PERSPECTIVES

PHOENIX CENTER FOR ADVANCED LEGAL & ECONOMIC PUBLIC POLICY STUDIES

The Vanishing Benefits of Fair Use: A Review of the Flynn-Palmedo Study on "User Rights" in Copyright Law

Dr. George S. Ford*

December 18, 2017

Introduction

After reviewing the copyright laws of twenty-one nations, Sean Flynn and Michael Palmedo—scholars at the American University Washington College of Law's Program on Information Justice and Intellectual Property ("PIJIP")—conclude in their recently released paper, *The User Rights Database: Measuring the Impact of Copyright Balance*, that "[f]ew countries, and almost no developing countries, have sufficient user rights most needed to support the digital economy."[1] The authors then gather and analyze data from twenty-one nations in search of a positive link between their own index of "user rights"—including fair use and safe harbors—and a few economic and scholastic outcomes.[2] Finding a correlation, the authors conclude that governments should weaken artists' and authors' copyright protections in order to increase the average size and profits of firms and economic sectors that, they suppose without explanation, profit from using others work without permission or compensation.[3]

In this PERSPECTIVE, I provide a review of the *Flynn-Palmedo Study*, focusing on the statistical analysis and construction of the "user rights" index. Like earlier studies on the economic benefits of fair use and "user rights," including Gibert (2015) and Ghafele and Gibert (2014), the statistical results of the *Flynn-Palmedo Study* are merely the consequence of basic errors in both the design and implementation of the empirical analysis, rendering spurious correlations.[4] As such, the *Flynn-Palmedo Study* is not relevant for policymaking. The paper requires substantial revision to address numerous and substantial errors.

My review begins by showing that the positive and statistically significant relationships touted by the *Flynn-Palmedo Study* vanish when standard methods replace their own poorly specified and improperly estimated models. Then, I address in less detail a number of statistical problems and inconsistencies, as well as problems with the authors' construction of the "user rights" index. Though wide-ranging, my review of the *Flynn-Palmedo Study* is not exhaustive; the errors are too numerous.

Though wide-ranging, my review of the Flynn-Palmedo Study is not exhaustive; the errors are too numerous.

Empirical Analysis

The *Flynn-Palmedo Study* is, in part, an introduction of the User Rights Database, a database on the exceptions and limitations in the copyright laws of nations.[5] From this database, the authors construct a single index of "user rights" by averaging the ordinal responses to over one hundred questions related to varying aspects of copyright laws. I discuss the obvious errors in this procedure in more detail later, but for now let's label the index of "user rights" as *OPEN*.

Having constructed the *OPEN* variable, the *Flynn-Palmedo Study* aims to test whether countries with greater "user rights," or higher values of *OPEN*, have different economic and scholastic outcomes. Regression analysis is used, and all the models in the *Flynn-Palmedo Study* employ the same basic format, which is as follows: The outcome of interest, *Y*, is regressed on the *OPEN* variable, along with a few other variables, and a positive and statistically significant coefficient on the *OPEN* variable is viewed as a "positive outcome."

Critically, even if the correlations withstand basic scrutiny, a mere correlation is uninformative for public policy—only causal relationships are useful for policy making, and the Flynn-Palmedo Study offers no plausible attempt to estimate a causal relationship.

The specification of these regression models is entirely ad hoc. No conceptual framework, formal or otherwise, is provided to explain why a positive correlation between the authors' chosen *Y* variables and *OPEN* variable is desirable. Nor is an identification strategy provided that would permit a causal interpretation to the results.[6] In fact, the authors offer strong evidence of selection bias (i.e., "user rights" are not randomly assigned), precluding a causal interpretation. My review of the *Flynn-Palmedo Study* does not aim to solve the numerous and severe problems to render valid statistical tests of the tradeoff between "creator rights" and "user rights," but instead demonstrates the consequences of these errors on the results reported in the study. Critically, even if the correlations reported in the study withstand basic scrutiny, a mere correlation is uninformative for public policy—only causal relationships are useful for policy making, and the *Flynn-Palmedo Study* offers no plausible attempt to estimate a causal relationship.

Regression Equation

The *Flynn-Palmedo Study* begins with panel data on twenty-one countries covering sixteen years. Considering missing data and other factors, the final sample sometimes involves an unbalanced panel with fewer than all the countries.

For a parsimoniously specified model and panel data (including multiple years of data for multiple countries) over this time period, the most obvious estimation procedure is to start with a two-way fixed effects regression with inference conducted using clustered standard errors.[7] A generalized specification of the estimated regression model somewhat consistent with that of the *Flynn-Palmedo Study* is:

$$\begin{aligned}\ln y_{i,t} &= \beta_0 + \beta_1 OPEN_{i,t} + \beta_2 \ln GDPCAP_{i,t} \\ &\quad + \beta_2 \ln POP_{i,t} + \lambda_i + \theta_t + \varepsilon_{i,t}\end{aligned}$$

where "ln" indicates the natural log transformation, $y_{i,t}$ is the outcome of interest for a particular sector in country *m* at time *t* (or scholarly output for the country), *OPEN* is the openness index for country *i*, *GDPCAP* is per-capita *GDP*, *POP* is the population, and ε is the econometric disturbance term. The two-way fixed effects model includes coefficients for the country fixed effects (λ_i) and yearly fixed effects (θ_t).

