COMPARISON OF PROXIES FOR FISH STOCK. A MONTE CARLO ANALYSIS

Antonio Álvarez
Department of Economics. University of Oviedo
Email: alvarez@uniovi.es
Phone: +34985104859
ORCID ID: https://orcid.org/0000-0001-7603-1047

Abstract

In fisheries, it is quite common to include a measure of fish stock as an argument of fishing production functions. Given that this information is not available in many cases, researchers use stock proxies. In this paper we show that catch-per-unit-effort, a common proxy based on the average catch of boats in a given period, introduces strong parametric restrictions when included as an explanatory variable in a linear production function. However, a proxy based on time dummy variables yields unbiased estimates of the effect of stock on catch. An empirical section that uses Monte Carlo simulation corroborates the theoretical results.

Keywords: Fish Stock, Production Function, Catch per unit Effort, Monte Carlo Simulation, Time Dummies

1. Introduction

It is common in fisheries research to include a measure of fish stock as an argument in fishing production functions. The idea is that the more fish in the sea, the higher the catch will be, so it seems logical to include an estimate of stock in the production function (Pascoe and Herrero, 2004; Solís et al., 2015). Given that this information is not available in many cases, researchers must resort to the use of stock proxies. This approach raises some interesting modelling issues, since the use of proxy variables can lead to biased and inconsistent parameter estimates due to measurement error (Gordon, 2014).

The most commonly-used proxy of fish stock is ‘catch per unit effort’ (CPUE), a concept that has a long tradition in fisheries (Schaefer, 1954). While CPUE can be calculated in different ways, a common procedure is to estimate the stock in a given period as the average catch of all (or some) boats in that period. Since measures of stock biomass are not available in many cases, CPUE has been widely used in empirical fisheries work as an index of stock abundance, especially in fish population assessments for fisheries management (Myers and Worm, 2003). Even though some papers in the fishing literature have been critical of the use of CPUE as a measure of stock abundance (Richards and Schnute, 1986; Wallace et al., 1998; Maunder et al., 2006), this variable continues to be frequently used in the empirical estimation of fishing production functions (Duy and Flaaten, 2016; Chavez Estrada et al., 2018). Hence, while many papers have shown concerns about the precision of CPUE as an estimate of abundance, we are interested in the performance of CPUE as a proxy of fish stock in the context of estimating fishing production functions.[1]

We compare CPUE with an alternative proxy for fish stock that uses time dummy variables to account for the temporal variation in the stock. The rational for this approach to proxying the unknown stock is that, in many fisheries, fish stock can be considered common to all boats but varying over time. For this reason, using time dummies to model the effect of an unobservable variable which only has time variation seems appropriate. This approach has also been frequently used in the empirical literature (Campbell and Hand, 1998; Fousekis and Klonaris, 2003).

The main contribution of the paper, therefore, is to analyze the statistical properties of both approaches to proxying fish stock: CPUE and time dummies. We do this in the framework of a model where there is a single catch (an aggregate of all species or just a single species without considering age-size classes), and stock is considered common to all boats. The individual vessel-level production function is linear with just two inputs (effort and stock), no individual boat effects and no technical change. We compare the performance of both proxies in this model using Monte Carlo simulation.[2] The main finding of the paper is that the proxy that estimates stock based on average catch is shown to impose strong parametric restrictions. In particular, the estimated coefficient of CPUE is always equal to 1, a result that holds true under different

[1] A different issue is the use of CPUE as a dependent variable in regression models. For example, CPUE has been used to evaluate the impact of economic policies (Chang, 2020).

[2] Some papers (Andersen, 2005; Duy and Flaaten, 2016) have also compared different ways of measuring fish stock but with the main objective of studying their effects on technical efficiency measurement.

scenarios. On the other hand, the time dummies allow for unbiased estimation of the production function parameters.

The paper is organized as follows. Section 2 discusses the general role of stock in a production function. The two different proxies for fish stock are presented in Section 3. Section 4 performs Monte Carlo simulations, which allow us to check the results of the theoretical section. Section 5 contains a discussion of the implications of the paper for empirical analysis, and Section 6 presents some conclusions.

2. The Role of Stock in Fishing Production Functions

The analytical framework is based on a simple production model where fishing output of boat i at time t (y_{it}) is a function of one variable input (Z_{it}) and fish stock (S_t), which is assumed to be common to all boats. Additionally, catches depend on luck and other stochastic effects (u_{it}). Therefore, the fishing production function can be written as:

$$y_{it} = f(Z_{it}, S_t) + u_{it} \tag{1}$$

If the model is linear, we have:[3]

$$y_{it} = \mu + \alpha Z_{it} + \beta S_t + u_{it} \tag{2}$$

While it is assumed that noise is uncorrelated with stock and with inputs, i.e., $\mathrm{Cov}(S_t,u_{it})=\mathrm{Cov}(Z_{it},u_{it})=0$, the correlation of stock and inputs is an empirical issue. If the inputs are time-invariant, as is the case with boat characteristics (length, gross tonnage, …), then $Z_{it}=Z_i$, and by construction, $\mathrm{Cov}(Z_i,S_t)=0$. On the other hand, if Z_{it} changes over time, then it can be correlated with S_t, as would be the case if fishers choose inputs depending on the stock level (e.g. bait…).[4]

[3] We recognize that this is a simple specification that is just used for illustration purposes. The specification of fishing production functions is far from simple and must deal with several complicated issues. While we concentrate on the use of proxies for fish stock, the measurement of effort also deserves special attention, particularly the different assumptions under which inputs are aggregated. A complete analysis of specification issues is beyond the scope of the paper. The interested reader may refer to Hannesson (1983).

