirical analysis of the welfare consequences of a large-scale industrial policy using micro-level data that cover the entire global industry. Second, methodologically, we non-parametrically estimate firm dynamic investment decisions with unobserved shocks (which are shown essential in determining firms' heterogeneous investment decisions). Third, our results are relevant to policy makers in both developing and developed countries and highlight the importance of taking distributional effects into consideration in policy designs.

The rest of the paper is organized as follows. Section 2 gives an overview of China's shipbuild-

ing industry and discusses the relevant industrial policy and our data sets. Section 3 incorporates industrial policy into a market equilibrium model of ship demand and supply. The estimation results are presented in Section 4. Section 5 quantifies the welfare impact of industrial policies and discusses various channels that could justify these preferential policies. Section 6 concludes.

2 Industry Background and Data

2.1 Industry Background

The scope and frequency of national policies issued since 2000's, especially after 2005, that aim at supporting the shipbuilding industry is unprecedented. After former Premier Zhu's inspection of China State Shipbuilding Corporation (CSSC), the National Marine Economic Development Plan in 2003 proposed constructing shipbuilding bases centered at the Bohai Sea area (Liaoning, Shandong, and Heibei), the East Sea area (Shanghai, Jiangsu, and Zhejiang), and the South Sea area (Guangdong).

China dubbed shipbuilding as a strategic industry in the 11th National Five-year Economic Plan (2006-2010). Since then, the shipbuilding industry, as well as the marine equipment industry and the ship repair industry, has received numerous supportive policies. During the Eleventh (2006-2010) and Twelfth (2011-2015) Five-Year plans, shipbuilding was identified as a pillar or strategic industry by twelve and sixteen provinces, respectively. Besides the Five-Year Plans, the central government issued a series of policy documents with specific production and capacity quotas. For example, in 2006 (as part of the The Medium and Long Term Development Plan of Shipbuilding Industry), the government set an annual production goal of 15 million deadweight tonne (DWT) to be achieved by 2010, and 22 million DWT by 2015. The country met this goal in 2008, several years in advance. Table 1 documents major national policies during this period that we have uncovered.

There is substantial heterogeneity in the extent of government support across provinces. Zhejiang was one of the first provinces that identified shipbuilding as a provincial pillar industry, with the first major policy released in 2004. Jiangsu is the close second, and set up special banks to help shipbuilding companies with buyer-financing etc. During the Twelfth Five-Year plan (2011-2015), Zhejiang and Jiangsu provinces reportedly invested 1.26 trillion RMB and 1.34 trillion RMB, respectively, in supporting local marine economy, including infrastructure investment (building roads and renovating ports and piers), modernizing the logistics system, developing and supporting large enterprises, etc.[2]

An often explicitly-stated goal of China's industrial policy is to create large successful firms that could compete against international conglomerates. In the aftermath of the economic crisis of

[2]Source: https://baike.baidu.com/item/haiyangjingji, last accessed Feb 26, 2018.

2008 that led to a sharp decrease in global new ship prices, the government promoted consolidation through filtering out inefficient firms. The Plan on Adjusting and Revitalizing the Shipbuilding Industry, implemented in 2009, resulted in an immediate moratorium on entry with increased investment subsidies to existing firms. The most crucial policy for achieving consolidation objectives is "Shipbuilding Industry Standard and Conditions" (2013), where the government periodically announces a list of selected firms that meet the industry standard and receive priority in subsidies and bank financing.[3] The so called "White List" included 60 firms in 2014 upon announcement, and has undergone a number of revisions since.

2.2 Data

Our empirical analysis draws on a number of datasets. The first data set comes from Clarksons, which is a leading ship brokerage firm based in the UK. This dataset contains quarterly data for all shipyards (including both Chinese and non-Chinese yards) between first quarter in 1998 and first quarter in 2014. We observe each yard's orders, deliveries, and backlog measured in Compensated Gross Tonnage (CGT), for a number of different ship types: bulk, tankers, and containers.[4] Finally we observe each shipyard's first year of delivery and the first year that it took an order. We use these two variables to construct the yard's entry year and its age.[5]

The second data source is the annual survey conducted by the National Bureau of Statistics (NBS) of Chinese manufacturing firms with annual sales above five million RMB. For each shipyard, we observe operating costs, fixed assets, sales revenue, wages, number of employees, etc. In addition, we observe the location of each shipyard (province and city) and its ownership status (state-owned enterprises, or SOEs, privately owned, or joint ventures with foreign investors). We follow the procedure described by Brandt et al. (2012) to link firms over time and construct each firm's real capital stock and investment. Our firm-level investment data ranges from 1998 to 2008 and from 2011 to 2012.[6]

Our third data source is China's economic census in 2004 and 2008 that covers all firms in China. This includes additional variables on sales, exports, employees by education, revenue,

[3] In practice, favorable financing terms and capital market access are often limited to firms on the white list.

[4] Backlog refers to undelivered orders. A large backlog could lead to economies of scale where shipyards gain efficiency by working on multiple ships at the same time, or diseconomies of scale when a shipyard hits capacity constraints and relies on overtime or multiple shifts to meet the delivery deadlines. Compensated Gross Tonnage, widely used in this industry, is a size measure that takes into consideration production complexities of different ship types and is comparable across types.

[5] The entry year is defined as the first year the yard took an order or two years prior to the first year the yard made a delivery, whichever is earlier. We subtract two years from the first year of delivery to account for the time to build.

[6] The Chinese manufacturing survey has some limitations. Most importantly, data for 2010 is missing. This prevents us from constructing the firm-level investment in either 2009 or 2010 – since investment is imputed from changes in the capital stock – though we can impute a firm's total investment in those years. Finally, investment in 2013 is missing because the sample ends in that year.

financial information (capital, assets, inventory, revenue, profit, tax, payroll, employment, etc.), total production, total sales, inventory, and depreciation. We merge together these datasets to obtain a quarterly panel of Chinese, Korean, and Japanese shipyards ranging from 1998 to 2013.[7]

In addition to these yard-level variables, we collect a number of aggregate variables for the shipbuilding industry, including quarterly global prices, measured in USD/CGT, for each of the three ship types.[8] We use the steel ship plate price to proxy for changes in the cost of building a new ship as steel is a major input in shipbuilding. Variables that reflect demand for new ships include the wheat price and Chinese iron ore imports for bulk carriers, Middle East refinery production for oil tankers, and world car trade for containerships.

Table 2 contains summary statistics on key variables relating to the shipbuilding industries in China, Japan, and Korea. There are a total of 14,455 firm-quarter observations, out of which 7186 are for Chinese yards, 5448 are for Japanese yards and 1821 are for Korean yards. Shipyards take new orders infrequently, less than 16% of the firm-quarter observations for bulk and lower for tanker and containerships. The infrequency of new orders and the fact that yards are often idle are both key features of the shipbuilding industry that inform our modeling choices in Section 3.

Like other manufacturing industries in China, the shipbuilding industry experienced exponential growth since early 2000s and became the largest shipbuilding country in terms of deadweight tons in 2009, taking over South Korea and Japan. Figure 1 plots the rise of China's shipbuilding industry from 1998 to 2013 and the stunning speed at which this 'catching-up' took place.

There is massive entry of shipyards along China's coastal area during the same period. Figure 2 plots the aggregate number of entering shipyards for China, Japan, and South Korea. Entry is on average 1.4 per year for Japan and 0.9 per year for South Korea. In contrast, the number of new shipyards in China varied from 30 to 45 during the boom years when entry subsidy was in place. Entry dropped to 15 in 2009 and became minimal after the 2009 entry moratorium was put in place.

The rise in entry was accompanied by a large and unprecedented increase in capital expansion in China's shipbuilding industry. Figure 3 shows that there was a steep increase of four-folds in investment starting from around 2006-2007 that was unparalleled in the industry's history. The expansion in capital was carried out not just by new entrants, but by incumbents as well.

The rapid rise in China's production, investment and entry documented here coincided with the introduction of China's industrial policies for the shipbuilding industry. The global shipbuilding industry went through a boom in the mid-2000s, roughly concurrent with China's initial expansion. As Figure 4 shows, ship prices began rising around 2003 and peaked in 2008, before collapsing in the aftermath of the financial crisis and remaining stagnant from 2009 through to 2013. China's

[7] For Korean and Japanese yards, we only have data from Clarksons on their orders, deliveries, backlog, as well as the number of docks and berths that serves as a proxy for capital.

[8] We convert these prices into real RMB/CGT for consistency with variables from the NBS manufacturing survey. Results are nearly identical if prices are measured in USD/CGT.

production and investment, on the other hand, continued expansion well after the financial crisis. In our empirical analysis, we account for the boom in ship demand when quantifying the impact of China's industrial policy on its domestic and global industry, as well as the extent to which the post-2009 stagnation in ship prices is attributable to China's expansion.

3 Model

In this section, we introduce a dynamic model of firm entry, exit, capital investment, and production. In each period, incumbent firms make static production decisions to maximize their period payoff taking global ship prices as given. They choose whether or not to exit, and conditional on staying, how much to invest. A pool of potential entrants make one-shot entry decisions based on their expected discounted stream of profits as well as the cost of entry. At the end of the period, entry, exit, and investment decisions are implemented and the states evolve.

Time is discrete and is a quarter. In period t there are $j = 1, ..., J_t$ firms in the world that produce ships. There are $m = 1, ..., M$ types of ships, such as containerships, dry bulkers, or oil tankers. Ships within a type are homogeneous.

Ship Demand In each period, the aggregate inverse demand for ships of type m is given by the function,

$$P_{mt} = P(Q_{mt}, d_{mt}) \tag{1}$$

for $m = 1, ..., M$, where P_{mt} is the market price of ship type m in period t, Q_{mt} is the total tonnage of type m demanded, and d_{mt} are demand shifters, such as the freight rates and the total fleet.

Ship Production Firm j produces q_{jmt} at cost:

$$C\left(q_{jmt}, s_{jmt}, \omega_{jmt}\right) = c_{\text{idle},m} + c_m\left(q_{jmt}, s_{jmt}, \omega_{jmt}\right) 1\left\{q_{jmt} > 0\right\}$$

where $c_{\text{idle},m}$ is a cost incurred even when shipyards have zero production, or a cost of "idleness". This cost is typically abstracted away from the literature, but it features prominently in later periods of our sample when the aggregate demand for new ships plummeted and many ships reported idle with zero production many periods in a row. It captures wages for workers who are not laid off during periods of minimal production, capital maintenance, land usage, costs of finding buyers (broker fees), etc.

The second term, $c_m(q_{jmt}, s_{jmt}, \omega_{jmt})$, is the standard production cost when $q_{jmt} > 0$. We use s_{jmt} to denote observed variables that affect firms' production cost, which includes firm characteristics (e.g. capital, backlog – the total ship tonnage by type ordered at the firm but not yet delivered,

age, location, ownership status) and aggregate cost shifters (e.g. government subsidies, steel prices) that affect all shipyards. With a slight abuse of notation, we also use s_{jmt} to denote all observed pay-off relevant variables (or state variables), which include ship prices in addition to cost shifters. Production cost also depends on a cost shock ω_{jmt}. The larger ω_{jmt} is, the less productive the firm is. The marginal cost of production is given by $MC\left(q_{jmt}, s_{jmt}, \omega_{jmt}\right)$.

Firms are price takers and choose how many tons to produce for ship type m, q_{jmt}, to maximize their profits:

$$\max_{q_{jmt}} \pi(s_{jmt}, \omega_{jmt}) = P_{mt} q_{jmt} - C\left(q_{jmt}, s_{jmt}, \omega_{jmt}\right)$$

If the optimal production tonnage for type m, q^*_{jmt}, is positive, it satisfies the following first order condition:

$$P_{mt} = MC_m\left(q^*_{jmt}, s_{jmt}, \omega_{jmt}\right) \tag{2}$$

Note that firms equate marginal cost with price instead of the marginal revenue because they are assumed as price takers. We discuss this assumption in detail later.

Let $s_{jt} = \{s_{j1t}, ..., s_{jMt}\}$ and $\omega_{jt} = \left\{\omega_{j1t}, ..., \omega_{jMt}\right\}$ denote the union across ship types of observed state variables and unobserved cost shocks, respectively. Firm's expected profit from all types, before the cost shocks are realized, is given by:

$$\pi(s_{jt}) = E_{\omega_{jt}} \sum_{m=1}^{M} \pi(s_{jmt}, \omega_{jmt})$$

Finally, in each period, the prevailing ship price, P_{mt}, equates aggregate demand and supply, where the aggregate supply is the sum of q^*_{jmt} defined in (2).

Investment and Exit Once firms make optimal production choice, they observe a scrap or sell-off value ϕ_{jt} that is distributed i.i.d. and decide whether to remain in operation or exit. If a firm chooses to exit, it receives a scrap value. If it remains active, it makes an investment i_{jt} by incurring an investment cost, $c^i(i_{jt}, v_{jt})$, where v_{jt} is an investment cost shock that is distributed i.i.d. and realized after the exit decision has been made.[9]

The value function for incumbent firm j is:

$$V(s_{jt}, \phi_{jt}) = \pi(s_{jt}) + \max\left\{\phi_{jt}, E_{v_{jt}}\left(\max_i\left(-c^i(i, v_{jt}) + \beta E\left[V(s_{jt+1})|s_{jt}, i\right]\right)\right)\right\} \tag{3}$$

Let $CV(s_{jt})$ denote the continuation value, which includes the expected cost of optimal invest-

[9]The timing assumption helps to separate the exit decision from the investment decision.

ment and the discounted future profit stream:

$$CV(s_{jt}) \equiv E_{\nu_{jt}}\left(-c^i(i^*, v_{jt}) + \beta E\left[V(s_{jt+1})|s_{jt}, i^*\right]\right)$$

where i^* denote the optimal investment $i^* = i^*(s_{jt}, v_{jt})$ and $E_{\nu_{jt}}$ is the expectation w.r.t. the random investment cost shock ν_{jt} that is observed by firms but unobserved by the econometrician.

The optimal exit policy is of the optimal threshold form: the firm exits the market if the drawn scrap value ϕ_{jt} is higher than its continuation value $CV(s_{jt})$. Since the scrap value is random, the firm exits with probability, $p^x(s_{jt})$, defined by,

$$p^x(s_{jt}) \equiv \Pr\left(\phi_{jt} > CV(s_{jt})\right) = 1 - F^{\phi}\left(CV(s_{jt})\right) \tag{4}$$

where F^{ϕ} is the distribution of ϕ.

Conditional on staying, firm j observes its investment shock. Its optimal investment $i^* = i^*(s_{jt}, v_{jt})$, if non-negative, satisfies the first-order condition,

$$\beta \frac{\partial E\left[V(s_{jt+1})|s_{jt}, i^*\right]}{\partial i} = \frac{\partial c^i(i^*, v_{jt})}{\partial i} \tag{5}$$

When the investment costs are prohibitively high, firm j opts for no investment and its capital depreciates at rate δ that is common to all firms.

Entry In each period t, there are $j = 1, ..., \bar{N}$ potential entrants in each Chinese region $r = 1, ...R$. To simplify notation, we suppress the region subscript r in the discussion below. It is understood that both the potential number of entrants and the entry cost distribution are region specific.

The potential entrants observe the pay-off relevant state variables and their own entry cost κ_{jt}, which is i.i.d., and make a one-time entry decision. If entrant j decides not to enter, it vanishes with a payoff of zero.[10] If it enters, it pays the entry cost and continues as an incumbent next period. In addition, the entrant is assumed to be endowed with and pay for a random initial capital stock that is realized next period when it becomes the incumbent and starts operation.

Potential entrant j solves,

$$\max\left\{0, -\kappa_{jt} + E\left[-c(k_{jt+1}) + \beta E\left[V(s_{jt+1})|s_t\right]\right]\right\}$$

where k_{jt+1} is entrant j's initial capital stock that is realized next period and $c(k_{jt+1})$ is the as-

[10]Here we follow the bulk of the empirical literature on industry dynamics (Ericson and Pakes, 1995; Ryan, 2012). Under this assumption, the entry decision involves a simple comparison between the value from entering the market and the random entry cost.

sociated building cost. The expectation is taking over entrant j's information set at time t, which includes all aggregate state variables.

The optimal entry policy is of the optimal threshold form; a potential entrant enters the market if the entry cost κ_{jt} drawn is lower than the value of entering, i.e. if

$$\kappa_{jt} \leq VE_{jt} \equiv E_t\left[-c(k_{jt+1}) + \beta E\left[V(s_{jt+1})|s_t\right]\right]$$

Since the entry cost, κ_{jt}, is random, the potential entrant enters with probability, p_t^e, defined by,

$$p_t^e \equiv \Pr\left(\kappa_{jt} \leq VE_{jt}\right) = F^{\kappa}\left(VE_{jt}\right) \tag{6}$$

where F^{κ} is the distribution of κ.

Industrial Policy Governmental interventions affect firms' decisions in several dimensions. We group policies that supported the Chinese shipbuilding industry into three categories: production, investment, and entry subsidies. Production subsidies lower the marginal cost of producing ships. Besides subsidized input material and export credits (Collins and Grubb 2008)[11], buyer financing in the form of collaborated loans provided by local banks constitutes another important component of production subsidies. Ships are the largest commercial products in the world. To help attract buyers, shipyards have traditionally offered loans and various financial services to facilitate purchasing payment. As documented in the 2010 OECD report, favorable terms on buyer financing from local banks fueled the production expansion among Chinese shipyards. Investment subsidies take the form of low-interest long-term loans that lower the per-unit cost of investment. Finally, reduced processing time and simplified licensing procedures as well as heavily subsidized coastal land greatly lower the cost of entry for potential entrants.

These policies affect costs of production, investment, and entry and are part of the payoff relevant state variables (s_{jt}). We assume that these policies are unexpected and perceived permanent.

Equilibrium The equilibrium of this model is defined as follows:

Definition 1. An equilibrium of this model consists of policies, $\{q_{jmt}\}_{m=1}^{M}$, $i^*(s_{jt}, v_{jt})$, $p^x(s_{jt})$, p_{rt}^e, value function $V(s_{jt})$ and prices P_{mt}, such that the production quantity satisfies (2) and maximizes period profit, the investment policy satisfies (5), the exit policy satisfies (4), the entry policy satisfies (6), and ship prices clear the market each period. Moreover, the incumbent's value function satisfies (3) and employs the above policies to form expectations.

[11] Until 2016, the Chinese government provided a range of subsidies for exporters, including reduced corporate income tax, refund of value-added-tax, etc. Shipbuilding companies benefit from export subsidies since most of their products are traded internationally.

Discussion We close this section with a discussion on our assumptions.

We begin by discussing the ship product market. We assume that shipyards are price-takers. This assumption is motivated by the large number of shipyards in the world. As discussed above, there are more than four hundreds shipyards globally, while the market share of the largest firm is less than 5%. The market is sufficiently unconcentrated that market power is not a first order consideration. In addition, as discussed in the empirical sections below, we consider a variation of our model where firms are Cournot competitors in the product market. In that case, the optimal production of firm j in period t (when positive) satisfies a variant of the first order condition (2) that includes a markup equal to $-q_{jmt}\left[\partial P\left(Q_{mt},d_{mt}\right)/\partial q_{jmt}\right]$. Empirically, this term turns out to be very small. Moreover, when the empirical analysis is replicated under the Cournot assumption, results are very robust. Finally, it is worth noting that on the shipowner side, monopsony power should not be a first-order issue. The dry bulk and tanker segments are unconcentrated. The containership segment may be concentrated among the operators, but a large number of small shipowning firms buy ships with the sole purpose of leasing them to the operators (and leasing rates have increased dramatically in the last decade).