The purpose of these fixed effects is to measure unobservable heterogeneity across countries and time that influence the dependent variable, thereby avoiding spurious correlations caused by, for instance, size differences across the countries or, say, inflation or broad economic shocks (like global recessions) over time. For the country fixed effects, this unobserved heterogeneity is assumed to be inter-temporally constant, so the dummy variable is a sufficient means to measure them. Year fixed effects account for temporal variations common to all the countries in the sample, and again a dummy variable is up to the task.

Given the panel data and the serial correlation in the time series component, statistical tests should be based on clustered standard errors. Research shows that conducting statistical tests using the regular or robust standard errors often leads to significant over-rejection of the null hypothesis when the data contains groups.[8] In this PERSPECTIVE, I use clustered standard errors unless otherwise indicated. The *Flynn-Palmedo Study*, alternately, uses robust standard errors, which do not account for group effects and are likely understated, thus producing statistically significant results too often.[9]

Overall, the empirical work has a willy-nilly, unprofessional feel to it.

As demonstrated here, a key problem with the *Flynn-Palmedo Study* is the failure to include the country fixed effects (that is, the λ_i in the equation above are assumed to be zero).[10] As a result, the large size and economic development differences among the countries in the sample are nearly sure to produce spurious correlations between the *OPEN* variable and other economic outcomes. The problem is particularly acute in the *Flynn-Palmedo Study* because the *OPEN* variable is larger for more developed countries (see Figure 1 in the *Flynn-Palmedo Study*). Not only does this imply selection bias, but it also ensures correlations between the *OPEN* variable and the scales of economic outcomes. The fixed effects, in practical terms, "de-scales" the data (through mean centering), thereby reducing (though not eliminating) the risk of spurious results. I will demonstrate the problem now.

Why Fixed Effects?

Before demonstrating the consequences of poor model specification, a review of the purpose of fixed effects may be useful. Say you are interested in the demand for pizza. You have data from three cities and by the law of demand expect that the quantity of pizza is inversely related to its price. You look at data from a single year and observe the following:

	Quantity (Q)	Price (P)
Chicago	200	18
San Francisco	150	15
Atlanta	100	12

where the quantity of pizza consumed rises in its price—an unexpected result.[11] An unskilled researcher might conclude the data rejects the law of demand. A possibly better explanation for this peculiar result is that people in Chicago really love pizza and demand more expensive, higher quality fare. Maybe San Francisco has a high demand for "organic" ingredients, which raises the cost of pizza. Whatever the reasons, given the law of demand it is sensible to presume there is *unobserved heterogeneity* across the cities.

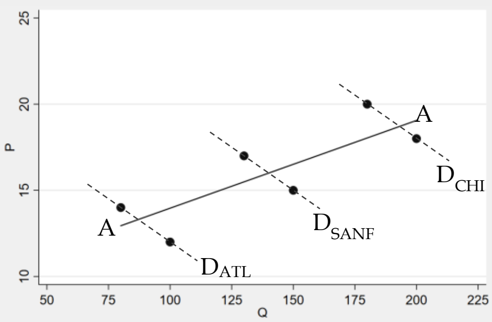
If we expand the data to include an additional year of data, we have:

	$(Q,P)_0$	$(Q,P)_1$
Chicago	(200, 18)	(180, 20)
San Francisco	(150, 15)	(125, 17)
Atlanta	(100, 12)	(85, 16)

where the consumption patterns begin to make sense. Within Chicago, for instance, 200 pizzas are consumed at a price of \$18 but only 180 pizzas are consumed at a higher price of \$20. Price and quantity are inversely related. In fact, the quantity consumed is inversely related to price for all three cities. Seeing this *within* variation among variables is the purpose of fixed effects regression. By accounting for the unobserved heterogeneity across the sampled units, a *causal interpretation* to relationships—

rather than just a spurious one—becomes more plausible.

Figure 1. Why Fixed Effects?



A graphical analysis may be helpful. The data on pizza prices and sales from above is illustrated in Figure 1. Naively fitting a line to the data points (using least squares regression as in the *Flynn-Palmedo Study*) produces the upward sloping line labeled AA (with coefficient 0.051). This positive relationship does not measure the true relationship between prices and quantities. What we know from the table above is that the actual relationship between price and quantity is illustrated by the downward sloping lines (demand curves) labeled D for each city. Accounting for fixed effects (adding a dummy variable for each city to the regression), the regression of P on Q produces a coefficient of -0.1, which is equal to the parameter used to create the data in the table. Figure 1 illustrates the basic problem with the *Flynn-Palmedo Study*, and I will employ this simple figure below using data from that study.

