Race, Diversity, and Productivity

American Whaling in the Nineteenth Century

Michele Baggio and Metin Coşgel

Department of Economics
University of Connecticut

March 26, 2021

UCONN
UNIVERSITY OF CONNECTICUT

Overview

Main question

- How did racial diversity affect productivity in American whaling industry?
- Part of a larger project on whaling industry.

Data sources

- Archival records on the crew and the output of the voyages that span the 19th century.
- over 15,000 voyages and 113,000 crewmembers.

Results

- Very preliminary.
- Comments and suggestions welcome.

Why important

Whaling was one of the most important industries in the United States in the 19th century.
In the middle of the 19th c.,

- The American fleet was the largest in the world (about 70% of world activity).
- Whaling was the third biggest business in Massachusetts.
- A handful of New England ports (Nantucket, MA; New Bedford, MA; New London, CT) dominated the industry.

Whaling industry provided essential raw materials:

- Oil used for illumination and lubrication of high-speed machinery
- whalebone (baleen: strong but flexible plates of cartilage) used in the production of various consumer products (e.g., corsets, whips, umbrellas)

Highly diverse crew

"A Multi-Racial Enterprise During most of the history of American whaling, ships drew their crews from men of varied racial and ethnic backgrounds.

...

A kind of racial harmony: Genuine integration did not exist on most American whaleships, and violence sometimes flared. In general, however, men who were packed into tight quarters for years at a time and subject to the nearly unlimited power of the captain and officers, usually found it wise to tolerate each other." [whalingmuseum.org]

Yes, but did diversity affect productivity?

- Whaling voyages depended on collaborative teamwork.
- Did diversity affect the efficiency of the team?

Related literature

Historical literature on American whaling

- Data collection on voyages (Joseph Dias; Hegarty, 1959; Lund et al, 2008; Starbuck, 1878)
- General history (Busch, 1994; Creighton, 1995; Dolin, 2007; Hohman, 1928; Nordhoff, 1856; Tower, 1907)
- Race and diversity
  - Natives (Mancini, 2006, 2015; Munro, 2017; Shoemaker, 2015)
  - Blacks (Bolster, 1997; Farr, 1983; Putney, 1987)

Economic history literature on whaling

- Productivity (Davis, et al, 1997)
- Labor market
  - Wage discrimination (Craig and Fearn, 1993)
  - Whaling norms (Ellickson, 1989)
  - Organization, risk and incentives (Hilt, 2006, 2008; Hohman, 1926)

Economics literature on diversity and productivity

Prediction by theory is unclear, implying different directions and channels

- Benefits
  - human capital complementarity (Lazear 1998)
  - creation of new ideas and knowledge spillover (Berliant and Fujita 2008)
- Adverse effects of diversity on performance:
  - linguistic and cultural barriers in skill-transfer among workers (Lazear 1999)
  - weakening social ties and trust (Glaeser et al. 2000)

Empirical evidence for macro units (cities, regions, or countries) for long-run outcomes:

- civil conflicts (Montalvo and Reynal-Querol 2005, Desmet et al. 2017, Arbatlı et al. 2020)
- economic growth (Glaeser et al. 1995, Easterly and Levine 1997, Alesina et al. 2003, Ashraf and Galor 2013)
- public goods provision (Goldin and Katz 1999, Alesina et al. 1999, Desmet et al. 2017, Yang 2019)

Distinct features of the industry

Each voyage was treated as a separate venture

- Voyage is an ideal unit of analysis, virtually as an analogue of firm.
- With the competitive nature of whaling industry (especially during the heyday), we can use economic models based on profit maximization.

Occupational structure and payment system

- The rank of crew members was well organized.
- Crewman's earnings consisted of a specified fractional share, known as a lay, of the total net proceeds of a voyage.

Crew diversity

- Significantly higher proportions of blacks, natives, and foreign workers in whaling crews than in other industries.

Whaling data

Voyages

- Nearly all of the voyages made during the 19th century are included.
- Includes vessel characteristics (name, tonnage, rigging type), agent, captain, departure and return dates, amounts of catch (sperm oil, whale oil, baleen), hunting ground.
- Over 15,000 voyages (about 4,000 used in this analysis)

Crew list

- Required by law (in 1803), prepared for the collector of the customs at the beginning of the voyage, now available in the National Archives
- Includes crew member's first and last name, residence city, citizenship, age, height, skin, hair, eye color
- Updated as crew changed during voyage

Whaling data

Whalemen's shipping papers

- Also required by law (in 1790), filed with the collector of the customs at the beginning of the voyage, now available at the New Bedford Free Public Library
- Serving as contract between the owners and the crew in case of dispute, they list departing crew's service (rank) and payment in the form of share (lay) of output
- Updated as necessary.

Logbooks (not used here)