Ships are assumed homogeneous within a type conditioning on size. In prior work (Kalouptsidi 2014, Brancaccio, Kalouptsidi and Papageorgiou, 2018), this assumption is substantiated for dry bulkers. These papers show that in secondary markets, the majority of price variation is accounted for by a ship's age and the aggregate macro variables (e.g., international trade flow). We have access to a small sample of spot-market contracts with detailed price information and ship attributes for new ships. Controlling for ship type and ship size, measured in Compensated Gross Tonnage, ship characteristics have limited explanatory power of the price variation, except the country fixed effects indicating that ships produced by Japanese and South Korean shipyards command a 5-8% price premium.

We assume away any dynamic consideration in production. In practice, producing a ship takes time and the production decision is in general dynamic: production amount today affects the backlog tomorrow, which affects tomorrow's operation costs and therefore production decisions. However, as documented in Kalouptsidi (2018), cost function estimates under static and dynamic assumptions are similar, especially the estimates that reflect the impact of policy intervention on firms' production costs. This is partly because the amount of drastic production expansion cannot be easily explained by the intertemporal considerations that arise with a dynamic production. We thus make the simplifying assumption that shipyards choose production to maximize their static profits. This keeps the model tractable and allows us to focus on entry, exit, and investment.

The production cost shock ω_{jmt} is assumed i.i.d. There are several reasons for our choice. First, the product cost shocks are estimated to be moderately persistent, with a serial autocorrelation of 0.15. Second, while it is straight-forward to estimate the persistence of these shocks using

observed quantity choices, incorporating a persistent time-varying unobserved state variable in a dynamic model raises considerable modeling and estimation challenges. For the same reason, the investment cost shocks are assumed i.i.d. Given that aggregate investment increased by more than four folds within a year upon the announcement of the 11th National Five-Year Plan (3) and that all firms expanded regardless of their efficiency level, the persistent investment shocks are unlikely a first-order contribution factor to the boom of the capital expansion observed in our sample.

Investment is assumed to be non-negative. Compared to the massive investment taken by Chinese shipyards, divestment is much less common and much smaller in magnitude. In addition, estimating the divestment decision with asymmetric adjustment cost introduces a kink in the cost function and makes the value function non-differentiable. Incorporating divestments makes estimation considerably harder but leads to similar estimates.

4 Estimation Strategy

In this section we present the empirical approach undertaken to estimate the model parameters. The key elements are the world demand function for new ships, the production cost function, the investment cost function, the distribution of scrap values, and the distribution of entry costs. There are few entry and exits and limited capital expansion among shipyards in Japan and South Korean. We estimate the heterogeneous production cost function for shipyards in all countries, but only estimate the dynamic decision for Chinese shipyards.

We proceed in three steps. In the first step, we estimate the demand curve for new ships as well as shipyard production costs. We use these estimates to construct firm profits in each period. In the second step, we estimate parameters governing investment costs and the scrap values for exiting firms, using a two-step estimator in the tradition of Hotz and Miller (1993) and Bajari et al. (2007). In the final step, we estimate entry costs.[12] Here, we describe each step in turn.

4.1 Demand

There are three main types of new ships, dry bulk carriers, oil tankers, and containerships. The demand curve (1) for type m is parametrized as follows:

$$log(Q_{mt}) = \alpha_{0m} + \alpha_{pm} P_{mt} + d'_{mt} \alpha_{dm} + \varepsilon^d_{mt} \tag{7}$$

The demand shifters d_{mt} includes freight rates, the total backlog of type m, and some other type-specific variables. Demand for new ships is higher when demand for shipping services is high,

[12] Another approach is to combine step two and three and estimate all of the dynamic parameters jointly. This is more efficient but computationally more complex.

reflected in higher freight rates. Conversely, high backlog implies that more ships will be delivered in the future, which reduces demand for new ships today. We also control for type-specific demand shifters: the wheat price and Chinese iron ore imports for bulk carriers, Middle East refinery production for oil tankers, and world car trade for containerships. In some specifications, we allow for time trends to capture the macro shocks. We also allow the price elasticity to change before and after 2006, the main policy year.

We instrument for prices using steel prices and steel production.[13] Steel is a major input in the shipbuilding process and accounts for about 13% of the cost of building a ship, according to Stopford (2009). The identification assumption is that steel prices and steel production are uncorrelated with new ship demand shocks ε^d_{mt}. This is a plausible assumption because only a small portion of steel produced is used in shipbuilding; as such, an increase in ship demand ($\varepsilon^d_{mt} > 0$) is unlikely to have much impact on steel prices and steel production.[14]

As there is a single global market for each ship type, demand for new ships is estimated using the time series variation. To improve the precision of parameter estimates, we restrict some parameters (e.g., the time trend) to be the same across ship types and estimate equation ((7)) jointly across the three types using GMM.[15]

4.2 Production Cost Function

We parameterize the marginal cost function, $MC_m\left(q_{jmt}, s_{jmt}, \omega_{jmt}\right)$, as follows:

$$MC_m(q_{jmt}, s_{jmt}, \omega_{jmt}) = \beta_{0m} + s_{jmt}\beta_{sm} + \beta_{qm} q_{jmt} + \omega_{jmt} \tag{8}$$

where q_{jmt} is tons of ship type m produced by firm j in period t and ω_{jmt} is normally distributed with mean zero and variance $\sigma^2_{\omega m}$.[16,17] Importantly, the vector of state variables s_{jmt} includes

[13] Other potential instruments include the aggregate number of shipyards, J_t, and the aggregate capital stock. These cost-side instruments shift the industry supply curve and are determined in period $t-1$, before demand shocks in period t are observed. Results are robust with or without these additional IVs.

[14] One potential concern is that steel production, if much of which is exported, could affect demand for dry bulkers. World trade in steel accounts for less than 4% in volume of goods transported by dry bulkers.

[15] Clarkson reports aggregate price and quantity for new ships back to 1990s. Our demand elasticity estimates, the most crucial ingredient for our empirical analysis, are similar if we extend the estimation sample back to the 1990s.

[16] It is understood here that s_{jmt} also includes firm specific state variables and polynomial terms of state variables. Note that we model shipyards' decision in tonnage, rather than number of ships. In practice, production is discrete rather than continuous. In other words, shipyards choose how many ships to produce and of what size. This optimization is substantially more complicated. Since our main focus here is on entry, exit, and investment we abstract away from this. It is worth noting that our estimated subsidies are consistent with those in Kalouptsidi (2018) that models the choice as number of ships and focuses on shipyards producing Handysize vessels (a type of dry bulkers).

[17] It is worth noting that because of time to build, there is a difference between the orders placed in a period, the deliveries and the production. Although we observe the first two, we lack data on the latter. Here, we use orders as a measure of q_{jmt}, following Kalouptsidi (2018). We do so because: (i) orders are the most important decision of the firms given the period's price, since orders are later deterministically transformed to production and orders; (ii)

dummies for policy intervention between 2006 and 2008 and then from 2009 onwards.

The optimal production q^*_{jmt} ,if positive, is given by the first order condition (2) that equates the marginal cost of production with new ship price:

$$q^*_{jmt} = \frac{1}{\beta_{qm}}\left(P_{mt} - \beta_{0m} - s_{jmt}\beta_{sm} - \omega_{jmt}\right) \tag{9}$$

The yard chooses positive production if and only if:

$$\omega_{jmt} < P_{mt} - \beta_{0m} - s_{jmt}\beta_{sm}$$

The parameters characterizing shipyards' production costs are $\theta^q \equiv \left\{\beta_{0m}, \beta_{sm}, \beta_{qm}, \sigma_{\omega m}, c_{\text{idle},m}\right\}_{m=1}^{M}$. One data feature is the frequency of zero orders in our sample: shipyards have zero orders in any given quarter 65% of the time. Besides observed shipyard attributes, zero orders are driven by unfavorable cost shocks: firms with high cost shocks, ω_{jmt}, are more likely choose zero production.

We estimate the cost function parameters θ^q using MLE. We construct the likelihood assuming that ω_{jmt} is i.i.d., but our estimates of production cost are also robust to serial correlation in ω_{jmt}. Computing the likelihood function in the Tobit model is cumbersome when ω_{jmt} is correlated over time. Robinson (1982) showed, however, that θ^c can be consistently estimated by maximizing the likelihood function that assumes (erroneously) that ω_{jmt} is i.i.d. Intuitively this is similar to how the OLS estimator in the standard linear regression model continues to be consistent (though no longer efficient) when we relax the assumption that the errors are i.i.d.

For each ship type m, the equation below illustrates the likelihood function of observed orders of shipyards over time, where we ignore any serial dependence in ω_{jmt} when forming the likelihood:

$$L_m = \prod_{q_{jmt}=0} Pr(q_{jmt} = 0 | s_{jmt}; \theta^c) \prod_{q_{jmt}>0} f_q(q_{jmt} | s_{jmt}; \theta^c)$$

Not every yard produces every ship type. A shipyard may choose $q_{jmt} = 0$ because its marginal cost is too high, or because it does not have the ability to produce type m. If we include the latter yards in the estimation, the estimates will be biased. Thus, before carrying out the estimation, we identify the yards that are potential participants in the market for type m, and which are not. We do so by assuming that every yard that ever produces type m is a potential participant in the market for type m for the entirety of its lifetime. When constructing the likelihood for type m, we only include observations for participant in the market for type m. Thus a yard that never took an order for type m is not going to be included in the estimation sample for type m.

it is not straightforward to smooth past orders into a production series, or match them to past prices; (iii) orders are directly observed. Moreover, we assume that production cost shocks ω_{jmt} are independent across types. Allowing for correlated cost shocks lead to very similar estimates.

In practice, estimating all parameters separately for each type leads to imprecise estimates. In our preferred specification, we estimate the parameters jointly by maximizing the product of the above likelihoods and impose the restriction that some of the coefficients (e.g. the effect of ownership, region, and subsidies) be the same across types. There are economically intuitive reasons behind the restrictions that we impose. For example, it is reasonable to suppose that subsidies are given to the firm as a whole and not necessarily earmarked for a particular ship type; this is especially because the same yard can produce many different kinds of ships at different times, depending on the prevailing market conditions.

Finally, it should be noted that a firm's production decisions provide no information on its idleness cost (or fixed cost of operation) $c_{\text{idle},m}$, since the firm incurs this cost regardless of whether or not it produces. We use accounting cost data to calibrate $c_{\text{idle},m}$. The idea behind the calibration is to select periods when the firm is not engaged in any production and use their reported costs during those periods to infer the average cost firms incur when they are idle. Further details of the calibration procedure are in Appendix A.

4.3 Estimation of investment cost and scrap value parameters

Our empirical analysis assumes that the cost of investment, $c^i(i, \nu_{jt})$, is a quadratic function of the amount invested and linear in the investment cost shock ν_{jt}. We allow the policies to affect the marginal cost of investment, with the parameter c_4 capturing the extent to which the policies affect the marginal cost of investment.

$$c^i(i_{jt}, \nu_{jt}) = c_1 i_{jt} + c_2 i_{jt}^2 + c_3 \nu_{jt} i_{jt} + c_4 T_t i_{jt}$$

Our strategy for estimating the dynamic model is to find the parameter vector θ that maximizes the joint likelihood function of the exit and investment decisions observed in the data. In order to do this, for each guess of θ, we need to solve for the value function of the shipyard using the Bellman equation. This is computationally costly and so to reduce the computational burden, we follow a two-step approach similar to Hotz and Miller (1993) and Bajari et al. (2007). We first flexibly estimate investment and exit policy functions from the data. Then, when solving the Bellman equation for each guess of θ, we plug in these estimated policy functions. This avoids the need to carry out policy function iteration within the inner loop for any guess of θ.

Because of the large dimension of the state space and the fact that several key state variables (such as capital) are continuous, it is not feasible to solve the value function at every possible realization of the state variables. We tackle this challenge in two ways.

First, we reduce the state space by using the structure of firm payoffs. In our setting, a subset of the state variables, s_{jt}^d, have deterministic transition processes and affect the firm's payoffs only

through a linear index $\beta' s^d_{jt}$. We reduce the state space by replacing this M-dimensional subset of states s^d_{jt} with a single-dimensional state $\bar{s}_{jt} = -\beta' s^d_{jt}$.[18] Appendix B discusses this procedure in more detail.

Second, we follow Barwick and Pathak (2015), Sweeting (2013) and Kalouptsidi (2018) and assume that the ex ante value function can be approximated by a set of B-spline basis functions of the state variables:

$$V(s_{jt}) = \gamma^0 f^0(s_{jt}) = \sum_{l=1}^{L} \gamma^0_l u_l(s_{jt}) \tag{10}$$

In principle, there is a very large set of B-spline basis functions that we could use to approximate $V(s_{jt})$. Appendix B.2 describes in more detail how we select the basis functions that we use to approximate $V(s_{jt})$.

4.3.1 Policy function estimation

We estimate the exit policy function by way of a probit regression:

$$Pr(\chi_{jt} = 1; s_{jt}) = \Phi(f(s_{jt}))$$

where $f(s_{jt})$ is a flexible polynomial of the states, and χ_{jt} equals 1 if firm j exits in period t. We denote the first-stage estimate of the probability of exit at each j,t by $\hat{p}^x(s_{jt})$.

Estimating the investment policy $I^*_{jt} = I(s_{jt}, \nu_{jt})$ is more challenging since optimal investment depends non-linearly on the unobserved investment cost shock ν_{jt}. Our model implies that $I(s_{jt}, \nu_{jt})$ is decreasing in ν_{jt}, since a higher ν_{jt} implies the marginal cost of investing is higher. We exploit this monotonicity property to estimate the investment policy function, following Bajari et al. (2007). We first estimate the distribution of investment at each state, $F(I_{jt}|s_{jt})$. We then obtain an estimate $\hat{i}(s_{jt}, \nu_{jt})$ of the investment policy function by inverting the CDF $F(.)$ at each point s_{jt} and evaluating this at the (known) distribution $G(.)$ of ν. A more detailed description of this procedure can be found in Appendix B.1.

4.3.2 State transitions

To compute the value function for incumbents and entrants, we need shipyards' expectations over s_{jt}. Among the state variables in s_{jt}, province, ownership, and size are fixed over time. The transition processes for age, capital and backlog are known, as we have already discussed in Section 3. As such we are left to specify firm beliefs over the transition process for: (i) dummy variables

[18] We include a negative sign so that $\bar{s}_{jt}$ can be interpreted as a measure of a firm's (observed) productivity.

for the subsidy policies (ii) ship and steel prices.[19]

We assume that shipyards are myopic regarding the industrial policies: whenever they are in a particular policy regime, they believe that the policy regime is permanent. In other words, before the policies are announced, firms believe they will not receive subsidies at any time in the future; once the policies are announced, firms believe the subsidies they are receiving will continue in any future periods. The assumption of myopia regarding government policy changes is standard in the literature (see, for example, Ryan (2012) and Kalouptsidi (2018)) and is made here partly out of convenience, as the task of modelling unobserved firm beliefs about government policies is a difficult one.

Equilibrium ship prices are determined each period by the intersection of the ship demand and supply curves. Thus the true transition process for ship prices is complex, as it requires shipyards to predict the evolution of several demand shifters as well as the production, investment, entry and exit decisions of all the other yards. Similar to Barwick and Pathak (2015) and Kalouptsidi (2018), we model shipyard beliefs about ship and steel prices as AR(1) processes with time trends. The introduction of Chinese government policies represents a permanent, unanticipated shock to the industry which can potentially change the evolution of firm beliefs over prices. To capture this, we allow all the AR(1) coefficients to change when the policies came into effect in 2006.

4.3.3 Estimating the structural parameters

Armed with first-stage estimates of the policy functions and firm beliefs about state transitions, we can now estimate the structural parameters using the following maximum likelihood procedure. At each guess of θ, we go through the following two steps.

The first step involves solving for $V(s_{jt};\theta) = \sum_{l=1}^{L} \gamma_l^0 u_l(s_{jt})$ by searching for parameters γ_l^o that ensure that the Bellman equation holds approximately in all the states. We find the vector $\gamma^o = \{\gamma_1^o, ..., \gamma_L^o\}$ that minimizes the least squares objective function $\sum_{j,t} \xi_{jt}^2$, where

$$\xi_{jt} = V(s_{jt};\theta) - \pi(s_{jt}) - \hat{p}^x(s_{jt})\sigma - CV(s_{jt};\theta)$$

and

$$CV(s_{jt};\theta) = E_{\nu_{jt}}\Big(- c_1\hat{i}(s_{jt},\nu_{jt}) - c_2\hat{i}(s_{jt},\nu_{jt})^2 - c_3\nu_{jt}\hat{i}(s_{jt},\nu_{jt}) - c_4 T_t\hat{i}(s_{jt},\nu_{jt}) + \beta E\left[V(s_{jt+1})|s_{jt},\hat{i}(s_{jt},\nu_{jt});\theta\right]\Big)$$

[19] An implicit assumption is that the dynamic model is stationary conditioning on state variables, including the policy shocks.

where $\hat{p}^x(s_{jt})$ and $\hat{i}(s_{jt}, \nu_{jt})$ are the estimated first-stage exit and investment policy functions, respectively. As discussed earlier, we use these estimated policy functions when solving the Bellman equation so as to avoid solving for new policy functions for every guess of θ.

Appendix B.3 describes in detail how we solve for the value function. In particular, Section B.3.1 describes how we select the set of basis functions with which we approximate the value function. Section B.3.2 discusses how we choose the sample of states on which we carry out the least squares regression to find the approximating parameters γ^o. Finally Section B.3.3 describes how to construct the continuation value given a set of basis functions and a parameter guess θ.

The second step involves computing the MLE objective function, which includes both the likelihood for the exit decisions and the likelihood for the investment decisions. The log-likelihood for exit over the full sample is:

$$L_1 = \sum_{j,t} log(f(\chi_{jt}; \theta))$$

where $log(f(\chi_{jt}; \theta))$ is defined as:

$$log(f(\chi_{jt}; \theta)) = -\chi_{jt} \frac{CV(s_{jt}; \theta)}{\sigma} + (1 - \chi_{jt}) log(1 - exp(-CV(s_{jt}; \theta)/\sigma))$$

Computing the likelihood $g(i_{jt}; \theta)$ over investment decisions requires the optimal investment rule implied by the investment first-order condition. Suppose that the optimal rule for investment is $i^* = i(\nu_{jt})$ (to keep the notation simple we drop the states). As long as $i(\nu_{jt})$ is strictly monotonic and differentiable in ν_{jt}, we can write the density function of i_{jt} as follows:

$$g(i_{jt}; \theta) = \frac{f_\nu(\nu_{jt})}{|i'(\nu_{jt})|} \tag{11}$$

where $f_\nu(\nu_{jt})$ is the density function of the cost shock ν_{jt} and $|i'(\nu_{jt})|$ is the absolute value of the derivative of investment with respect to ν_{jt}.

Appendix B.4 describes how we can construct both the numerator and denominator of (11). The necessary conditions are that the value function be twice differentiable in capital, which we already have from the fact that we are approximating the value function using smooth polynomial/spline basis functions, and that investment be strictly monotonic in the investment cost shock.

The log-likelihood over the full sample is then:

$$L_2 = \sum_{j,t} log(g(i_{jt}; \theta))$$

Since ϕ_{jt} and ν_{jt} are independent of each other, the joint log-likelihood for exit and investment

decisions is the sum of the two log-likelihoods:

$$L = L_1 + L_2 = \sum_{j,t} log(f(\chi_{jt};\theta)) + \sum_{j,t} log(g(i_{jt};\theta))$$

4.4 Solving for optimal policy functions

We use a fixed point algorithm to solve for the policy functions $p^*(s)$ and the value function coefficients, γ^* that are consistent with the estimated investment cost and scrap value parameters $\hat{\theta}$. At each iteration j, starting with a guess of policy functions $p^j(s)$, we solve the Bellman equation to update the value function coefficients to γ^{j+1}. We then use the first-order condition for optimal investment and exit to update the policy function to $p^{j+1}(s)$. We keep iterating until the policy functions converge ($||p^{j+1}(s) - p^j(s)|| < \varepsilon$), and verify that the value function coefficients also converge. Appendix C.1 contains further details on the algorithm.

