

# Time to Build and Fluctuations in Bulk Shipping

## Online Appendix

Myrto Kalouptsi\*<sup>\*</sup>

Department of Economics, Princeton University

December 2012

### APPENDIX A

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

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

To construct the backlog vector, we take the following steps. We use the orderbook dataset to identify a quarter's entrants and compute the average time to build offered to these entrants. We then use the timeseries of entrants and total backlog, as well as the estimated time to build function (using the observed total backlog) and create the backlog vector for each quarter of our sample,  $\mathbf{b}_t \in \mathbb{R}^T$ .

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

---

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

## APPENDIX B

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

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

where the expression above results from the fact that  $Z_t$  follows the Binomial distribution. We find  $\mu_\phi = -7.4823$  and  $\sigma_\phi = 5.6096$ . Figure 1 compares the Poisson exit rate to the exit rate under the assumption of a normal scrap value distribution.

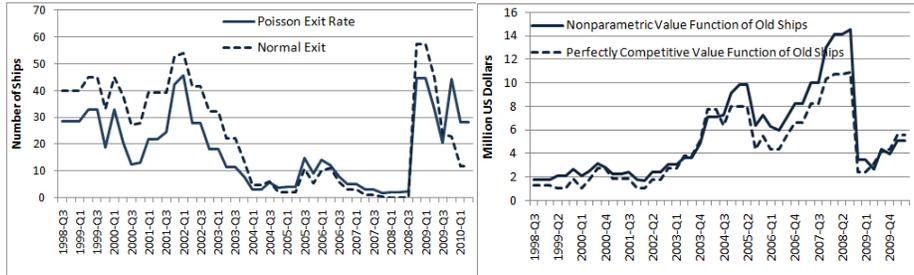


Figure 1: Poisson exit rate and exit rate under normal scrap value density. Value function of old ships estimated nonparametrically and computed from perfectly competitive profits.

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

$$\mathbf{V}_{80} = \boldsymbol{\pi}_{20+} + \beta (1 - \tilde{\boldsymbol{\mu}}) \mathbf{P}\mathbf{V}_{80} + \beta \tilde{\boldsymbol{\mu}} E[\phi | \phi > \mathbf{P}\mathbf{V}_{80}]$$

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

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

Figure 1 compares the nonparametrically estimated value function of old ships to the one

that solves the above fixed point and shows that the two are close, as suggested by the good fit of the perfectly competitive behavior in old ships.

Once  $\mathbf{V}_{80}$  is known, we recover all remaining value functions sequentially from:

$$\begin{aligned}
\mathbf{V}_{79} &= \boldsymbol{\pi}_{10-20} + \beta \mathbf{P} \mathbf{V}_{80} \\
\mathbf{V}_{78} &= \boldsymbol{\pi}_{10-20} + \beta \mathbf{P} \mathbf{V}_{79} \\
&\vdots \\
\mathbf{V}_{39} &= \boldsymbol{\pi}_{0-10} + \beta \mathbf{P} \mathbf{V}_{40} \\
&\vdots \\
\mathbf{V}_0 &= \boldsymbol{\pi}_{0-10} + \beta \mathbf{P} \mathbf{V}_1
\end{aligned}$$

Finally, using the retrieved  $\mathbf{V}_0$  the entry costs that correspond to the perfectly competitive environment are given by (14) of the paper.<sup>1</sup>

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

1. Update exit  $\boldsymbol{\mu}^* = s^{80} (1 - F_\phi [\mathbf{P}_{\boldsymbol{\lambda}_{n-1}, \boldsymbol{\mu}_{n-1}} \mathbf{V}_{80, n-1}])$
2. Compute the new transition matrix  $\mathbf{P}_{\boldsymbol{\lambda}_{n-1}, \boldsymbol{\mu}^*}$  as described in Section 5.1.5.