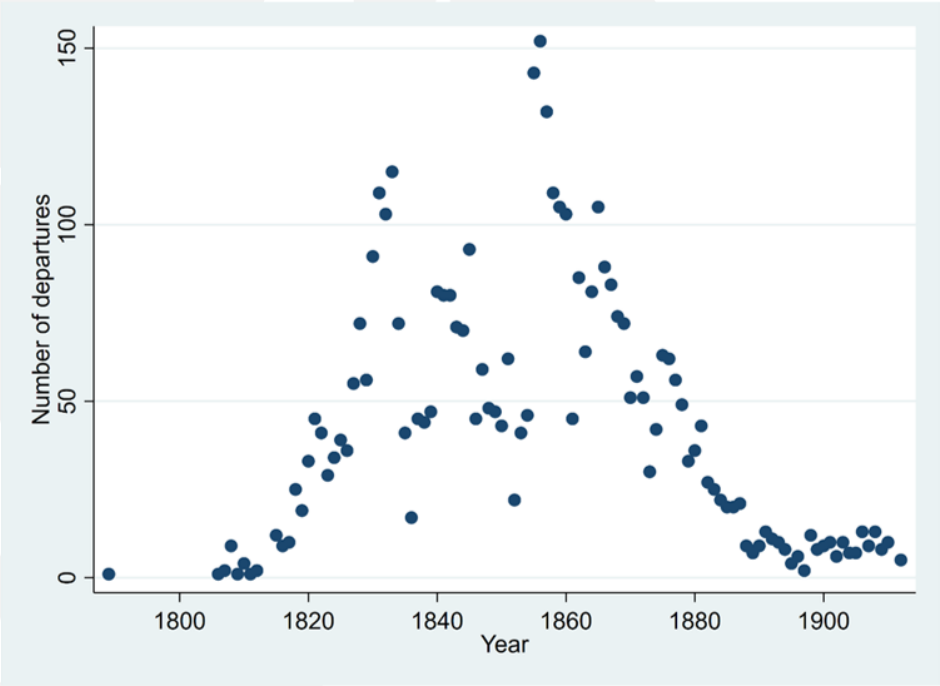
Data challenges

- Missing information on key variables (e.g., skin, rank)
- Spelling errors/omissions in names complicate the merging of records between crew lists and shipping papers
- Cannot fully detect the change of crew members during the voyage. Such changes were recorded as "remarks", but difficult to process systematically.
- Race information is based on skin complexion and hair color (working to obtain census records with names).