Spurious correlations, particularly between variables over time, appear everwhere. The divorce rate in Maine, for instance, is positively related to the consumption of margarine.[12] And, one herbal supplement company claims that root canals cause cancer because a doctor reported that 97% of his cancer patients have had the procedure.[13] Sound statistical methods typically eliminate evidence of such spurious correlations. As detailed below, the *Flynn-Palmedo Study* might become another popular example of spurious correlation, but one based on the failure to use fixed effects to account for unobserved heterogeneity across countries.

Spurious Results from Poor Specification

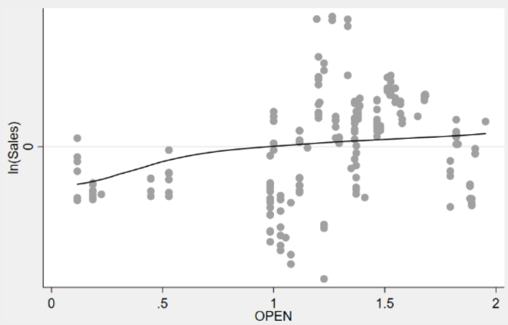
In order to demonstrate the effects of the mis-specification of the model, I replicate the dataset used for the regression results reported in Tables 3 and 5 of the *Flynn-Palmedo Study*. Data on multinational corporations and their affiliates is freely available from the U.S. Bureau of Economic Analysis ("BEA") and scholarly output is available from SCImago, so replication is mostly straightforward.[14] GDP and population data, as well as some data on other variables used later in parts of my analysis, is from the World Bank.[15]

In the *Flynn-Palmedo Study*, the dependent variables from the BEA data include the Total Sales, Net Income, and Value Added for the affiliates of U.S. (majority owned) multinational corporations for NAICS 54 ("Professional, Scientific, and Technical Services") in the countries of interest. No explanation is given as to why this sector is chosen or why it would be affected much by fair use, or why increasing the sectors profits suggests expanding fair use is good policy. For scholarship, the *Flynn-Palmedo Study* uses data on "citable documents" and the "H-Index," which is based on citable documents. For replication purposes, I limit my attention to Total Sales from the BEA data and citable documents from the SCImago data.

Affiliate Total Sales

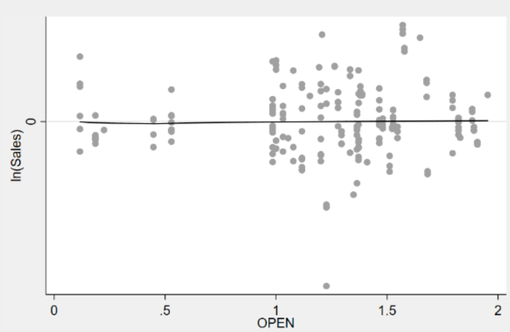
It is perhaps easiest to see the effect of excluding the country fixed effects on the *Flynn-Palmedo Study* results using a visual analysis like that above. In Figure 2, the data on multinational affiliate sales is illustrated, where the data has been pre-filtered for year fixed effects and a line fit to the data.[16] This figure is not unlike Figure 4 in the *Flynn-Palmedo Study*, though here the relationship between sales and the *OPEN* variable is not required to be linear and the data is pre-filtered for time. Figure 2 shows a positive relationship between aggregate affiliate sales and the *OPEN* variable, which is consistent with that reported in the *Flynn-Palmedo Study*.

Figure 2. Affiliate Total Sales
(Time Fixed Effects)



The data in Figure 2 has not been adjusted for the large differences in the average sales across countries using country fixed effects. Figure 3 illustrates the data after filtering it for the both the country and time fixed effects. The positive relationship has vanished. The figure shows that the positive effect reported in the *Flynn-Palmedo Study* is likely a product of model misspecification and that the reported correlation is spurious.

Figure 3. Affiliate Total Sales
(Time and Country Fixed Effects)



In Table 1, I summarize the results of the regression analysis, including and excluding the country fixed effects. Column (1) of the table is the Flynn-Palmedo specification—no country fixed effects and robust errors. In Column (2), the same model is used (and the coefficients unchanged) but the reported standard errors account for the clustering of the data. Finally, in Column (3), the results are from a model including both country fixed effects and clustered standard errors.

Table 1. Affiliate Sales

	(1)	(2)	(3)
OPEN	0.583^{a}	0.583^{b}	0.516
	(0.104)	(0.232)	(0.417)
ln*GDPCAP*	1.105^{a}	1.105^{a}	1.00
	(0.102)	(0.251)	(0.671)
ln*POP*	-0.764^{a}	-0.764^{a}	-3.20
	(0.052)	(0.142)	(5.24)
Time FE	Yes	Yes	Yes
Cross FE	No	No	Yes
SE	Robust	Clust.	Clust.
N	190	190	190

Stat. Significance: ^{a} 1% , ^{b} 5%, ^{c} 10%.