---

<sup>1</sup>As discussed in Section 6.1, the estimated perfectly competitive profits are less variable than the nonparametric profits, especially in the 10 to 20 years old ships. As a consequence, the entry costs resulting from parametric profits are considerably smoother than the nonparametric entry costs, which compensates for keeping the same entry costs in all counterfactual worlds. Indeed, entry costs are a function of the state (in contrast to current practice where entry costs are iid draws from some distribution) and are taken as a primitive in counterfactuals. The fact that they exhibit low variability, in some sense deals with the difficult question of how entry costs would be different in a constant or no time to build world.

3. Update the value functions:

$$\begin{aligned}
\mathbf{V}_{80,n} &= \boldsymbol{\pi}_{20+} + \beta(1 - \boldsymbol{\mu}^*) \mathbf{P}_{\lambda_{n-1}, \boldsymbol{\mu}^*} \mathbf{V}_{80,n-1} + \beta \boldsymbol{\mu}^* E[\phi | \phi > \mathbf{P}_{\lambda_{n-1}, \boldsymbol{\mu}^*} \mathbf{V}_{80,n-1}] \\
\mathbf{V}_{79,n} &= \boldsymbol{\pi}_{10-20} + \beta \mathbf{P}_{\lambda_{n-1}, \boldsymbol{\mu}^*} \mathbf{V}_{80,n} \\
&\vdots \\
\mathbf{V}_{0,n} &= \boldsymbol{\pi}_{0-10} + \beta \mathbf{P}_{\lambda_{n-1}, \boldsymbol{\mu}^*} \mathbf{V}_{1,n}
\end{aligned}$$

4. Update entry  $\lambda^*(\mathbf{X}) = \lambda_{n-1}(\mathbf{X}) \left( \frac{\beta^{T(\mathbf{X})} (\mathbf{P}_{\lambda_{n-1}, \boldsymbol{\mu}^*, X})^{T(\mathbf{X})} \mathbf{V}_0}{\kappa(\mathbf{X})} \right)^{1/2}$ , all  $X$ , where  $\mathbf{P}_{\lambda_{n-1}, \boldsymbol{\mu}^*, X}$  is the  $X^{th}$  row of  $\mathbf{P}_{\lambda_{n-1}, \boldsymbol{\mu}^*}$ .

5. Set  $\boldsymbol{\lambda}_n = \alpha_n \boldsymbol{\lambda}^* + (1 - \alpha_n) \boldsymbol{\lambda}_{n-1}$  and  $\boldsymbol{\mu}_n = \gamma_n \boldsymbol{\mu}^* + (1 - \gamma_n) \boldsymbol{\mu}_{n-1}$  where  $\alpha_n$  and  $\gamma_n$  go to zero as  $n$  increases.

## APPENDIX C

In this Appendix we explore the robustness of our results with respect to the state aggregation, the estimator of the value function, the demand curve specification and the first stage policy functions.

### State Aggregation and Value Function Estimation

We now examine how robust our estimation results are to the aggregation employed in Sections 5.1.3 and 5.1.5, by reestimating the model using a different aggregation; instead of the baseline aggregation:

$$\{S_t^1, S_t^2, S_t^3, B_t, d_t\} = \left\{ \sum_{i=0}^{39} s_t^i, \sum_{i=0}^{79} s_t^i, s_t^{80}, \sum_{i=0}^{\bar{T}} b_t^i, d_t \right\}$$

we now use

$$\{S_t^{0-79}, MS_t, VS_t, S_t^{80}, B_t, d_t\}$$

where

$$S_t^{0-79} = \sum_{i=0}^{79} s_t^i, \text{ total fleet younger than 20 yrs old}$$

$$MS_t = \text{mean age of fleet younger than 20 yrs old}$$

$$VS_t = \text{standard deviation of age of fleet younger than 20 yrs old}$$

$$S_t^{80} = s_t^{80} = \text{fleet older than 20 yrs old}$$

In other words, instead of three age groups, we use the fleet (younger and older than 20 years old), as well as the mean and standard deviation of the age of the fleet. We distinguish between ships younger than 20 years and ships older than 20 years, in order to avoid bias in the mean and standard deviation of fleet age, resulting from assigning the same age to all ships in the “20 plus” category.

