



NOAA
FISHERIES

Estimation of Commercial Fishing Trip Costs using Sea Sampling Data

Di Jin, Andrew Kitts, and Geret DePiper

International Institute of Fisheries Economics and Trade

Aberdeen, Scotland

July 14, 2016

Motivation

- Survey is stratified random sampling
 - Based on biological considerations
- Traditionally assumed to be MCAR
- Data actually either:
 - MAR
 - MNAR

Two Competing Models

- If Missing at Random
 - Must condition on stratified variables
 - Inverse Probability Weights
- If Missing Not at Random
 - Must fully control for selection
 - Heckman's incidental truncation

Data

- Observer data on costs
 - Coverage highly variable
 - High = 2014 NH Gillnet with 32.7% coverage
 - Low = 2010 MD Gillnet with 0.5% coverage
- Logbook data on vessel & trip characteristics
 - “Census” of Federally Permitted vessel trips

Data

- Cost is sum of Bait, Damage, Food, Fuel, Ice, Oil, Supply, & Water
- Independent Variables
 - Vessel Characteristics (Gross Tonnage)
 - Trip Duration (Days Absent)
 - Time (Year, Season)
 - State of Departure
 - Fuel Prices

Is there selection?

- Probability of observation suggests yes
 - Probit indicates significance across all three gear
- DuMouchel and Duncan (1983)

$$\text{Test } \hat{\beta}_{WLS} - \hat{\beta}_{OLS} = 0$$

$$\text{Test statistic } MS_W / \hat{\sigma}^2 \sim F(p, n - 2p)$$

Weighted estimates are significantly different

Source	df	Sum of Squares	Mean Square	F	Significance (p)
<i>Trawl</i>					
Regression	16	20,848.00	1,303.00	2,925.503	<.0001
Weights	17	85.00	5.00	11.226	<.0001
Error	8,781	3,911.00	0.45		
Total	8,814	24,844.00			
<i>Dredge</i>					
Regression	6	3,919.64	653.27	862.596	<.0001
Weights	7	14.12	2.02	2.664	0.010
Error	1,759	1,332.15	0.76		
Total	1,772	5,265.91			
<i>Gillnet</i>					
Regression	17	2,627.64	154.57	375.266	<.0001
Weights	18	21.85	1.21	2.947	<.0001
Error	7,083	2,917.39	0.41		
Total	7,118	5,566.89			

Weighted estimates are significantly different

Source	df	Sum of Squares	Mean Square	F	P-Value
<i>Trawl</i>					
Regression	16	20,848.00	1,303.00	2,925.503	<.0001
Weights	17	85.00	5.00	11.226	<.0001
Error	8,781	3,911.00	0.45		
Total	8,814	24,844.00			
<i>Dredge</i>					
Regression	6	3,919.64	653.27	862.596	<.0001
Weights	7	14.12	2.02	2.664	0.010
Error	1,759	1,332.15	0.76		
Total	1,772	5,265.91			
<i>Gillnet</i>					
Regression	17	2,627.64	154.57	375.266	<.0001
Weights	18	21.85	1.21	2.947	<.0001
Error	7,083	2,917.39	0.41		
Total	7,118	5,566.89			

Selection on Unobservables?

- Heckman Selection model

$$\ln(\text{cost}_i) = x_i' \beta + \rho \sigma \lambda(w_i' \gamma) + v_i$$

- If ρ is significant, bias exists

Trawl Model

Cost function

Variable	Coefficient	Std Error	p-value
Intercept	5.886	0.106	<.0001
Gross ton	0.875	0.033	<.0001
Days absent	0.729	0.008	<.0001
Days absent^2	-0.029	0.001	<.0001
Days absent*Gross ton	-0.064	0.006	<.0001
ln(fuel price)	1.063	0.086	<.0001
CT	0.247	0.061	<.0001
MA	-0.373	0.021	<.0001
MD	-0.870	0.069	<.0001
ME	-0.184	0.047	<.0001
NC	0.282	0.073	0.0001
NH	-0.400	0.056	<.0001
NY	0.443	0.037	<.0001
Summer	0.074	0.019	0.0001
Year 2010	-0.069	0.033	0.0407
Year 2011	-0.239	0.039	<.0001
Year 2012	-0.182	0.041	<.0001
Year 2013	-0.185	0.039	<.0001
Year 2014	-0.255	0.043	<.0001
σ	1.227	0.016	<.0001