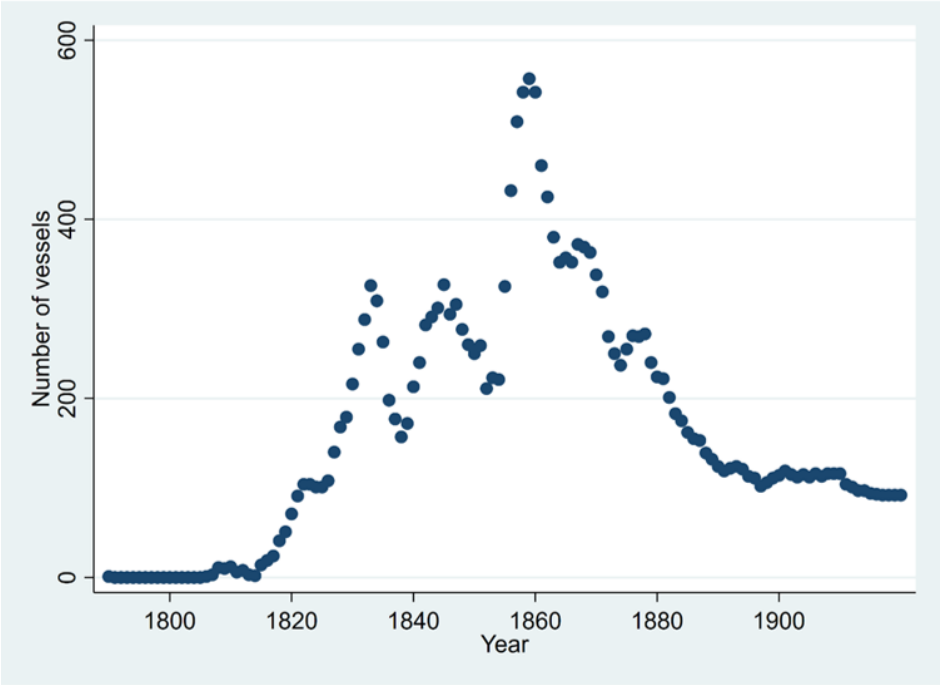
A dynamic industry: rise and fall

American whaling rose to world dominance by mid-century and then collapsed

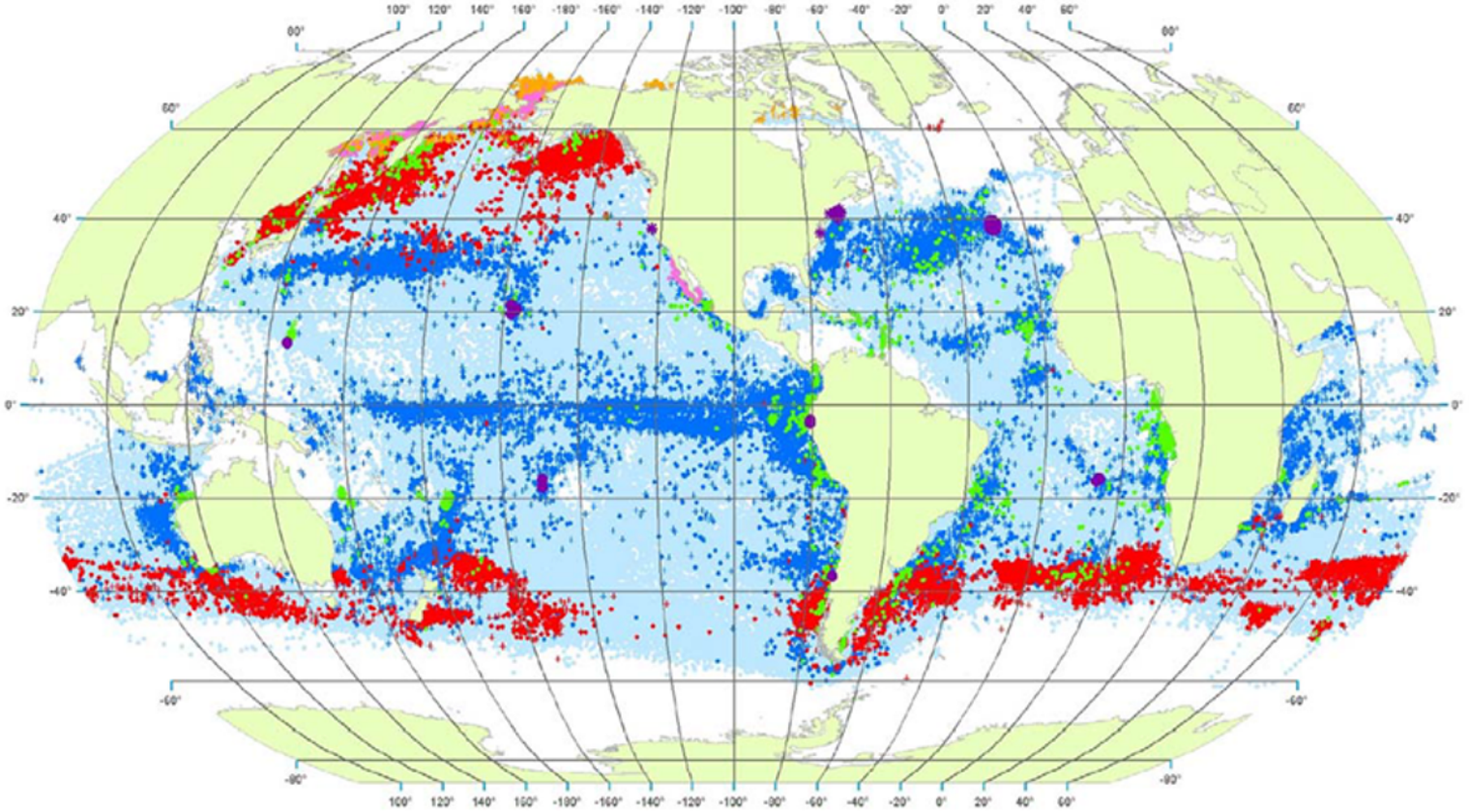
Departures



Total vessels out



Places whales were sighted or caught

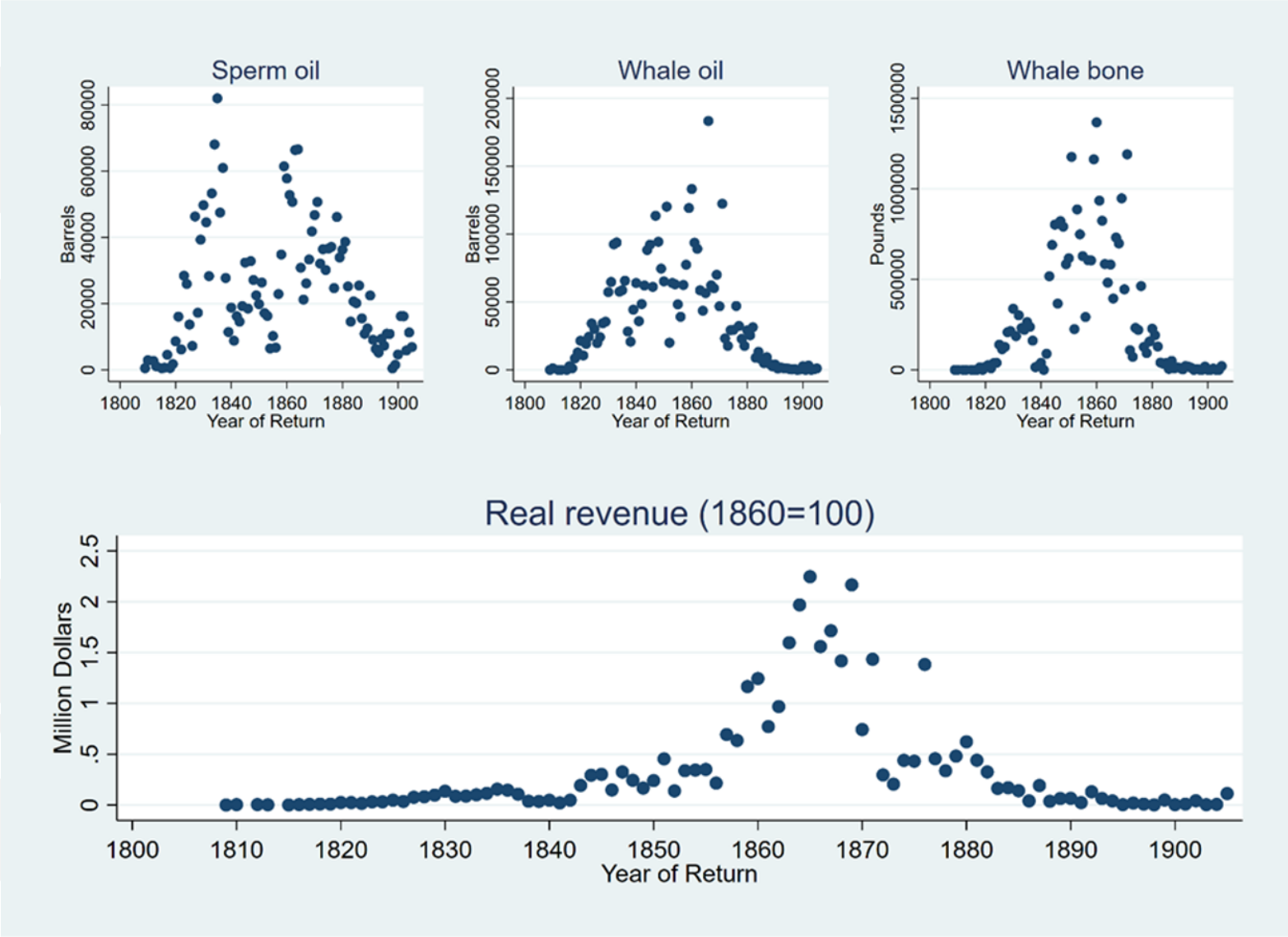


Light Blue	Days when none of these 5 species of whales were sighted or caught (•)
Dark Blue	Days when one or more sperm whales were sighted (•) or caught (+)
Red	Days when one or more right whales were sighted (•) or caught (+)
Orange	Days when one or more bowhead whales were sighted (•) or caught (+)
Green	Days when one or more humpback whales were sighted (•) or caught (+)
Pink	Days when one or more gray whales were sighted (•) or caught (+)
Violet	Locations of home ports (*) and more frequently used ports (•)

Measuring output

- Three main products:
  - sperm oil,
  - whale oil,
  - whale bone (baleen).