Before discussing the results, we can check whether the country fixed effects are unnecessarily included in the regression model. The null hypothesis of the test is that the cross section fixed effects are redundant, and the F-Statistic of the test is 21.4, so the null hypothesis is rejected at better than the 1% level.[17] The cross section fixed effects are omitted variables in the *Flynn-Palmedo Study* model (either with or without the other regressors). The null of Wooldridge's panel serial correlation test ("no serial correlation") is also rejected at better than the 1% level.[18] So, statistical tests indicate that both the country fixed effects and clustered standard errors are called for with this data.

Without either the country fixed effects or clustered standard errors (the mis-specified Flynn-Palmedo model), the coefficient on the *OPEN* variable is positive and statistically significant (see Column 1), as reported in the *Flynn-Palmedo Study*. In Column 2, we see that switching to clustered standard errors does not render the *OPEN* variable statistically insignificant. However, as expected, the clustered standard error is about twice as large as the robust standard error, portending the standard problem of over-rejection from a failure to account for clustering.

With cross section fixed effects and clustered errors, the coefficient on the *OPEN* variable is not statistically different from zero (Column 3). The positive and statistically significant relationship between affiliate sales and the *OPEN* variable disappears. Plainly, the positive relationship reported in the *Flynn-Palmedo Study* is spurious—the result of an improperly estimated model (at a minimum), as the visual analysis above illustrates.

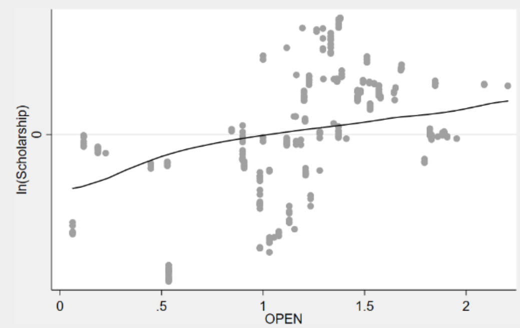
Critically, the mere inclusion of the country fixed effects and the use of clustered standard errors, while eliminating the statistically significant effect of the *OPEN* variable, does not remedy all the problems with the models in the *Flynn-Palmedo Study*. All of their models are ad hoc, quite parsimonious, and suffer from selection bias through the *OPEN* variable. My analysis here merely demonstrates that the results of the *Flynn-Palmedo Study* are flimsy, vanishing upon the most basic adjustments to model specification and estimation.

Critically, the mere inclusion of the country fixed effects and the use of clustered standard errors ... does not remedy all the problems with the models in the Flynn-Palmedo Study. All of their models are ad hoc, quite parsimonious, and suffer from selection bias.

Citable Research Papers

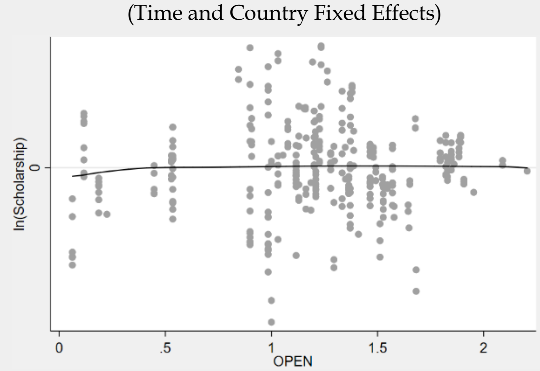
Another outcome analyzed in the *Flynn-Palmedo Study* is scholarly output, measured as the number of citable documents published by researchers living in a country. Repeating the analysis above, Figure 4 shows the relationship when ignoring the country fixed effects. A positive relationship is found once more, as reported in the *Flynn-Palmedo Study*.

Figure 4. Scholarly Output
(Time Fixed Effects)



The data filtered for the time and country fixed effects is illustrated in Figure 5. As before, the positive relationship vanishes once the country fixed effects are included.

Figure 5. Scholarly Output
(Time and Country Fixed Effects)



Regression analysis confirms what the figures demonstrate. Testing the citable documents model for redundant fixed effects, the F-Statistic is 978, with a probability well below 1%. Serial correlation is present ($F = 26.7$, prob < 0.01), so clustering (or some other remedy) is required.

In Column 1 of Table 2, we see the results of the Flynn-Palmedo model—the coefficient on the *OPEN* variable is positive and statistically significant. Unlike the Total Sales data above, for citable documents simply switching to clustered standard errors renders the *OPEN* variable statistically insignificant, as shown in Column 2 of the table. Once the fixed effects and clustered standard errors are used, none of the regressors is statistically different from zero, and the δ coefficient shrinks considerably.

Table 2. Scholarly Output

	(1)	(2)	(3)
OPEN	0.778[a]	0.778	0.232
	(0.164)	(0.617)	(0.175)
ln*GDPCAP*	0.796[a]	0.796[a]	0.406
	(0.056)	(0.204)	(0.249)
POP	3.4e-09[a]	3.4e-09[a]	2.7e-09
	(1.4e-10)	(4.9e-10)	(1.9e-09)
Time FE	Yes	Yes	Yes
Cross FE	No	No	Yes
SE	Robust	Clust.	Clust.
N	319	319	319

Stat. Significance: [a] 1% , [b] 5%, [c] 10%.