Observation function (probit)

Variable	Coefficient	Std Error	p-value
Intercept	-1.861	0.026	<.0001
Gross ton	0.349	0.022	<.0001
Gross ton^2	-0.053	0.009	<.0001
CT	-0.279	0.039	<.0001
MA	0.315	0.024	<.0001
MD	0.312	0.045	<.0001
ME	0.674	0.033	<.0001
NC	-0.141	0.048	0.0036
NH	0.469	0.039	<.0001
NY	-0.404	0.022	<.0001
Spring	-0.046	0.012	0.0002
Summer	-0.174	0.015	<.0001
Fall	-0.042	0.012	0.0007
Year 2010	0.102	0.024	<.0001
Year 2011	0.223	0.025	<.0001
Year 2012	0.158	0.025	<.0001
Year 2013	0.210	0.024	<.0001
Year 2014	0.330	0.028	<.0001
Observation effort	0.204	0.030	<.0001
ρ	-0.947	0.002	<.0001

Root Mean Square Error

Gear	Observations	OLS	WOLS	Heckman
Trawl	8,815	4,212	4,521	4,064
Dredge	1,773	8,503	8,656	8,322
Gillnet	7,119	546	589	545

Root Mean Square Error

Gear	Observations	OLS	WOLS	Heckman
Trawl	8,815	4,212	4,521	4,064
Dredge	1,773	8,503	8,656	8,322
Gillnet	7,119	546	589	545

Difference in Estimated Costs

Gear	Ton Class	Model	Estimated Costs	Difference
Trawl	1	WOLS	512,233	-1.02%
		Heckman	507,078	
	2	WOLS	20,588,078	1.08%
		Heckman	20,811,991	
	3	WOLS	164,368,356	-0.96%
		Heckman	162,809,983	
	4	WOLS	117,775,062	-5.28%
		Heckman	111,871,533	
Dredge	1	WOLS	276,711	-3.30%
		Heckman	267,878	
	2	WOLS	12,460,785	-0.51%
		Heckman	12,397,436	
	3	WOLS	118,077,941	-1.58%
		Heckman	116,242,607	
	4	WOLS	134,031,558	-2.74%
		Heckman	130,456,695	
Gillnet	1	WOLS	2,716,228	-7.00%
		Heckman	2,538,512	
	2	WOLS	21,431,954	-3.41%
		Heckman	20,726,189	
	3	WOLS	2,239,613	2.59%
		Heckman	2,299,103	

Difference in Estimated Costs

Gear	Ton Class	Model	Estimated Costs	Difference
Trawl	1	WOLS	512,233	-1.02%
		Heckman	507,078	
	2	WOLS	20,588,078	1.08%
		Heckman	20,811,991	
	3	WOLS	164,368,356	-0.96%
		Heckman	162,809,983	
	4	WOLS	117,775,062	-5.28%
		Heckman	111,871,533	
Dredge	1	WOLS	276,711	-3.30%
		Heckman	267,878	
	2	WOLS	12,460,785	-0.51%
		Heckman	12,397,436	
	3	WOLS	118,077,941	-1.58%
		Heckman	116,242,607	
	4	WOLS	134,031,558	-2.74%
		Heckman	130,456,695	
Gillnet	1	WOLS	2,716,228	-7.00%
		Heckman	2,538,512	
	2	WOLS	21,431,954	-3.41%
		Heckman	20,726,189	
	3	WOLS	2,239,613	2.59%
		Heckman	2,299,103	

Questions?

This research was supported by the National Atmospheric and Oceanic Administration (NOAA) through the Cooperative Institute for the North Atlantic Region (CINAR) under NOAA Cooperative Agreement No. NA14OAR4320158.