- For preliminary analysis, we aggregate the tonnage information for these products with prices to calculate the total revenue of each voyage.
- Use CPI to find real value.

Whaling revenues over time



Measuring diversity

How to infer race from information regarding skin complexion and hair color

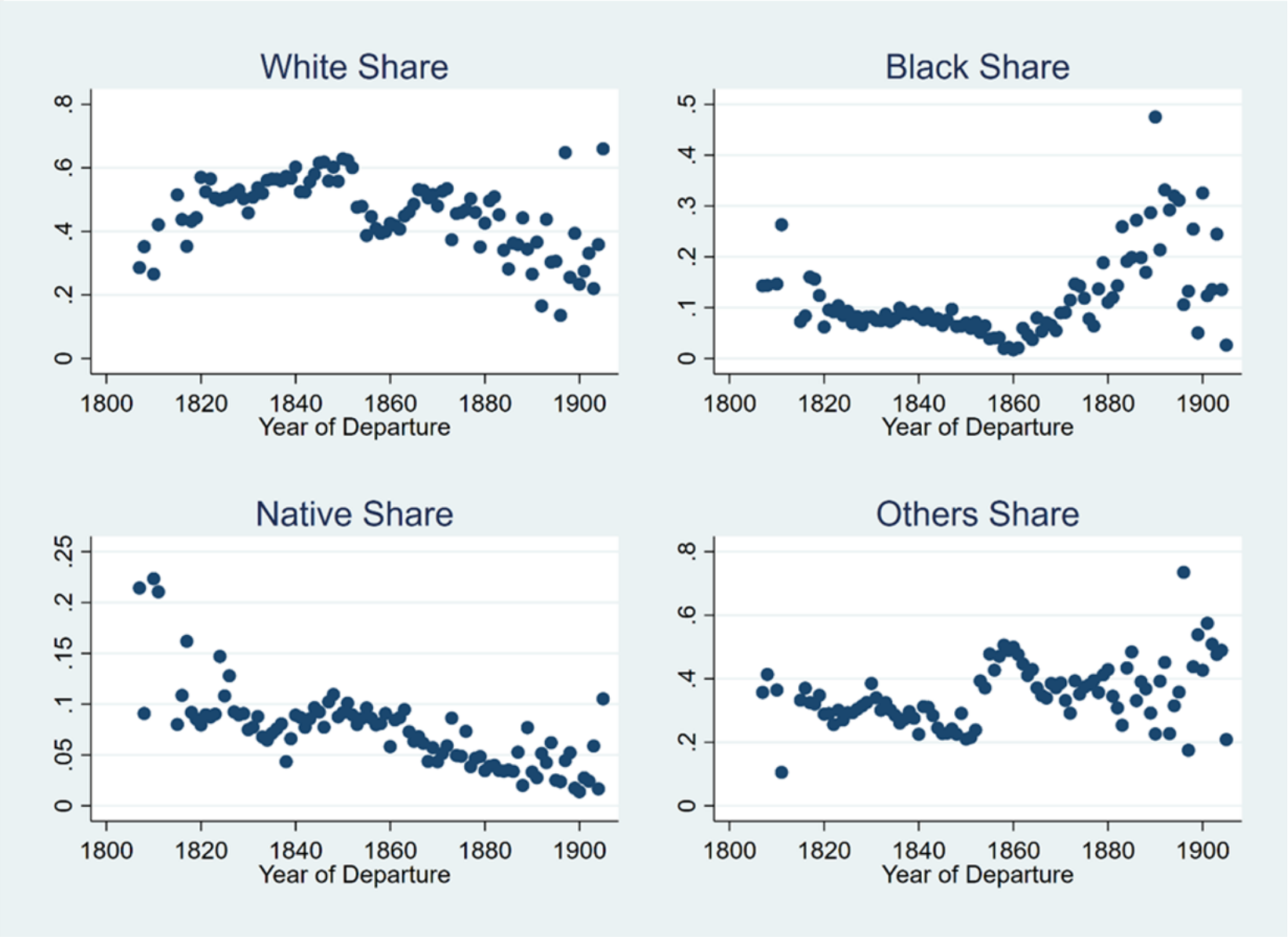
- The crew lists include numerous terms and notations used to differentiate among types of “skin” and ”hair”
  - 213 for skin
  - 220 for hair
- To simplify, we grouped them into
  - 17 categories of skin (e.g., black, colored, light, brown, dark, fair, copper, sandy, Indian)
  - 20 categories of hair (e.g., black, light, brown, woolly, curly, gray, white, red, straight)
- To construct simple (preliminary) categories of race, we used this information (in combination with other information from Mancini and Shoemaker regarding natives) to generate 4 broad categories:
  - native, black, white, other
  - current calculations are conservative in the first three categories