Table 2 illustrates another careless aspect of the *Flynn-Palmedo Study*. For most of the regression models, the combination of per-capita GDP (*GDPCAP*) and population (*POP*) are intended by the authors to control for the "size of the national markets."[19] For the firm-level regressions, employment is also included (*EMP*) as a regressor. In every regression, the dependent variable is subjected to the natural log transformation. For unexplained reasons, the treatment of the right-hand side variables is inconsistent.

In the *Flynn-Palmedo Study*'s firm-level regressions (Table 1 and 4), the natural log transformation applies to the dependent variable and the *EMP* variable, but the included *GDPCAP* and *POP* variables are not transformed. For the regressions summarized in Table 3 of the *Flynn-Palmedo Study* (on the multinational affiliates), the authors apply the log transformation to both *GDPCAP* and *POP*. Then, when looking at scholarly output, *GDP* is logged but *POP* is not. No explanation is provided for the inconsistent treatment of the regressors. There is no theoretical reason for it, and no statistical analysis is provided to support the treatment. In fact, statistical analysis suggests the natural log transformation of *POP* is preferable.[20]

Also, while I do not evaluate the Net Income (i.e., profit) regressions, the authors err in applying the natural log transformation to this data. Net income can be negative, and from what I can tell often is in the sample. The natural log transformation is not applicable to non-positive numbers, though Flynn and Palmedo apply it anyway and thus eliminate valid data from the sample. Overall, the empirical work has a willy-nilly, unprofessional feel to it.

More Spurious Results

As anyone familiar with regression analysis knows, ignoring unobservable size effects when using panel data can render all sorts of spurious relationships (e.g., the pizza discussion above). In fact, the analysis above indicates that all the results reported in the *Flynn-Palmedo Study* and replicated here are spurious. Such spurious outcomes are not limited to outcomes studied in the *Flynn-Palmedo Study*.

Table 3. More Spurious Relationships

	(1)	(2)
Fisheries Production	-0.67[a]	-0.04
	(0.27)	[0.20]
Rainfall	-0.40[a]	…
	(0.13)	
Gas Prices	0.18[a]	0.05
	(0.05)	[0.07]
Birth Rate	-2.47[a]	0.42
	(0.45)	[0.67]
Tourism Expenditures	0.22[a]	-0.10
	(0.07)	[0.13]
Rural Population	9.87[a]	-0.58
	(2.20)	[0.91]
Time FE	Yes	Yes
Cross FE	No	Yes
SE	Robust	Clust.

Stat. Significance: [a] 1% , [b] 5%, [c] 10%.

In Table 3, the relationship between a number of outcomes in the sample countries and the *OPEN* variable is illustrated. The first column includes time fixed effects and the second both time and country fixed effects. As shown in Column 1, the lack of the country fixed effects and the use of robust standard errors leads to statistically significant relationships (at the 1% level or better) between the *OPEN* variable and fisheries production, gas prices, rainfall, the birth rate, tourism expenditures, and rural population.[21] None of these spurious results survives a better-specified model, as shown in Column 2.

The BEA data also permits a look at other industry sectors, some of which have no plausible nexus to fair use. Take the mining industry, for instance. Using the Flynn-Palmedo model, the coefficient on the *OPEN* variable is -1.25 in the total sales regression, which is statistically different from zero at better than the 1% level (using robust errors). Adding in the country fixed effect makes this spurious result disappear. The same pattern is seen for the chemicals industry, for which copyright plays little role, with a statistically significant effect on total sales without country fixed effects but a statistically insignificant effect with them.

The Openness Score (OPEN)

Given the analysis above, it is clear that the *Flynn-Palmedo Study* should be dismissed on statistical errors alone. Still, a brief analysis of other problems, some of a more conceptual nature, is worthwhile. To begin, let's look at the construction of the *OPEN* variable.

Empirical research on the effects of copyright law is very limited. Partly to blame is the complexity of copyright law, which is a combination of many interdependent statutory provisions often mixed with conflicting judicial opinions and inconsistent enforcement. Each nation's copyright regime is unique, complicating the specification and testing of copyright's impact on economic outcomes.

Naturally, researchers and policy advocates would like to condense the complexity of copyright law into a single index that could be evaluated using standard statistical methods. It is a hopeless endeavor. Nonetheless, it has been tried.

