Analyzing the Temporal Attribution Problem Associated with Aquaculture R&D and Production

Nicholas Keohan

MASTER OF SCIENCE IN
Applied Economics
BOSTON COLLEGE
WOODS COLLEGE OF ADVANCING STUDIES

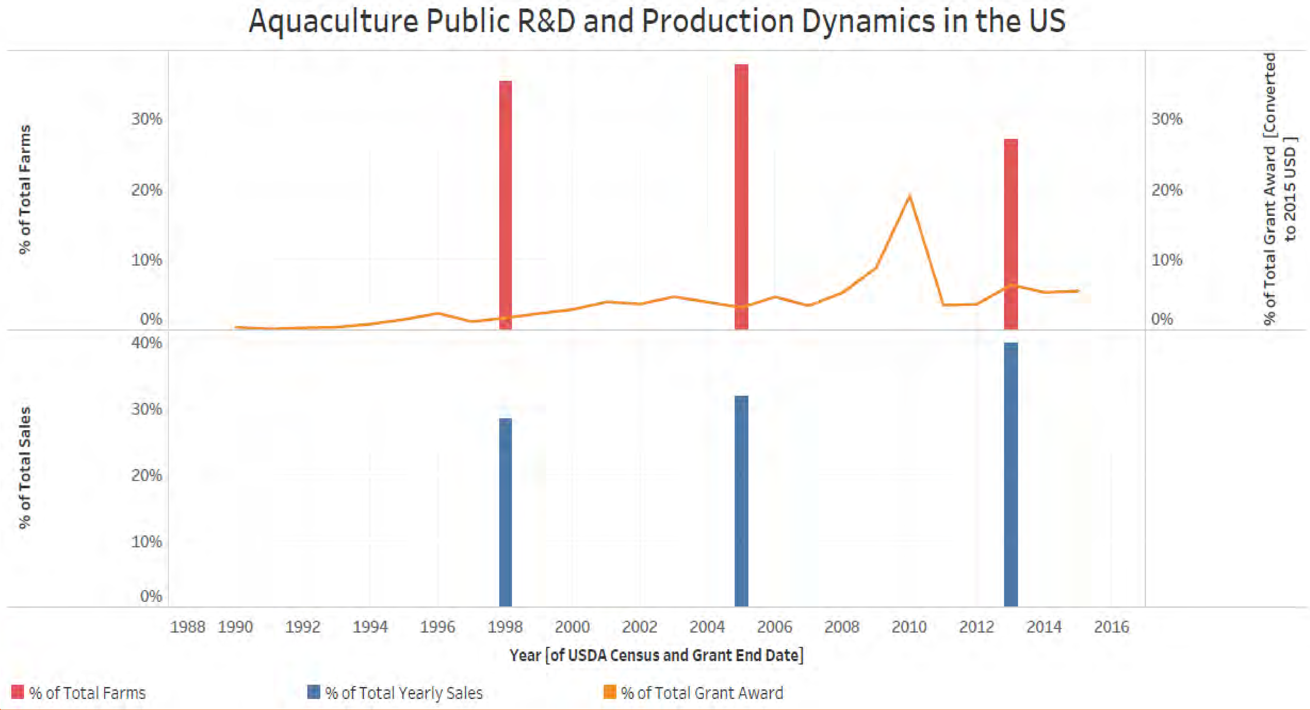
Introduction

Aquaculture is considered "one of the most resource-efficient ways to produce animal protein and has helped improve nutrition and food security in many parts of the world" (NOAA 2018). Such a potential has led the US Federal Government to make significant investments into its research and development through the form of grants. From 1990- 2015, the US Federal Government has invested $919 Million USD into Aquaculture research (Love et al. 2017).

While investment in aquaculture research and development (R&D) has been a relatively new endeavor, agriculture R&D's impact on productivity measures have been widely documented. Research into this relationship has been so extensive that it is credited with contributing to the transformation of economics (Alston 2009). Although total annual use of agricultural inputs has changed little since 1948, the mix of inputs has changed significantly. Agricultural chemicals and purchased services have increased, while land and labor inputs have decreased (USDA ERS 2020). This has led analysts to intuitively attribute much of the growth in the sector to public and private R&D investments. In the early 2000s, spending for private agricultural R&D nearly doubled between 2003 and 2013 while public R&D spending declined (USDA ERS 2020). In order to further analyze this trend, multiple econometric studies have modeled the impact of public versus private agricultural R&D on agricultural production measures.