We use these optimal policy functions $p^*(s)$ and the associated value function coefficients γ^* both for entry estimation (Section 4.5) and for counter-factual analysis. In principle it is also possible to use the first-stage approximate policy functions $\hat{p}(s)$ when constructing the value function for entry estimation, since $\hat{p}(s)$ is a consistent estimate of $p^*(s)$. However using $p^*(s)$ for entry estimation has the advantage of minimizing the extent to which noise in the estimated first-stage policy function affect the entry cost estimates.

4.5 Entry Costs

Estimating entry costs is straightforward once we have estimated the investment cost and scrap value parameters. Recall that equation (6) characterizes the entry decision. A potential entrant chooses to enter if their value of entry, which is the LHS of the aforementioned equation, exceeds their random draw of the entry cost.

We first construct the value of entry, VE_{jt}, for each potential entrant, using (6). We then estimate the mean of κ_{jt} using MLE, using the observed entry decisions. Since the mean of κ_{jt} is allowed to differ both by region and by policy period, we carry out the MLE estimation separately for each region and each policy period. Each draw of κ_{jt} is independent, so there are no gains from joint estimation.

5 Results

5.1 Demand estimates

Table 3 reports estimates of the demand curve in Equation (7). In all specifications, we allow for a different price coefficient before and after 2006, to capture changes in the ship demand curve after the introduction of Chinese subsidies. We include the logarithm of the type-specific aggregate backlog to capture dynamics in demand, and include the type-specific freight rate and other demand shifters to control for shifts in the demand curve. The freight rate measures we use are the Balitc Exchange Freight Index for bulk shipping, Baltic Exchange Clean Tanker Index for tankers and Containership Timecharter Rate Index for containerships. The demand shifters we include are the US wheat price and total Chinese iron ore imports for bulk demand, Middle East refinery production for tanker demand, and world car trade for containership demand.

We impose the following coefficient restrictions. Firstly, we restrict the coefficient on the logarithm of the aggregate backlog to be the same across types. This means that the elasticity of bulk demand with respect to bulk backlog is identical to the elasticity of tanker demand with respect to tanker backlog, and the elasticity of container demand with respect to container backlog. Secondly, we restrict the Price*Post-2006 coefficient to be the same across types.

We instrument for prices using steel prices and steel production, and estimate the demand system using 1-step GMM. The weighting matrix is $(Z'Z)^{-1}$, where Z collects all instruments. In the absence of any coefficient restrictions, this would be equivalent to estimating the demand curve separately for each type using two-stage least squares (2SLS). Because we restrict the coefficient on both the logarithm of the backlog and the change in the price coefficient post-2006 to be the same for all three types, 1-step GMM estimates are different from 2SLS estimates since they utilize the additional information provided by the coefficient restrictions.

Column (1) presents the simplest specification where the only demand shifter we include is the type-specific freight rate. In Column (2), we add in the type-specific demand shifters. Column (3) adds a time trend to the demand curve, and Column (4) allows the time trend to differ before and after 2006 (the purpose being again to control for possible effects of the policy on demand evolution). In Columns (3) and (4), the coefficients on the time trend are constrained to be the same across all types.

Adding the demand shifters improves the fit for the demand equations, as we can see by comparing columns (1) and (2). Adding the time trend (column (3)) and allowing the time trend to be different before and after 2006 (column(4)) makes little difference to the fit. As such, and because the time trends are challenging to deal with when extrapolating beyond our sample period when conducting the counter-factual analyses, we use specification (2) for counter-factual analyses.

The price coefficient is negative in all 4 specifications. We find that demand becomes sig-

nificantly less elastic in 2006. In specification (2), the elasticities pre-06 equal 1.8 for bulk and tanker and 3.4 for containers. The elasticities post-06 equal 0.3 for bulk, 0.6 for tanker, and 1.7 for containers.

The coefficient on the backlog is negative and statistically different from 0 only when we control for type-specific demand shifters. Omitting the demand shifters can bias the backlog coefficient upwards since backlog is positively correlated with demand shifters. In our preferred specification (2), a 1% increase in the backlog leads to a 1% decrease in quantity demanded.

5.2 Cost function estimates

Table 4 shows the estimated cost function parameters and bootstrapped standard errors. We allow the coefficients β_q, σ and the coefficients on capital, backlog and the steel price to be different across types, but restrict the coefficients on the subsidy dummies, age variable, ownership and region to be the same across types. The type-specific coefficients reflect the effect of scale/economies from holding a larger backlog, the return to capital and input intensity which are likely to be quite different across types. The common coefficients, on the other hand, reflect characteristics that are more likely to affect all types equally. For example, in the absence of evidence that subsidies are ear-marked for particular ship types, we would expect production subsidies to proportionally reduce the cost of producing all ship types rather than differentially affect the cost of producing different types of ships.

Because the Chinese policies came into effect in 2006 and underwent major changes in 2009, we include separate dummies for Chinese yards in the 2006-2008 and the 2009+ periods in order to capture any changes in production subsidies from 2009 onwards. The subsidy between 2006-2008 is about 2,790 RMB/CGT which is between 19-24% of the average price. The subsidy in the 2009+ period is estimated to be considerably smaller, at 900 RMB/CGT.

β_q is the increase in MC (in 1000 RMB/CGT) from taking an additional order of 100,000 CGT. The larger β_q is, the more convex the cost function is, and the less responsive supply is to changes in prices. On average, a 10% price increase causes bulk/ tanker/container production to increase by 17%/34%/42% respectively. $\beta_q Q$ measures the convex portion of MC: at the average order, $\beta_q Q$ accounts for 25-60% of the firm's marginal cost.

The coefficient on capital is negative and significant; the coefficient on K^2 is positive (suggesting diminishing returns) but not statistically significant and small in absolute terms. Over the range of capital we observe, the effect of increasing capital is always to reduce the marginal cost. Increasing K by 100 million RMB (mean K =400 million RMB) reduces marginal costs by 2.1% for bulk, 1.8% for tanker and 1.7% for container.

The coefficient on own-backlog b_t is negative. This is consistent with scale economies where

yards producing several ships at once are able to produce each of them more cheaply than a yard forced to work on one ship at a time. It is also possible that backlog could be a proxy for productivity: more productive yards build larger backlogs and take more orders. However the coefficient on b_t^2 is positive and in contrast to what we saw with capital, both statistically significant and economically meaningful. When the backlog is sufficiently large, increasing the backlog leads to increased marginal costs. This reflects the fact that beyond a certain point shipyards become capacity constrained and can no longer increase their backlog without incurring higher costs. Finally, the total backlog of other types is positive, and is statistically significant for containers at the 5% level; however the effect of this is small compared to the effect of the own-type backlog.

Large firms have significantly lower marginal costs: the point estimate is -4.17, which means their marginal costs are lower by 25-32%. Jiangsu, Zhejiang and Liaoning (three of the major shipbuilding regions) have lower marginal costs, by about 11-14% for Zhejiang, 13-16% for Liaoning and 20-25% for Jiangsu. We have less strong evidence that the ownership structure matters. While we find that private firms have the highest marginal costs, followed by small SOEs, CSSC/CSIC owned SOEs and finally foreign JVs, none of these coefficients are statistically significant.

Increases in the steel price increase the marginal cost for all types, as expected (since steel is an input in shipbuilding). The coefficient is significant for bulk and containers. The effect of age is also positive though the coefficient is economically small in magnitude, amounting to a 1% increase in costs every year.

Finally, the fixed cost of production c_0 is estimated to be 0.25/0.27/0.54 mill RMB for container/tanker/bulk respectively. The magnitude of the fixed cost is negligible given that it accounts for less than 1% of profits, and we find that the cost function estimates are very similar if we assume firms face no fixed costs. By contrast, the cost of idleness that we calibrate equals 15 mil RMB per quarter, which at the estimated cost function parameters accounts for 15% of total industry profits per year.

Table 5 illustrates how sensitive the key coefficients of interest are to different specifications. Column (1) shows the same estimates as Table 4. The estimates are almost identical if we add in a time trend (Column (2)). In Column (3), we show the cost function estimates when our sample does not include any Chinese yards that entered after 2004. We find that incumbent yards appear to enjoy somewhat larger returns to capital. The estimated subsidy for the 2006-08 period decreases to 2,080 RMB/CGT but remains both economically meaningful and statistically significant. Column (4) allows for learning-by-doing by allowing the marginal cost to depend directly on a firm's cumulative production. We find the coefficients on cumulative past production to be positive, which is inconsistent with there being significant learning-by-doing. The estimated subsidy during the 2006-2008 period is about 2,430 RMB/CGT in this specification. Finally in Column (5), we have allowed a firm's marginal cost to depend on industry cumulative output, as a crude test of possible

aggregate learning-by-doing. Without instrumenting for industry cumulative output, this exercise is likely to over-estimate spillover effects: if there are common unobserved shocks that raise the output of all firms, it will look as though there are positive spillover effects when there aren't any. Even so, we find very limited evidence for spillover effects. Marginal costs decrease with cumulative industry production only for bulk (though the effect is not significant), and increase with cumulative industry production for tanker and container.

Because the time trend does not appear to be important and there is little evidence of learning effects, we use the baseline specification in Column (1) for dynamic estimation and counter-factual analysis.

5.3 Policy Functions and State Transitions

This section begins by describing how we construct the basis functions in practice. We then discuss how we estimate exit and investment policy functions, and state transition processes.

5.3.1 Construction of basis functions

We begin by constructing the state $\bar{s}_{jt}$ that captures how the backlog, size, region, ownership and age affect the firm's profits, where the weights on each term are the estimated cost function coefficients (see Section 4.3 for details). $\bar{s}_{jt}$ can be thought of as a measure of a firm's productivity: the higher $\bar{s}_{jt}$ is, the lower the firm's marginal costs, and hence the more profitable the firm is.

The state variables we work with in the dynamic estimation are the firm's capital stock, ship and steel prices, $\bar{s}_{jt}$ and two policy dummies for the periods 2006-08 and 2009+ respectively. We construct basis functions from them as follows. First, we include each linear term. Second, to construct the higher-order terms, we use B-splines of order 3 (which are piecewise polynomials of order 2) in capital with knots at 0 and 1 billion RMB respectively. Our estimates are robust to inclusion of a larger number of knots in capital. For $\bar{s}_{jt}$ and ship and steel prices, we include higher-order polynomial terms instead of using B-splines; we have found that including B-splines for these variables makes little difference to the estimates and as such adopt the most parsimonious specification. Finally we include every interaction term between states in the set of basis functions we use to approximate the value function. The total number of basis functions we use is 44.

5.3.2 Exit policy function

We estimate the exit policy function by a probit regression. We can reliably identify exits from 1998 through to 2011, but not in 2012 and 2013. As such we only use observations from 1998 to 2011 when estimating the exit policy function, and later when constructing the likelihood over exit decisions. We observe a total of 47 exits during this period.

Ideally, we would like to include all the basis functions with which we approximate V in the probit regression. However the limited number of exits we observe calls for a more parsimonious approach. In Table 6, we include two sets of estimates. In column (1), we include linear terms in all of the states as well as a squared term in capital. In column (2), we include region fixed effects.

Firms with larger capital are more likely to exit, though this effect is statistically insignificant and the sign is reversed for firms with capital exceeding 1 billion RMB. (About 10% of observations in the sample are for firms with capital exceeding 1 billion RMB). We find that the higher $\bar{s}_{jt}$ is, the lower the probability of exit which is intuitive as $\bar{s}_{jt}$ is a measure of how productive firms are; however the coefficient is again statistically insignificant. Exit probabilities are lower during the periods when subsidies are in place (between 2006-2008 and then from 2009 onwards).

Finally, the effect of ship prices appears strange: with exit probabilities increasing in some prices and decreasing in others. This is likely due to the fact that ship prices are highly correlated with one another: if all ship prices rise by 1000 RMB/CGT, the net effect on exit is close to zero. Exit probabilities are decreasing in the steel price, though again the effect is statistically insignificant.

In column (2), when we add in region fixed effects, we find that exit probabilities are somewhat higher in Zhejiang, Jiangsu and Liaoning than in other provinces in China. The coefficients on the remaining state variables retain the same signs though the magnitudes are different for capital and the two subsidies.

Column (2) fits the data somewhat better than Column (1), as we can see from looking at the Pseudo-R^2 or the log-likelihood. Despite this, we use fitted exit probabilities from the estimates in Column (1) for the dynamic estimation. This is because Column (1) is consistent with the way we do the dynamic estimation, with the effect of regions on the firm's value function captured by the state variable $\bar{s}_{jt}$. We carry out robustness checks to verify that the dynamic model estimates are not sensitive to the particular choice of exit policy function.

The low values of the Pseudo-R^2 suggests that it is generally difficult to predict when firms will exit, and that firms often exit for idiosyncratic reasons that are not picked up by their observed state. We find little evidence of systematic behavior suggesting that exits are planned several periods in advance: for example, we do not see firms stop taking new orders and dispose of their capital stock in the periods prior to their exit. In our model, we allow for firms to randomly draw a scrap value which helps account for the large variation in exit.

The fitted exit probabilities from Column (1) of Table 6 are shown in Figure 5.

5.3.3 Investment policy function

We assumed that investment is additively separable in observed states s: $i(s,\nu) = h_1(s) + h_2(\nu)$. Estimating the investment policy function thus involves first regressing investment on s, which

gives us estimates of $h_1(s)$, and then inverting the CDF of the residuals, which gives us $h_2(\nu)$.

Just as with exit, we have to choose which basis functions of the state variables to include in the regression equation. Recall that investment is characterized by the FOC given by equation 14. The FOC shows that investment depends on the subsidies T_t and $\frac{\partial E[V(s_{jt+1}|s_{jt},i_{jt}]}{\partial i}$; as such we regress investment on T_t as well as any basis functions that enter into $\frac{\partial E[V(s_{jt+1}|s_{jt},i_{jt}]}{\partial i}$.

Our sample of investment decisions excludes the years 2009, 2010 and 2013 during which we do not observe firm-level investment. We also drop 221 outliers which are observations with very large investments (exceeding 0.25 billion RMB) or very large capital stock stocks (exceeding 4 billion RMB). The resulting sample has 4286 observations. Finally, it is important to note that as we do not model firm divestment decisions, we treat any divestments as the firm deciding not to invest, so that $i = 0$.

In Table 7 we present estimates of $h_1(s)$: that is, how investment depends on the observed states. There are several key findings. First, the coefficients on both the B-spline terms in capital are negative, but this is not straightforward to directly interpret. Figure 6 shows the relationship between capital and the fitted value of investment implied by these estimates. We can see that for firms with small capital stocks, investment increases in the capital stock, but for firms with large capital stocks, investment decreases in the capital stock.

Second, both the 2006-08 and 2009+ policy dummies are positive: investment goes up after the policies were introduced, after controlling for other states. The 2009+ dummy however is not statistically different from zero. Third, intuitively, firms with higher $\bar{s}_{jt}$ (which are therefore more productive) invest more.

Fourth, the bulk price is positive and statistically significant, while the tanker and container prices are insignificant. The estimates indicate that if all the ship prices increase, investment increases. By contrast if the steel price (which is the input price) increases, investment decreases.

The R^2 from the investment is 0.24, indicating that a significant amount of the variation in investment is explained by idiosyncratic factors that in our model are captured by the investment cost shock ν.

5.3.4 State transitions

Table 8 presents estimates of the transition processes for bulk, tanker, container and steel prices. These assume that each price follows an AR(1) process with a time trend, with all coefficients changing when the policy came into effect in 2006. Allowing for a break-point is important because the subsidy policy represents a permanent unanticipated shock to the industry which is likely to cause industry participants to permanently change their beliefs. We find that the own-price coefficients are between 0.86-0.92 for bulk/tanker/container. Including the trend terms is important: without the trend, the series behave like unit roots. Basically the two trend terms pick up the fact

that ship prices are mostly trending up before 2006, and mostly trending down after 2006.

We have also tried estimating the state transitions by VAR, or with an AR(1) without any trend terms/breakpoints. We tend to have issues with potential unit roots in those specifications, with own-price coefficients that are very close to 1 (and sometimes exceeding 1).

5.4 Estimates of Dynamic Model

5.4.1 Investment and exit

In implementing the estimation procedure described in Section 4.3, we abstract away from divestment decisions and assume that $c_1 = 1$, since I is directly a measure of spending on investment, following Robert et al 2004.

Table 9 reports the parameter estimates. Subsidies existed in both 2006-2008 and from 2009 onwards. Between 2006-2008, the subsidy is 0.40 RMB/RMB of investment, meaning that 40% of the per-unit cost of investment (excluding adjustment costs) is subsidized. The subsidy increases to 0.81 RMB/RMB of investment in 2009. The increase in the investment subsidy in 2009 is consistent with the government policy change in that period that shifted the focus towards consolidating the industry and supporting large firms. As we shall see later, the increase in the investment subsidy coincided with a reduction in entry subsidies.

In addition, c_3 is high implying that the ν shock plays a large role in explaining investment. The R^2 (i.e. how much investment is predicted by the observed states alone) is around 0.16. The coefficient on quadratic investment, c_2, which determines adjustment costs, equals 24. This implies adjustment costs account on average for 28% of investment costs, and exceed 50% of the cost of investment for large investments exceeding 50 million RMB. Finally, the mean of the scrap value distribution is estimated to be 0.69 billion RMB. This is significantly lower than the average value of a firm in our sample, which is consistent with the fact that exit is a rare event and occurs in only 1% of observations.

We used our estimates and a randomly drawn set of investment cost shocks ν to simulate investment. Figure 7 plots the distribution of observed as well as simulated investment. Our estimated distribution of investment is reasonably similar to the actual distribution, but the actual distribution is more heavy-tailed than our predicted distribution, with more large investments, fewer medium-sized investments and more investments that are very close to or equal to 0.

We also simulated the entire industry from 2000 onwards to look at how well we match exit rates. Table 10 compares the actual number of exits with the number of exits when we simulate the industry. There is somewhat less exit in the simulation (38 in total) than in the actual data (48 in total).

5.4.2 Entry cost estimates

Just as with production cost and investment cost/scrap value estimation, we allow for three different policy periods that affect entry: a "Pre" period which covers the years before any policies were implemented, followed by one policy regime until 2008, and a separate policy regime beginning from 2009 onwards. The actual policy implemented involved encouraging entry from 2006 onwards in most regions, followed by a moratorium on entry imposed in 2009. From 2009 onwards, no new yards were approved and the only entries we see were approved before 2009.[20]

The province of Zhejiang began encouraging entry earlier than other provinces. As such we allow for entry subsidies to begin in the year 2004 for Zhejiang, so that the three policy periods for Zhejiang are 1998-2003, 2004-2008 and 2009+. For all other provinces, entry subsidies are assumed to begin in the year 2006.

One way to allow for different capital size among entrants is to assume that entrants draw a capital size from the observed distribution. We assume that the cost of the initial investment in capital, K_0, equals $c(K_0) = c_1 K_0 + c_4 T_t K_0$. Unlike the cost of investment, the initial cost of capital is assumed to be deterministic: since the initial capital stock is assumed to be a random draw in our model, there is nothing to be gained by including a stochastic component, since it will average out to 0. We also assume that firms face no adjustment costs when making their initial investment in capital. The reason for this is that the initial choice of K_0 is typically much larger than typical choices of investment that we observe. If we include adjustment costs in the cost of initial capital and use the estimated parameter c_2 to estimate these adjusment costs, the total cost of K_0 becomes implausibly large, and the value of entry becomes negative in many periods solely because of the large adjustment costs. It is unlikely that yards actually incur such large adjustment costs because the initial capital stock K_0 was probably built over many periods, unlike the rest of the investments that we see. We also do not observe how many periods it takes for firms to build their initial capital stock making it difficult to know how costly the adjustment was.