The *Flynn-Palmedo Study* is a second attempt to create an index of copyright law's "openness" or "flexibility" in an attempt to get policymakers to expand copyright's exceptions and limitations. Gibert's 2015 study, published by the Lisbon Council, also constructs such an index and claims to document some positive effects of openness on broad economic outcomes.[22] In an earlier and detailed PERSPECTIVE, I revealed Gibert's study to be "a showcase of methodological blunder."[23] Despite the obvious and well-documented mistakes in the Gibert study, the Lisbon Council did not retract it or repair it.[24]

In many ways, the *Flynn-Palmedo Study* follows in the footsteps of Gibert (2015). The first task of the *Flynn-Palmedo Study* is to construct a measure of the breadth and scale of exceptions to copyright protection—an Openness Score. They do so using the User Rights Database maintained by PIJIP, which relies on a survey of "friends" in the Global Expert Network on Copyright User Rights interpreting the copyright laws of different nations.[25]

The survey seeks responses on the application, and the degree of application, of twenty-one limitations and exceptions to copyright over the period 1970 through 2016.[26] Strength is measured (we are told) on a four-point ordinal scale:

(a) Not Included;

(b) Mostly/Probably Not Included;

(c) Mostly/Probably Included; and

(d) Clearly Included.

In the *Flynn-Palmedo Study*, these categorical answers are arbitrarily assigned a numeric value ranging from 0 to 3, respectively. With over 100 survey questions on the 21 topics, the Openness Score is then constructed as an "unweighted average of the coded answers for each year."[27] If a country's score is 3.0, then every user right is clearly included in the law.

Clearly, this approach to construct the Openness Score is problematic. First, the survey is an experiment in itself, and the need to converse with respondents (as the authors' indicate they did) suggests there may be some divergences in the interpretation of the questions and of copyright law.[28] Ideally, multiple responses to the survey for each country could be compared to check for variations in respondent interpretations of the statutes.

Second, the claim that the responses are given values of 0, 1, 2, or 3 is untrue. In some cases, the scores are 1.5 and 2.5, illustrating potential problems with the survey design and interpretation. These answers are inconsistent and incompatible with the survey design, and indicate that the respondents are somewhat confused by the survey. The unacceptable results also indicate carelessness in the construction of the database and the *OPEN* variable.

Third, the responses to the survey questions are *ordinal* in nature; the numerical values assigned by Flynn and Palmedo are purely arbitrary. No harm is done in replacing a 0-3 response with an A-D response, with the exception being you cannot take an average of letters and insert that into a regression model. Or, the responses could be coded 0, 10, 100, and 110, perhaps to reflect the non-linear scale of the response categories.

Taking a simple mean of this large number of varied questions is dubious, at best. There is significant debate about computing descriptive statistics of an ordinal scale and using them in a parametric regression, and there are rules to follow when doing so.[29] Despite obvious theoretical problems, numerical calculations of ordinal responses is sometimes done under the assumptions of "linear approximation" and normality, and the median is often preferred to the mean. That said, these assumptions and choice of statistic warrant some attention, especially at the design phase, and some testing, neither of which is done in the *Flynn-Palmedo Study*.

Linearity seems clearly unsatisfied—a move from "Mostly Not" to "Mostly So," a huge difference, is not the same as a move from "Mostly So" to "Cleary," which is quite small in contrast. The fact that some intermediate values (e.g., 1.5 and 2.5) are entered by the respondent "friends" indicates some confusion in this regard and a dissatisfaction by respondents with the four-point ordinal scale. Also, it's the intermediate values point to careless handling of the data by Flynn and Palmedo.

For many reasons, the Flynn-Palmedo Study's use of the User Rights Database to construct the OPEN variable is improper. But the real problem is that there is no single index of "openness." Errors will accompany any attempt to create one.

The authors also average across responses to items covering many different topics, including the disparate concepts of fair use and safe harbor. While practitioners hold that it is generally best to average over multiple items measuring the same thing when treating ordinal data as cardinal (i.e., thus moving from a "Likert item" to a "Likert scale"), the items averaged over should reflect the same outcome. Yet, the differences across topics in the User Rights Database are many; a fact recognized by the authors.

For instance, the authors distinguish between limitations and exception relevant to the "digital economy" and limitations and exceptions relevant for the analog economy. And, some questions go to safe harbor and some to fair use. The incentives created by broad fair use and no safe harbor might be very different than those from no fair use and broad safe harbor, or between the rights for news reporting and data mining. Finally, from a pure mathematical perspective, asking more questions about one type of user right than another dilutes the contribution of the latter to the average across all responses.

For many reasons, the *Flynn-Palmedo Study's* use of the User Rights Database to construct the *OPEN* variable is improper. But the real problem is that there is no single index of "openness." Errors will accompany any attempt to create one.

Selection Bias

The *Flynn-Palmedo Study* provides no theoretical motivation (formal or informal) for its empirical models—the regression models are entirely ad hoc. Nor do the authors offer an identification strategy for quantifying a causal effect (rather than mere correlation). For there to be a causal effect, the treatment—here, the *OPEN* variable—must be independent of the potential outcomes.[30] Yet, the set of "user rights" in a nation's copyright law arises over time for many reasons including, perhaps, broader economic activity. If the causal variable of interest is not randomly assigned (or as good as randomly assigned once all relevant factors are accounted for), then there is a selection bias and the estimated coefficients of the regression do not measure true effects.

