DEAKIN UNIVERSITY AUSTRALIA

Faculty of Business and Law

School of Accounting, Economics and Finance

School Working Paper

**ECONOMICS SERIES**

**SWP 2011/5**

# How Large is Large? Preliminary and relative guidelines for interpreting partial correlations in economics

**Hristos (Chris) Doucouliagos**

# How Large is Large?
# Preliminary and relative guidelines for interpreting partial correlations in economics

**Hristos (Chris) Doucouliagos***

douc@deakin.edu.au
School of Accounting, Economics and Finance
Deakin University

June 2011

## Abstract

An essential part of empirical economics research is the identification of the size of an empirical effect. Partial correlations offer a convenient statistically based measure of the strength of an economic relationship. A key question arises in their interpretation: When is a partial correlation large? This paper draws upon the observed distribution of 22,000 partial correlations from a diverse group of economics fields. The median absolute partial correlation from these fields is 0.173, which under Cohen's (1988) conventional guidelines for zero order correlations is a small to moderate effect. The paper develops new guidelines for key qualitative categories (small, medium and large). According to the new guidelines, partial correlations that are larger than ± 0.33 can be deemed to be large. This is considerably different to Cohen's guideline of ±0.50 for zero order correlations. Researchers and meta-analysts should exercise caution when applying Cohen's guidelines to describe the importance of partial correlations in economics.

Keywords: partial correlations; guidelines; empirical economics; meta-analysis

JEL Code: C01 and C50

Word count: 7,245

* Acknowledgements: The paper benefitted greatly from comments and feedback received from Geoff Pugh, Adnan Effendic, Margaret Giles, Debdullal Mallick, Patrice Laroche and especially from Tom Stanley. They helped to improve my thoughts and this paper. I take full responsibility for all limitations and errors.

# How Large is Large? Preliminary and relative guidelines for interpreting partial correlations in economics

"A medium effect size is conceived as one large enough to be visible to the naked eye".

Cohen (1988, p. 26).

## 1. Introduction

Effective policy is guided by reliable empirical evidence.[1] As evidence accumulates, associations can become clearer, when analysed by the appropriate tools. Meta-analysis (and meta-regression analysis) has been developed to make sense of the global "flood of numbers" and to clarify underlying associations and, hence, draw valid statistical inferences (Stanley, 2001). One of the main aims of meta-analysis is to quantify the size of an empirical effect. This is achieved by combining comparable estimates from independent empirical studies and then calculating averages from these estimates. Whether this involves an unweighted average, a weighted fixed or random effects average, or a selection bias corrected average, interest centres on the *size* of the effect.[2] Whether we are dealing with a single econometric study or a meta-analysis of a group of econometric studies, the size of an effect is of particular relevance for policy. It is simply insufficient to establish a statistically significant association (McCloskey, 1985 and 1995).[3] An association must *also* be of *practical* importance. If an effect is small, then perhaps alternative interventions might prove more effective.

[1] Perhaps this is a bold statement. Policy *should* be guided by evidence. Sadly, the responsiveness of policy to empirical evidence is probably closer to zero than many would care to admit. Policy is probably more likely to be driven by political considerations than it is by economic policy imperatives.
[2] See Hunter and Schmidt (2004) on the construction of averages in meta-analysis and Stanley (2008) on selection bias corrections. Another major aim of meta-analysis is to identify the source of heterogeneity in empirical estimates. Here too it is important to have an understanding of how large are the various dimensions of heterogeneity. While some meta-analyses state that their focus is only on explaining heterogeneity, their MRA coefficients can be used to inform on the size of an effect.
[3] McCloskey unfortunately takes the argument too far when she argues that statistical significance does not matter. A large practical effect that is not statistically significant is of very dubious policy relevance.