The number of potential entrants is rarely observed in empirical analysis. In our sample, the highest number of entrants ever observed in a region in a quarter is 17, so we assume that $\bar{N} = 20$, meaning there are 20 potential entrants in each period in each quarter. The estimated average entry costs for observed entrants are robust to this assumption.

Table 11 has the entry cost estimates. The estimated mean entry costs are quite high in magnitude, ranging from 54 bn RMB to 630 bn RMB in the "Pre" period, and are several times higher than the value of entry. This is because with the number of potential entrants equal to 20 in each region each quarter, the rate of entry is only about 2.5%. Only firms with favorable draws of the entry cost actually end up entering the industry. Conditional on entering, the mean entry cost paid

[20] The specific text of the policy is: "No new shipyard building project shall be approved for the next three years, except the ones mentioned in the Medium and Long Term Development Plan of Shipbuilding Industry."

across the sample is 2.3 bn RMB, which is close to accounting estimates of the value of firms in the shipbuilding industry.

In all regions, there is evidence of significant entry subsidies in the first policy period from 2006-2008 (2004-08 for Zhejiang). In Jiangsu, Zhejiang and Other provinces, close to 60% of entry costs were subsidized, while in Liaoning province about 44% of entry costs were subsidized.

Entry costs increased substantially in 2009 when the entry moratorium was put in place and entry was severely discouraged: entry costs were higher than pre-policy levels in all regions by 19% in Jiangsu, 54% in Other regions and 111% in Zhejiang. For Liaoning, we were unable to estimate the entry cost post-2009 as we see zero entries in that province from 2009 onwards.

The entry cost estimates are sensitive to the assumed number of potential entrants, $\bar{N}$, which of course cannot be observed. The higher $\bar{N}$ is, the higher the estimated mean of the entry cost estimate. However what is crucial for estimating the magnitude of entry subsidies actually paid out is the *conditional mean* of entry costs, conditioning on the event that entry occurs. The conditional mean is comparatively robust to $\bar{N}$. As a robustness exercise we repeat the entry cost estimation with different assumptions on $\bar{N}$ and show that the estimated entry subsidies are similar.

To get an idea of how well the estimated entry costs can rationalize the observed entry patterns, we simulate the model from the year 2000 with randomly drawn entry costs for each potential entrant. Table 12 shows that the simulated number of entries is close to the actual number of entries in each of the policy periods.

6 Evaluation of China's Industrial Policies

In this section, we carry out counter-factual exercises to evaluate the effects of China's industrial policies. We first begin by evaluating the various subsidy policies. We then study consolidation policies implemented by the government towards the end of our sample period, in particular the White List policy.

We implement the counter-factuals as follows. We first solve for the value function coefficients and the associated policy functions using a policy function algorithm. Details of this algorithm can be found in Appendix C.1. We then begin from a particular starting date (which is 2004 in our counter-factuals), and simulate industry outcomes starting from that date. In each period, Chinese firms make optimal production, investment, exit and entry decision, taking both prices and government policies as given. Korean and Japanese firms make optimal production decisions taking prices as given: note that we do not model investment, entry and exit decisions by Korean and Japanese firms, because there is limited capacity expansion by these firms and we lack their micro-level data on capital. Equilibrium prices are determined each period by the intersection of the industry demand curve and supply curves. For each counter-factual, we carry out 30 simulations

and take the average of these simulations.[21] More details of the implementation of the counterfactuals can be found in Appendix C.

6.1 Evaluation of Subsidy Policies

We first quantify the effect of subsidies on the industry. We do so by simulating the industry from 2004 onwards under two scenarios. In the first scenario, we simulate the industry keeping all subsidies in place; this closely matches the acual evolution of the industry. In the second scenario, we simulate the industry shutting down all the subsidies. We can quantify the impact of the subsidies by comparing these two scenarios.

6.1.1 Effect of subsidies on evolution of industry

Table shows that the total magnitude of subsidies handed out to Chinese shipbuilders was close to 600 billion RMB (or 100 billion USD) between 2006 and 2013. The subsidies were large in relation to the size of the industry: for example, China's total revenue from shipbuilding was around 1700 billion RMB in this period. Entry subsidies accounted for more than half of the subsidies handed out.

We find that the subsidies had a large effect on production, investment and entry. Subsidies increased China's market share between 2006-13 by nearly 40%, from 30% to 42%. We find that China's production (measured in the volume of its orders in CGT) experienced a large increase as a result of the subsidies. In the absence of subsidies, China's production still increases in 2006 as a result of the increase in shipbuilding demand in that boom period, but the increase is much less pronounced.

In absolute terms, only about one-fourth of China's increased production translated into higher industry output; the remaining three-fourths constituted business-stealing whereby Chinese production expanded at the expense of competing firms in Japan and Korea. We find that Chinese subsidies caused Korea's market share during the 2006-13 period to decrease from 47% to 38% and Japan's market share from 23% to 20%.

Table 14 shows that the subsidies led to substantial reductions in global ship prices: prices of bulk ships, oil tankers, and containers are reduced by 8.3%, 6.2%, and 3.3% from 2006 to 2008, respectively. The price effect is largest for bulk because Chinese shipyards account for a bigger world market share of bulk than the other two types and that bulk demand appears the most inelastic. The price drops are more pronounced from 2009 to 2013 and reached 16.2% for bulk, 10.1% for tankers, and 3.9% for containers. As the effect of past subsidies accumulates through an increased production capacity for existing firms and a larger number of firms, the long-run effect

[21] We have tried increasing the number of simulations to 50 and found that it makes very little difference to the results.

of subsidies is bigger than the short-run effect. Because of the reduction in ship prices, shippers benefit by 230 bill RMB from Chinese subsidies, though only a small proportion of these gains accrue to Chinese shippers.[22]

Figure 8 compares the number of firms in the Chinese shipbuilding industry with and without subsidies by year. The number of entries with subsidy is more than twice as high as the number of entries otherwise, 145 firms vs. 68 firms entered from 2006 to 2013. The number of exits is the same in the two scenarios (around 40), but the composition of the firms that exit is different. In the scenario with subsidies, incumbent firms exit less often since they receive production and investment subsidies that boost their profit. On the other hand, many of the entrants under the subsidy are small and relatively inefficient and are more likely to exit once ship prices slump after 2009.

Figure 9 illustrates that investment skyrocketed post 2006 and nearly tripped as a result of subsidies. Total investment during 2006-2013 is 133 bill RMB with subsidies, compared to 45 bill RMB without subsidies.

As a result of the flood of entry, the industry became much more fragmented post the policy period. The Herfindahl-Hirschman index (HHI) plummeted from 1,500 in 2004 to less than 500 in 2013 (Figure 10), while the 4-firm concentration ratio is also significantly lower. All three subsidies (entry, production, and investment subsidies) induce entry, which tends to be small in size. In addition, these subsidies (especially production and investment subsidies) boost firms' variable profit and reduce the exit probability of small firms.[23] The subsidies handed out between 2006 and 2013 led to an estimated 159 billion RMB increase in discounted lifetime profits for Chinese shipbuilders.

Despite a sizeable increase in China's production, overall capacity utilization was significantly lower as a consequence of massive investments induced by subsidies, particularly in the 2009-2013 when demand was low. For example, during the 2009-2013 recession, the ratio of production to capital (which proxies for capacity utilization) would have been 20% higher if China had not subsidized the shipbuilding industry.

6.1.2 Welfare Effect of Subsidies

Subsidies led to a substantial increase in the number of firms, capacity and production in China's shipbuilding industry, but resulted in lower prices and a more fragmented industry structure. How do the benefits of these policies in terms of raising social welfare (firm profit and consumer surplus) compare to the cost of the subsidies?

[22] According to Clarkson World Shipyard Monitor, China's order of ships accounted for under 10% of world total orders from 2010 to 2013. China's market share of total fleet is considerably smaller than 10% of world fleet.

[23] Large firms also experience a reduction in the exit probability, but the effect is larger for small firms.

Since there is no market failure that these subsidies address, we know from economic theory that the subsidies are necessarily welfare-reducing. What is less well-understood is how distortionary subsidies are, and the magnitude of the welfare losses. Our goal in this section is to quantify the welfare losses from subsidies. An alternative view of the subsidies is that they were put in place to achieve national objectives (such as the prestige from becoming the no. 1 shipbuilder in the world and national security) whose benefits cannot be quantified in economic terms. Under this alternative view, our analysis can be interpreted as illustrating the economic cost of achieving such objectives through subsidies. Finally, in addition to quantifying the welfare effects of subsidies, we also shed light on the effectiveness of different types of instruments, in particular the relative efficacy of production, investment, and entry subsidies.

The benefits of the subsidies accrue beyond 2013, since the accumulated capital remains productive and new firms can continue operation for a long period of time. To assess the welfare effect of subsidies, we simulate the industry from 2006 to 2099.[24] We then look at the discounted stream of industry revenues and profits and compare them to the discounted stream of subsidies the government hands out. Note that we ignore consumer surplus in this calculation because domestic consumer surplus accounts for a small fraction of the world consumer surplus and is modest compared to industry profit.

In order to compare different types of subsidy policies, we carry out five different counterfactual analysis: all subsidies, only production subsidy, only investment subsidy, only entry subsidy, and no subsidy. When simulating the industry beyond 2013, we assume that the policy environment that is in place in 2013 is continued forever. For example, in the scenario with all subsidies, entry subsidies run from 2006 to 2008 whereas production and investment subsidies run from 2006 to 2099.[25]

The welfare effects of the subsidies are summarized in Table 15, which reports the discounted sums of industry revenue and profit, scrap value, investment/entry costs, as well as the magnitude of different subsidies from 2006 to 2099, averaged across 30 simulations. We use a discount factor of 0.98 per quarter, or 0.92 annually, taking into account the high interest rates (averaged 6% from 1996 to 2018) in China.[26]

The last 3 rows, "ΔRevenue/Subsidy", "ΔProfit/Subsidy", "Δ(Profit-Inv. Cost)/Subsidy", and "ΔNet Profit/Subsidy", report different measures of the effectiveness of the subsidies. "ΔRevenue" is the revenue difference between having some or all subsidies and no subsidy. Many official documents that we have uncovered quoted explicit production targets for the domestic shipbuilding industry. Here we report the ratio between revenue increase and subsidy cost to evaluate the ef-

[24] At our discount factor, the present value of 1$ in 2099 is 0.0004$ in 2006.

[25] We have carried out a large number of simulations, including varying length of different subsidies. Results differ in magnitude but are qualitatively similar.

[26] Source https://tradingeconomics.com/china/interest-rate,

fectiveness of subsidies at promoting industry revenue. "ΔProfit" is the change in variable profit (revenue minus production cost). "Δ(Profit-Inv. Cost)" is the difference in variable profit minus investment costs. Finally, "ΔNet Profit" is the difference in net profit, which is total revenue minus all costs including production costs, investment cost, entry and exit costs. We use "ΔNet Profit/Subsidy" to measure the net rate of return on different subsidies. A rate lower than 100% indicates that the cost of the subsidies exceeds the net benefits to the shipbuilding industry.

When all subsidies are in place, their rate of return is only 21%, suggesting the subsidies are highly wasteful and distortionary. This is driven by the fact that at the margin, subsidies lead to inefficient entry, production, and investment whose benefit in terms of lifetime discounted revenue does not warrant the associated cost.

While empirical evaluation of the rate of return of public policies are few, this number appears low relative to the estimates in the existing literature using data from developed countries.[27] As discussed above, subsidies lead to higher fragmentation, partly driven by small and unproductive firms that are more likely to be idle. Firms incur idleness costs to stay in business when they receive no orders from buyers. If idleness costs are zero, then the rate of return on subsidies increase by nearly a third from 21% to 28%.

There is a strong evidence that the distortions induced by subsidies are convex. When the subsidies are combined with one another, there is an increased take-up of the subsidies and the subsidies are more distortionary in per-dollar terms. For example, when implemented as stand-alone policies, production subsidies would have been 219bn RMB, instead of 260bn RMB when investment and entry subsidies are also in place. The return to production, investment, and entry subsidy as a standalone policy is 51%, 82% and 13%, respectively. If we add the net profit increase from each of these three scenarios and divide by the total cost of subsidies, the average rate of return is 42%, twice as high as the net return of 21% when all three subsidies are jointly implemented. Distortions are convex because when different subsidies interact with one another, they lead to more inefficient firm decisions that are away from the margin. For example, entry subsidy attracts inefficient entry and increases the number of firms in the industry. With the introduction of production subsidy, the number of firms in operation increases even more due to a lift in firm revenue induced by the production profit. More inefficient firms entry, which further reduces these subsidies' rate of return.

Using industry revenue as the criterion, both production and investment subsidies appear to be effective. On average, one RMB increase in the production and investment subsidy increases the industry's revenue by 1.6 RMB and 1.1 RMB respectively. This could potentially justify the popularity of these subsidies, especially since quantity and revenue targets (as well as GDP targets) are often linked to local officials' promotion in China (add a citation).

[27] Cite my paper on Walmart and a few other studies on city subsidies that try to attract firm entry. Jerry Xing also reports estimates for the federal subsidy to electric cars.

Entry subsidy is the most inefficient policy instrument among the three by a large margin. Both production and investment subsidies favor incumbents that are larger, receive more orders and more likely to invest, both of which lead to scale economies that drive down current and future production costs. In contrast, the entry subsidy predominantly attract entry of small and high-cost firms that are unable to exploit the scale economies that larger firms benefit from. As such the large number of additional entries contributes little to industry profits.

6.1.3 Comparison of Production and Investment Subsidies

The investment subsidy appears less distortionary than the production subsidy (Table 15), though this comparison is confounded by the fact that the magnitude of the production subsidy is significantly higher than that of the investment subsidy, and in general larger subsidies are associated with more distortion (convexity). Thus, to conduct a fair comparison of the two, we reduce the per-unit production subsidy by 50% (Table 16). Total production subsidy reduces to 97bn RMB, comparable to total investment subsidy of 93bn RMB. As expected, the return to subsidy in net profit terms is raised from 51% to 62%. The rate of return in terms of revenue increases further from 161% to 192%.

When we adjust for the amount of subsidy, the return to investment subsidy remains ahead leading to a higher increase in profits than production subsidy. On the other hand, production subsidy is nearly twice effective at increasing revenue: the increase in revenue per RMB of subsidy is 192% for production subsidies, versus only 106% for investment subsidies. Thus, depending on the policymaker's objectives, there is potentially a trade-off between production and investment subsidies. If the policymaker's goal is to maximize industry profits given a particular subsidy budget, investment subsidies are superior. However, if the policymaker's goal is to achieve a production/revenue target, then production subsidies might be preferred. Finally, for a government that cares about both industry revenue and profit, a mix of production and investment subsidies may be most effective.

What explains the differences in the performance of production and investment subsidies? It is straightforward to see why production subsidies are better at increasing output/revenue, since production subsidies directly target output, whereas investment subsidies only indirectly raises future output by increasing capital. Papers studying subsidies in other industries have found similar results. For example, Aldy, Gerarden and Sweeney (2018) find that wind farms claiming output subsidies produced 10-11% more power than wind farms claiming investment subsidies.

It is a priori not obvious why investment subsidies are less distortionary than production subsidies and lead to a higher return on net industry profits. In our context, it turns out that the wedge arises due to a differential dynamic impact on exits. When there are no production or investment subsidy, firms exit when the scrap value exceeds the continuation value of staying in the industry. Subsidies distort exit decisions by artificially increasing the continuation value of staying in busi-

ness. Production subsidy leads to fewer exits and retains a larger number of inefficient firms that create little industry profit. One evidence is that the industry average productivity is higher with investment subsidies than with production subsidies, as Figure 11 shows.[28]

There is thus a selection effect at work: production subsidy leads to fewer exits of inefficient firms, lowering the average productivity of the industry. At the same time, by pushing down ship prices with a larger industry output, these firms lower the profit earned by the rest of the industry as well. Together these factors make production subsidy less efficient than investment subsidy.

This comparison of production and investment subsidies suggests a couple of implications for policy-makers considering subsidies in other settings. First, production subsidies are likely to be more effective at achieving output or revenue targets. Second, it is important to consider the dynamic implications of subsidies when deciding between alternative policy instruments. In particular, subsidies that lead to a bigger distortion on the entry and exit margins can cause bigger welfare losses and lead to lower average productivity.

Our evaluation of these subsidies so far is mixed. They have benefited large firms and enhanced China's market share at the expense of South Korea and Japan. However much of the entry subsidy, which accounts for more than 40% of total subsidy, went toward small and unproductive firms. These subsidies contributed to the development of a fragmented industry and a large amount of under-utilized capital and to a large extent dissipated through inefficient entry/exit and production decisions.

6.2 Evaluation of Consolidation Policies

One explicit goal of China's industrial policies is to create large successful firms that could compete against international conglomerates. Chinese policymakers tried to promote greater consolidation and industry concentration through filtering out inefficient firms, especially in the aftermath of the economic crisis. The Plan on Adjusting and Revitalizing the Shipbuilding Industry, implemented in 2009, resulted in a moratorium on entry as well as increased investment subsidies to existing firms. The 12th Five-Year Plan for the Development of Shipbuilding Industry (2011-2015) set a number of targets for achieving increased industrial concentration, including 70% of the country's shipbuilding to be carried out by the top ten domestic firms, and at least five Chinese firms to be included in the world's top ten largest firms.

The most crucial policy for achieving these consolidation objectives is one titled "Shipbuilding Industry Standard and Conditions" that is released in 2013, where the government announces periodically a list of selected firms that meet the industry standard and receive priority in subsidies and bank financing.[29] This "White List" was first announced in 2014 and included 60 firms (out of

[28] Firm productivity is measured by $\bar{s}_{jt}$, the negative of the observed components of marginal costs.

[29] In practice, favorable financing terms and capital market access are often limited to firms on the white list.

which 56 firms can be matched to our sample) and has undergone a number of revisions since.

In this section, we assess the government's White List policies through the following questions. First, does targeting subsidies towards efficient firms improve the return on subsidies, and if so by how much? Second, did the government pick the right set of firms for the White List, or could they have done better?

6.2.1 Are there gains from targeting subsidies towards efficient firms?

In order to evaluate the effect of targeting towards efficient firms, we carry out the following counter-factual exercise. We rank firms in the industry in 2013 based on their expected profits π_{jt} in that year, and select the top 56 firms with the highest profitability into the White List. These firms receive production and investment subsidies beyond 2013 while firms that do not make it into the White List are allowed to keep operating but receive no subsidies post 2013. The results from this targeted policy is shown in the second column of Table 17, labeled "Subsidize White List firms After 2013". We compare this policy to the one that subsidizes all firms after 2013 (the first column) and the one that removes subsidies completely after 2013 (the last column).

There are considerable gains from directing subsidies toward the best set of firms. The net rate of return to targeted production and investment subsidies from 2014 onwards (when the subsidies are targeted) is 72%, whereas the return is only 45% if all firms are subsidized. This comparison doesn't change if we look at revenue instead of profits.

6.2.2 Did the government pick the right set of firms?

Our analysis above shows that subsidies are less distortionary when targeted towards the most efficient firms. However it is unclear a priori that the government would target the most efficient firms. Information asymmetry might prevent it from identifying firms that are more profitable in the future. In addition, regulatory capture might bias the process of picking winners to favor interest groups (Lane, 2017).

Not all of the firms selected into the White List thrived. By the end of 2016, seven of the white-list firms declared bankruptcy and 20 of them were idle for more than a year.[30] The ex-post performance of the White Listed firms in itself is not evidence that the government selected the wrong set of firms, since successful firms can fail in the future. To evaluate the government choice, we simulate the industry and compare the ex-ante outcomes (industry profit and revenue) averaged over different simulation runs.