In fact, Flynn and Palmedo present strong evidence of selection bias.[31] In discussing the Openness Score, the authors note,

> [t]he high income countries in our study have more open user rights in their laws, and the gap between them and developing countries has been growing since the 1990's. As one participant from a developing country at a workshop of ours remarked on seeing the data, "we [developing countries] are 30 years behind."[32]

The statistical implications of this acknowledgement are quite substantial. Figure 1 of the *Flynn-Palmedo Study* provides good evidence that this selection problem is present.

The *Flynn-Palmedo Study* looks at the effects of "openness" on economic outcomes, and the data show a systematic disparity between the Openness Score and economic development. When the outcomes of interest also establish the treatment, which Flynn and Palmedo argue they do, the regression disturbance and the *OPEN* variable are correlated. This selection bias ensures that the coefficients on the *OPEN* variable are a biased measure of its effect on the outcome (that is, the estimated coefficients do not measure the true effect).[33]

"User Rights" and Market Power?

While Flynn and Palmedo claim their results speak to "innovation and creativity," they do not.[34] In fact, the outcomes studies include mostly the revenues and profits of firms presumably exploiting copyright's exceptions. I have no doubt that weaker copyright enforcement will increase the size and profits of some firms, and that these firms will encourage governments to expand fair use and safe harbor.

Whether increasing average firm size and raising industry profits is a good public policy is entirely left to the reader to determine. The *Flynn-Palmedo Study* offers no guidance as to why these outcomes should be pursued or how they correlate with general social well being. In fact, a plausible interpretation of the *Flynn-Palmedo Study* is that expansive "user rights" increase the market power of some firms. While these results are claimed to be "positive outcomes," expanding market power through public policy is generally thought to be undesirable. Why the authors describe these findings as a "positive outcome" is testament to the mostly thoughtless approach to their empirical research.

…a plausible interpretation of the Flynn-Palmedo Study is that expansive "user rights" increase the market power of some firms. While these results are claimed to be "positive outcomes," expanding market power through public policy is generally thought to be undesirable.

Conclusion

Many nations are reviewing their copyright laws, but prudent changes to these laws requires sound theoretical arguments and robust empirical research. The *Flynn-Palmedo Study* is neither. First, it offers no theoretical basis for its regression models. The choice of outcomes is ad hoc, and the measure of "user rights" improperly constructed. Second, the regression analysis does not meet minimum professional standards. Improper specification of the model and selection bias are apparent, a combination that leads invariably to spurious and biased results. Indeed, Flynn and Palmedo's Openness Score is shown to be positively correlated with all sorts of economic and social outcomes across the countries in the sample, including the amount of rain and the birth rate.

In light of the flimsy results, the Flynn-Palmedo Study is not relevant for policymaking. In fact, given the magnitude of the errors, the draft study should be retracted and substantially revised.

With even marginal improvements in model specification, which cannot and do not resolve all the defects in the regression analysis, the results of the study vanish. In light of the flimsy results, the *Flynn-Palmedo Study* is not relevant for policymaking. In fact, given the magnitude of the errors, the *Flynn-Palmedo Study* should be retracted and substantially revised.

NOTES:

* **Dr. George S. Ford** ***is the Chief Economist of the Phoenix Center for Advanced Legal and Economic Public Policy Studies. The views expressed in this Perspective do not represent the views of the Phoenix Center or its staff. Dr. Ford may be contacted at ford@phoenix-center.org.***

1 S. Flynn and M. Palmedo, *The User Rights Database: Measuring the Impact of Copyright Balance*, Draft Paper, Program on Information Justice and Intellectual Property (2017) (available at: http://infojustice.org/wp-content/uploads/2017/11/Flynn-and-Palmedo-v1.pdf) at p. 17 (hereinafter the "*Flynn-Palmedo Study*"); "User Rights" database available at: http://infojustice.org/survey.

2 *Id*.

3 The authors are using the draft paper in an attempt to influence the NAFTA negotiations with the aim of broadening fair use and safe harbor policies (http://infojustice.org/wp-content/uploads/2017/11/Washington-Principles-on-Copyright-Balance-in-Trade-Agreements-November-15-2017.pdf).

4 G.S. Ford, *The Lisbon Council's 2015 Intellectual Property and Economic Growth Index: A Showcase of Methodological Blunder*, PHOENIX CENTER POLICY PERSPECTIVE No. 15-03: (June 29, 2015) (available at: http://phoenix-center.org/perspectives/Perspective15-03Final.pdf); G.S. Ford, *The Economic Impact of Expanding Fair Use in Singapore: More Junk Science for Copyright Reform*, PHOENIX CENTER POLICY PERSPECTIVE No. 16-01 (February 16, 2016) (available at: http://phoenix-center.org/perspectives/Perspective16-01Final.pdf).

5 *Supra* n. 1.

6 *See*, *e.g.*, G.W. Imbens and J.M. Wooldridge, *Recent Developments in the Econometrics of Program Evaluation*, 47 JOURNAL OF ECONOMIC LITERATURE 5-86 (2009).