Many researchers rely on Cohen's (1988) guidelines for the practical significance of a simple (zero order) correlation.[4] According to Cohen, the absolute value of a correlation is small if it is 0.10, 0.30 is a medium effect and 0.50 is large. Cohen's guidelines have been developed for zero order correlations. However, in economics, interest in zero order correlations is rare.[5] Instead, interest lies on the partial effect of an explanatory variable X on a dependent variable Y, conditional on other factors. The ideal economics measure is an elasticity (the percentage change in Y resulting from a percentage change in X). However, in a large and significant proportion of empirical economics studies there is a failure to report sufficient information from which to calculate elasticities. For studies using the double-log form, the regression coefficients are obviously direct measures of elasticity. However, studies using the log-lin, or the linear functional form, are problematic. Reporting standards in economics are such that descriptive statistics are often not reported.[6] The meta-analyst may use outside information, e.g. from studies employing similar datasets or directly from the source from which the study extracted its data. This, however, injects some element of measurement error, especially since authors often clean and/or transform data, or use a sub-sample of a dataset.

One solution to this is to use partial correlations (Fisher, 1954).[7] Partial correlations measure the correlation between Y and X, holding all other variables constant. That is, the effects of all other factors are partialled out, leaving only the contribution of X. This, of course, is the familiar *ceteris paribus* criterion, quintessential for economic analysis. The

[4] Cohen (1988) also provides corresponding guidelines for other effect sizes, but the interest in this paper is solely on correlations.
[5] Some exceptions to this include meta-analyses by Fidrmuc and Korhonen (2006) and Tosi, Werner, Katz and Gomez-Mejia (2000).
[6] There is also the issue of scaling. There are cases where descriptive statistics are reported but the scaling used does not match the scaling used to derive regression coefficients. Using these statistics can thus result in rather strange and unreliable elasticities.
[7] In order to keep notation simple, in this paper the partial correlation coefficient is denoted as *r*. In a model with two explanatory variables, most authors denote the partial correlations as $r_{YX.Z}$, where Y is the dependent variable and X and Z are the explanatory variables. This is the correlation between Y and X, holding Z constant.

advantage of partial correlations is that they can be calculated directly from routinely reported regression output using the associated t-statistics: Averages for the dependent and independent variables are not needed (Greene, 2000).[8] Thus, partial correlations enable a more comprehensive dataset to be compiled. Their key disadvantage is that they are a statistical measure rather than the ideal economic measure of practical significance. Doucouliagos (1995) presented the first application of partial correlations to meta-analysis (to the literature on ownership, participation and productivity). Since then, partial correlations have been used to analyse various relationships, including privatisation programs (Djankov and Murrell, 2002), the growth effects of foreign aid (Doucouliagos and Paldam, 2008), the effects of institutions on growth and income levels (Efendic and Pugh, 2009), and alcohol price responsiveness (Wagenaar, Salois and Kormo, 2009).

Unfortunately, there is no direct relationship between zero order correlations and partial correlations: Partial correlations can be larger or smaller than zero order correlations.[9] Hence, it will only be by chance that Cohen's guidelines will apply to partial correlations. *So, how do we determine whether a partial correlation is small or large?* Is it possible to assess whether a partial correlation signifies a practically significant effect?

Obviously, the closer a partial correlation is to ± 1, the larger is the effect. However, experience shows that such large partial correlations are not common in economics. The vast majority of partial correlations in economics are much smaller than this (see section 2

[8] The calculation of the partial correlation coefficient, *r*, from regression output is straightforward: $r = \frac{t}{\sqrt{t^2 + df}}$, where *t* denotes the t-statistic of the appropriate multiple regression coefficient, and df denotes the degrees of freedom of this t-statistic. The partial correlation can also be calculated directly from raw data using a routine such as Stata's **pcorr**. This routine also reports the partial correlation squared discussed in section 4.2 below. Note, however, that if there is dependence in the data, than robust or clustered standard errors should be used in the regression analysis and this will generate lower (and more accurate) partial correlations than the **pcorr** routine.

[9] In field studies, partial correlations are typically found to be smaller than zero order correlations but theoretically this need not be the case.

below).[10] This raises the question of 'how large is large?' While we can say that a partial correlation of +0.81 is larger than +0.62, and that +0.43 is larger than +0.24, can we conclude that +0.18 is a small partial correlation?