Observer Coverage

Scallop Dredge

Calendar Year	Total Trips	Total Observed	Coverage
2009	4,298	111	2.6%
2010	9,017	333	3.7%
2011	9,902	377	3.8%
2012	9,357	406	4.3%
2013	8,548	457	5.3%
2014	4,064	188	4.6%

Observer Coverage

Gill Net

Calendar Year	Total Trips	Total Observed	Coverage
2009	11,384	313	2.7%
2010	16,977	1,986	11.7%
2011	17,822	2,226	12.5%
2012	15,623	1,663	10.6%
2013	11,076	653	5.9%
2014	4,859	386	7.9%

Observer Coverage

Bottom Trawl			
Calendar Year	Total Trips	Total Observed	Coverage
2009	13,318	759	5.7%
2010	22,289	1,763	7.9%
2011	20,756	2,170	10.5%
2012	19,604	1,831	9.3%
2013	17,732	1,591	9.0%
2014	7,151	799	11.2%

Dredge Model

<i>Cost function</i>				<i>Observation function (probit)</i>			
Variable	Coefficient	Std Error	p-value	Variable	Coefficient	Std Error	p-value
Intercept	9.517	0.277	<.0001	Intercept	-2.588	0.064	<.0001
Gross ton	1.714	0.194	<.0001	Gross ton	-0.216	0.087	0.0124
Gross ton ²	-0.253	0.094	0.007	Gross ton ²	0.108	0.038	0.0051
Days absent	0.275	0.026	<.0001	Days absent	0.183	0.011	<.0001
Days absent ²	-0.005	0.002	0.0046	Days absent ²	-0.008	0.001	<.0001
Days absent*Gross ton	-0.072	0.011	<.0001	Spring	0.088	0.038	0.0212
ln(fuel price)	0.665	0.191	0.0005	Summer	0.146	0.037	<.0001
Spring	-0.196	0.081	0.0156	Fall	0.199	0.041	<.0001
Summer	-0.295	0.079	0.0002	Year 2010	0.167	0.051	0.0011
Fall	-0.383	0.087	<.0001	Year 2011	0.229	0.050	<.0001
Year 2010	-0.416	0.110	0.0002	Year 2012	0.250	0.050	<.0001
Year 2011	-0.587	0.123	<.0001	Year 2013	0.418	0.050	<.0001
Year 2012	-0.503	0.125	<.0001	Year 2014	0.410	0.060	<.0001
Year 2013	-0.782	0.123	<.0001	ρ	-0.978	0.002	<.0001
Year 2014	-0.769	0.135	<.0001				
σ	1.827	0.047	<.0001				

Gillnet Model

Cost function

Variable	Coefficient	Std Error	p-value
Intercept	3.046	0.248	<.0001
Gross ton	0.994	0.073	<.0001
Days absent	0.861	0.025	<.0001
Days absent^2	-0.057	0.003	<.0001
ln(fuel price)	0.336	0.112	0.0028
MA	0.345	0.068	<.0001
MD	-0.387	0.104	0.0002
ME	0.810	0.075	<.0001
NH	0.381	0.074	<.0001
RI	0.170	0.049	0.0006
Summer	-0.124	0.019	<.0001
Fall	-0.046	0.022	0.0358
Year 2010	0.265	0.066	<.0001
Year 2011	0.440	0.078	<.0001
Year 2012	0.389	0.075	<.0001
Year 2013	0.355	0.069	<.0001
Year 2014	0.476	0.082	<.0001
σ	0.655	0.014	<.0001

Observation function (probit)

Variable	Coefficient	Std Error	p-value
Intercept	-2.753	0.052	<.0001
Gross ton	1.127	0.145	<.0001
Gross ton^2	-0.957	0.210	<.0001
Days absent	-0.072	0.022	0.0008
Days absent^2	0.007	0.003	0.0083
MA	0.741	0.047	<.0001
MD	-0.198	0.070	0.0046
ME	1.057	0.044	<.0001
NC	-0.143	0.057	0.0115
NH	1.090	0.038	<.0001
NJ	0.107	0.039	0.0059
NY	-0.156	0.049	0.0014
RI	0.086	0.043	0.0444
Spring	-0.065	0.023	0.0056
Summer	-0.055	0.020	0.0064
Fall	0.092	0.021	<.0001
Year 2010	0.657	0.036	<.0001
Year 2011	0.722	0.037	<.0001
Year 2012	0.674	0.035	<.0001
Year 2013	0.683	0.034	<.0001
Year 2014	0.962	0.040	<.0001
Observation effort	0.325	0.032	<.0001
ρ	0.222	0.106	0.0371