[30] See https://www.hellenicshippingnews.com/chinese-shipyards-watching-the-white-list/ and https://www.maritime-executive.com/article/chinese-shipyard-white-list-to-shrink#gs.QcYLB=k/.

To carry out this evaluation, we need to make an assumption on how the White Listed firms are treated differently from the firms that do not make it to the White List. To make the comparison stark and avoid confounding effects from subsidies, we assume all subsidies are discontinued after 2013. In addition, to minimize impact on ship prices from other firms, we assume that only the firms on the White List are allowed to operate beyond 2013 while other firms are forced to exit from the industry.

The first scenario includes the actual White List while the second scenario (the simulated white list) includes firms with the highest expected profits in 2013 (that is, firms whose observable attributes are most favorable). Our selection rule is 'feasible' since it is based on observed firm attributes and is something the government could implement. Note, however, this simple criterion focuses on short-run profitability and is unlikely to be optimal in the long run. Our goal is to provide a relatively weak test: if the government picked the right set of firms with high profitability in the long-run, then their selected firms should do at least as well as the set of firms we pick.

Figure 12 shows that industry profits are significantly higher with the simulated White List than the actual White List. Industry profits and revenue (the discounted sum from 2014 to 2099) would be 10% and 8.5% higher, respectively, if the government had picked the set of firms with the highest short-run profitability instead of the set of firms on the White List. A simple logit regression on the probability of being selected into the White

Why did the government not get the selection of firms right? While a complete answer to this question requires more information on the government's policymaking process than we have, our analysis suggests that there was a bias towards picking state-owned firms (SOEs) in the government's selection rule. This is illustrated by Figure 13, which plots the share of industry profits earned by SOEs in each counter-factual. We can see that SOE's earn a higher share of profits in the scenario with the actual White List, implying that SOEs were more likely to be selected in the actual White List than if firms were picked purely based on short-run expected profitability as in the "Simulated White List".

This analysis illustrates the hazards of selecting winners: it is not straightforward to pick the right set of firms, the process for selecting firms is in practice influenced by factors (such as whether the firm is state-owned) that are unrelated to how productive the firms are, and not picking the right set of firms can lead to significantly lower profits.

It is important to note that even when targeted well, the subsidies are still very distortionary with a rate of return well below 1. If the government's goal is to maximize the net industry profit, the best subsidy would be no subsidy.

6.3 Discussion

Our counter-factual analysis shows that industrial subsidies are a very costly means of increasing industry profits. The particular policy regime adopted in China - such as the heavy reliance on entry subsidies - was also sub-optimal if the objective were purely to maximize shipbuilding profits net of the cost of the subsidies. In this section, we assess various other arguments for subsidies and the extent to which the industrial policies were effective in achieiving these objectives.

Suppose that the government's goal is to let the industry grow for a few years, before consolidating by picking the best firms and forcing the other firms to exit. The initial phase when the industry grows is akin to a horse-race where firms compete with one another to be the best. A possible justification for subsidizing the industry during the initial horse-race is to create a larger, more developed pool of firms from which to select the eventual set of industry leaders.

We assess this argument for subsidies by simulating two scenarios. One ("White List, Subsidies") is to subsidize entry, production and investment from 2006 to 2013 in order to accelerate industry growth during the horse-race phase, and then pick the 51 best firms from that pool. A second policy ("White List, No Subsidies") would be to just let the industry evolve for a few years without any intervention, and then pick the best firms. In both scenarios, once the White List firms have been selected and the remaining firms removed, the industry is allowed to evolve from 2014 onwards with no further government intervention.

Figure 12 shows that subsidizing the industry prior to consolidation does in fact lead to higher long-run industry profits. Industry profits after consolidation in the scenario "White List, Subsidies" are 18% higher than in the scenario "White List, No Subsidies". Since the number of firms is the same once consolidation has happened, this profit difference is driven by the fact that in the scenario with subsidies from 2006 to 2013, the top 51 firms are bigger and more productive than in the scenario with no subsidies.

However the long-run increase in industry profits is only 29% of the cost of the subsidies from 2006 to 2013. If we also account for the costs of increased entry during 2006 to 2013, the net industry benefits from the policy of subsidies followed by consolidating are close to zero, whereas the costs of the subsidies are very large. Thus the horse-race argument does not provide a compelling justification for subsidizing the industry. In particular, entry subsidies are particularly costly when combined with consolidation because a lot of the entry costs go to waste when the industry eventually consolidates.

A classic justification for subsidies is the presence of spillover effects across firms: if firms benefit from being located close to one another or if firms learn from one another, then there is a theoretical basis for subsidizing firms. (This is akin to the usual justification behind infant industry arguments for protectionist trade policies). Our benchmark specification rules out spillover effects since we assumed that a shipbuilder's costs are independent of costs faced by other shipbuilders.

As we discussed in Section 5, though, in robustness analyses we have allowed a firm's marginal cost to depend on industry cumulative output. We found limited evidence for any spillover effects: marginal costs decrease with cumulative industry production only for bulk (though the coefficient is not significant), and increase with cumulative industry production for tanker and container. This appears inconsistent with a world where spillovers across firms are important; in particular, one would have thought that any learning effects would be more important for the more advanced ship types (tankers and containerships), yet it is precisely for these ship types that we see the least evidence of learning spillovers.

Our analysis focuses on the shipbuilding sector and does not directly account for potential spillover benefits to other sectors. Liu (2018) shows that in the presence of market imperfections (such as financial frictions), there is an economic rationale for subsidizing upstream industries with lots of linkages towards downstream sectors. However, it is unlikely such a justification for subsidies holds for the Chinese shipbuilding industry. Kalouptsidi (2018) shows that the benefit to shippers from Chinese production subsidies was small relative to the cost of the subsidies. The benefit to Chinese shippers is smaller still since China is not a dominant buyer in the global shipping market, and accounted for slightly less than 10% of the total value of the global ship orderbook from 2010-2013 (the number is lower in earlier years). Not all the ships produced by Chinese shipbuilders go to Chinese shippers: a large share is exported which reduces the share of benefits from subsidies that can be captured domestically. While we do not see exports directly in our dataset, according to the 2012 Input-Output Table, at least 43% of the output from the shipbuilding industry that was not re-used within the sector itself was exported.[31]

There are also potentially spillovers from shipbuilding to upstream sectors. 63% of the value of the output produced by shipbuilding is accounted for by intermediate inputs from other sectors. In particular, steel is a major input used in shipbuilding and the cost of steel accounts for 12% of the value of shipbuilding output in the year 2012, according to the Input-Output table. However steel used in shipbuilding accounts for only a small proportion of the total amount of steel produced in China. This makes it unlikely that subsidizing shipbuilding is an effective way of countering any distortions in the upstream steel sector. Even if the subsidies to shipbuilding were large enough to have an appreciable effect on steel demand, they would lead to additional distortions since the increase in steel prices would reduce the amount of steel used in other industries.

Subsidizing output increases employment, which could have welfare benefits if there are labor market distortions that lead to depressed employment. However shipbuilding is not a particularly labor-intensive activity and wages account for only 5% of the costs of a shipyard on average, which suggests that the employment benefits from the subsidies are modest.

[31] This number is possibly an under-estimate of the true share of exports. 48% of shipbuilding output net of own use is used for "Capital Formation", which includes firms that purchase ships and add ships to their stock of fixed assets but may also include "work-in-progress", or ships that the yard has begun to build but not yet completed.

Our model of shipbuilding abstracts away from market power by assuming perfectly competitive firms. In the presence of market power, there are in theory strategic trade benefits from subsidizing output. However, our main findings are unchanged if we allow for firms to exercise market power by assuming Cournot competition: indeed we find the returns to subsidies are marginally lower. The shipbuilding industry is a somewhat unlikely target for strategic trade policy because the industry is unconcentrated to begin with (as illustrated, for example, in Figures 10 and **??**). Moreover the strategic trade literature suggests subsidizing the most efficient domestic firms, whereas entry subsidies (which form the largest component of subsidies according to our estimates) primarily go to small, less productive firms.

7 Conclusion

We evaluate the effectiveness of Chinese industrial policies in the shipbuilding sector. We carry out our evaluation using a simple model of the industry where price-taking firms make static production decisions and dynamic investment, entry and exit decisions. Starting from 2006 the government subsidizes three of these decision margins: production, investment and entry. In 2013, the government set up a White List of "top" firms that were prioritized for financing.

We find the subsidies led to a large increase in production, investment and entry in China's shipbuilding industry. However the cost of subsidies is very large relative to the benefits domestic producers receive from being subsidized: the rate of return on subsidies is only 21%. One reason for the low ratio of benefits to costs is the sheer scale of the subsidies, which amounted to around 100 billion USD from 2006 to 2013. We find that the larger the magnitude of subsidies, the larger the per-dollar distortion caused by the subsidies.

In order to understand the sources of these policy-induced distortions, we compare the effectiveness of different types of subsidies. Entry subsidies are the most ineffective among the three policies as they result in a fragmented industry with a large number of small unproductive firms, many of whom end up remaining idle most of the time. Comparing production and investment subsidies, we find that while production subsidies are more effective for raising industry output, investment subsidies are more effective for increasing industry profits. This latter effect is driven by the fact that production subsidies lead to fewer exits of inefficient firms, and shows the importance of taking dynamic considerations into account when evaluating industrial policies.

Our finding that industrial policies were costly relative to the benefits contrasts with a recent literature that argues that industrial policy can be beneficial under certain conditions. Lane (2017) empirically analyses the effect of industrial policy in South Korea, finding that targeted sectors (among them shipbuilding) grow significantly more than non-targeted sectors, and there were spillovers benefits to downstream industries connected to the targeted industries. China's industrial policy,

however, crucially differed from that of South Korea in that it led to the formation of a fragmented industry with many small, inefficient firms which stayed idle for long periods of time. Our analysis shows that this fragmentation is a key reason why Chinese industrial subsidies were costly relative to their benefits.

The increased fragmentation and idleness resulting from subsidies suggests alternate designs of industrial policy with more targeted subsidies would work better. Using counter-factual simulations, we find that subsidies can be more effective when targeted towards the most productive firms. In practice, though, the most productive firms may not be selected, as shown by the experience of the White List policy, where the government ended up picking a disproportionately high number of state-owned firms. We show that not selecting the most productive firms significantly reduces the potential benefits from targeting subsidies.

References

Amsden, Alice H., *Asia's Next Giant: South Korea and Late Industrialization*, Oxford University Press, 1989.

Bajari, Patrick, Lanier Benkard, and Jon Levin, "Estimating Dynamic Models of Imperfect Competition," *Econometrica*, 2007, *75* (5), 1331–1370.

Baldwin, Robert E., "The Case against Infant-Industry Tariff Protection," *Journal of Political Economy*, 1969, *77* (3), 295–305.

Barwick, Panle Jia and Parag Pathak, "The Cost of Free Entry: An Empirical Analysis of Real Estate Agents in Greater Boston," *The RAND Journal of Economics*, 2015, *46* (1), 103–145.

Brandt, Loren, Johannes Van Biesebroeck, and Yifan Zhang, "Creative Accounting or Creative Destruction? Firm-level Productivity Growth in Chinese Manufacturing," *Journal of Development Economics*, 2012, *97*, 339–351.

Ericson, R. and A. Pakes, "Markov-Perfect Industry Dynamics: A Framework for Empirical Work," *Review of Economic Studies*, January 1995, *62*, 53–82.

Greenwald, Bruce and Joseph E. Stiglitz, "Industrial Policies, the Creation of a Learning Society, and Economic Development," *The Industrial Policy Revolution I: the Role of Government Beyond Ideology, edited by Joseph Stiglitz and Justin Lin*, 2013, *1*, 43–71.

Harrison, Ann and Andres Rodriguez-Clare, "Trade, Foreign Investment, and Industrial Policy for Developing Countries," *Handbook of Development Economics*, 2010, pp. 4039–4214.

Hirschman, Albert O., *The Strategy of Economic Development*, Yale University Press, 1958.

Hotz, J. and R. Miller, "Conditional Choice Probabilties and the Estimation of Dynamic Models," *Review of Economic Studies*, 1993, *60*, 497–529.

Ito, Takatoshi, *The Japanese Economy*, MIT Press, 1992.

Itskhoki, Oleg and Benjamin Moll, "Optimal Development Policies with Financial Frictions," *Econometrica*, 2018.

Johnson, Chalmers, *MITI and the Japanese Miracle: The Growth of Industrial Policy*, Stanford University Press, 1982.

Kalouptsidi, Myrto, "Detection and Impact of Industrial Subsidies: The Case of Chinese Shipbuilding," *Review of Economic Studies*, 2018, *85* (2), 1111–1158.

Krueger, Anne O., "Government Failures in Development," *Journal of Economic Perspectives*, September 1990, *4* (3), 9–23.

Krugman, Paul, *Geography and trade*, The MIT Press, 1991.

Lane, Nathan, "Manufacturing Revolutions: Industrial Policy and Networks in South Korea," *Working paper of Institute for International Economic Studies, Stockholm*, 2017.

Liu, Ernest, "Industrial Policies in Production Networks," *Princetone University Working Paper*, 2018.

Rodrik, Dani, "The Return of Industrial Policy," *Project Syndicate*, April 12 2010.

Ryan, Stephen, "The Costs of Environmental Regulation in a Concentrated Industry," *Econometrica*, 2012, *80* (3), 1019–1061.

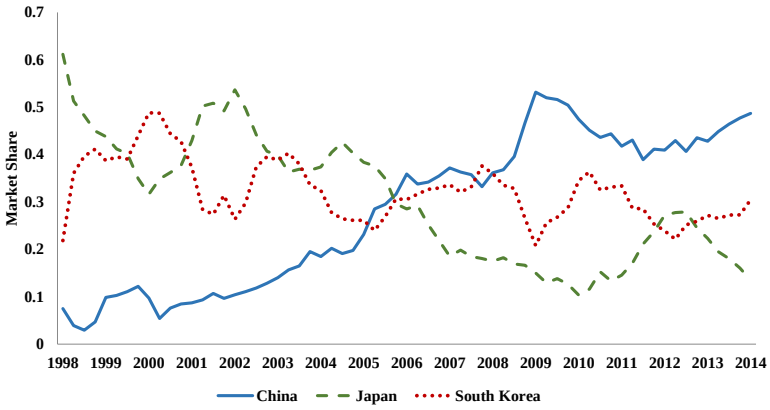
Stiglitz, Joseph E. and Justin Yifu Lin, *The Industrial Policy Revolution I: The Role of Government Beyong Ideology*, PALGRAVE MACMILLAN, 2013.

Stopford, Martin, *Maritime Economics*, New York: Routledge, 2009.

Sweeting, Andrew, "Dynamic Product Positioning in Differentiated Product Markets: The Effect of Fees for Musical Performance Rights on the Commercial Radio Industry," *Econometrica*, 2013, *81* (5), 1763–1803.

Wade, Robert, *Governing the Market: Economic Theory and the Role of Government in East Asian Industrialization*, Princeton University Press, 1990.

Figure 1: China's Market Share Expansion



Source: Clarkson.

Figure 2: Entry of New Shipyards

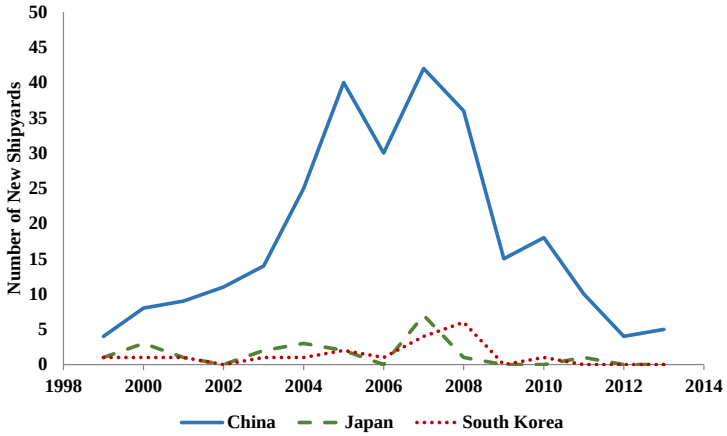


Figure 3: Investment by Chinese yards

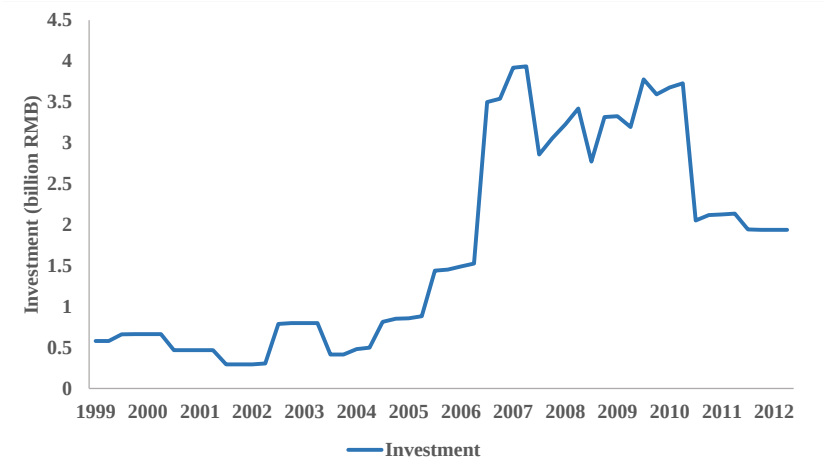


Figure 4: Ship Prices

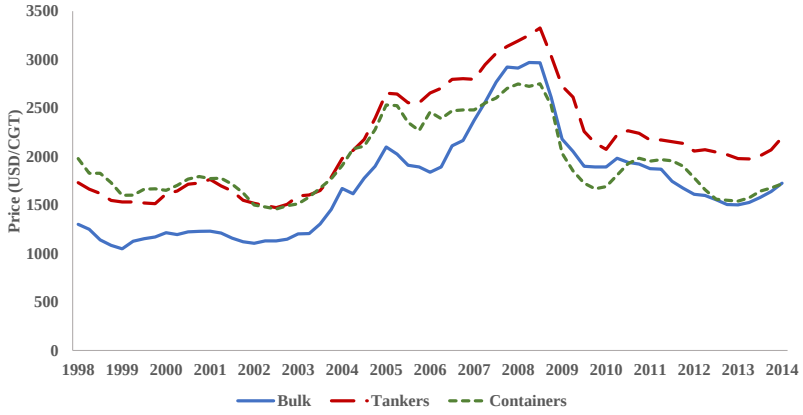


Figure 5: Fitted exit probability from 1st-stage policy function estimates

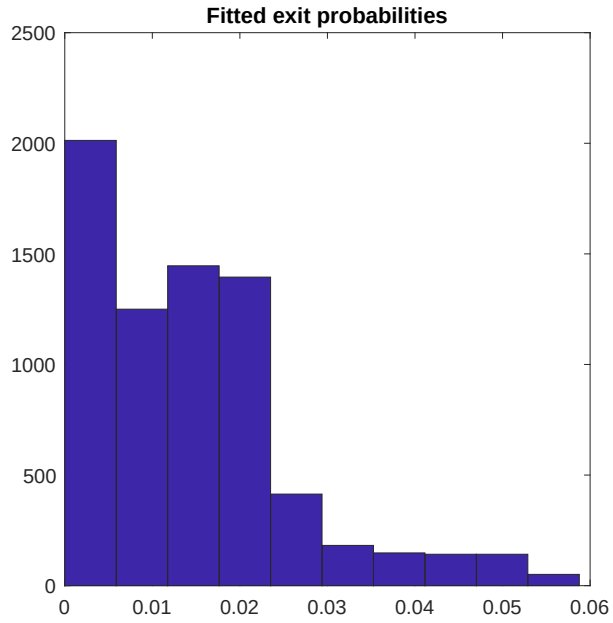


Figure 6: Fitted investment against capital, from 1st-stage policy function estimates