The aim of this paper is to provide a set of guidelines for the interpretation of the practical significance of partial correlations. This is achieved by exploring the actual distribution of partial correlations found in empirical economics and then deriving 'reasonable' thresholds from this distribution. The guidelines are derived from the observed distribution of more than 22,000 partial correlations. Hence, the guidelines are *relative* to what the literature has found. As this observed distribution might not be representative of all empirical economics, the guidelines should be deemed to be *preliminary*. Section 2 discusses the data used to construct the distribution of partial correlations. Section 3 presents the guidelines. Section 4 discusses the usefulness of the guidelines, together with some limitations and cautions. Section 5 provides a summary of the paper.

## 2. Data and Approach

This paper takes a pragmatic approach to the issue, by considering the size of an effect in the context of what is found throughout empirical economics. That is, rather than taking an absolute position on effect size, the focus is on a relative position: Instead of asking is a partial correlation big or small, the question asked is whether a partial correlation is big or small *relative to what is typically found in empirical economics*.

The guidelines are constructed from the distribution of partial correlations found across a wide range of empirical economics literatures. This involved collecting the findings

[10] Certainly some individual studies report regression results that indicate very large (e.g. greater than 0.90 in absolute terms) partial correlations. However, meta-analyses of the entire body of evidence within an empirical literature report averages that are far smaller than this (see Table 1).

from 41 meta-analyses in economics.[11] Most of the studies were chosen from the meta-meta-analysis conducted by Doucouliagos and Stanley (2008) who identified 87 fields. In addition, I included some studies that became available since that survey. In some cases, authors focus on a measure that was not the partial correlation, e.g. the rate of beta convergence (Abreu, de Groot and Florax, 2005) and the dollar value of a statistical life (Bellavance, Dionne and Lebeau, 2009). However, where they provided sufficient information, I converted study results into partial correlations. This was done in order to increase the coverage of studies included in the dataset. From these 41 fields, I was able to derive 22,141 partial correlations. While each field represents a separate line of inquiry, the partial correlations can be pooled for the purposes of this paper. Recall that the aim here is to identify the distribution of the size of empirical effects found in empirical economics. The 41 fields are listed in Table 1. Column 1 lists the number of partial correlations for each field. Column 2 reports the median (absolute) partial correlation.[12] Column 3 evaluates the size of the empirical effect using Cohen's guidelines. Finally, column 4 presents the median partial correlation squared (discussed in section 4.2 below).

[11] To be included in this analysis, a meta-analysis had to report data on the actual empirical studies included, as well as t-statistics and sample size from which the partial correlations were calculated. This information is not available from most meta-studies. An appendix is available with a full reference list for the meta-analyses used. The first meta-analysis was made available in 1995 and the most recent in 2011.
[12] When identifying effect sizes, meta-analysts prefer to focus on the weighted average effect, rather than the median. Here the median is chosen in order to identify the 50$^{th}$ percentile in the distributions.