[4] The production function could have been specified with individual fixed effects (Kirkley et al., 1998; Squires and Kirkley, 1999), but since the fixed effects are independent of stock by construction, they cannot alter the estimation of the parameters of interest. We have therefore opted for a more parsimonious specification.

If there are data on fish stock, then S_t should be included as an explanatory variable.[5] Since stock is not known in most cases, if the researchers ignore the effect of stock then model (2) becomes:

$$y_{it} = \mu + \alpha Z_{it} + v_{it} \quad where\ v_{it} = \beta S_t + u_{it} \tag{3}$$

If the objective is to estimate the effect of inputs on output, ordinary least squares applied to equation (3) will yield an unbiased estimate of α only if $Cov(Z_{it},S_t)=0$. Otherwise, failing to control for stock can lead to biased and inconsistent estimates of the remaining parameters (see Gordon, 2015). Alternatively, α can be estimated consistently even in the case where inputs are correlated with the stock. Subtracting time means in model (2), yields:

$$y_{it} - \bar{y}_t = \alpha(z_{it} - \bar{z}_t) + (u_{it} - \bar{u}_t) \tag{4}$$

This procedure, which eliminates variables with no variation across boats, allows α to be consistently estimated by ordinary least squares under the assumption that inputs and noise are uncorrelated.

However, if interest lies in estimating the effect of the stock on catches (β), the researcher must use a proxy for the unknown stock. In general, there are two cases. If boats fish in the same area, then a variation in fish stock will affect all fishermen in the same way. In this case, the problem calls for a stock measure common to all boats. As such, the effect of a change in the fish stock is similar to Hicks-neutral technical change, since an increase in the stock allows fishers to catch more fish at any level of input use. On the other hand, if boats fish on different grounds, the stock measure should take this factor into account. In the next section, we study two proxies for fish stock considering that stock is common to all boats.

3. Proxies for Fish Stock

The issue at hand is how to control for the effect of biomass stock when data for this are not available. As stated in the Introduction, two main proxies for fish stock have been used in empirical work. One is average catches in a given period, which is known as 'catch per unit of

[5] Few papers include a measure of fish stock obtained as an independent estimate from external sources. See, for example, Grafton et al., (2000), Pascoe et al., (2001), Jin et al., (2002), or Mulazzani and Marlogio (2013).

effort' (CPUE), while the other proxy uses time dummies to account for variations in the stock over time.

a) Catch per unit of effort

In this approach, the stock in period *t* is estimated as the average catch of all (or some) boats in that period. That is,

$$CPUE_t = \bar{y}_t = \frac{\sum_i y_{it}}{N} \qquad (5)$$

where y is catch and N is the number of boats.

Not all computations of CPUE are identical. For example, Eggert (2001) calculated CPUE as "the overall average landing value per unit effort on a monthly basis". Comitini and Huang (1967) referred to the need to include the "density of the fish population existing at the points where the boat happens to be fishing" as an input, which they measured as 'catch per skate'. Andersen (2005) calculated a CPUE stock index for each species as the sum across boats of the catches per day at sea, while Chavez et al. (2018) measured CPUE as "aggregate landings level adjusted by the aggregate effort made by the fleet per day" and referred to it as 'fish availability'.

Some authors also calculate a CPUE measure that is not common to all boats. For example, Pascoe and Coglan (2002) used a stock index calculated as the geometric mean of the "value of catch per hour fished for the boats that operated in the same month in the same area using the same gear". Kirkley et al. (1998), on the other hand, computed CPUE as "the geometric mean of baskets of scallops caught per hour from the last tow of all vessels participating in the stock monitoring program and fishing the same area using the same dredge size".

In the case of multispecies fisheries, the computation of CPUE as the average of all catches implies that there is a single stock. This assumption can be circumvented. For example, Andersen (2005) included an index of fish stock for each of the primary species in his Data Envelopment Analysis of Danish seiners, which he computed as the average catch of each species per day-at-sea.

The rationale for CPUE is based on the traditional specification of a production function in fisheries, where catch (y) is a function of fishing effort (E) and stock (S). That is:

$$y = qES \qquad (6)$$

where q is the coefficient of catchability. Since q is usually considered constant, it is clear that catch per unit of effort (y/E) is proportional to the stock.

Moving $\bar{y}_t$ in equation (4) to the right-hand side, and using the definition of CPUE in (5), we have:

$$y_{it} = \alpha(Z_{it} - \bar{Z}_t) + CPUE_t + (u_{it} - \bar{u}_t) \qquad (7)$$

Equation (7) can now be compared to a production function where CPUE is used as a proxy for the stock, which is:

$$y_{it} = \mu' + \alpha' Z_{it} + \beta' CPUE_t + v_{it} \qquad (8)$$

Comparing these last two equations it is easy to see that the estimated β’ should be equal to 1 if CPUE is independent of Z and the random term. This is because the construction of the stock index based on catches implies that any variation in the stock will be captured by the boats.

Criticism of CPUE

CPUE has been criticized on several grounds. Richards and Schnute (1986) used a submersible to visually obtain an index of resource density and compared it with CPUE, finding that CPUE is not a good index of stock abundance. Kleiber and Maunder (2008) simulate fish abundance data and find that CPUE “cannot be relied on as an index of aggregate abundance”. In particular they find that under CPUE is usually hyper-responsive, that is, the change in CPUE will be larger than that of stock abundance.

The direct proportionality between CPUE and stock abundance implied by equation (6) has been questioned in the literature (see, for example, Harley et al., 2001).