Measure of diversity

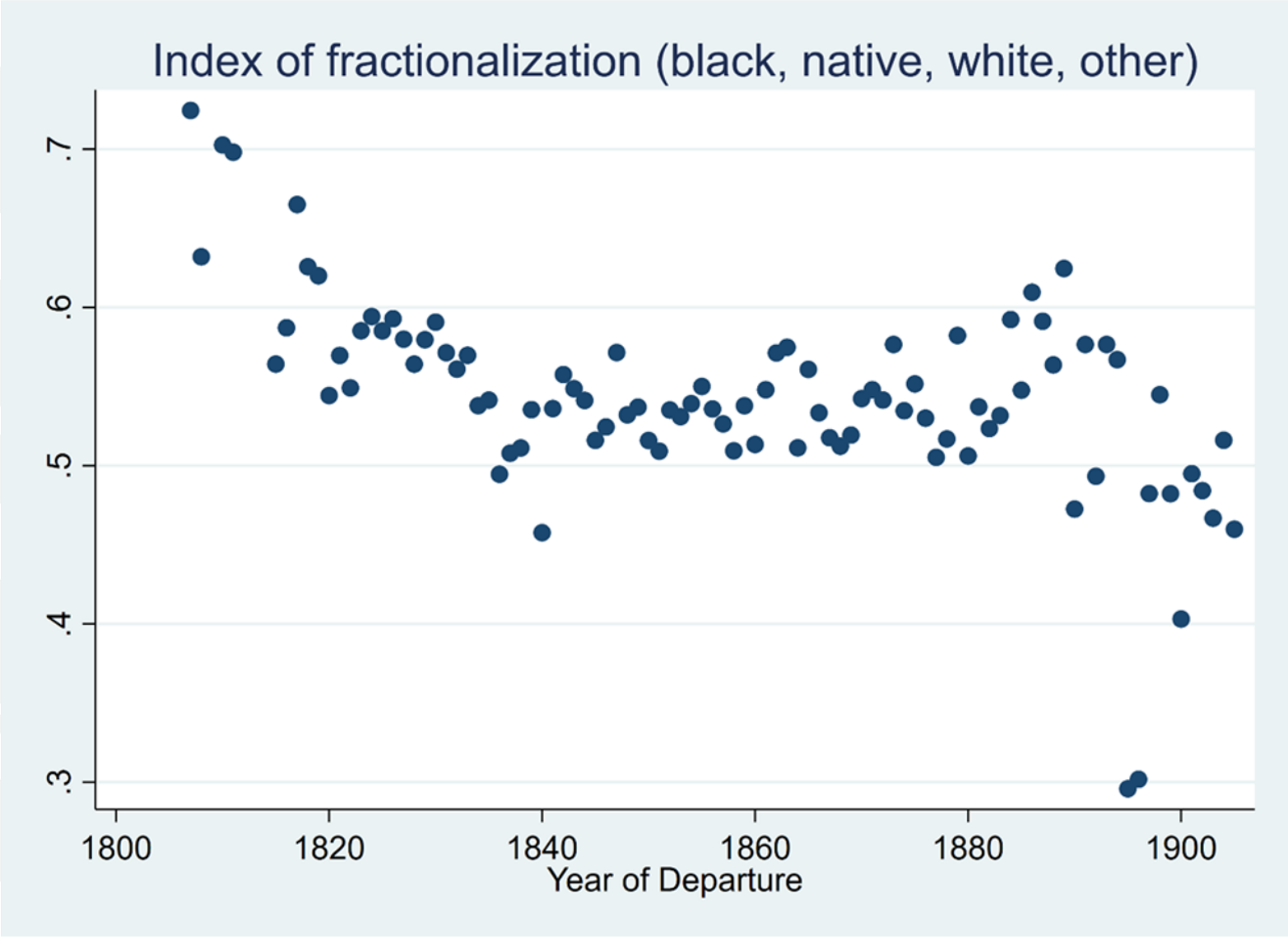
For preliminary analysis, we use these categories to generate two simple measures of diversity:

- Index of fractionalization:
  - 1- sum of squared shares
  - measure of the probability that two randomly drawn individuals within a vessel are not from the same racial group
  - an index number of 0 corresponds to an entirely homogenous crew
- Fraction non-white (white the largest category)