**Table 1: Fields Covered in the Analysis**

| Field | Number of partial correlations (1) | Median absolute $r$ (2) | Cohen's guideline (3) | Squared partial correlations (4) |
|---|---|---|---|---|
| Aid and growth | 1,243 | 0.107 | Small | 0.011 |
| Beta convergence | 610 | 0.327 | Medium | 0.107 |
| Board composition and performance | 67 | 0.100 | Small | 0.010 |
| Board duality and performance | 30 | 0.055 | Small | 0.003 |
| Business cycle correlations | 460 | 0.225 | Medium | 0.051 |
| Capital and growth | 1,671 | 0.364 | Medium | 0.132 |
| CEO pay and firm performance (UK) | 511 | 0.084 | Small | 0.007 |
| CEO pay and firm size (UK) | 265 | 0.293 | Medium | 0.086 |
| Commerce and aid allocation | 747 | 0.149 | Small | 0.022 |
| Demand for water | 110 | 0.280 | Medium | 0.078 |
| Democracy and growth | 483 | 0.159 | Small | 0.025 |
| Education and growth | 2,513 | 0.201 | Medium | 0.040 |
| Education and inequality | 847 | 0.162 | Small | 0.026 |
| Exchange rate volatility and trade | 1,255 | 0.140 | Small | 0.020 |
| FDI and growth | 876 | 0.214 | Medium | 0.046 |
| FDI spillovers | 24 | 0.154 | Small | 0.024 |
| Government and growth | 799 | 0.262 | Medium | 0.069 |
| Human rights and aid allocations | 493 | 0.090 | Small | 0.008 |
| Income and aid allocations | 1,030 | 0.242 | Medium | 0.059 |
| Inertia and aid allocations | 204 | 0.501 | **Large** | 0.251 |
| Inequality and growth | 677 | 0.248 | Medium | 0.062 |
| Inflation and central bank independence | 384 | 0.156 | Small | 0.024 |
| Institutions and growth | 112 | 0.371 | Medium | 0.138 |
| Inter-government competition | 622 | 0.086 | Small | 0.007 |
| Market orientation and performance | 47 | 0.295 | Medium | 0.087 |
| Military interests and aid allocations | 1,143 | 0.168 | Small | 0.028 |
| Minimum wage employment effects | 1,528 | 0.100 | Small | 0.010 |
| Participation and job satisfaction | 41 | 0.334 | Medium | 0.112 |
| Pensions and savings | 583 | 0.111 | Small | 0.012 |
| Pharmaceutical demand | 60 | 0.079 | Small | 0.006 |
| Politics and taxes | 410 | 0.037 | Small | 0.001 |
| Population and aid allocations | 738 | 0.214 | Medium | 0.046 |
| Population and growth | 486 | 0.216 | Medium | 0.047 |
| Social responsibility | 82 | 0.146 | Small | 0.021 |
| Technical efficiency and gender | 16 | 0.059 | Small | 0.003 |
| Transport noise | 31 | 0.165 | Small | 0.027 |
| Unions and intangible capital | 71 | 0.126 | Small | 0.016 |
| Unions and productivity | 77 | 0.108 | Small | 0.012 |
| Unions and profits | 532 | 0.123 | Small | 0.015 |
| Value of a statistical life | 39 | 0.078 | Small | 0.006 |
| Wage curve | 208 | 0.056 | Small | 0.003 |
| **Total** | **22,141** | **0.173** | Small to Medium | 0.030 |

The fields listed in Table 1 are a representative sample of the meta-analyses that have been conducted (for details see Doucouliagos and Stanley, 2008). However, it is not clear whether they are a representative sample of empirical economics in general. Hence, the guidelines drawn from these data should be interpreted as *preliminary*.[13] They might well change as more data becomes available.

Figure 1 is a histogram of the partial correlations included in the dataset. This distribution was used to identify (the admittedly arbitrary) threshold levels. The first quartile is deemed to capture *small* effects, the median captures *moderate* effects, and *large* effects are identified by the third quartile. While these thresholds are certainly arbitrary, they do capture the essential feature of the distribution of reported partial correlations from a diverse range of fields. A partial correlation that is smaller than the first quartile of all partial correlations found in economics is of little (small) practical significance. A partial correlation that falls near the median can reasonably be regarded as moderate. It seems appropriate to identify this threshold at the location of 50% of the data. Similarly, a partial correlation that is larger than the third quartile of all partial correlations found in economics can be reasonably considered to be a relatively large effect.