While most of the criticism to CPUE has been done in the area of fish stocks assessment, not much has been said about its use in fishing production functions. An exception is Sharma and Leung (1998), who indicate that "CPUE figures are commonly used as indicators of stock abundance. However, because of its dependence on other inputs (crew size, fuel, and gear type) CPUE is not suitable to include as an input variable in production function analyses".

Moreover, on econometric grounds, the introduction of a proxy for fish stock based on average catches can sometimes lead to endogeneity problems. This has already been recognized in the literature. For example, Felthoven and Morrison (2004) indicate that the use of CPUE may lead to endogeneity bias since "catch levels are present in both the dependent variable and the aggregate, period-specific CPUE measures". As a possible solution, they suggest to calculate CPUE from vessels outside the sample. Pascoe and Coglan (2002) also warn about the endogeneity of CPUE since "the observed value per unit of effort is a function of both the actual stock abundance and the stochastic element (e.g., "luck") associated with fishing. The ultimate estimate of the stock abundance may therefore be influenced by the average "luck" of the observed vessels, which may vary from period to period". Even though average "luck" is most likely to be close to zero in the sample, it is possible that some unobserved (to the researcher) characteristics of the stock vary on average over time. For example, the stock of some mobile species can in a particular year pass closer to the coast than in other years. This factor will affect catches, and if not accounted for in the effort variable it will be picked up by the random term implying that $\bar{u}_t$ is correlated with catches, yielding biased and inconsistent estimates of the estimated parameters.

The cause of endogeneity can also be in the denominator (number of boats). This is probably common in many fisheries since when stock increases, more boats are attracted to fishing that stock. This doesn't mean that new boats are built, rather that when a mobile species arrives to a certain area (say, mackerel) boats that are specialized in other species may decide to spend some time catching mackerel. The larger the stock of mackerel, the larger the number of boats deciding to participate in this fishery.

The issue of endogeneity in fishing production functions has been discussed at length by Gordon (2015), who suggests the use of Instrumental Variable estimation methods in the presence of endogeneity (see also Ekerhovd and Gordon, 2013).

A seldom discussed problem has to do with the more general question of explaining individual behavior by the group average, which has been frequently used in other areas of empirical research, such as education (i.e., explaining students' grades by the average performance of their class). Manski (1995) analyzes the specification and estimation problems of this approach.

b) Time dummies

This approach has been used in the empirical literature by, among others, Campbell and Hand (1998) and Alvarez et al. (2020). If stock is common to all boats, one can write:

$$y_{it} = \alpha Z_{it} + \sum_{t=1}^{T} \gamma_t D_t + u_{it} \qquad (9)$$

where γ_t are parameters (time effects) to be estimated as coefficients of time dummy variables. D_t is a dummy variable that takes on value 1 if year is t and 0 otherwise. If the panel data set is short in the time dimension, the time effects will probably pick up only the effect of stock changes, otherwise they may also incorporate neutral technical change or any other unobserved effect which is common to all boats. This is for example the case of environmental conditions (weather, sea state), which clearly affect catches, although many times researchers do not have variables that represent them.[6]

An alternative and much more common way to specify equation (9) is to include an overall constant and exclude one of the time dummies, which acts as the reference category.

$$y_{it} = \mu' + \alpha Z_{it} + \sum_{t=2}^{T} \gamma_t D_t + u_{it} \qquad (10)$$

In this setting, the coefficients of the time dummies should be interpreted as the differential effect of each year on output with respect to the omitted category.[7] Note that the constant in (10) includes the time effect of the omitted category (year 1). That is, $\mu' = \mu + \gamma_1$. Equation (10)

[6] Some examples of articles including variables about environmental conditions are Torres and Felthoven (2014) who used the surface air temperature, and Alvarez et al. (2020) who included the chlorophyll concentration in the sea.

[7] Suits (1983) has a very clear exposition of the different interpretation of dummy variables when including all the categories and when omitting one.

can be estimated by ordinary least squares and the time effects are recovered using the typical expression of the least squares estimator for a constant:

$$\hat{\gamma}_t = \bar{y}_t - \widehat{\mu'} - \hat{\alpha}\bar{Z}_t \qquad (11)$$

Now we proceed to compare the two proxies for fish stock, namely catch per unit effort ($CPUE_t$) and time effects ($\hat{\gamma}_t$). The difference between them can be seen by calculating their respective expected values. The expected value of CPUE in equation (5) is:

$$E(CPUE_t) = E(\bar{y}_t) = \mu + \alpha\bar{Z}_t + \beta S_t \qquad (12)$$

This implies that $CPUE_t$ is a biased estimator of S_t, since average catch depends not only on the existing stock but also on the average fishing effort ($\bar{Z}_t$). This is the one of the reasons why previous papers have indicated that CPUE is not a good measure of stock abundance (Sharma and Leung, 1998).

On the other hand, the expected value of $\hat{\gamma}_t$ in equation (11) is:

$$E(\hat{\gamma}_t) = E(\bar{y}_t) - E(\hat{\mu}') - \bar{Z}_t E(\hat{\alpha}) = E(\bar{y}_t) - \mu' - \alpha\bar{Z}_t \qquad (13)$$

Substituting equation (12) into equation (13) yields:

$$E(\hat{\gamma}_t) = \beta S_t - \gamma_1 \qquad (14)$$

Equation (14) shows that the expected value of the estimated time effects ($\hat{\gamma}_t$) equals the effect of stock on output with respect to the omitted category (year 1). Therefore, the time effects can be considered unbiased estimators of the effects of fish stock on catch.