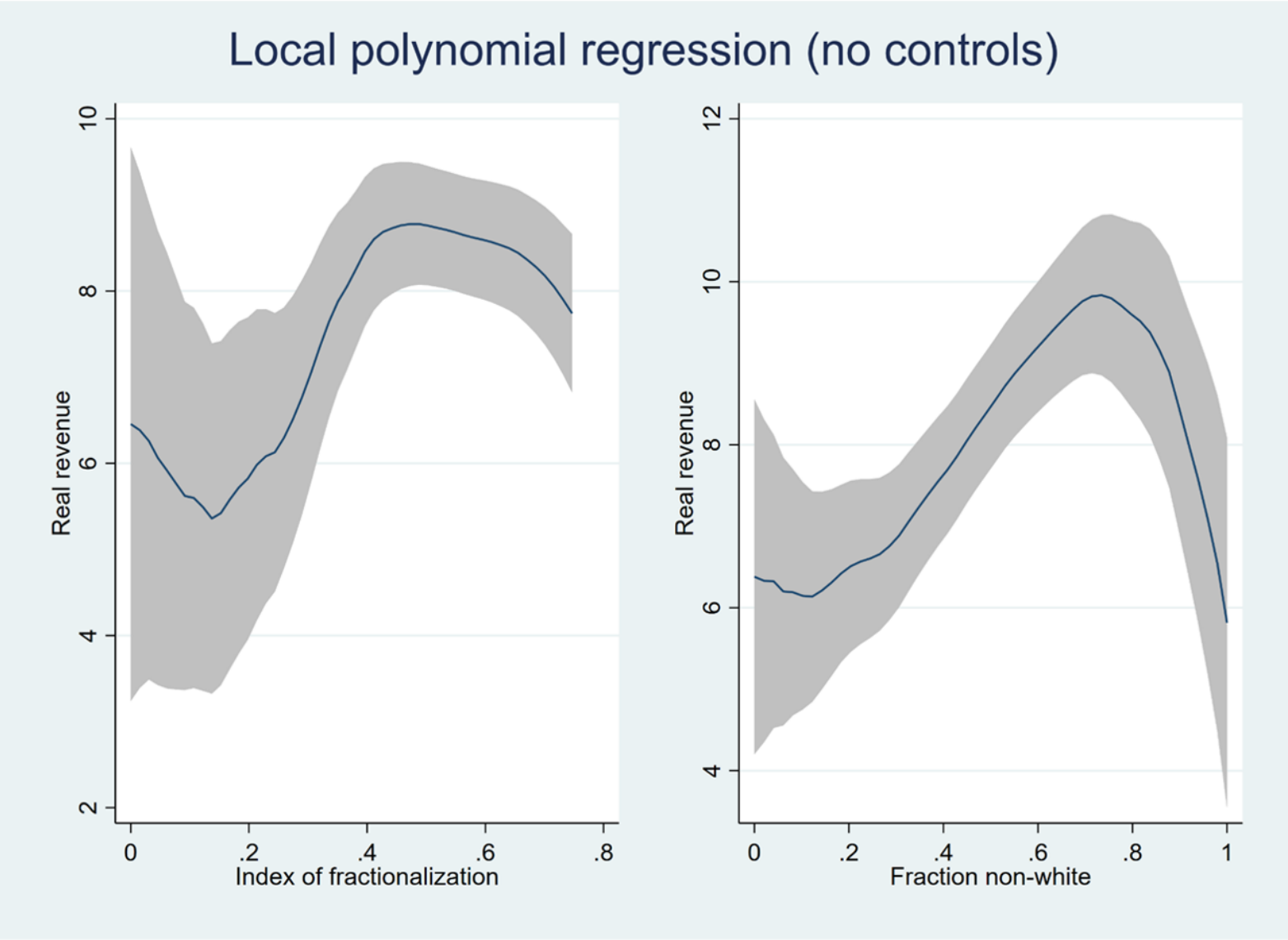
Crew diversity over time



Average fractionalization over time



How did diversity affect revenues?



Firm Behavior

Assume that a whaling firm's technology can be represented by a production function $f(\cdot)$ so that $Y = f(A, X)$, where Y denotes output, X inputs, and A is the efficiency level of the firm.

The fishing vessel (the firm) is free to choose its revenue-maximizing output bundle for its voyage given quasi-fixed inputs, weather and resource abundance constraints, and relative product prices.

Some inputs on the vessel are quasi-fixed during a voyage due to the inability to make major adjustment within a voyage, e.g., tonnage or rig. Other inputs, such as duration of voyage and, to some extent, crew size may vary during a voyage.

Revenue maximization subject to a single quasi-fixed input is a reasonable assumption for a multispecies fishing firm making output decisions over such a limited production period (e.g., Squires and Kirkley, 1991).

Empirical Approach

We estimate a revenue function that models the revenue-maximizing vessel-level production process for an unregulated fishing trip using:

- Generalized quadratic revenue function:

$$R(X,P) = \alpha_0 + \sum_{j=1}^{J} \alpha_j x_{it}^j + \sum_{j=1}^{J} \beta_j p_t^j + \frac{1}{2} \sum_{j=1}^{J} \sum_{k=1}^{K} \alpha_{jk} x_{it}^j x_{it}^k + \sum_{j=1}^{J} \sum_{k=1}^{K} \delta_{jk} p_t^j x_{it}^k + \frac{1}{2} \sum_{j=1}^{J} \sum_{k=1}^{K} \beta_{jk} p_t^j p_t^k$$

- Generalized leontief revenue function:

$$R(X,P) = \alpha_0 + \sum_{j=1}^{J} \alpha_j x_{it}^j + \sum_{j=1}^{J} \beta_j p_t^j + \sum_{j=1}^{J} \sum_{k=1}^{K} \alpha_{jk} (x_{it}^j)^{\frac{1}{2}} (x_{it}^k)^{\frac{1}{2}} + \sum_{j=1}^{J} \sum_{k=1}^{K} \delta_{jk} (p_t^j)^{\frac{1}{2}} (x_{it}^k)^{\frac{1}{2}} + \sum_{j=1}^{J} \sum_{k=1}^{K} \beta_{jk} (p_t^j)^{\frac{1}{2}} (p_t^k)^{\frac{1}{2}}$$

Generalized Quadratic

$$\log(R_{it}) = \alpha_0 + \sum_{j=1}^{J} \alpha_j x_{it}^j + \sum_{j=1}^{J} \sum_{k=1}^{K} \alpha_{jk} x_{it}^j x_{it}^k + \sum_{s=1}^{S} \beta_s race_{it}^s + \gamma_1 frac_{it} + \gamma_2 frac_{it}^2 + \Omega + \eta_{it} \tag{1}$$

where i denotes vessel, t voyage of the vessel, and

- x are inputs, with $x = \{\text{time at sea, crew size}\}$; ships characteristics are excluded because of *vessel* fixed effects
- Ω include the following fixed effects *vessel*, *voyage*, *return* are vessel, voyage, and year of return; *ground* $\times$ *departure* are hunting ground by year of departure.
- Unobserved productivity: $\sum_{s=1}^{S} \beta_s race_{it}^s + \gamma_1 frac_{it} + \gamma_2 frac_{it}^2 + \Omega$
- *race* represents the share of crew by race: white, black, native, other.
- *frac* is the fractionalization index
- The elements including price are suppressed because of *return* fixed effect.