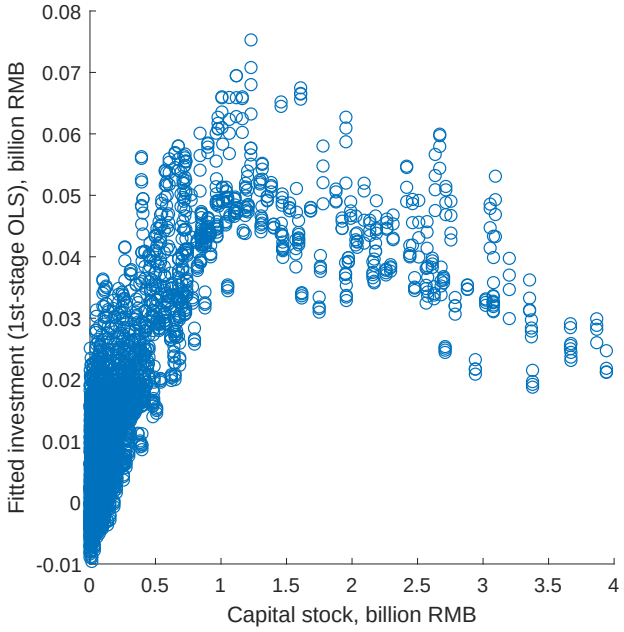


Figure 7: Simulated vs. actual investment

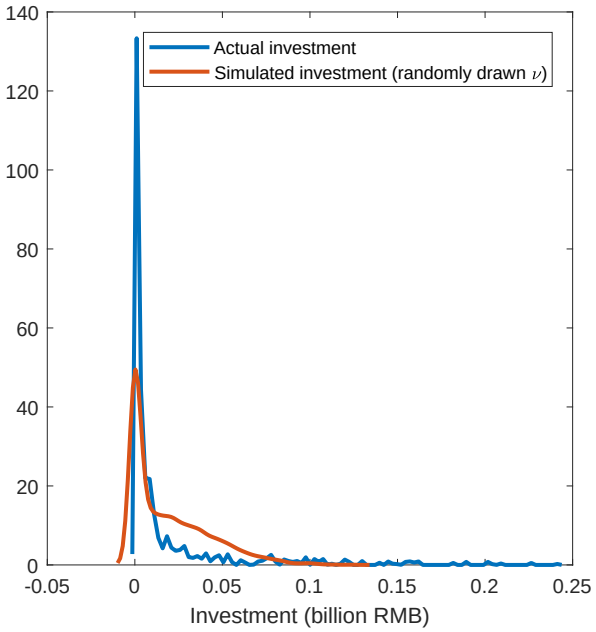


Figure 8: Number of firms, with and without subsidies

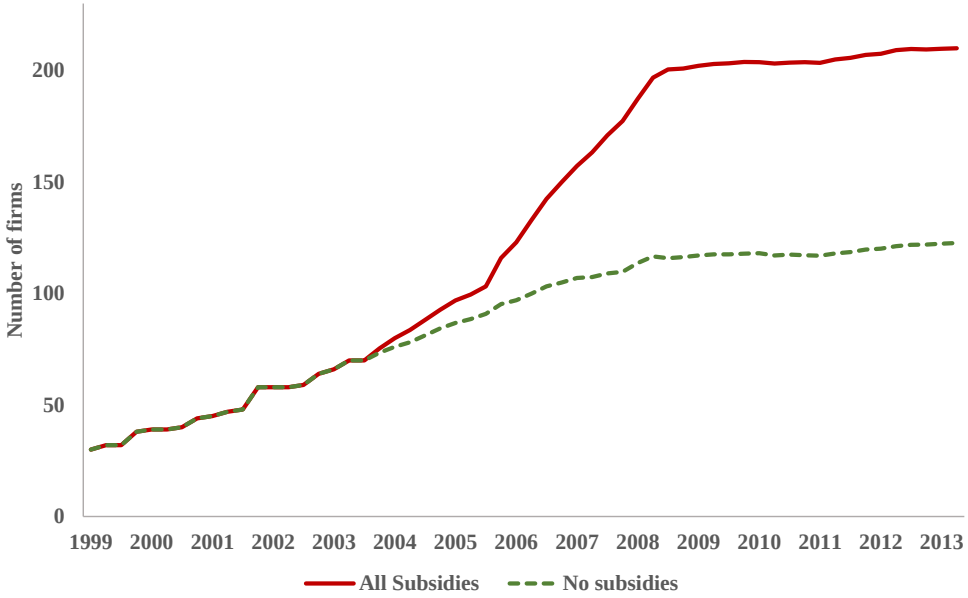


Figure 9: Investment, with and without subsidies

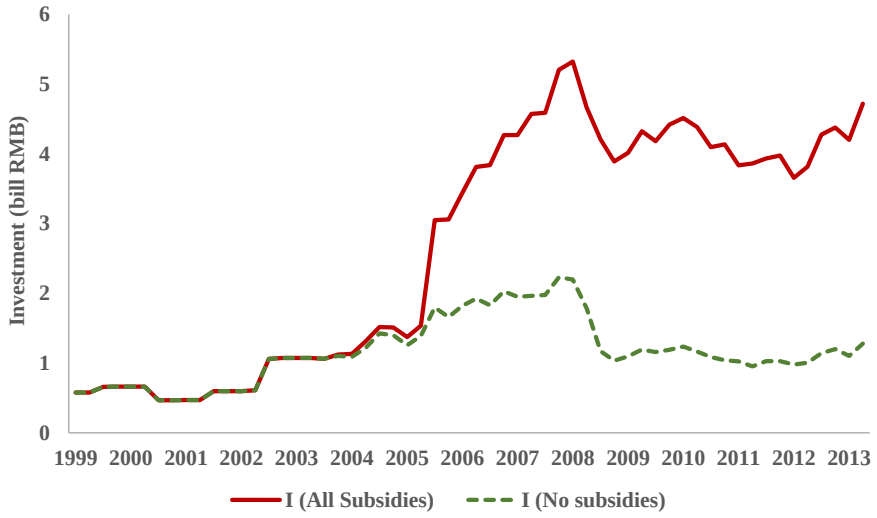


Figure 10: HHI, with and without subsidies

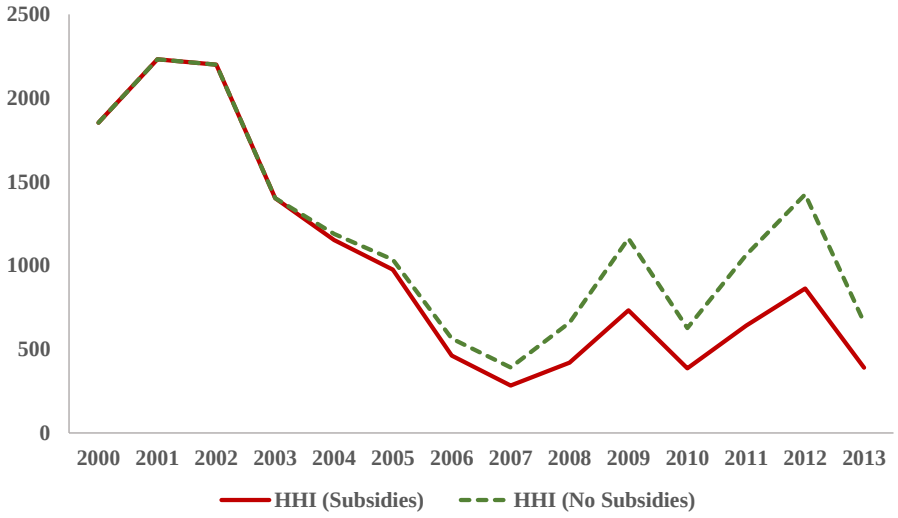


Figure 11: Average $\bar{s}_{jt}$ by year, in the scenario with 50% production subsidies vs. the scenario with investment subsidies

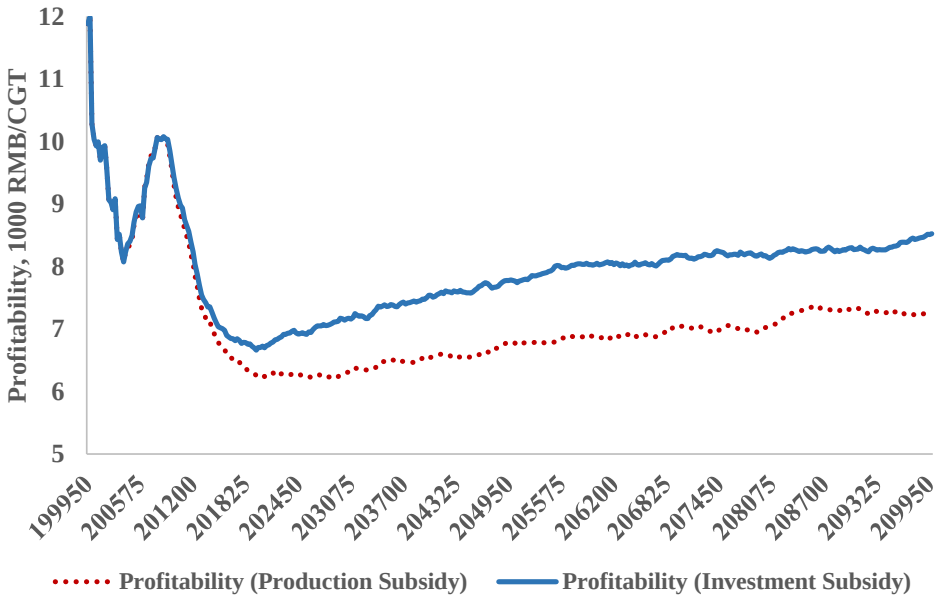


Figure 12: Industry Profits in Various White List scenarios

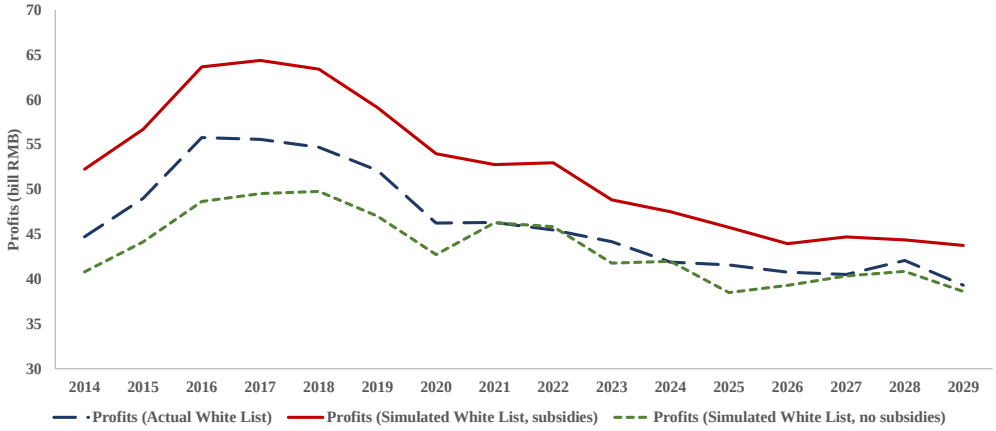


Figure 13: Profit share of state-owned firms (SOEs), actual White List vs. simulated White List

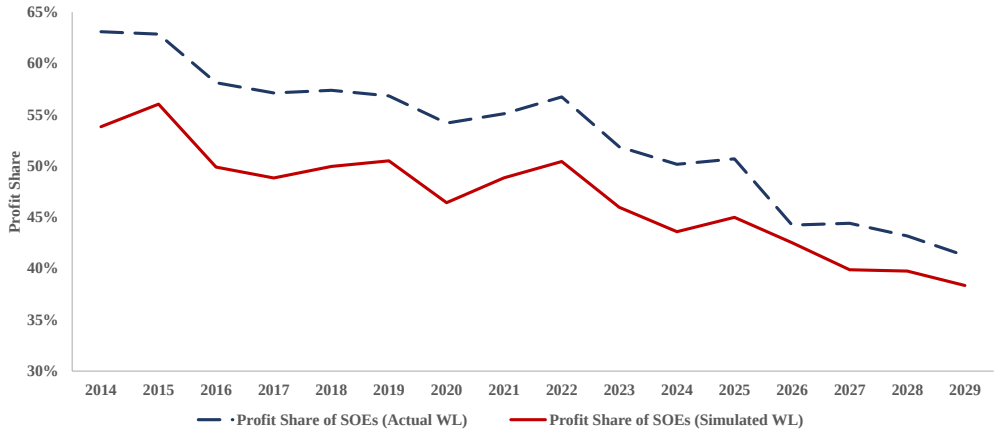


Table 1: Shipbuilding National Industrial Policies

Year	Shipbuilding National Industrial Policies	Plan Period
2003	National Marine Economic Development Plan	2001-2010
2006	The 11th Five-Year Plan for National Economic and Social Development	2006-2010
2006	The Medium and Long Term Development Plan of Shipbuilding Industry	2006-2015
2007	The 11th Five-Year Plan for the Development of Shipbuilding Industry	2006-2010
2007	The 11th Five-Year Plan for the Development of Shipbuilding Technology	2006-2010
2007	11th Five-Year Plan for the Development of Ship Equipment Industry	2006-2010
2007	Guideline for Comprehensive Establishment of Modern Shipbuilding (2006-2010)	2006-2010
2007	Shipbuilding Operation Standards	2007-
2009	Plan on the Adjusting and Revitalizing the Shipbuilding Industry	2009-2011
2010	The 12th Five-Year Plan for National Economic and Social Development	2011-2015
2012	The 12th Five-Year Plan for the Development of the Shipbuilding Industry	2011-2015
2013	Plan on Accelerating Structural Adjustment and Promoting Transformation and Upgrading of the Shipbuilding Industry	2013-2015
2013	Shipbuilding Industry Standard and Conditions	2013-

Table 2: Summary Statistics

Variable	Obs	Mean	S.D.	Min	Max
All Observations					
Bulk orders (1000 CGT)	14,455	11.9	44.1	0.0	968.2
Tanker orders (1000 CGT)	14,455	7.0	39.8	0.0	1,119.0
Container orders (1000 CGT)	14,455	6.3	54.9	0.0	1,644.1
Bulk backlog (1000 CGT)	14,455	119.8	286.3	0.0	2,830.5
Tanker backlog (1000 CGT)	14,455	72.1	273.1	0.0	3,840.8
Container backlog (1000 CGT)	14,455	68.8	398.9	0.0	7,362.8
Investment (mill RMB)	4,386	18.5	88.9	-1,240.5	1,770.7
Capital (mill RMB)	6,157	392.0	806.9	0.3	8,203.3
Observations With Positive Orders					
Bulk orders (1000 CGT)	2,316	74.6	86.5	3.9	968.2
Tanker orders (1000 CGT)	1,436	70.4	107.1	0.05	1,119.0
Container orders (1000 CGT)	625	145.3	222.7	2.3	1,644.1

[1.] The data on orders and backlog is for yards in China, Japan and Korea. Out of the 14,455 observations, 7,186 are for Chinese yards, 5,448 are for Japanese yards and 1821 are for Korean yards.
[2.] We observe investment and capital only for Chinese yards.

Table 3: Demand estimates

Dependent variable:	(1) Orders	(2) Orders	(3) Orders	(4) Orders
Price (bulk)	-2.34***	-1.67***	-2.07***	-2.12***
	(0.77)	(0.64)	(0.69)	(0.75)
Price (tanker)	-2.66***	-1.46*	-1.80**	-1.76**
	(0.60)	(0.88)	(0.78)	(0.89)
Price (container)	-4.85***	-2.44***	-3.39***	-3.39***
	(0.91)	(0.85)	(1.01)	(0.99)
Price*Post2006	1.34***	1.00***	1.15***	1.34**
	(0.18)	(0.14)	(0.15)	(0.55)
Backlog (log)	0.34	-1.00***	-0.78**	-0.81**
	(0.25)	(0.33)	(0.38)	(0.37)
Freight Rate	Yes	Yes	Yes	Yes
Demand shifters		Yes	Yes	Yes
Trend			Yes	Yes
Trend*Post2006				Yes
R^2, bulk	0.68	0.71	0.71	0.71
R^2, tanker	0.26	0.33	0.35	0.36
R^2, container	0.44	0.52	0.51	0.51

* N equals 64 for bulk/container, 61 for tankers. The freight rate we use is the Baltic Exchange Freight Index for bulk ships, Baltic Exchange Clean Tanker Index for tankers and the Containership Timecharter Rate Index for containerships. The demand shifters are the US wheat price and total Chinese iron ore imports for bulk demand, Middle East refinery production for tanker demand and world car trade for containership demand. All parameters are estimated using GMM. We instrument for ship prices using steel production and the steel ship plate price.

Table 4: Cost function estimates

	Bulk		Tanker		Container	
Type-specific	Coefficient	T-stat	Coefficient	T-stat	Coefficient	T-stat
MC (thousand RMB / CGT)						
β_q	9.10	12.94	11.04	9.36	5.06	7.22
σ_ω	10.36	13.78	15.12	11.11	13.36	9.20
China (1000 RMB/CGT)	19.40	17.11	32.30	12.99	29.79	12.68
Japan (1000 RMB/CGT)	12.57	16.67	30.92	12.78	25.02	11.45
Korea (1000 RMB/CGT)	15.92	14.88	23.07	12.01	19.83	11.13
Steel Price (1000 RMB/Ton)	2.12	14.20	2.46	8.80	1.89	6.25
Capital (bill RMB)	-2.92	-3.06	-2.06	-1.55	-1.41	-1.33
Capital^2	0.23	0.89	-0.01	-0.05	-0.03	-0.14
Backlog	-2.09	-6.63	-4.50	-6.38	-3.06	-4.40
Backlog^2	0.10	4.83	0.24	4.15	0.19	2.48
Backlog of Other Types	0.11	0.76	0.32	1.47	0.47	2.66
Common						
China 2006-2008	-2.79	-4.57				
China 2009+	-0.90	-1.56				
Large firms	-4.17	-6.84				
Jiangsu	-2.93	-4.81				
Zhejiang	-1.57	-2.79				
Liaoning	-1.87	-1.88				
CSSC/CSIC	-0.93	-1.28				
Private	0.13	0.24				
Foreign JV	-0.92	-1.61				
Age	0.26	6.14				
N	10013		10429		4661	

* Standard errors bootstrapped using 500 bootstrap samples.

Table 5: Cost function estimates: robustness analysis

	Baseline		*Trend*		*Existing yards*		*Learning*		*Industry Learning*	
	(1)		(2)		(3)		(4)		(5)	
	Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat
Bulk										
Capital	-2.92	-3.06	-2.91	-3.11	-3.33	-3.18	-1.94	-2.12	-2.04	-2.16
Backlog	-2.09	-6.63	-2.09	-6.41	-3.16	-6.38	-1.45	-5.05	-1.35	-4.58
Cumul. Q.							0.07	4.72	0.08	4.76
Cumul. Q., China									-0.03	-1.60
Tanker										
Capital	-2.06	-1.55	-2.06	-1.56	-2.49	-1.41	-1.95	-1.52	-3.41	-1.69
Backlog	-4.50	-6.38	-4.50	-6.42	-5.15	-6.13	-4.41	-5.06	-6.60	-4.02
Cumul. Q.							0.09	5.59	0.13	4.39
Cumul. Q., China									0.18	1.47
Container										
Capital	-1.41	-1.33	-1.41	-1.33	-0.44	-0.40	-1.11	-1.22	-2.04	5.77
Backlog	-3.06	-4.40	-3.06	-4.40	-3.66	-4.25	-0.90	-1.30	-1.58	1.91
Cumul. Q.							0.01	4.00	0.02	1.33
Cumul. Q., China									0.33	2.94
Common										
China 2006-2008	-2.79	-4.57	-2.78	-4.19	-2.08	-2.72	-2.43	-4.43	-3.10	-1.60
China 2009+	-0.90	-1.56	-0.89	-1.35	-1.01	-1.25	-0.95	-1.73	-0.08	0.80
Trend			-0.02	-0.03						

* Standard errors bootstrapped using 500 bootstrap samples. Column (1) is the same specification reported in Table 4 and includes the same set of marginal cost shifters. Column (2) adds a time trend. Column (3) estimates the cost specification as Column (1) but drops any Chinese yards that entered after 2004. Column (4) takes the specification in Column (1) and includes the variable "Cumul. Q" which refers to the firm's cumulative past production of that ship category. Column (5) takes the specification in Column (4) and adds the variable "Cumul. Q, China" which refers to China's cimulative past production.