We follow precisely the same estimation procedure on the new aggregated states and compare the estimated primitives, i.e. profits, entry cost and scrap value distribution. The demand curve, as well as the transition of the demand state variable  $d_t$  are the same under both aggregations. We reestimate the time to build function and the entry and exit policy functions. The value function is estimated via local linear regression on the new state space (the same clustering techniques are employed on the new aggregated states). The transition matrix is calculated following the same procedure and is then combined with the value functions to produce the primitives of interest, as in the baseline state aggregation. We next present and compare the results of our estimation procedure performed on the new aggregation to those obtained in the baseline aggregation.

	constant	$S^{0-79}$	$MS$	$VS$	$S^{80}$	$B$	$d$
Param	3.01	-0.00076	0.0053	-0.0187	$-8.66e - 005$	$4.1e - 005$	0.0305
s.e.	(1.335)*	(0.000386)	(0.0101)	(0.0114)	(0.00041)	( $9.9e - 005$ )	(0.0219)

Table 1: Time to Build regression estimates. Standard errors based on 500 bootstrap samples. Single star indicates significance at the 0.1 level.

<b>Entry</b>						
	constant	$S^{0-79}$	$MS$	$VS$	$S^{80}$	$d$
Param	7.27	-0.0017	-0.026	-0.303	-0.0037	0.81
s.e.	(8.36)	(0.00136)	(0.041)	(0.105)**	(0.0027)	(0.216)**
<b>Exit</b>						
Param	21.59	0.00112			0.00117	-1.52
s.e.	(6.44)**	(0.0013)			(0.0012)	(0.304)**

Table 2: Entry and exit regression estimates. Standard errors based on 500 bootstrap samples. Double star indicates significance at the 0.05 level.

Tables 1 and 2 depict the estimated time to build function, as well as the entry and exit policy functions (specifications chosen based on the same criteria as in the baseline case). Figure 2 compares the profits under the baseline aggregation and the current aggregation.

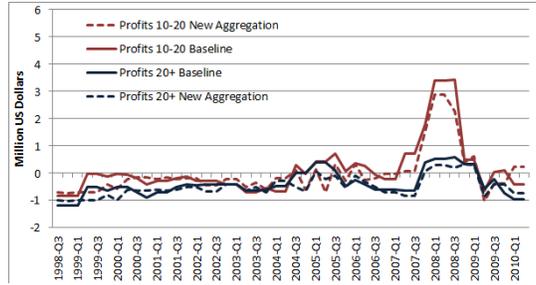


Figure 2: Profits under baseline and new aggregation.

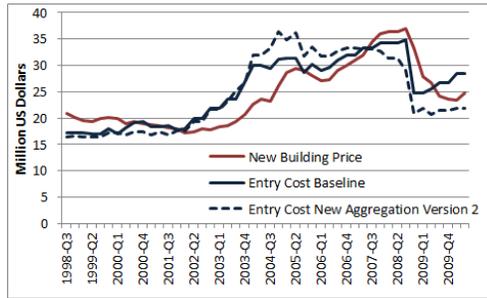


Figure 3: Entry costs under baseline and new aggregation.

Figures 3 and 4 compare the entry costs and scrap value distribution respectively. The two aggregation schemes lead to similar results. Some differences may arise from different bandwidths in the local linear regression, as this aggregation entails six variables, while our baseline consists of five. In addition, it is difficult to compare parametric profits, as the assumptions on the freight market (namely cost function and equilibrium behavior) that lead to the current state consisting of the fleet and moments of the age distribution are not straightforward.

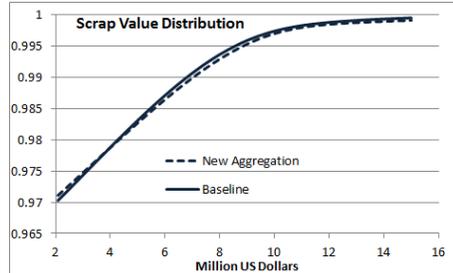


Figure 4: Scrap value distribution under baseline and new aggregation.