4. Monte Carlo simulation

In this section, we will show the empirical implications of the theoretical results obtained above. However, the exact results of Section 3 will be difficult to reproduce in most data sets due to problems caused by multicollinearity between inputs and the stock proxy, as well as to the presence of noise. For this reason, we have decided to use a simulated data set.[8]

We construct a panel data set consisting of 50 boats and 5 time periods that are assumed to be years.[9] This boats/years panel data structure is the most typical in fisheries. While it is also common to have several years of monthly or quarterly data (Campbell and Hand, 1998; Kirkley et al., 1998), the presence of both year and monthly dummies would only complicate the presentation of results but would not affect the main results of the paper.

Our data generating process is the model in (2), i.e., a linear production function. The random term *u* is generated from a N(0,1). Since we estimate the models using ordinary least squares, it should be noted that the variance of *u* does not affect the estimated parameters on average. The population parameters take the following values in the simulation: μ=1, α=0.2, β=0.8.

The input Z, effort, is generated in two different ways. In the first case, it is generated as a random variable from a uniform distribution U(1,2) in order to assure that it is always positive. In this way, effort varies across boats and time without following any particular pattern. This case reflects the typical situation where effort depends on activity (days at sea), which may not follow a particular pattern. In the second case it is generated as a time-invariant variable. This implies that effort varies across boats but is constant over time, representing the situation where effort is associated with the capacity of the boats (length, gross tonnage), which is time-invariant. It should be noted that, in both cases, effort is generated independent of the stock.

The stock variable is also generated in two ways. In the first case, it is generated as a deterministic function of time (S_t=0.5*t),.[10] While in the second case it is generated as a random term that was simulated by independent draws from a uniform distribution. In both cases,

[8] Other papers that use simulated data sets in fisheries are Pascoe and Robinson (1996), Herrero and Pascoe (2004), or Gordon (2015).

[9] The results are not sensitive to the number of boats and/or years. We have performed the empirical analysis for N=10, and 100, as well as for T=10, and the results are almost identical.

[10] We also allowed for some random variability in the stock around the deterministic trend. The results are very similar to the ones obtained using a deterministic stock and are not presented here.

therefore, the stock variable was generated as a time-varying variable that is common to all boats. Finally, output (catch) is computed using the production function in (2).[11]

We estimate the production function using the true stock (S_t), and the two proxies: CPUE and time dummies (γ_t). In order to infer the sampling distribution of the estimators, we perform a Monte Carlo analysis with 1000 replications. In each new replication the random term *u*, the output, and therefore CPUE, are newly generated but the values of stock and effort are held fixed.

5. Results

The production function is first estimated without correlation between effort and stock. We then allow for different degrees of correlation between these two variables. The estimation was conducted using Limdep V10.

a) Inputs and stock are uncorrelated[12]

Table 1 summarizes the results of the Monte Carlo exercise using the two ways to generate effort and stock. The table shows the means (over the 1000 replications) of the estimated coefficients using (i) the true stock, and (ii) CPUE as the stock variable in the regressions. The results illustrate some interesting points. As expected, the use of the true stock, in both the deterministic and the stochastic versions, allows all the technology parameters to be recovered. The estimates when using CPUE are in line with the results predicted by the theoretical discussion in Section 3. First, the estimated coefficient for CPUE is not statistically different from one. Second, the estimation also allows the coefficient of effort to be recovered, since effort is generated independent of the stock. Another interesting result is that the use of CPUE results in a biased estimate of the constant term. The reason is that CPUE incorporates the effect

[11] Even though the data has the structure of a balanced panel, the independence of the draws implies that we are treating it as a 'pooled' model. That is, there is independence over time. Boat effects are likely to be important for empirical analysis with real data since they are usually interpreted as indicators of skipper skill (Squires and Kirkley, 1999). However, they are not relevant for the issues discussed here.

[12] The lack of correlation is driven by the independence of the draws for both variables. In fact, the variables are not orthogonal, although the correlation is not statistically different from zero.

of effort on catches. Comparing equations (7) and (8), it is clear that the estimated constant when using CPUE is equivalent to $-\alpha\overline{Z}_t$. Since the mean of effort is 1.5 (both in the time-varying and time-invariant cases)[13] and α=0.2, we have that -0.2*1.5=-0.30, which is the estimated constant. Finally, it should be noted that the same results are obtained with the two versions of effort. Therefore, the predictions of the theoretical model hold when using empirical data.

We now turn to the estimation of the model using time dummies. We are interested in checking whether the dummy variables capture the effect of stock on output. In the deterministic version of the stock, the true effect is given by βS_t=0.8*0.5*t=0.4*t. Therefore, the true effect for year 1 will be 0.4*1=0.4 , and similarly for the remaining years. In the random version of stock, the true values of the stock are generated from a uniform distribution U(1,10). Since the estimated coefficients of dummy variables in linear models show the differential effect with respect to the omitted category, we need to compare the estimated time effects with the true effect of stock in each year, which is given by $\beta S_t - \beta S_1$,: as shown in equation (14). Table 2 shows the true values of the two versions of the stock, as well as the true time effects with respect to the omitted category for both versions of the stock.