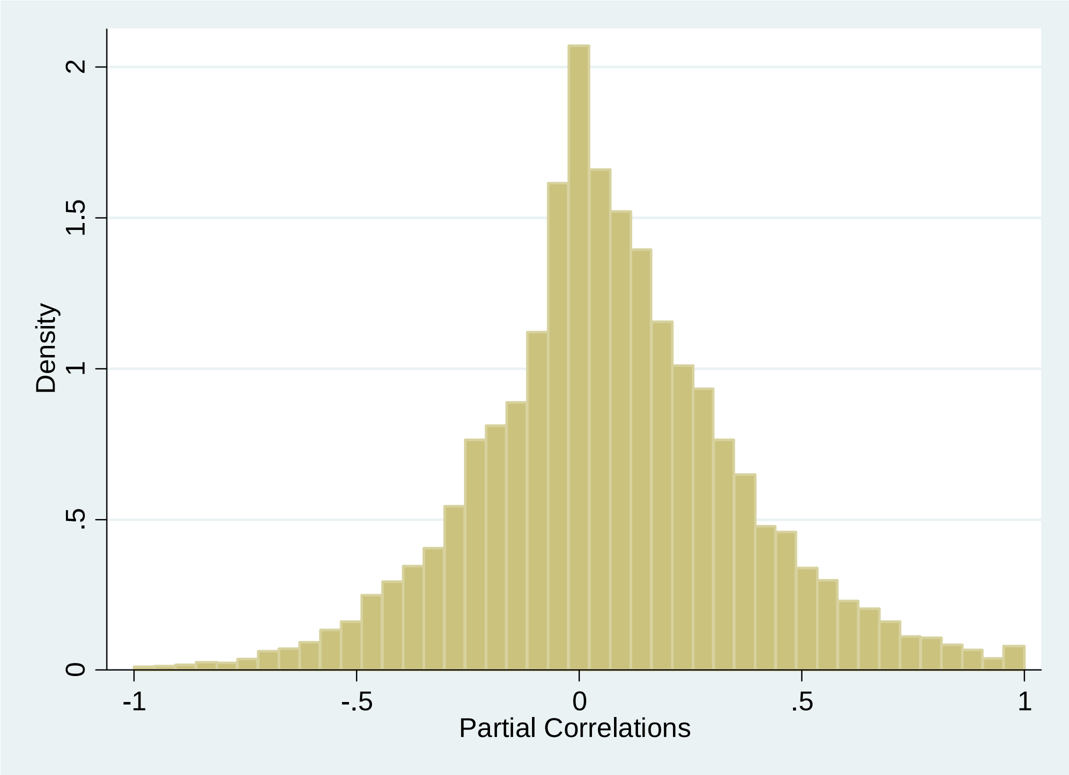
Figure 2 presents the histogram of the Fisher z-transformed partial correlations.[14] Though it is no longer truncated to ±1, the distribution is similar to the unadjusted partial correlations. Both figures highlight the very large proportion of near zero correlations detected in economics. This is heartening! It shows that economists do indeed report zero effects. There is a sizeable literature showing that many fields in empirical economics are severely distorted by selection bias. The direction of this bias tends to be upwards. That is,

[13] The most notable absence from the dataset is studies from experimental economics and studies from environmental economics.

[14] This is given by $Z_r = 0.5\ln\left[\frac{1+r}{1-r}\right]$.

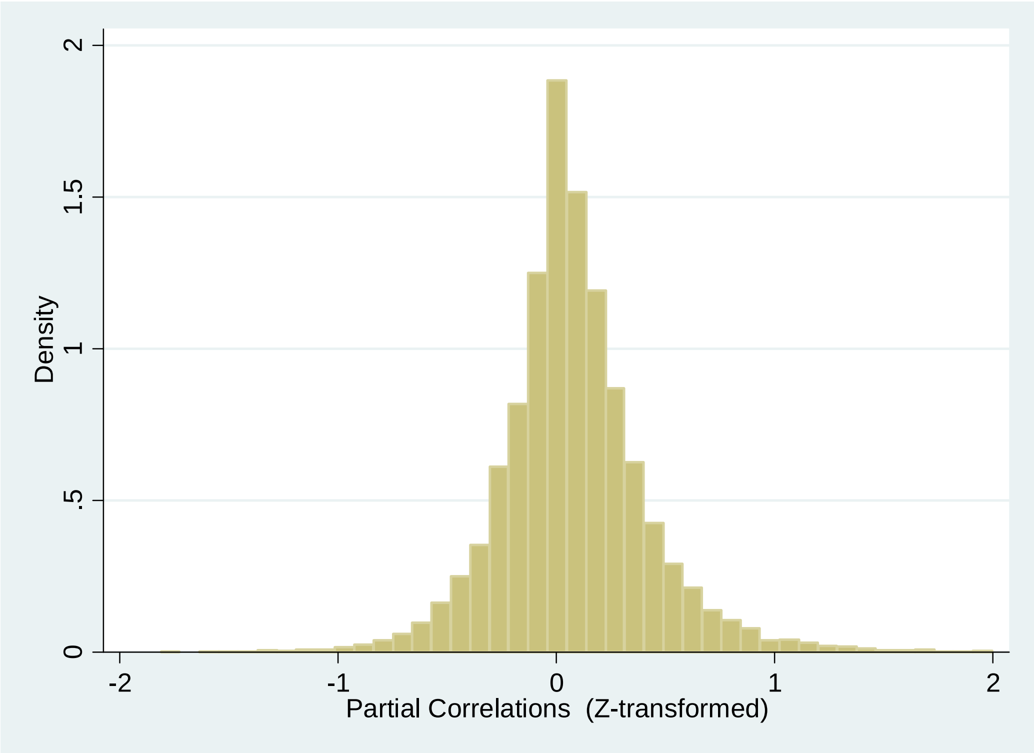
the bias favours, in most cases, a rejection of the hypothesis that there is no effect (rejection of the null).[15] Figures 1 and 2 show that there are actually a zero effects reported. Needless to say, zero effects are useful, as it is important to know what works and what does not.