Two main issues arise within these models in identifying the research lag structure (temporary attribution problem) as well as the treatment of knowledge spillovers (the spatial and institutional-cum-sectoral attribution problem) (Alston et al. 2009).The lag structure analyzed here will represent the publicly funded invention lag, or lag between the end of a grant award and impact on production. Emphasis was only placed on the temporal attribution problem and applying these agricultural models to the aquaculture sector in this analysis.

Aquaculture Public R&D and Production Dynamics in the US



Methods

The dataset for this analysis was created by merging a time series from the USDA Census of Aquaculture and panel data compiled by Love et al on US public aquaculture grants. States with three or less aquaculture farms were not required to report their total sales volume for proprietary reasons and were dropped from the dataset (Reed 2020). The Census data was reshaped into a long format for it to align with the US Federal Funding for Aquaculture data.

In order to study this temporal attribution problem in its most basic form, a nonparametric model was utilized. Prior research into the temporal relationship between research and production measures has assumed structured distributions on this lag which has been argued to constrain the analysis unnecessarily. Specifically, nonparametric models are advantageous in that they have no presumptive restrictions on possibilities among inputs and allow disaggregated inputs to be used to jointly estimate production measures. They also allow for the flexibility of examining different lengths and shapes of lag distributions, requiring only a standard linear program. The nonparametric estimates, on the other hand, are not statistically based and do not allow for hypothesis testing (Chavas et al. 1992). This fact along with my desire to test different models made the fixed effects estimation method more appealing for this project.

To estimate this nonparametric model, a set of dummies was composed for each state. The four states without census data and states with less than two observations were not given dummies because they would not contribute to the coefficient on the invention lag. This is due to this model utilizing within state variation in order to garner its estimates. By creating this exhaustive list of dummies, the estimation method was essentially a fixed effects regression.

Fixed effects was also chosen as the estimation method because of its ability to control for time invariant heterogeneity. This model needed to control for the unobserved effects associated with where grant dollars flow. Even though log transformation were utilized to reel in the outliers in the three nonbinary variables, total grant awards had a much larger scale than sales volume and number of farms. Fixed effects controls for this potential source of heterogeneity if it is time invariant, making our estimates more precise.

$$(1) lnSales_{it} = a + \beta_1 lnFarms_{it} + \beta_2(\textstyle\sum_k^{k-3} lnGrant_{i,t-k}) + \beta_3 AL_i + \cdots + \beta_{31} WV_i + \varepsilon_{it}$$

The nonparametric dummy variable model here represents how each of the seven estimates were garnered and fixed effects was utilized without an α_i. The subscript t represents one of the four census years and the subscript k represents the difference in years between the sales volume and starting year of the aggregated total grant awards. For example, when k is 3 and t is 2005, this model is analyzing the state level total grant awards from 2002, 2001, and 2000 on 2005's corresponding states' total aquaculture sales volume. Aggregation was necessary not only to denote a pseudo- knowledge stock, but also because grants are idiosyncratic and associated with substantial noise. Aggregating helps to mitigate this noise.

Results

	(1) lag2to4	(2) lag3to5	(3) lag4to6	(4) lag5to7	(5) lag6to8	(6) lag7t9	(7) lag8t10
lnFarm	0.260	0.200	0.296	-0.0174	-0.186	1.244***	0.303
	(0.211)	(0.234)	(0.292)	(0.232)	(0.249)	(0.0872)	(0.335)
grant2t4	0.0112						
	(0.0417)						
grant3t5		0.0370					
		(0.0368)					
grant4t6			0.0660				
			(0.0486)				
grant5t7				0.102*			
				(0.0428)			
grant6t8					0.0657		
					(0.0476)		
grant7t9						-0.154*	
						(0.0588)	
grant8t10							0.122
							(0.0713)
N	66	62	60	68	68	61	45
R-sq	0.955	0.970	0.965	0.966	0.949	0.713	0.971
adj. R-sq	0.921	0.943	0.932	0.936	0.903	0.703	0.921

Standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001

Due to the small number of degrees of freedom within each model, it is difficult to draw any statistically significant conclusions. This is a major downside of using a nonparametric dummy variable model. The high R^2 values for each model is expected when performing a dummy variable regression. This is attributed to a dummy variable being included for each cross-sectional unit or state, which can explain much of the variation in the data (Wooldridge 2016 pg 438). Since the number of observations was not consistent for each model but the number of parameters were, adjusted R^2 was the preferred statistic of choice for evaluating each model.

A point of interest is the varying of the statistical significance on the number of farms and their impact on sales. Upon further analysis, the model with a seven-year lag only had observations for the state of Alaska which had one of the smallest variance in number of farms between 1998 (38 farms) and 2018 (22 farms). Alaska also did not have many grants awarded to it within the scope of this analysis making the number of farms intuitively responsible for its range in sales. When this estimation was carried out with aggregated lags of more than three years, the number of farms was never statistically significant but had larger coefficient. This could be attributed to larger lag windows incorporating more states into the analysis. With more states in the analysis the explained variation appeared to increasingly be explained by the fixed effects or dummies.

For models 1, 2, 3, 5, and 7 the coefficients on lag structure and number of farms were not found to be significant or of a large magnitude. These models also had observations for all of the states with three and four matched observations. Meaning they had the most matching observations to work with and usually the most grants awarded. In models 4 and 6 where the aggregated lag was deemed significant, the lag had the opposite sign of the number of farms variable. This could be attributed to these two explanatory variables sharing the same portion of the explained variation. With the majority of the explained variance in these models being given to the state binary variables, what was left over was divided amongst the lag structure and number of farms. The fact that the state dummies consistently explained the most variation in these models indicates that more control variables need to be added in order to understand what has been absorbed by these fixed effects. As well as the lag structure involved in aquaculture R&D's impact on sales.

Conclusion

Understanding the relationship between aquaculture R&D and production is essential for establishing future policies in the sector that remove market inefficiencies. While analysis into the impact of agricultural R&D on production measures has yielded a significant influence, this analysis requires more control variables in order to parse out what is contributing to the statistical significance of the state level dummy variables.

References

- Alston, Julian & Pardey, Philip & James, Jennifer & Andersen, Matthew. 2009. The Economics of Agricultural R&D. Agribusiness. 1. 10.1146/annurev.resource.050708.144137.
- USDA, October 19, 2019. "US Census of Aquaculture." https://www.nass.usda.gov/Surveys/Guide_to_NASS_Surveys/Census_of_Aquaculture/index.php.
- Chavas, J. -P and T. L. Cox. 1992. "A Nonparametric Analysis of the Influence of Research on Agricultural Productivity." *American Journal of Agricultural Economics* 74 (3): 583-591. doi:10.2307/1242571. www.scopus.com.
- Fisheries, NOAA. 2008. "Farming in Water." NOAA. 2008. https://www.fisheries.noaa.gov/story-map/farming-water.
- Love, David. 2017. "Data from: An Analysis of Nearly One Billion Dollars of Aquaculture Grants made by the United States Federal Government from 1990 to 2015." Ag Data Commons. https://doi.org/10.15482/USDA.ADC/1349872. Accessed 2020-12-23.
- Love, D. C., Gorski, I., & Fry, J. P. 2017. An Analysis of Nearly One Billion Dollars of Aquaculture Grants Made by the US Federal Government from 1990 to 2015. *Journal of the World Aquaculture Society*, *48*(5), 689-710. https://doi.org/10.1111/jwas.12425 http://pubag.nal.usda.gov/catalog/10113/5808435
- Reed, Story. 2020. *Permitting & Statistics Program Manager.* Massachusetts Division of Marine Fisheries. Email Message to Nicholas Keohan.
- Wooldridge, Jeffrey M., Jenny Lye, and Mokhtarul Wadud. 2016 "Advanced Panel Data Methods." Essay. In *Introductory Econometrics*, 438–44. Boston, MA: Cengage Learning Australia, 2013.