Table 3 shows the estimation of the model using time dummy variables, leaving the year 1 dummy out. We find that the estimated time dummies are almost identical to the true yearly effects for both stock variables, which proves that in situations where stock is common to all boats, it can be proxied by time dummies. The estimation also recovers the constant for both models, which is equal to $\mu + \gamma_1$. In the case of deterministic stock, since $\mu = 1$ and $\gamma_1 = \beta S_1 = 0.8 * 0.5 * 1 = 0.4$, the true constant is 1.4. When stock is random, the generated S_1 is 3.10, which implies that the true constant is $\mu + \beta S_1 = 1 + 0.8 * 3.10 = 3.4$. The coefficient of effort is also recovered with precision, and the results hold for both effort variables.

b) Inputs and stock are correlated

[13] The reason for the mean of both versions of effort to be the same is that the time-invariant version was generated as the mean over time of the time-varying variable.

We now allow for correlation between effort and stock. This assumption seems quite logical in fisheries since it implies that when the stock grows, boats increase fishing effort. In particular, we assume a linear relationship between the two:

$$Z_{it} = \theta S_t + \varepsilon_{it} \tag{15}$$

where θ is a parameter and ε is a random term which is assumed to be independent of stock. The correlation between Z and S is given by:

$$Corr(Z, S) = \theta \frac{\sigma_s}{\sigma_z} \tag{16}$$

where σ_S and σ_Z are the standard deviations of S and Z respectively. In the simulations, we considered the correlation of time-varying effort with both definitions of stock (deterministic and stochastic). The time-invariant effort is orthogonal to both measures of stock since they vary over time but are constant across boats and therefore cannot be used to simulate correlation.

Three cases were considered: low (0.20), medium (0.50), and high (0.80) correlation. The results of the estimation using the true stock and CPUE can be seen in Table 4. As expected, the correlation between effort and stock does not affect the estimation when the true stock is used. The linear model is able to track the theoretical parameter values with precision. When using CPUE, however, the population parameters are estimated with less precision since the main consequence of multicollinearity is to increase the standard errors of the estimators. In any case, the estimated coefficients are still very close to the true parameters, as in the case of the stock variable, which ranges between 0.92 and 0.99. The loss of precision increases with the level of correlation. Since with real data it is always the case that at least some inputs are correlated with the stock level, it is not surprising that the empirical papers that have used CPUE have not estimated a value of 1 for the coefficient of this variable.

In Table 5 we present the estimation of the production function using time dummies as proxies for the stock under the three degrees of correlation between effort and stock. In this case, the estimated coefficients for all the variables (constant, time dummies, and effort) are very similar to those obtained when there is no correlation between effort and stock (see Table 3).

Overall, therefore, our simulation exercise shows that when there is no correlation between the inputs and the stock, the use of CPUE as a proxy for stock in a linear fishing production function yields estimates of the coefficient for stock equal to one, although it is possible to obtain a consistent estimate of the parameter associated with effort. When inputs are correlated with the stock, in the model with CPUE the parameters are estimated with less precision, while the model with time dummies still recovers the true parameters with precision.

6. Discussion

The size of the stock elasticity is of great interest to fisheries scientists because of its important implications for policy recommendations. The previous sections contribute to showing that CPUE is not a good proxy of the effect of fish stock when included in a linear fishing production function. In fact, the finding that CPUE - computed as the average catch of all vessels - always results in an estimated stock parameter equal to 1, regardless of the species, must be of some concern to fisheries researchers since the conventional wisdom in the profession maintains that the elasticity is close to zero for pelagic species while it should be close to 1 for demersal fish stocks (Ekerhovd and Gordon, 2013). On the other hand, in the framework of this paper (single species, stock common to all boats), the use of time dummies is preferable to CPUE since they are able to recover the effect on output of the different stock levels.

The use of time dummy variables in this context is not without problems. One of these concerns their interpretation as proxies for stock effects. In our data generating model, the only time-varying variable which was later unobserved in the estimated models was the stock. Therefore, it makes sense that the time dummies can track down those effects. However, in real situations it is rather common to have other unobserved time-varying variables which are constant across boats, such as environmental changes, regulations, or weather, to name just a few. In summary, the interpretation of time dummies as indicators of fish abundance must be made with care since the estimated effects will be mixed with those of other omitted time-varying variables. Obviously, this argument does not imply that CPUE does not suffer from these types of problems, since those omitted variables will most likely be correlated with CPUE and the estimation will therefore be biased.

The time dummies may also be confounded with neutral technical change. In empirical applications it is possible to include both a time trend and time dummies. For example, Campbell and Hand (1998) claimed that the time trend was expected to pick up the effect of technical change, while the time dummies would pick up variations of the stock over time. However, in this specification there are identification problems and the interpretation of the time dummies as pure measures of stock effects is questionable. Of course, the specification of technical change can be made more flexible than just a linear function of time. For example, Solis et al. (2014) used both a linear and a quadratic trend.

We have also performed a Monte Carlo simulation including in the data generating process a linear trend representing technical change. In the models with true deterministic stock, there is perfect correlation between the true stock and the trend since deterministic stock was generated as a linear function of time, and the model with true stock cannot be estimated. When using CPUE, the multicollinearity between CPUE and the linear trend (since CPUE is computed from a stock which is a linear function of time) doesn't allow to identify the parameters with precision. In the models with true random stock, all the parameters (including the technical change) can be recovered. If CPUE is used instead of true stock we obtain again the result that the estimated coefficient for CPUE equals 1. For both versions of the stock, the model with time dummies cannot be estimated because the time trend can be written as a linear combination of the time dummies.

Another problem has to do with the periodicity of the dummies. For example, even if the data are at a monthly level, it is not necessary to include monthly dummies, with seasonal (quarterly) dummies usually being enough (Campbell and Hand, 1998). An additional problem may arise if the data set comprises years and months (or any time frequency lower than years). In this case, the specification will include both yearly and seasonal dummies (Alvarez et al., 2020). If there are many years, the yearly dummies can be substituted by a function of time. Additionally, Pascoe and Coglan (2002) point out that if the seasonal pattern of the fish stock is not the same over years, the seasonal dummy variables have to be interacted with the year dummies and the number of interactions needed could be substantial.