Generalized Leontief

$$\log(R_{it}) = \alpha_0 + \sum_{j=1}^{J} \alpha_j x_{it}^j + \sum_{j=1}^{J} \sum_{k=1}^{K} \alpha_{jk} (x_{it}^j)^{\frac{1}{2}} (x_{it}^k)^{\frac{1}{2}} + \sum_{s=1}^{S} \beta_s race_{it}^s + \gamma_1 frac_{it} + \gamma_2 frac_{it}^2 + \Omega + \eta_{it} \tag{2}$$

- x are inputs, with $x = \{\text{time at sea, crew size}\}$; ships characteristics are excluded because of *vessel* fixed effects
- Ω include the following fixed effects *vessel*, *voyage*, *return* are vessel, voyage, and year of return; *ground* $\times$ *departure* are hunting ground by year of departure.
- Unobserved productivity: $\sum_{s=1}^{S} \beta_s race_{it}^s + \gamma_1 frac_{it} + \gamma_2 frac_{it}^2 + \Omega$
- *race* represents the share of crew by race: white, black, native, other.
- *frac* is the fractionalization index
- The elements including price are suppressed because of *return* fixed effect.

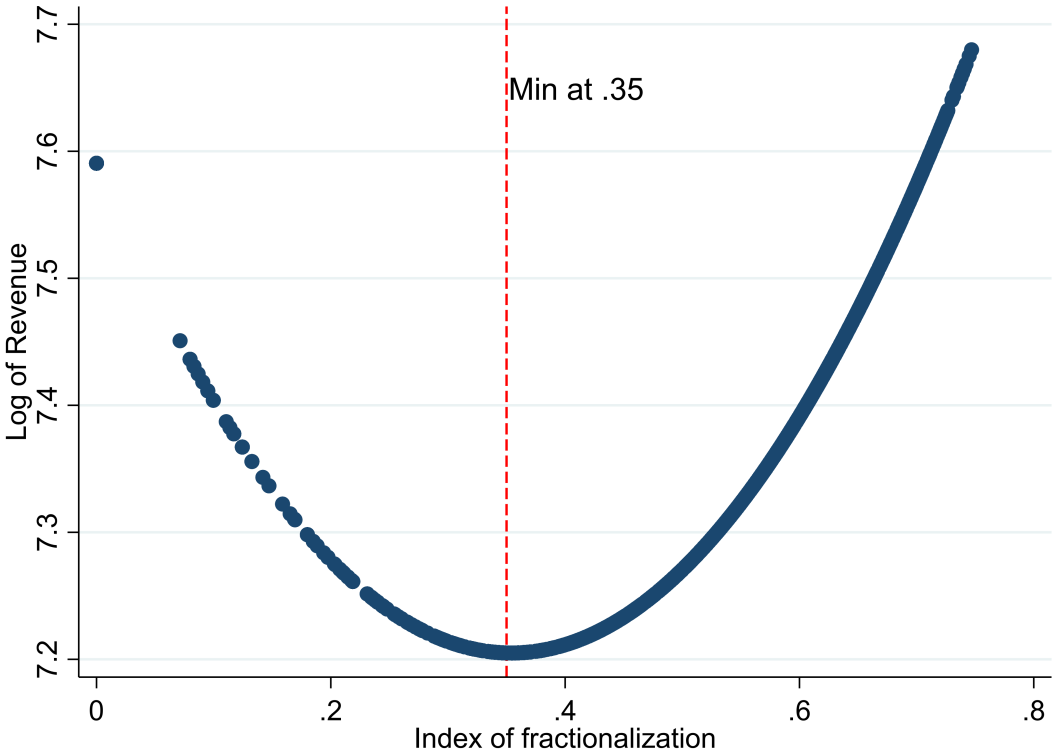
Results: Generalized Revenue Function

	Quadratic		Leontief	
Time at sea	0.0132*	0.0151	-0.0304**	-0.0650***
	(0.00774)	(0.0171)	(0.0115)	(0.00929)
Time at sea squared	-0.000192***	-0.000406***		
	(0.0000506)	(0.0000898)		
Crew size	0.00365	0.0133	-0.0319***	-0.0538***
	(0.0258)	(0.0284)	(0.00891)	(0.00860)
Time crew interaction	0.000690***	0.000676***		
	(0.000220)	(0.000179)		
Crew size squared	-0.000114	-0.000269		
	(0.000274)	(0.000284)		
Time sqrt - Crew sqrt			0.102***	0.145***
			(0.0237)	(0.0250)
Share of black crew	-1.226***	-1.207***	-1.255***	-1.209***
	(0.207)	(0.274)	(0.216)	(0.286)
Share of native crew	-1.012***	-0.981**	-1.033***	-0.996*
	(0.183)	(0.460)	(0.178)	(0.496)
Share of others crew	-0.515***	-0.467***	-0.520***	-0.437***
	(0.0804)	(0.109)	(0.0844)	(0.116)
Fractionalization	-1.035	-2.142**	-1.151	-2.054**
	(1.054)	(0.988)	(1.067)	(0.951)
Fractionalization squared	1.972*	3.035***	2.102*	2.977***
	(1.116)	(0.978)	(1.136)	(0.946)
Constant	6.966***	7.238***	6.728***	7.316***
	(0.689)	(0.976)	(0.305)	(0.555)
Observations	3713	3693	3713	3693
vessel FEs	X	X	X	X
voyage FEs	X	X	X	X
year of return FEs	X	X	X	
Ground $\times$ year departure		X		X

UCONN
UNIVERSITY OF CONNECTICUT

The effect of Racial Diversity on Revenue

At sample mean (0.54), a 1% increase in Fractionalization is associated with a 0.63% increase in Revenue.