We next turn to the value function estimation procedure and compare our baseline procedure based on local linear regression with sieve estimation via both OLS, as well as LASSO (Tibshirani (1996)). In particular we approximate the value function using polynomials of the aggregate states  $\mathbf{X}_t = [S_t^1, S_t^2, S_t^3, B_t, d_t] \in \mathbb{R}^5$  of varying degrees. Clearly, estimation cannot be performed on the initial 94 state variables as complexity is prohibitive. Some aggregation is, therefore, necessary. In addition to the baseline aggregation, we try out a finer one- in particular, we experiment with the LASSO method and the age distribution in years (21 terms), the total backlog based on years away from delivery (3 terms) and the demand state variable. Note that this aggregation of 27 terms would not be possible with local linear regression, but is possible with the LASSO method.

Figure 5 compares the estimated value function for three ages (0, 10 and older than 20 years old) resulting from the baseline estimation method, as well as the ones obtained by 3<sup>rd</sup>, 4<sup>th</sup>, 5<sup>th</sup> and 6<sup>th</sup> degree polynomial approximation. The resulting approximations are obtained by the Post-LASSO method.<sup>2</sup> The LASSO and Post-LASSO methods are appropriate for sparse regression problems, i.e. when there is a large number of potential regressors but in reality only a small subset of them is important in capturing the regression function accurately. LASSO identifies the relevant regressors by performing a modified OLS procedure which penalizes a large number of nonzero coefficients through regularization by a penalty based on the  $\mathcal{L}_1$  norm. Post-LASSO simply performs OLS on the chosen coefficients. In our application, the regressors are products of monomials of the states in  $[j, \mathbf{X}]$ . We found the Post-LASSO to significantly improve LASSO results. In principle, adjusting the bandwidth in the local linear regression and the “lambda” parameter in LASSO, which determines the trade-off between estimation error and sparsity of parameters (as lambda increases, fewer non-zero coefficients remain), the two value functions can become arbitrarily close. Interestingly, though, in the LASSO results depicted

<sup>2</sup>For details, see for instance Belloni and Chernozhukov (2011).

here  $\lambda$  is chosen via cross validation.<sup>3</sup> This is encouraging for both the results in this Appendix, as well as our baseline results in the main text.