Another important issue relates to the functional form of the production function. This paper has shown that if the production function is linear, the coefficient of CPUE will be equal to 1. However, many papers use more flexible production functions such as Cobb-Douglas (Comitini and Huang, 1967; Grafton et al., 2000; Mulazzani et al., 2015; Alvarez et al., 2020) or translog (Sharma and Leung, 1998; Pascoe and Coglan, 2002; Fousekis and Klonaris, 2003). In the case of the Cobb-Douglas production function it is possible to get an estimated coefficient of 1 for CPUE in the case that CPUE is calculated as the mean of the log of catches by all boats. The analytics for the translog production function are more complicated than in the linear and Cobb-Douglas functions and it does not seem possible to get such an exact result (coefficient of CPUE equal to 1). Still, even if there is no exact result one would think that CPUE will continue to impose some parametric restrictions even if they are not as evident as in the other cases.

The use of CPUE or time dummies is not the only way to account for stock in fishing production models. Some alternatives to deal with the problem of controlling for unmeasured fish stock have been suggested in the literature. For example, some papers use a time trend to account for the temporal pattern of the unobserved stock (Chowdhury et al., 2010). This approach implies fewer parameters need to be estimated than in the case of time dummies, but it imposes a rather rigid time pattern on a variable which is subject to strong random variability. We also did Monte Carlo simulations using a time trend as a proxy for stock under the same scenarios and, as expected, we are able to recover with precision the effect of stock when stock is generated as a linear function of time. In the case that stock is random, the time trend as a proxy doesn’t work well since the constant growth rate implied by the linear function of time doesn’t fit the random process of stock. Therefore, a time trend will only be an appropriate proxy in the case of stocks that increase or decrease over time in a monotonous way.

In a similar vein, Orea et al. (2005) use a flexible function of time to model the seasonal pattern of stock in the framework of an output-oriented distance function. The advantage of this approach is that fewer parameters need to be estimated and the time trend can be interacted with other inputs in an easier way than with time dummies. They interact the parameters of a cubic polynomial time curve with output and gear variables.

Another alternative to account for stock can be seen in Pascoe and Herrero (2004), who use Data Envelopment Analysis to estimate an index of stock abundance. Their measure is in fact a 'stock effect' rather than a stock index. That is, it is an indicator of how much the changes in stock have affected the catch of each of the boats. This method was used by Herrero and Pascoe (2003) to normalise the dependent variable (catch) of a fishing production function instead of including it as an additional input.

7. Conclusions

This paper compares two alternative proxies for fish stock in the framework of the estimation of a fishing production function. The theoretical analysis shows the estimated coefficient of CPUE is always equal to 1 in the case of a linear production function, while time dummy variables are a better proxy for stock. In the empirical section we use Monte Carlo simulations to show that the predictions of the theoretical model are correct. Concretely, the result that the estimated coefficient od CPUE equals 1 holds under different scenarios defined by several ways to generate the stock variable (deterministic or stochastic) and the effort variable (time-varying or time-invariant). The conclusion of the paper is that, at least for the production setting of this paper (single species, stock common to all boats), the use of CPUE introduces strong parametric restrictions and, therefore, it seems better to use time dummy variables as proxies for the unknown fish stock in the estimation of fishing production functions.

Interestingly, the problem with CPUE studied in this paper is not due to measurement error or endogeneity and therefore cannot be dealt with in the framework of instrumental variable estimation methods, instead being a specification problem, which calls for a different proxy.

References

Alvarez, A, Couce, L. Trujillo, L., 2020. Does Specializacion Affect the Technical Efficiency of Fishing Boats? Mar. Pol. 113, 4–17. https://doi.org/10.1016/j.marpol.2019.103796.

Andersen, J., 2005.The Inclusion of Stocks in Multi-species Fisheries: The Case of Danish Seiners. Mar. Resour. Econ. 20, 163–184. https://doi.org/10.1086/mre.20.2.42629467.

Campbell, H.F., Hand, A.J., 1998. Joint Ventures and Technology Transfer: the Solomon Islands Pole-and-line Fishery. J. Dev. Econ. 57, 421–442. https://doi.org/10.1016/S0304-3878(98)00095-9.

Chang, H.L., 2020. Economic impacts of Papahanaumokuakea Marine National Monument expansion on the Hawaii longline fishery. Mar. Pol. 115:103869. https://doi.org/10.1016/j.marpol.2020.103869.

Chavez Estrada, G., Quiroga Suazo, M., Dresdner Cid, J., 2018. The effect of collective rights-based management on technical efficiency: the case of Chile's common Sardine and Anchovy fishery. Mar. Resour. Econ. 33, 87–112. https://doi.org/10.1086/696130.

Chowdhury, N.K., Kompas, T., Kalirajan, K., 2010. Impact of control measures in fisheries management: evidence from Bangladesh´s industrial trawl fishery. Econ. Bull. 30(1), 765–773. http://www.accessecon.com/Pubs/EB/2010/Volume30/EB-10-V30-I1-P71.pdf .

Comitini, S., Huang, D., 1967. A Study of Production and Factor Shares in the Halibut Fishing Industry. J. Polit. Econ. 75, 366–372. https://doi.org/10.1086/259292.