**Figure 1: Histogram of Partial Correlations (n= 22,141)**




[15] See Roberts and Stanley (2005). Doucouliagos and Stanley (2008) present several examples of literatures afflicted with bias, as well as several fields that are free of such bias.

**Figure 2: Histogram of z-transformed Partial Correlations (n= 22,141)**




*Note:* There are a small number of observations that are outside the frame of the figure.

*No 'oomph' in empirical economics?*

If we accept Cohen's guidelines, then it is clear from column 3 of Table 1 that effect sizes in economics are small: 25 of the 41 fields have small (some very small) median effect sizes. In only one instance is there a large effect size.[16] Hence, in Ziliak and McCloskey's (2008) nomenclature, empirical economics lacks findings with 'oomph'. The situation is probably worse than this, as it is well documented that there is significant selection bias in empirical economics, resulting in a preference for large effects which are more likely to be statistically significant (Roberts and Stanley, 2005), i.e. there are some small effects that go unreported.

[16] This is the literature that explores the effect of aid allocations on the basis of past aid allocations, with the effect size based on the coefficient on a lagged dependent variable.

Table 2 explores the distribution of effect sizes in economics using Cohen's guidelines. All the partial correlations are used to construct column 1. Column 2 presents the results for the largest sub-set of data, the determinants of economic growth studies. The inevitable conclusion from Table 2 is that effects are small in economics. That may very well be the case. However, it is possible to mount a case that perhaps Cohen's guidelines should be reviewed and revised.

**Table 2: Size Distribution of Empirical Economics, Using Cohen's Guidelines**

| Cohen's guidelines | All estimates (1) | Determinants of growth (2) |
|---|---|---|
| Small (0.10) | 34% | 25% |
| Medium (0.3) | 38% | 38% |
| Large (0.50) | 28% | 37% |
| N | 22,141 | 22,141 |

*Notes*: Cells report the percentage of reported partial correlations that fall into each one of Cohen's categories. N denotes the number of partial correlations.

## 3. New guidelines

Table 3 presents the new guidelines. Column 1 presents the guidelines drawn by using all data. A partial correlation that is less than ±0.07 can be regarded as small (25$^{th}$ centile), *even if it is statistically significant*.[17] A partial correlation is large if it is greater than ±0.33 (75$^{th}$ centile). Half of the observed partial correlations are smaller than ± 0.17 (50$^{th}$ centile). Note

[17] Note however that the finding of a small effect does not mean that the research topic is trivial. Perhaps it becomes important to find out why the effect is small and how it can be made larger. This is the case, for example, for the aid (in)effectiveness literature (Doucouliagos and Paldam, 2008).

that this is smaller than what is found in other literatures (see, for example, Lipsey and Wilson, 1993).

Table 3 suggests that Cohen's guidelines are too restrictive when applied to economics. Recall, that Cohen's guidelines state that a correlation needs to be 0.50 to be considered to be large. Instead, the proposal here is that this threshold should be revised downwards to 0.33. Researchers should be cautious about applying Cohen's guidelines to economics research.