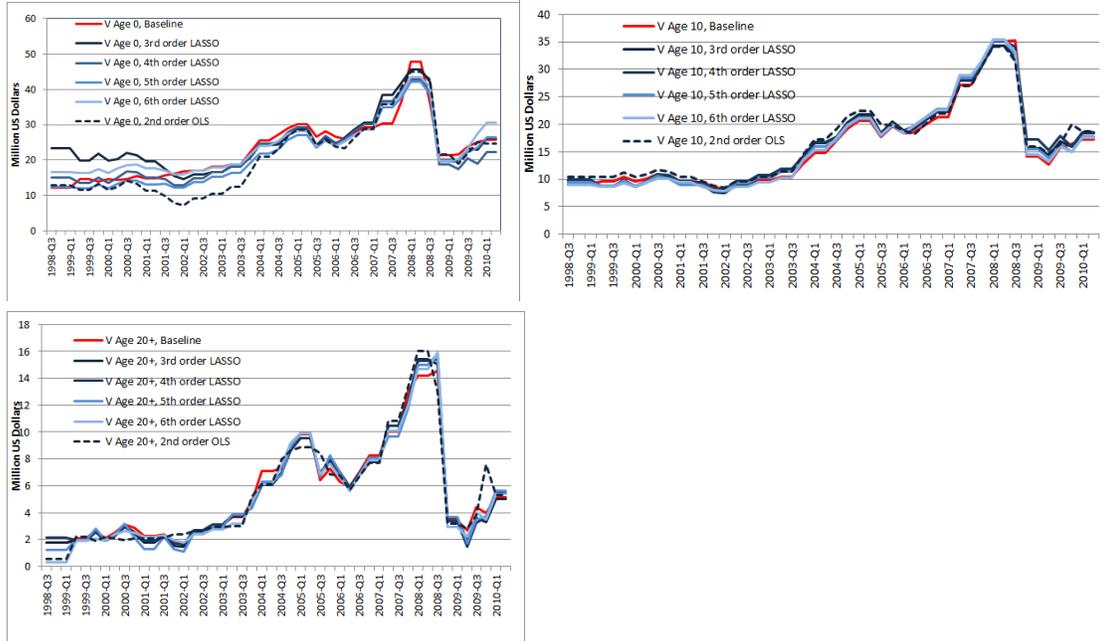


Figure 5: Obtained value function using local linear regression (baseline) and sieve approximation via Post-LASSO of several orders and OLS.

The  $3^{rd}$  degree polynomial approximation involves 84 terms, while LASSO (at the  $\lambda$  chosen via cross validation) retains 20. The only surviving linear term is the individual age, while all other remaining terms are  $3^{rd}$  order. The  $4^{th}$  order polynomial approximation involves 210 terms, while LASSO retains 16, all  $4^{th}$  order. The  $5^{th}$  order polynomial approximation involves 462 terms, of which LASSO retains 17, all  $5^{th}$  order except for one  $4^{th}$  order term. The Post-LASSO r-square is 0.81 in all these specifications. The  $6^{th}$  order polynomial approximation involves 924 terms, of which LASSO retains 25, all  $6^{th}$  order. The Post-LASSO r-square is 0.82. Value functions differ more for young ships, for which fewer observations are available (again, if we don't use cross-validation but choose  $\lambda$  manually, these differences can in theory vanish). In contrast, value functions for older ships are more robust to the estimation method, as well as the number of approximating terms. Figure 5 also depicts a  $2^{nd}$  degree polynomial approximation estimated via OLS, rather than LASSO. This does somewhat worse, as it doesn't allow for

<sup>3</sup>We also experimented with the rules for  $\lambda$  found in Belloni and Chernozhukov (2011), which are adjustable by the user. No significant changes were observed.

higher powers. The r-square is still 0.81 and 23 out of 28 terms are statistically significant.

To further assess our aggregation and discretization, we estimate the value function on a more finely aggregated state via the LASSO method. In particular, we use the yearly age distribution consisting of 21 terms (number of ships 0 years old, 1 year old, ..., older than 20 years old), the yearly backlog (number of ships to be delivered in 1, 2 or 3 and up years) and demand. We employ Post-LASSO on linear and  $2^{nd}$  order terms; 378 terms in total. It is impossible to estimate such a model with local linear regression or OLS. Figure 6 compares the obtained value functions to our baseline specification. Lambda is chosen via cross-validation. Note that further differences between the value functions arise due to state inconsistencies- the baseline specification is depicted on grid states (i.e. the states from the grid that are closest to the observed states), while the new value functions are depicted on actual data (no discretization has taken place here). Out of 378 terms, LASSO keeps 19, all  $2^{nd}$  order. The Post-LASSO r-square is 0.8.<sup>4</sup>

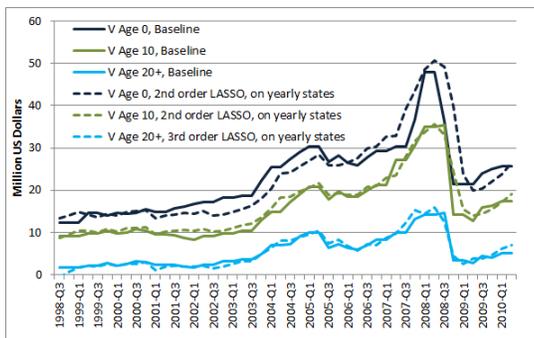


Figure 6: Obtained value function from yearly age distribution and backlog via LASSO.