Duy, N.N., Flaaten, O., 2016. Efficiency Analysis of Fisheries using Stock Proxied. Fish. Res. 181, 102–113. https://doi.org/10.1016/j.fishres.2016.04.006.

Eggert, H., 2001. Technical Efficiency in the Swedish Trawl Fishery for Norway Lobster. Working Papers in Economics. 53, Göteborg University. https://core.ac.uk/download/pdf/9125959.pdf

Ekerhovd, N-A., Gordon, D.V., 2013. Catch, Stock Elasticity, and an Implicit Index of Fishing Effort. Mar. Resour. Econ. 28(4), 379–395. https://doi.org/10.5950/0738-1360-28.4.379.

Felthoven, R., Morrison Paul, C., 2004. Multi-Output, Nonfrontier Primal Measures of Capacity and Capacity Utilization. Am. J. Agric. Econ. 86(3), 619-633.

Fousekis, P., Klonaris, S., 2003.Technical efficiency determinants for fisheries: a study of trammel netters in Greece. Fish. Res. 63, 85–95. https://doi.org/10.1016/S0165-7836(03)00019-5.

Gordon, D.V., 2014. The Endogeneity Problem in Applied Fisheries Econometrics: A Critical Review. Environ Resource Econ. 61, 115–125. https://doi.org/10.1007/s10640-013-9740-1.

Grafton, R.Q., Squires, D., Fox, K.J., 2000. Private Property and Economic Efficiency: A Study of a Common-Pool Resource. J. Law Econ. 43(2), 679–713. https://www.jstor.org/stable/10.1086/467469.

Hannesson, R., 1983. Bioeconomic production function in fisheries: theoretical and empirical analysis. Can. J. Fish. Aquat. Sci. 40, 968–982. DOI:10.1139/F83-123.

Harley, J.H., Myers, R.A., Dunn, A., 2001. Is catch-per-unit-effort proportional to abundance?. Can. J. Fish. Aquat. Sci. 58, 1760–1772. https://doi.org/10.1139/f01-112.

Herrero, I., Pascoe, S., 2003. Value *versus* Volume in the Catch of the Spanish South-Atlantic Trawl Fishery. J. Agric. Econ. 54(2), 325–341. https://doi.org/10.1111/j.1477-9552.2003.tb00066.x.

Herrero, I., Pascoe, S., 2004. Analysing the Effect of Technical Change On Individual Outputs Using Modified Quasi-Malmquist Indexes. J. Oper. Res. Soc. 55(10), 1081–1089. DOI:10.1057/palgrave.jors.2601787.

Jin, D., Thunberg, E., Kite-Powell, H., Blake, K., 2002. Total Factor Productivity Change in the New England Groundfish Fishery: 1964-1993. J. Environ. Econ. Manage. 44, 540–556. https://doi.org/10.1006/jeem.2001.1213.

Kirkley, J., Squires, D., Strand, I., 1998. Characterizing Managerial Skill and Technical Efficiency in a Fishery. J. Prod. Anal. 9(2), 145–160. https://www.jstor.org/stable/41769964.

Kleiber, P., Maunder, M., 2008. Inherent bias in using aggregate CPUE to characterize abundance of fish species assemblages. Fish. Res. 93, 140-145.

Manski, C., 1995. Identification Problems in the Social Sciences. Harvard University Press. Chapter 7.

Maunder, M.N., Sibert, J.R., Fonteneau, A. Hampton, J. Kleiber, P. Harley, S.J. 2006. Interpreting catch per unit effort data to assess the status of individual stocks and communities ICES J. Mar. Sci. 63, 1373–1385. DOI:10.1016/J.ICESJMS.2006.05.008.

Mulazzani, L., Malorgio, G., 2013. Regional management of multi-species fisheries on the basis of shared stocks and property rights: a Mediterranean case. Sci. Mar. 77(3), 439– 448. DOI:10.3989/scimar.03693.05B

Mulazzani, L., Manrique, R., Trevisan, G., Piccinetti, C., Malorgio, G. 2015. The relationship among catch, fishing effort, and measures of fish stock abundance: implications in the Adriatic Sea. Can. J. Fish. Aquat. Sci. 72(3), 410–421. DOI:10.1139/cjfas-2014-0337.

Myers, R.A., Worm, B., 2003. Rapid worldwide depletion of predatory fish communities. Nature. 423, 280–283. doi: 10.1038/nature01610.

Orea, L., Alvarez, A., Morrison-Paul, C., 2005. Modeling and Measuring Production Processes for a Multi-Species Fishery. Nat. Resour. Model. 18(2), 183–214. https://doi.org/10.1111/j.1939-7445.2005.tb00154.x.

Pascoe, S., Andersen, J.L., Wilde, J. de, 2001. The Impact of Management Regulation on the Technical Efficiency of Vessels in the Dutch Beam Trawl Fishery. Eur. Rev. Agric. Econ. 68(2), 187–206. https://doi.org/10.1093/erae/28.2.187.

Pascoe, S., Coglan, L., 2002. The Contribution of Unmeasurable Inputs to Fisheries Production: An Analysis of Technical Efficiency of Fishing Vessels in the English Channel. Am. J. Agric. Econ. 84(3), 585–597. https://doi.org/10.1111/1467-8276.00321.

Pascoe, S., Herrero, I., 2004. Estimation of a Composite Fish Stock Index Using Data Envelopment Analysis. Fish. Res. 69(1), 91–105. https://doi.org/10.1016/j.fishres.2004.03.003.

Pascoe, S., Robinson, C., 1996. Measuring Changes in Technical Efficiency over Time using Catch and Stock Information. Fish. Res. 28, 305–319. https://doi.org/10.1016/0165-7836(96)00502-4.