Table 6: Estimates of exit policy function

	(1)		(2)	
	Coefficient	SE	Coefficient	SE
Constant	-0.57	(0.97)	-0.56	(1.02)
K	0.05	(0.35)	0.54	(0.43)
K^2	-0.05	(0.12)	-0.16	(0.15)
2006-2008	-0.57	(0.41)	-0.64	(0.43)
2009+	-0.47	(0.41)	-0.72	(0.44)
$\bar{s}_{jt}$	-0.01	(0.01)	-0.04	(0.02)
Bulk price	0.36	(0.12)	0.36	(0.12)
Tanker price	-0.18	(0.11)	-0.16	(0.11)
Container price	-0.22	(0.10)	-0.25	(0.11)
Steel price	-0.06	(0.07)	-0.10	(0.08)
Jiangsu			0.77	(0.24)
Zhejiang			0.58	(0.19)
Liaoning			1.04	(0.28)
N(exit)	47		47	
N	4605		4605	
Residual deviance	478.59		459.47	
Log-likelihood	-239.30		-229.74	
Pseudo-R2	0.09		0.12	

Table 7: Estimates of investment policy function

	(1)		
	Coefficient	SE	t-stat
Constant	-0.0001	(0.0149)	0.00
B-spline 1 capital	-0.0697	(0.0395)	-1.76
B-spline 2 capital	-0.0748	(0.0314)	-2.38
2006-2008	0.0064	(0.0026)	2.46
2009+	0.0027	(0.0042)	0.64
$\bar{s}_{jt}$	0.0007	(0.0002)	3.74
Bulk price	0.0021	(0.0007)	3.04
Tanker price	0.0005	(0.0014)	0.34
Container price	-0.0012	(0.0012)	-1.03
Steel price	-0.0025	(0.0009)	-2.63
R^2	0.2363		
N	4286		
N($I > 0$)	3301		
N($I = 0$)	985		

Table 8: AR(1) estimates for state transition processes

	Bulk	Tanker	Container	Steel
Constant	0.88	0.70	1.25	-0.023
	(0.87)	(0.94)	(1.11)	(0.37)
Post	3.44	3.63	1.80	2.32**
	(2.28)	(3.43)	(3.33)	(1.01)
Price (t-1)*Pre	0.86***	0.92***	0.88***	0.89***
	(0.12)	(0.086)	(0.090)	(0.19)
Price (t-1)*Post	0.86***	0.86***	0.88***	0.69***
	(0.072)	(0.095)	(0.10)	(0.11)
Trend*Pre	0.042	0.038	0.029	0.024
	(0.033)	(0.028)	(0.024)	(0.027)
Trend*Post	-0.058**	-0.054	-0.040	-0.022*
	(0.027)	(0.041)	(0.040)	(0.013)
N	57	57	57	57
R^2	0.95	0.97	0.96	0.80

* The dependent variable is the price in quarter t, or $p(t)$."Pre" refers to 2005Q4 or earlier. "Post" refers to 2006Q1 or later. The sample ranges from 1999 Q4 to 2013Q4. *,** and *** imply the coefficients are statistically significant at the 10%, 5% and 1% level respectively.

Table 9: Estimates of investment cost and scrap value parameters

	Coeff.	T-stat
σ	0.69	10.60
c1	1.00	
c2	24.16	16.40
c3	1.45	10.16
$c4_{2006-08}$	-0.40	-3.05
$c4_{2009+}$	-0.81	-7.61
N	4286	
R^2 for investment	0.16	

* Standard errors bootstrapped using 200 block bootstrap samples.

Table 10: Actual vs. Simulated Exits

	1999-2005	2006-2013	Total
Actual exits	5	43	48
Simulated exits	8	30	38

Table 11: Entry Cost Distribution (Mean), billion RMB

	κ_{pre}	$\kappa_{post,08}$	% of pre-06 costs	$\kappa_{post,09+}$	% of pre-2006 costs
Jiangsu	101	41	40%	121	119%
Zhejiang	54	23	42%	114	211%
Liaoning	603	340	56%	-	-
Other	56	24	43%	86	154%

[1] κ_{pre} refers to the mean of the entry cost distribution prior to 2004 for Zhejiang, and prior to 2006 for Jiangsu, Liaoning and Other regions.

[2] $\kappa_{post,08}$ refers to the mean of the entry cost distribution between 2004 and 2008 for Zhejiang and between 2006 and 2008 for Jiangsu, Liaoning and Other regions.

[3] $\kappa_{post,09+}$ refers to the mean of the entry cost distribution from 2009 onwards.

Table 12: Actual vs. Simulated Entries

	Pre	Post, Until 2008	Post, 2009+	Total
Actual entries	83	122	39	244
Simulated entries	84	121	36	241

[1] "Pre" refers to the period prior to 2004 for Zhejiang, and prior to 2006 for Jiangsu, Liaoning and Other regions.

[2] "Post, Until 2008" refers to the period between 2004 and 2008 for Zhejiang and between 2006 and 2008 for Jiangsu, Liaoning and Other regions.

[3] "Post, 2009+" refers to the period from 2009 onwards.

Table 13: Magnitude of Subsidies, 2006-13, in billion RMB

	Subsidies, 2006-13
Production	160
Investment	99
Entry	330
Total	590

Table 14: Impact of Subsidies on Ship Prices

	Bulk	Tanker	Container
Subsidies, 2006-08	16.3	19.8	17.1
No subsidies, 2006-08	17.6	21.0	17.7
% difference	8.3%	6.2%	3.3%
Subsidies, 2009-13	8.8	8.2	9.1
No subsidies, 2009-13	10.2	9.1	9.4
% difference	16.2%	10.1%	3.9%

[1] Prices in 1000 RMB/CGT

Table 15: Comparison of Different Subsidies

	All Subsidies	Only Production	Only Investment	Only Entry	No Subsidies
Lifetime Revenue 2006-	2336	2104	1851	1898	1752
Lifetime Profits 2006-	876	783	641	609	581
Scrap Value	271	210	231	308	231
Investment cost	25	55	-1	37	31
Entry cost	329	183	152	212	137
Production subsidies	260	219	0	0	0
Investment subsidies	167	0	93	0	0
Entry subsidies	300	0	0	188	0
Δ Revenue/Subsidy	80%	161%	106%	77%	
Δ (Profit-Inv. Cost)/Subsidy	41%	82%	98%	12%	
Δ Net Profit/Subsidy	21%	51%	82%	13%	

[1] Each element in the table refers to the discounted sum from 2006 to 2099, averaged across all of the simulations. For example, "Scrap Value" refers to the discounted sum of scrap values earned by exitting firms from 2006 to 2099.

[2] ΔRevenue/Subsidy equals the discounted sum of revenue in the scenario minus the discounted sum of revenue in the scenario with no subsidies, divided by the discounted sum of subsidies.

[3] (Profit-Inv.Cost) refers to profits net of the cost of investment. Δ(Profit-Inv.Cost)/Subsidy equals the discounted sum of (profits-investment cost) in the scenario minus the discounted sum of (profits-investment cost) in the scenario with no subsidies, divided by the discounted sum of subsidies.

[4] Net Profit = (Profits-Investment Cost+Scrap Value-Entry Cost). ΔNet Profit/Subsidy equals the discounted sum of net profits in the scenario minus the discounted sum of net profits in the scenario with no subsidies, divided by the discounted sum of subsidies.

Table 16: Comparing Production and Investment Subsidies

	100% Production subsidies	50% Production subsidies	Investment subsidies
Lifetime Revenue 2006-	2104	1937	1851
Lifetime Profits 2006-	783	682	641
Production subsidies	219	97	0
Investment subsidies	0	0	93
Scrap value	210	219	231
Investment cost	55	42	-1
Entry cost	183	156	152
Δ Revenue /Subsidies	161%	192%	106%
Δ (Profit-Invest Cost)/Subsidies	82%	93%	98%
Δ Net Profit/Subsidies	51%	62%	82%

[1] Each element in the table refers to the discounted sum from 2006 to 2099, averaged across all of the simulations. For example, “Scrap Value" refers to the discounted sum of scrap values earned by exitting firms from 2006 to 2099.

[2] ΔRevenue/Subsidy equals the discounted sum of revenue in the scenario minus the discounted sum of revenue in the scenario with no subsidies, divided by the discounted sum of subsidies.

[3] (Profit-Inv.Cost) refers to profits net of the cost of investment. Δ(Profit-Inv.Cost)/Subsidy equals the discounted sum of (profits-investment cost) in the scenario minus the discounted sum of (profits-investment cost) in the scenario with no subsidies, divided by the discounted sum of subsidies.

[4] Net Profit = (Profits-Investment Cost+Scrap Value-Entry Cost). ΔNet Profit/Subsidy equals the discounted sum of net profits in the scenario minus the discounted sum of net profits in the scenario with no subsidies, divided by the discounted sum of subsidies.

Table 17: Targeting subsidies to White List firms

	Subsidize All Firms After 2013	Subsidize White List firms After 2013	No Subsidies After 2013
Lifetime Revenue 2006-	972	918	804
Lifetime Profits 2006-	439	404	299
Production subsidies	134	89	0
Investment subsidies	96	30	0
Scrap value	171	184	199
Investment cost	-9	8	5
Entry cost	69	47	47
Δ Revenue/Subsidies	73%	96%	
Δ (Profit-Invest Cost)/Subsidies	67%	85%	
Δ Net Profit/Subsidies	45%	72%	

[1] Each element in the table refers to the sum from 2014 to 2099, discounted back to 2006, averaged across all of the simulations. For example, “Scrap Value" refers to the discounted sum (in 2006) of scrap values earned by exitting firms from 2014 to 2099.

[2] ΔRevenue/Subsidy equals the discounted sum of revenue in the scenario minus the discounted sum of revenue in the scenario with no subsidies after 2013, divided by the discounted sum of subsidies.

[3] (Profit-Inv.Cost) refers to profits net of the cost of investment. Δ(Profit-Inv.Cost)/Subsidy equals the discounted sum of (profits-investment cost) in the scenario minus the discounted sum of (profits-investment cost) in the scenario with no subsidies after 2013, divided by the discounted sum of subsidies.

[4] Net Profit = (Profits-Investment Cost+Scrap Value-Entry Cost). ΔNet Profit/Subsidy equals the discounted sum of net profits in the scenario minus the discounted sum of net profits in the scenario with no subsidies after 2013, divided by the discounted sum of subsidies.

Appendix

A Calibrating the fixed cost of operation (idleness)

As we discuss in Section 4.2 a firm's production decisions provide no information on its idleness cost (or fixed cost of operation) c_{idle}, since the firm incurs this cost regardless of whether or not it produces. Instead, we calibrate the fixed cost of operation by using accounting costs reported by the firms even during periods when it is not engaged in any production.

The firm's costs equal:

$$C_{jt} = c_{idle} + C(q_{jt}) \tag{12}$$

where c_{idle} is a fixed cost of operation (or an idleness cost) that the firm incurs whether or not it engages in production, and $C(q_{jt})$ is the cost of taking q_{it} orders. We estimate $C(q_{jt})$ from the first-order condition, as discussed in Section 4.2.

To calibrate c_{idle}, we utilize the accounting costs reported by each firm in the NBS manufacturing survey, which we call C_{jt}^{NBS}. However C_{jt}^{NBS} is generally larger than C_{jt}, because C_{jt}^{NBS} partly reflects repair activities: even yards that have zero orders have positive revenue and positive costs (that are not part of their fixed cost) since they may be engaged in repair activity.

We deal with the issue of repairs using the following procedure. Let $R_j^{NBS} = \sum_t R_{jt}^{NBS}$ denote the yard's lifetime revenues in the NBS data. Let $R_j^{ships} = \sum_t R_j^{ships}$ denote the yard's lifetime revenues from building new ships. Let R_j^{ships}/R_j^{NBS} denote the share of the yard's revenues that go towards shipbuilding.

We next assume that each yard's costs are assigned to shipbuilding and ship repairs in proportion to the revenue they receive from the respective activities. Thus, we can now form an estimate $\hat{C}_{jt}$ for C_{jt} as follows:

$\hat{C}_{jt} = C_{jt}^{NBS} * (R_j^{ships}/R_j^{NBS})$

Plugging in $\hat{C}_{jt}$ into (12), we can now use two approaches to estimate the fixed cost of operation. The first approch is to plug in our estimated production cost function, $C(q_{jt})$, into (12), and take the average of both sides of the equation across the lifetime of the yard (to deal with the fact that accounting costs are not always reported during the period when the ship was ordered, due to time-to-build). This gives us the following equation which we can use to calculate the fixed cost of operation:

$$[\hat{C}_{jt} - C(q_{jt})] = c_0$$

A second approach is to use the subsample of observations where the shipyard engages in zero production (so that $C(q_{jt}) = 0$) and use the accounting costs (after adjusting for repairs) in those

periods. This gives us the following equation: Take observations where ship production equals 0, so we know $C(q_{jt}) = 0$ and that $\hat{C}_{jt} = c_0$.

Both approaches give rise to similar predictions for the average fixed cost of operation. Notice that we restrict the calibrated fixed cost of operation to be the same for all firms. While in principle either of these approaches could be allowed to estimate fixed costs that vary by firm characteristics, our data in practice is not rich enough to allow that.

B Estimation of investment cost and exit scrap value parameters

B.1 Estimation of the first-stage investment policy function

Section 4.3.1 provided an overview of how we estimate the first-stage policy functions. Here we provide further details on estimation of the investment policy function $I^*_{jt} = I(s_{jt}, \nu_{jt})$. We assume that $I(s_{jt}, \nu_{jt}) = h_1(s_{jt}) + h_2(\nu_{jt})$, meaning that investment is additively separable in s_{jt} and ν_{jt}.

Following Bajari et al. (2007), we estimate the investment policy function as follows:

- First, estimate the distribution of investment at each state, $F(I_{jt}|s_{jt})$. Relying on the assumption that investment is additively separable in s_{jt} and ν_{jt}, we regress I_{jt} on a set of spline basis functions in s_{jt}, and obtain the residuals $\hat{h}_2(\nu_{jt}) = I - \hat{h}_1(s_{jt})$. Then the distribution $F(I_{jt}|s_{jt})$ is the empirical distribution of the residuals $\hat{h}_2(\nu_{jt})$, shifted by the conditional mean $\hat{h}_1(s_{jmt})$.[32]

- Second, observe that $I(s_{jt}, \nu_{jt})$ is weakly monotone in ν_{jt}: the higher ν_{jt} is, the lower is optimal investment $I(s_{jt}, \nu_{jt})$. Suppose that for some i and ν, $i = I(s_{jt}, \nu)$, meaning that if we hold the state fixed at s_{jt}, when the cost shock equals ν, the shipyard's optimal investment equals i. Then:

$$\begin{aligned} F(i|s_{jt}) &= Pr(I_{jt} \leq i|s_{jt}) \\ &= Pr(\nu_{jt} \geq \nu|s_{jt}) \\ &= 1 - G(\nu|s_{jt}) \\ &= 1 - G(\nu) \end{aligned}$$

[32] A non-parametric way to estimate $F(I_{jt}|s_{jt})$, that does not involve making the separability assumption, would be to (i) estimate the joint distribution of investment and the states $f_1(I_{jt}, s_{jt})$ using kernel methods (ii) estimate the joint distribution of the states $f_2(s_{jt})$ using kernel methods (iii) the conditional pdf of investment given states is then $\hat{f}(I|s_{jt}) = \frac{f_1(I_{jt}, s_{jt})}{f_2(s_{jt})}$, and we can integrate this to obtain the CDF.

where $G(\nu)$ is the (known) CDF of ν_{jt} (in our case, the standard normal CDF) evaluated at ν. We can now invert this to obtain:

$$i|s_{jt} = F^{-1}\big(1 - G(\nu)\big) \tag{13}$$

- The intuition behind this is that by estimating the CDF of $i|s$ and comparing it to the known CDF of ν, we can infer the relationship between i and ν at every s.
- For example, if for some state s, investment exceeds $\bar{i}$ 5% of the time in the data, and the 5th percentile of ν is $\bar{\nu}$, it must be the case that the optimal level of investment when the shock equals $\bar{\nu}$ is $\bar{i}$, or $\bar{i} = I(s, \bar{\nu})$. Repeating this for every ν in the support allows us to trace out the function $I(s, \nu)$.

B.2 Reduce number of states using structure of profit function

Our specification of production costs allows for a rich set of characteristics to affect a shipyard's marginal costs. Each of these then becomes a state variable in the firm's dynamic problem of choosing exit and investment. This can potentially lead to a dynamic problem with a large number of states, which can be challenging to solve.

In this section, we show that we can use the structure of the profit function to reduce the number of state variables. Under certain conditions, a set of state variables can be replaced by a single index that is a linear combination of these states. This reduces the computational burden of solving a dynamic programming problem having many state variables.

Recall that s_{jt} is the vector of states, and s_{jmt} is the subset of states that linearly enter the firm's marginal cost function for type m. We split the states into two groups. Let s^d_{jmt} collect all state variables that (i) affect the value function only through their effect on payoffs π_{jt} (ii) have deterministic, exogenous transition processes (for example, they are fixed over time). Let s^1_{jmt} collect all other state variables: this includes states such as subsidies which affect both payoffs and the cost of investment, as well as state variables such as capital that have non-trivial transition processes. We show that instead of working with s^d_{jmt}, we can collapse s^d_{jmt} into a single index and treat that index as the state variable in the firm's problem.

We begin by decomposing s_{jmt} as: $s_{jmt} = \left\{s^1_{jmt}, s^d_{jmt}\right\}$. Similarly we can also decompose the marginal cost coefficients as $\beta_{sm} = \left\{\beta^1_{sm}, \beta^d_{sm}\right\}$. We can then rewrite the equation for the firm's marginal costs, (8), as:

$$\begin{aligned} MC_m(q_{jmt}) &= \beta_{0m} + s_{jmt}\beta_{sm} + \beta_{qm} q_{jmt} + \omega_{jmt} \\ &= \beta_{0m} + s^1_{jmt}\beta^1_{sm} + s^d_{jmt}\beta^d_{sm} + \beta_{qm} q_{jmt} + \omega_{jmt} \end{aligned}$$

Let $\bar{s}_{jmt} = -s^d_{jmt}\beta^d_{sm}$. (We add the negative sign so that $\bar{s}_{jmt}$ can be interpreted as a measure

of the yard's productivity). Notice that in period t, knowledge of s^1_{jmt} and $\bar{s}_{jmt}$ is sufficient for constructing the firm's marginal costs, and hence its expected profit function, in period t. Observing each separate element of s^d_{jmt} does not tell us anything new about firm's profits that we did not already know from observing $\bar{s}_{jmt}$.

If in addition each element of s^d_{jmt} is fixed over time (which applies to state variables such as region and ownership status), then knowledge of s^1_{jmt} and $\bar{s}_{jmt}$ is sufficient to construct expectations of firm's profits not just in period t but in all future periods as well, and there is no additional information gained from observing each element of s^d_{jmt} separately. As such, we can equivalently write the firm's dynamic problem in terms of the reduced state space $(s^1_{jmt}, \bar{s}_{j1t}, .., \bar{s}_{jMt})$.