An Alternative Approach: Analysis on Total Factor Productivity (TFP)

An alternative way to investigate how diversity affects TFP is proposed by Parrotta, Pozzoli, and Pytlikova (2014). They estimate a two-stage model using a Cobb–Douglas function, calculate the implied total factor productivity (first stage) and relate the latter to diversity (second stage). A similar approach has been used by

- Irarrazabal, Moxnes and Ulltveit-Moe (2013) to investigate to what extent worker heterogeneity explains the wage and productivity exporter premium.
- Fernandes and Paunov (2012) to study the impact of foreign direct investment inflows on the total factor productivity (TFP) of Chilean manufacturing firms.
- De Loecker and Warzynski (2012) to estimate markups and investigate the link between markups and exporting behavior.

Estimating a Gross Revenue Function

Assuming using a Cobb-Douglas technology,

$$y_{it} = \beta_0 + \sum_k \beta_k x_{it}^k + \varepsilon_{it}$$

where y_{it} and x_{it}^k denote the log of output (or gross revenue) and the log of the k input for firm i at period t, respectively.

$\ln(A_{it}) = \beta_0 + \varepsilon_{it}$, and $\varepsilon_{it} = \omega_{it} + \eta_{it}$ model the two unobservable firm's productivity ω_{it} and error term η_{it}. Productivity is observed only by the firm and the error term in not observable.

Estimating a Gross Revenue Function

When subject to productivity shocks, firms respond by expanding/contracting their level of output and by demanding more/less input.

This implies that a positive correlation exists between the observable input levels and the unobservable productivity shocks is a source of bias in OLS when estimating TFP.

This issue has been addressed using various approaches: fixed-effects (FE), instrumental-variables (IV), and control function.

- FE techniques do not solve the simultaneity problem with time-varying productivity shocks.
- IV methods are limited by the difficulty of finding appropriate instruments.

A Control Function Approach

The endogeneity problem can be addressed using the demand for a proxy variable such as investment (Olley and Pakes, 1996) or an intermediate input (Levinsohn and Petrin, 2003; Ackerberg Caves and Frazer, 2015), to define the unobserved productivity as a function of observables (control function).

This method addresses the potential lack of identification due to the fact that the choice of variable inputs correlated with unobserved time-varying productivity shock.

Empirical Specification: First Stage

Following the literature the empirical specification is as follows:

$$y_{it} = \beta_0 + \beta_1 k + \beta_2 l + \omega_{it} + \eta_{it}$$

where y_{it} denotes the log of output revenue for vessel i in voyage t, k is log of capital, tonnage and l is log labor, crew size, respectively.

The unobservable firm's productivity is expressed as an inverse function of the inputs $\omega = f^{-1}(k, l, m)$, where m denotes the log of the intermediate input time at sea, which allows us to identify the parameters for k and l.

First Stage: Results

Log of Revenue	
Crew size	0.870***
	(0.0000107)
Tonnage	1.349***
	(0.00000803)
Time at sea	0.782***
	(0.0000145)
Observations	3928

Empirical Specification: Second Stage

Using the estimates of the production function parameters obtained in the first stage, the estimated TFP for vessel i in voyage t is defined as

$$TFP_{it} = y_{it} - \beta_1 k + \beta_2 l$$

Then, the relationship between TFP and the measure of racial diversity can be estimated with the following equation:

$$TFP_{it} = \sum_{s=1}^{S} \beta_s race_{it}^s + \gamma_1 frac_{it} + \gamma_2 frac_{it}^2 + \\ vessel + voyage + return + ground \times departure$$

where as before

- *race* represents the share of crew by race
- *frac* is the fractionalization index
- *vessel*, *voyage*, *return* are vessel, voyage, and year of return fixed effects; *ground* $\times$ *departure* are hunting ground by year of departure fixed effects.

The effect of Racial Diversity on TFP

At sample mean (0.54), a 1% increase in Fractionalization is associated with a 0.59% increase in TFP.

Log of TFP	(1)	(2)
Share of black crew	-1.598***	-1.167***
	(0.346)	(0.288)
Share of native crew	-0.758***	-0.937**
	(0.269)	(0.452)
Share of others crew	-0.418***	-0.386***
	(0.111)	(0.120)
Fractionalization	-2.156*	-2.048**
	(1.119)	(0.981)
Fractionalization squared	3.135**	2.875***
	(1.149)	(0.970)
Constant	-4.769***	-4.785***
	(0.331)	(0.231)
Observations	3708	3686
vessel FEs	X	X
voyage FEs	X	X
year of return FEs	X	X
Ground $\times$ year departure		X

Conclusions

- Whaling was one of the most important industries in the United States in the 19th century
- The crew on whaling voyages was highly diverse. An ideal setting to test the relationship between diversity and productivity.
- Regression analysis shows a U-shaped relationship between crew diversity and productivity.