Richards, L.J., Schnute, J.T., 1986. An Experimental and Statistical Approach to the Question: Is CPUE an Index of Stock Abundance? Can. J. Fish. Aquat. Sci. 43, 1214–1227. https://doi.org/10.1139/f86-151.

Schaefer, M.B., 1954. Some aspects of the dynamics of population important to the management of the commercial marine fisheries. Bull. Inter-Am. Trop. Tuna Comm. 1(2), 27–56. https://doi.org/10.1016/S0092-8240(05)80049-7.

Sharma, K.R., Leung, P., 1998. Technical Efficiency of the Longline Fishery in Hawaii: An Application of a Stochastic Production Frontier. Mar. Resour. Econ. 13(4), 259–274. https://core.ac.uk/download/pdf/6522859.pdf.

Solís, D., Agar, J., del Corral, J., 2015. IFQs and total factor productivity changes: The case of the Gulf of Mexico red snapper fishery. Mar. Pol. 62, 347–357. https://doi.org/10.1016/j.marpol.2015.06.001.

Solís, D., del Corral, J., Perruso, L., Agar, J., 2014. Individual fishing quotas and fishing capacity in the US Gulf of Mexico red snapper fishery. Aust. J. Agric. Resour. Econ. 59, 288-307.

Squires, D., Kirkley, J., 1999. Skipper Skill and Panel Data in Fishing Industries. Can. J. Fish. Aquat. Sci. 56(11), 2011–2018. https://doi.org/10.1139/f99-135.

Suits, D., 1983. Dummy Variables: Mechanics vs. Interpretation. Rev. Econ. Stat. 66(1), 177–180. https://www.jstor.org/stable/1924713.

Torres, M., Felthoven, R., 2014. Productivity growth and product choice in catch share fisheries: The case of Alaska pollock. Mar. Pol. 50, 280–289. https://doi.org/10.1016/j.marpol.2014.07.008.

Wallace, I.F., Lindner, R.K., Dole, D.D., 1998. Evaluating stock and catchability trends: annual average catch per unit effort is an inadequate indicator of stock and catchability trends in fisheries. Mar. Pol. 22(1), 45–55. https://doi.org/10.1016/S0308-597X(97)00028-6.

Table 1. Estimates of the production function using CPUE as a stock proxy

Generated Effort	Generated Stock	Stock Variable	Constant (μ=1.0)	Stock (α=0.8)	Effort (β=0.2)
Time varying	Deterministic	True Stock	0.999	0.799	0.200
		CPUE	-0.310	1.004	0.201
	Stochastic	True Stock	1.000	0.800	0.199
		CPUE	-0.302	1.000	0.199
Time invariant	Deterministic	True Stock	0.999	0.799	0.201
		CPUE	-0.300	1.000	0.201
	Stochastic	True Stock	0.999	0.799	0.200
		CPUE	-0.298	1.000	0.200

Means of 1000 replications

Table 2. True values of the effects of the time dummies

Stock		Year 2	Year 3	Year 4	Year 5
Deterministic	Generated S_t	1.000	1.500	2.000	2.500
	Yearly effect	0.400	0.800	1.200	1.600
Stochastic	Generated S_t	9.728	2.051	8.794	9.397
	Yearly effect	5.300	-0.841	4.553	5.035

Means of 1000 replications

Table 3. Estimates of the production function with time dummies as stock proxies

Generated Effort	Generated Stock	Constant	Year 2	Year 3	Year 4	Year 5	Effort
Time varying	Deterministic	1.399	0.399	0.798	1.199	1.599	0.200
	Stochastic	3.482	5.301	-0.840	4.554	5.035	0.199
Time invariant	Deterministic	1.398	0.399	0.799	1.200	1.598	0.201
	Stochastic	3.482	5.300	-0.841	4.553	5.035	0.200

Means of 1000 replications

Table 4. Estimates of the production function with CPUE when inputs are correlated

Generated Stock		Constant (μ=1.0)	Stock (α=0.8)	Effort (β=0.2)
Corr(Z,S) = 0.20				
Deterministic	True Stock	1.000	0.799	0.200
	CPUE	0.040	0.970	0.194
Stochastic	True Stock	1.002	0.800	0.198
	CPUE	0.014	0.994	0.196
Corr(Z,S) = 0.50				
Deterministic	True Stock	1.000	0.799	0.200
	CPUE	0.059	0.950	0.194
Stochastic	True Stock	1.005	0.798	0.189
	CPUE	0.025	0.986	0.188
Corr(Z,S) = 0.80				
Deterministic	True Stock	1.000	0.799	0.200
	CPUE	0.083	0.923	0.198
Stochastic	True Stock	0.998	0.801	0.190
	CPUE	0.046	0.981	0.185

Means of 1000 replications

Table 5. Estimates of the production function with time dummies when inputs are correlated

Generated Stock	Constant	Year 2	Year 3	Year 4	Year 5	Effort
Corr(Z,S) = 0.20						
Deterministic	1.400	0.399	0.800	1.199	1.599	0.200
Stochastic	3.496	5.289	-0.861	4.537	5.029	0.199
Corr(Z,S) = 0.50						

Deterministic	1.400	0.399	0.799	1.198	1.599	0.200
Stochastic	3.473	5.230	-0.822	4.556	5.023	0.189
Corr(Z,S) = 0.80						
Deterministic	1.399	0.399	0.799	1.199	1.599	0.200
Stochastic	3.499	5.298	-0.867	4.540	5.046	0.188

Means of 1000 replications