7 *See, e.g.,* A.C. Cameron and D.L. Miller, *A Practitioner's Guide to Cluster-Robust Inference*, 50 JOURNAL OF HUMAN RESOURCES 317-372 (2015) (draft available at: http://cameron.econ.ucdavis.edu/research/Cameron_Miller_JHR_2015_February.pdf); J.M. Wooldridge, ECONOMETRIC ANALYSIS OF CROSS SECTION AND PANEL DATA (2010), at Chs. 10-11; B.H. Batalgi, ECONOMETRIC ANALYSIS OF PANEL DATA (2013); J.D. Angrist and J. Pischke, MOSTLY HARMLESS ECONOMETRICS (2009), at Ch. 8.

8 Cameron and Miller, *id.*; Bertrand, E. Duflo, and S. Mullainathan, *How Much Should We Trust Differences-in-Differences Estimates?*, NBER WORKING PAPER No. 8841 (March 2002) (available at: http://www.nber.org/papers/w8841); J.D. Angrist and J. Pischke, MASTERING 'METRICS: THE PATH FROM CAUSE TO EFFECT (2015) at Ch. 5.

9 Even clustered errors may lead to over-rejection of the null hypothesis (of "no effect") when there are few clusters (countries in this case), which may be a problem here.

10 It is not always the case that fixed effects for the cross sections are required, but standard procedure is to include them unless theory or statistical tests indicate otherwise. Random effects for the cross section may be preferred in some cases, but tests indicate not for this data.

11 This example is borrowed from Matt Bogard with some modifications (available at: http://econometricsense.blogspot.com/2014/04/intuition-for-fixed-effects.html?m=1).

12 http://www.tylervigen.com/spurious-correlations.

13 https://www.snopes.com/medical/toxins/rootcanal.asp.

14 Bureau of Economic Analysis, U.S. Direct Investment Abroad (USDIA) (https://www.bea.gov/international/di1usdop.htm; https://www.scimagojr.com).

15 https://data.worldbank.org.

16 The filtered data is the residual of a regression of the dependent variable on year fixed effects, thus demeaning the data over time. The figure is of the kernel density function using the Epanechnikov kernel with a bandwidth of 0.5 and degree zero.

NOTES CONTINUED:

17 The critical F(14, 157) at the 1% level is 2.20. Random effects are also rejected by the Wooldridge Test at the 1% level. Wooldridge (2010), *supra* n. 7 at pp. 290-1.

18 *Id*., at pp. 282-3; *see also* D.M. Drukker, *Testing for Serial Correlation in Linear Panel-Data Models*, 3 STATA JOURNAL 168-177 (2003) (available at: http://www.stata-journal.com/sjpdf.html?articlenum=st0039).

19 *Flynn-Palmedo Study*, *supra* n. 1 at p. 20.

20 The R^2 is larger and the Akaike Information Criteria ("AIC") is smaller with ln*POP* instead of *POP*.

21 All data is provided by the World Bank (https://data.worldbank.org). All regressors except for the birth rate and rural population (as a percent of total population) are expressed as natural logs. Robust standard errors are used for the models summarized in Column 1 to match the *Flynn-Palmedo Study*. Given a lack of data, the rainfall regression cannot be estimated with fixed effects.

22 B. Gibert, *2015 Intellectual Property and Economic Growth Index: Measuring the Impact of Exceptions and Limitations in Copyright on Growth, Jobs and Prosperity*, Lisbon Council (May 2015) (available at: http://www.lisboncouncil.net/publication/publication/122-the-2015-intellectual-property-and-economic-growth-index.html).

23 *Id*.

24 *Id*.

25 *Flynn-Palmedo Study*, *supra* n. 1 at pp. 10-11.

26 The survey is available at: http://infojustice.org/survey.

27 *Flynn-Palmedo Study*, *supra* n. 1 at p. 14.

28 *Id.* at p. 13.

29 The debate is lengthy, though suitably outlined in S. Westland, *The Dangers of Likert Scale Data*, COLOURWARE.ORG (February 18, 2014) (available at: http://colourware.org/2014/02/18/the-dangers-of-likert-scale-data). *See also,* I.E. Allen and C.A. Seaman, *Likert Scales and Data Analyses*, QUALITY PROGRESS (July 2007) (available at: http://asq.org/quality-progress/2007/07/statistics/likert-scales-and-data-analyses.html); G.M. Sullivan and A.R. Artino, Jr., *Analyzing and Interpreting Data From Likert-Type Scales*, 5 JOURNAL OF GRADUATE MEDICAL EDUCATION 541-542 (2013); G. Normal, *Likert Scales, Levels of Measurement and the "Laws" of Statistics*, 15 ADVANCES IN HEALTH SCIENCES EDUCATION 625–632 (2010).

30 For an analysis in the context of regression, *see, e.g.,* MOSTLY HARMLESS ECONOMETRICS, *supra* n. 7 at Ch. 3.

31 *Flynn-Palmedo Study*, *supra* n. 1 at Figure 1.

32 *Id.* at p. 15.

33 Imbens and Wooldridge, *supra* n. 6.

34 *Flynn-Palmedo Study*, *supra* n. 1 at p. 4.