A simple example should make this clear. Suppose there are two separate states x_1 and x_2, which are deterministic and fixed over time. Both these states affect the shipyard's marginal cost, which equals $MC = \alpha + \alpha_1 x_1 + \alpha_2 x_2$. Suppose that the states x_1 and x_2 do not affect the firm's value function in any other way except through the effect on its marginal cost.

One way to solve the firm's dynamic programming problem is to treat x_1 and x_2 as two separate state variables. However, because x_1 and x_2 enter into the firm's discounted sum of payoffs only through the linear combination $\bar{s} = -(\alpha_1 x_1 + \alpha_2 x_2)$, and because $\bar{s}$ is constant over time, we can equivalently express the firm's value function in terms of $\bar{s}$ and treat $\bar{s}$ as the single state variable.

In our empirical analysis, we make two additional assumptions in order to reduce the state space even further. First of all, we include the firm's backlog and age into the state $\bar{s}_{jmt}$, even though these are not constant over time. Secondly, we aggregate over the types to construct a single state variable $\bar{s}_{jt} = \sum_m \bar{s}_{jmt}$, instead of having a separate state variable $\bar{s}_{jmt}$ for each ship type. Inclusion of the backlog and age in $\bar{s}_{jt}$ means that the state variable $\bar{s}_{jt}$ is no longer constant over time, and we instead model the transition process of $\bar{s}_{jt}$ as a simple linear process $\bar{s}_{jt+1} = \alpha_0 + \alpha_1 \bar{s}_{jt}$. We have found that using $\bar{s}_{jt}$ to predict $\bar{s}_{jt+1}$ gets the state transition process almost exactly right, and therefore there are unlikely to be major gains from treating backlog and age as distinct state variables. Moreover $\bar{s}_{jt}$ turns out to be a very strong predictor of a firm's expected profitability, and so there are limited gains from having separate indices for each type.

B.3 Additional details on value function approximation

We now provide some additional details on how we approximate the value function (equation (10)).

B.3.1 Selection of basis functions via LASSO

Given the large number of state variables, the full set of B-spline basis functions that can be used to approximate the value function is very large. It is computationally challenging to include every possible basis function. Moreover not all of these basis functions are likely to be important, and the coefficients γ^0_l on some basis functions might be close to 0. To deal with this dimensionality

problem, we follow Kalouptsidi (2018) and assume that the value function is sparse in the basis functions; in other words, only a subset of elements of the vector of coefficients γ^0 are different from zero. We then use LASSO to identify which of these coefficients are non-zero.

Our approach builds on Barwick and Pathak (2015) who suggest using basis functions that fit the payoff function π_{jt}: these will preserve the shape of the payoff function, and hence are good candidates for approximating the value function. In addition, we also include basis functions that fit the observed level of investment, since this regression yields additional information about the derivative of V with respect to K. We use the following LASSO-based procedure that takes advantage of information contained in both profits and investment:

1. Run a LASSO regression of profits on basis functions of states. Let $U_1(s)$ denote the set of basis functions with non-zero coefficients.

2. Run a LASSO regression of investment on basis functions of states. Let $U_2(s)$ denote the set of basis functions with non-zero coefficients.

   From the investment first-order condition (Equation (14)), we know investment will depend on $U_2(s)$ only if $\frac{\partial E(V(s'|s,i))}{\partial i}$ depends on $U_2(s)$. This in turn implies that the basis functions formed by interacting the basis functions in $U_2(s)$ with capital K are relevant for explaining V. Let $U_{2,K}(s)$ denote the resulting set of basis functions.

3. Then the set of basis functions of s that we use to approximate the value function is given by the union of $U_1(s)$ and $U_{2,K}(s)$, or $U_1(s) \cup U_{2,K}(s)$.

Note that it is important to include Step 2, and include terms that explain investment. Omitting Step 2 risks throwing away basis functions that are important for capturing the shape of V in K. In our empirical analysis, we have found that if we omit Step 2, and rely only on the LASSO regression of profits on states to determine which polynomial terms to include, the fit for investment becomes significantly worse.

The number of basis functions selected depends on the LASSO penalty parameter: the larger the penalty, the fewer the number of basis functions selected in both Steps 1 and 2 of the above procedure. We carry out robustness analyses to look at how the results depend on the penalty parameter. In addition, we also provide robustness checks where we drop the assumption of sparsity and instead include every basis function when approximating the value function.

B.3.2 Regression sample when solving Bellman equation

Solving the value function for any guess of θ requires solving for the γ^0 coefficients via a least squares regression, as described in the first part of Section 4.3.3. When doing this, we have to

choose the sample space on which we run the regression model. Because many of the states are continuous, and because discretization is impractical given the number of state variables we work with, the regression sample cannot include every possible realization of the state variables, and must necessarily be a subset of the full state space.

We construct the sample in a way that ensures there is sufficient variation in each of the state variables, so that each of the γ^0 coefficients is well-identified. First, we include every state observed in the data in the sample; this gives us N observations. Second, we randomly draw a large number of additional state vectors that we do not see in the data, or N_{add}. These states are drawn so as to span the full range of the state variables. We then solve the Bellman equation on the sample of $N + N_{add}$ states formed by combining the observed states with the additional randomly drawn sample of states. Our approach of combining observed states with additional randomly drawn states is similar to that of Sweeting (2013), who also uses a parametric approximation of the value function and solves the approximation on a set of states that includes both the observed states and duplicates of these states with perturbations of the states that can vary over time.

We have found that adding the unobserved states is important for getting a good approximation of the value function. There are two issues with constructing the sample space based only on the observed set of states. Firstly, some of the state variables may be highly correlated with one another: in our setting, the prices of different ship types are very correlated with one another and with variables such as capital stock and backlog. This can make it challenging to separately identify the coefficients on basis functions formed from these state variables, if we use only the observed set of states. Secondly, there may be only a limited number of observations in some regions of the state space. This can make it difficult to identify the coefficients on basis functions that mostly span those particular regions of the state space. For example, there are only a few observations with very large capital stocks which makes it difficult to identify the coefficients on some of the B-spline basis functions that take positive values only for large values of capital. Both of these problems can be mitigated by the solution of randomly drawing additional states, since that both kills the multicollinearity between states and ensures that there is sufficient data across all regions of the state space.

B.3.3 Deriving continuation values and estimating approximating parameters γ^0

In this section, we describe how we derive the continuation value $CV(s_{jt};\theta)$ given a particular guess of the parameters θ. We then describe how to estimate the approximating parameters γ^0 via a least squares procedure. Our discussion here elaborates on the discussion on how to solve the value function in Section 4.3.3.

As discussed in the text, we approximate $V(s_{jt})$ via polynomials/spline basis functions as fol-

lows:

$$V(s_{jt}) = \sum_{l=1}^{L} \gamma_l^0 u_l(s_{jt})$$

How do we compute $EV(s_{jt+1}|s_{jt}, i(s_{jt}, \nu_{jt}))$? Note that:

$$\begin{aligned} EV(s_{jt+1}|s_{jt}, i(s_{jt}, \nu_{jt})) &= E\big[\sum_{l=1}^{L} \gamma_l^0 u_l(s_{jt+1})|s_{jt}, i(s_{jt}, \nu_{jt})\big] \\ &= \sum_{l=1}^{L} \gamma_l^0 E\big(u_l(s_{jt+1})|s_{jt}, i(s_{jt}, \nu_{jt})\big) \end{aligned}$$

As such, $EV(s_{jt+1}|s_{jt}, i(s_{jt}, \nu_{jt}))$ is a linear function of the L terms $E\big(u_l(s_{jt+1})|s_{jt}, i(s_{jt}, \nu_{jt})\big)$, which are essentially expectations of the basis functions conditional on today's states and investment levels. How do we construct these expectations? Observe that s_{jt+1} includes both yard-level states s^y_{jt+1}, and aggregate states s_{t+1}: that is, $s_{jt+1} = (s^y_{jt+1}, s_{t+1})$. When choosing investment in period t, yards face uncertainty only about the aggregate state s_{t+1}. The investment cost shock ν_{jt} is known at the point when the shipyard chooses investment. Other than the aggregate states, every other term of s_{jt+1} (such as backlog, capital etc.) is a deterministic function of s_{jt} and i, and investment i is itself a function of the states and the cost shock ν_{jt}. The aggregate states s_{t+1} evolve via the following AR(1) process:

$$s_{t+1} = R_0 + Rs_t + \xi_t$$

where $\xi_t \sim N(0, \Sigma_\xi)$. We can use this AR(1) process to construct expectations of various basis functions of the states. The form of this expectation depends on the relationship between the basis function and the aggregate state. Some examples are provided below:

- Suppose the basis function $u_l(.)$ does not include any aggregate states. In this case $Eu_l(.) = u_l(.)$
- Suppose that $u_l(.)$ includes a single aggregate state (e.g. bulk price). In that case it can be written as $u_l(s_{jt}) = g(s^y_{jt})s^k_t$ for some $k \in \{1, ..., R\}$ where R denotes the number of aggregate states. Then $E\big(u_l(s_{jt+1})|s_{jt}, i\big) = g(s^y_{jt+1})E(s^k_{t+1}|s_t)$, where $E(s^k_{t+1}|s_t) = R^k_o + R^k s_t$ and R^k_0 and R^k are the k-th rows of the matrices R_0 and R respectively.
- Suppose that $u_l(.)$ includes two aggregate states, k and n. Then it can be written as $u_l(s_{jt}) = g(s^y_{jt})s^k_t s^n_t$. Therefore $E\big(u_l(s_{jt+1})|s_{jt}, i\big) = g(s^y_{jt+1})E(s^k_{t+1}s^n_{t+1}|s_t)$, where:

$$\begin{aligned} E(s^k_{t+1}s^n_{t+1}|s_t) &= (R^k_o + R^k s_t)(R^n_0 + R^n s_t) + E(\xi^k_t \xi^n_t) \\ &= (R^k_o + R^k s_t)(R^n_0 + R^n s_t) + (\Sigma_\xi)_{k,n} \end{aligned}$$

Here R_0^k, R_0^n and R^k, R^n are the k-th and n-th rows of the matrices R_0 and R respectively, and $(\Sigma_\xi)_{k,n}$ is the (k,n) entry of the matrix Σ_ξ.

Once we have constructed $EV(s_{jt+1}|s_{jt}, i(s_{jt}, \nu_{jt}))$, we can evaluate the firm's continuation value. Unlike EV, which is evaluated after observing the cost shock ν_{jt}, the continuation value must be evaluated *before* the yard sees ν_{jt}, and as such we need to integrate out ν_{jt}. The continuation value is given by:

$$\begin{aligned}
CV(s_{jt};\theta) = E_{\nu_{jt}}\Bigg(& -c_1 i(s_{jt},\nu_{jt}) - c_2 i(s_{jt},\nu_{jt})^2 - c_3 \nu_{jt} i(s_{jt},\nu_{jt}) \\
& - c_4 T_{jt} i(s_{jt},\nu_{jt}) + \beta E\left[V(s_{jt+1})|s_{jt}, i(s_{jt},\nu_{jt});\theta\right]\Bigg) \\
= & -c_1 E_{\nu_{jt}} i(s_{jt},\nu_{jt}) - c_2 E_{\nu_{jt}} i(s_{jt},\nu_{jt})^2 - c_3 E_{\nu_{jt}} \nu_{jt} i(s_{jt},\nu_{jt}) \\
& - c_4 T_{jt} E_{\nu_{jt}} i(s_{jt},\nu_{jt}) + \beta \sum_{l=1}^{L} \gamma_l^0 E_{\nu_{jt}}\Bigg(\gamma_l^0 E\left(u_l(s_{jt+1})|s_{jt}, i(s_{jt},\nu_{jt})\right)\Bigg)
\end{aligned}$$

Notice that apart from the unknown parameters c_1, c_2, c_3, c_4 and γ_l^0, all the terms that determine $CV(s_{jt};\theta)$ can be pre-computed prior to estimation, using the estimated investment policy function. $CV(s_{jt};\theta)$ is a linear function of the unknown parameters.

Finally we discuss how we can estimate the approximating parameters γ^0 by a least squares procedure. The Bellman equation implies that:

$$V(s_{jt};\theta) = \pi(s_{jt}) + p^x\sigma + CV(s_{jt};\theta)$$

Approximating $V(s_{jt};\theta)$ by basis functions, and letting ξ_{jt} denote approximation error, we get the following equation:

$$\begin{aligned}
\xi_{jt} &= \tilde{V}(s_{jt};\theta) - \pi(s_{jt}) - p^x\sigma - \tilde{CV}(s_{jt};\theta) \\
&= \sum_{l=1}^{L} \gamma_l^0 u_l(s_{jt}) - \pi(s_{jt}) - p^x\sigma + c_1 E_{\nu_{jt}} i(s_{jt},\nu_{jt}) + c_2 E_{\nu_{jt}} i(s_{jt},\nu_{jt})^2 + c_3 E_{\nu_{jt}} \nu_{jt} i(s_{jt},\nu_{jt}) \\
&+ c_4 T_{jt} E_{\nu_{jt}} i(s_{jt},\nu_{jt}) - \beta \sum_{l=1}^{L} \gamma_l^0 E_{\nu_{jt}}\Bigg(E\left(u_l(s_{jt+1})|s_{jt}, i(s_{jt},\nu_{jt})\right)\Bigg) \\
&= Y_{jt} - \sum_{l=1}^{L} \gamma_l^0 X_{l,jt}
\end{aligned}$$

where:

$$\begin{aligned}
Y_{jt} &= -\pi(s_{jt}) - p^x\sigma + c_1 E_{\nu_{jt}} i(s_{jt},\nu_{jt}) + c_2 E_{\nu_{jt}} i(s_{jt},\nu_{jt})^2 + c_3 E_{\nu_{jt}} \nu_{jt} i(s_{jt},\nu_{jt}) + c_4 T_{jt} E_{\nu_{jt}} i(s_{jt},\nu_{jt}) \\
X_{l,jt} &= -u_l(s_{jt}) + \beta E_{\nu_{jt}}\Bigg(E\left(u_l(s_{jt+1})|s_{jt}, i(s_{jt},\nu_{jt})\right)\Bigg)
\end{aligned}$$

Thus, for any guess of θ, we can recover γ^0 from the linear regression of Y_{jt} on $X_{l,jt}$.

B.4 Constructing the likelihood function for investment

This section describes how we can construct both the numerator and denominator of the investment density function in (11) (see Section 4.3.3). Recall that we can write the density function of i_{jt} as follows:

$$g(i_{jt};\theta) = \frac{f_\nu(\nu_{jt})}{|i'(\nu_{jt})|}$$

where $f_\nu(\nu_{jt})$ is the density function of the cost shock ν_{jt} and $|i'(\nu_{jt})|$ is the absolute value of the derivative of investment with respect to ν_{jt}.

The FOC, which is satisfied at any i_{jt}, is:

$$\beta \frac{\partial E\left[V(s_{jt+1}|s_{jt},i_{jt}\right]}{\partial i} = c_1 + c_3\nu_{jt} + c_4 T_t + 2c_2 i_{jt} \tag{14}$$

We can solve for ν_{jt} in terms of i_{jt} and the states and parameters as follows:

$$\nu_{jt} = \frac{1}{c_3}(\beta \frac{\partial E\left[V(s_{jt+1}|s_{jt},i_{jt}\right]}{\partial i} - c_1 - c_4 T_t - 2c_2 i_{jt})$$

We can then use the above equation to form the numerator for equation (11).

Next, we need the denominator, $|i'(\nu_{jt})|$. Differentiating ν_{jt} with respect to i_{jt}:

$$\frac{d\nu_{jt}}{di_{jt}} = \frac{1}{c_3}(\beta \frac{\partial^2 E\left[V(s_{jt+1}|s_{jt},i_{jt}\right]}{\partial i^2} - 2c_2)$$

Then, we have:

$$\begin{aligned} |i'(\nu_{jt})| &= |\frac{di_{jt}}{d\nu_{jt}}| \\ &= \frac{1}{|\frac{d\nu_{jt}}{di_{jt}}|} \end{aligned}$$

So we are able to solve for both the numerator and denominator of (11) without analytically solving for investment. All we need is that the value function be twice differentiable in capital, which we obtain from the fact that we are approximating the value function using polynomial/spline basis functions, and that investment be strictly monotonic in the investment cost shock.

C Counter-factual routine

The counter-factual routine involves two steps (i) solve the new Bellman equation and policy functions (ii) simulate the industry under the counter-factual assumptions.

C.1 Solving for the policy functions and Bellman equation

We first describe how we solve for the policy function and value function coefficients, given a particular set of parameters θ. The broad idea is to start with a guess of the policy function, update the value function and then update the policy function, until the policy function converges. We approximate $V(.)$ with a sparse polynomial approximation $V(s) = \sum_l \gamma_l f_l(s)$.

When solving the Bellman equation, we do it for the states we observe in the data as well as a set of additional states that we randomly draw. The procedure for drawing these additional states was discussed in Appendix B.3.2.

The specific routine is:

1. Compute all objects that do not require knowledge of the value function parameters:
   - Expected profits $\pi(s)$ at each state. In this counter-factual, we compute expected profits under the assumption that firms do not receive any production subsidies; thus, after 2006, expected profits are lower in this counterfactual than in the benchmark scenario where firms do receive production subsidies.
   - Use the VAR processes for ship and steel prices and demand shifters to compute polynomials of expected future ω, steel prices/demand shifters.
2. Start with an initial guess of the three policy functions.
3. Update the guess of the policy functions. At each iteration j, starting with a guess of policy functions $p^j(s)$, we will update the guesses as follows:
   - Update the value function parameter vector to γ_l^{j+1} using the equation $V^{j+1}(s) = \pi(s) + p^{j,x}\sigma + CV^{j+1}(s)$.
     - When doing this, we have to compute $EV^{j+1}(s'|s,i)$ for every possible i, using the jth iteration of the investment policy function.
   - Using the new guesses of V^{j+1} and CV^{j+1}, update the investment policy function to $\hat{i}^{j+1}$ by solving the investment FOC. Because we assume that the value function can be approximated by B-splines of order at most 3 (i.e. quadratic piecewise polynomials), we can solve analytically for the investment policy function.
   - Update the exit policy function to $p^{j+1,x}$ using the new guesses of V^{j+1} and CV^{j+1}.

- Check if the objective function $||p^{j+1}(s) - p^j(s)|| < tol$, where tol is a pre-assigned tolerance level.

C.2 Simulate the model

To simulate the model, given that we already have $V(s)$, $EV(s'|s,i$, $CV(s)$ and the policy functions, we do the following:

- Start with an initial set of yards and an initial set of states in Q1 2004.
- Draw steel prices from the AR(1) process.
- Draw ω for each incumbent yard from the estimated distribution.
- Solve for equilibrium ship prices P by equating demand and supply. Each incumbent then chooses orders q taking the price P as given.
- Draw investment cost shocks ν and exit scrap value shocks for incumbents. Each incumbent now decides first whether or not to exit (by comparing $CV(s)$ and the scrap value), and second (conditional on staying) how much to invest (by maximizing $\beta EV(s'|s,i) - c(i,\nu)$).
- Draw entry cost shocks for each potential entrant. Each potential entrant now decides whether or not to enter, by comparing the entry cost draw with the value of entry. The value of entry differs across regions, but is the same within the region.
- Assign an initial state to each entrant.
  - The initial capital stock, the size (large or small) and whether or not the firm produces bulk/tanker/container (or a combination of these) are all drawn from the empirical distribution $F(region, policy)$ of these states that we observe for entering yards. The empirical distribution is estimated for each region separately for the pre-2006, 2006-08 and post-2009 periods.
- Implement entry, exit and investment.
- Update all states, and move on to the next period.