### Demand Curve Specification

We next examine robustness of the estimation results under a linear (rather than log-linear) demand curve. The empirical analogue of (1) in the paper in this case becomes:

$$P_t = \eta_0 + \eta_1 \mathbf{H}_t + \eta Q_t + \varepsilon_t$$

where all variables are as in the main text. Table 3 presents the IV regression estimates for the first and second stage demand curve estimation. The implied price elasticity is somewhat lower than the log-linear case.

Table 4 reports the entry and exit policy functions using the new demand state variable, while Table 5 reports the parameters governing the transition of  $d_t$ ; under the linear

<sup>4</sup>We don't proceed to  $3^{rd}$  order, as it involves 3654 terms.

	<b>1<sup>st</sup> stage, dep var <math>Q_t</math></b>		<b>2<sup>nd</sup> stage, dep var <math>P_t</math></b>	
	param	s.e.	param	s.e.
const	-38873	(15636)**	$-1.2 \times 10^7$	$(3 \times 10^6)$ **
WIP	-20.1	(25.96)	137693	(29294)**
agr raw mat P	3.09	(5.71)	9753	(7178)*
mineral P	0.85	(2.36)	-11056	(3007)**
food P	-2.772	(3.44)	7242	(4210)**
China steel	5.32	(4.5)	9144	(6604)*
handymax	-1.48	(1.07)*	-2943	(737)**
fleet	1.79	(1.08)*		
mean age fl	326.4	(124.2)**		
std age fl	716.03	(306.1)**		
$\widehat{Q}_t$			-644.3	(565)

Table 3: Inverse demand curve for freight transport: IV regression results. Double star indicates significance at the 0.05 level, single star at the 0.1 level.

demand curve specification, the demand state variable is somewhat less persistent and exhibits lower variance. Finally, we turn to profits and entry costs in Figures 7 and 8.

<b>Entry</b>					
	constant	$S^1$	$S^2$	$S^3$	$d$
Param	-8.1083	-0.00215	-0.00127		0.973
s.e.	(2.38)	(0.0027)	(0.00078)		$(2e - 007)$ **
<b>Exit</b>					
Param	20.1	0.00858	0.002	0.00064	-1.81
s.e.	$(4.04)$ **	$(0.0029)$ **	(0.0012)	(0.001)	$(3e - 007)$ **

Table 4: Entry and exit regression estimates. Standard errors based on 500 bootstrap samples. Double star indicates significance at the 0.05 level.

	$c$	$\rho$	$a$	$\gamma$
Param	1.39	0.91	1.568	59.40
s.e.	(0.57)	(0.041)	(0.035)	(23.48)

Table 5: Demand transition parameters. Standard errors based on 500 bootstrap samples.

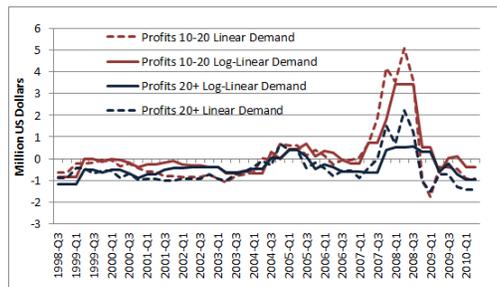


Figure 7: Profits under log-linear and linear demand.

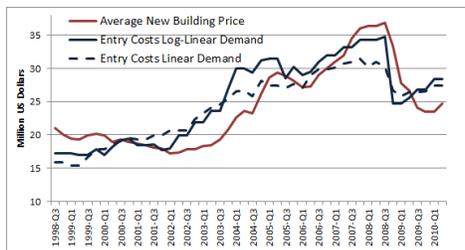


Figure 8: Entry costs under log-linear and linear demand.

## References

- [1] Belloni, Alexandre and Victor Chernozhukov. 2011. “High-Dimensional Sparse Econometric Models, an Introduction”. Springer Lecture Notes.
- [2] Tibshirani, Robert. 1996. “Regression Shrinkage and Selection Via the Lasso”. Journal of the Royal Statistical Society Series B. 58 (1): 267-288.