**Table 3: Effect Size Guidelines for Partial Correlations in Empirical Economics**

| Centile | All estimates (1) | z-transformed (2) | Determinants of growth (3) | Labour economics (4) | Industrial organization (5) |
|---|---|---|---|---|---|
| 25 (small) | 0.070 | 0.070 | 0.104 | 0.048 | 0.031 |
| | (0.068 to 0.072) | (0.068 to 0.072) | (0.100 to 0.108) | (0.045 to 0.050) | (0.025 to 0.040) |
| 50 (moderate) | 0.173 | 0.175 | 0.226 | 0.112 | 0.106 |
| | (0.170 to 0.176) | (0.171 to 0.178) | (0.221 to 0.232) | (0.107 to 0.116) | (0.090 to 0.117) |
| 75 (large) | 0.327 | 0.338 | 0.386 | 0.234 | 0.205 |
| | (0.322 to 0.331) | (0.333 to 0.344) | (0.379 to 0.391) | (0.224 to 0.242) | (0.185 to 0.228) |
| N | 22,141 | 22,141 | 9,934 | 4,718 | 872 |

*Notes*: Figures in brackets are 95% confidence intervals. The table reports absolute values. N denotes the number of partial correlations.

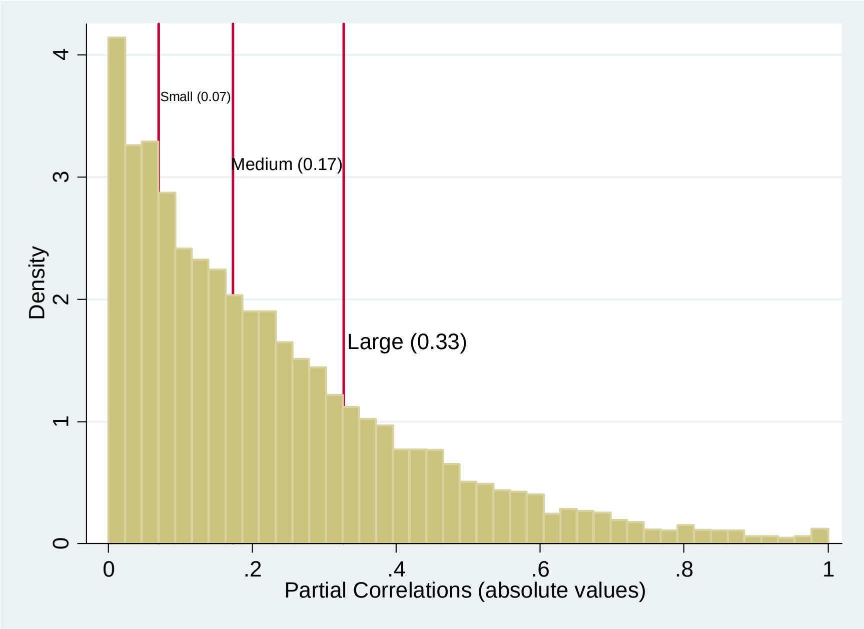
The partial correlation is truncated at -1 and 1 and, hence, this might distort the relative size of the underlying effect. To accommodate this, the Fisher z-transform was applied to all partial correlations. Using this transformation produces essentially the same guidelines (see column 2, Table 3). This is not surprising, as it is clear from the actual distribution of the correlations (see Figures 1 and 2) that the transformation is actually unnecessary, as there are relatively few larger correlations.

Columns 1 and 2 use all data from all areas of economics. However, it might be more relevant to derive guidelines for sub-disciplines. To accommodate this, the data was partitioned into four groups. First, we grouped all studies that related to the empirical growth literature. As already noted, this is the largest sub-group within the dataset. These guidelines are presented in column 3. Next, we considered all studies relating to labour economics (see column 4). Finally, we considered all studies relating to industrial organisation (see column 5). The guidelines are fairly similar for labour economics and industrial organisation. However, larger thresholds apply for the empirical growth literature, where typically larger effects are found.

Figure 2 reproduces the partial correlations, this time using their absolute value, showing the location of the guidelines (from column 1 of Table 3). Figure 3 presents the histogram associated with only the empirical growth studies. Note that when all observations are used, the distance between small and medium (0.10) is smaller than the distance between medium and large (0.16). In the case of economic growth studies, the distance between small and medium (0.13) is similar to that between medium and large (0.16). While Cohen kept these distances the same, there is no necessary reason to do so.

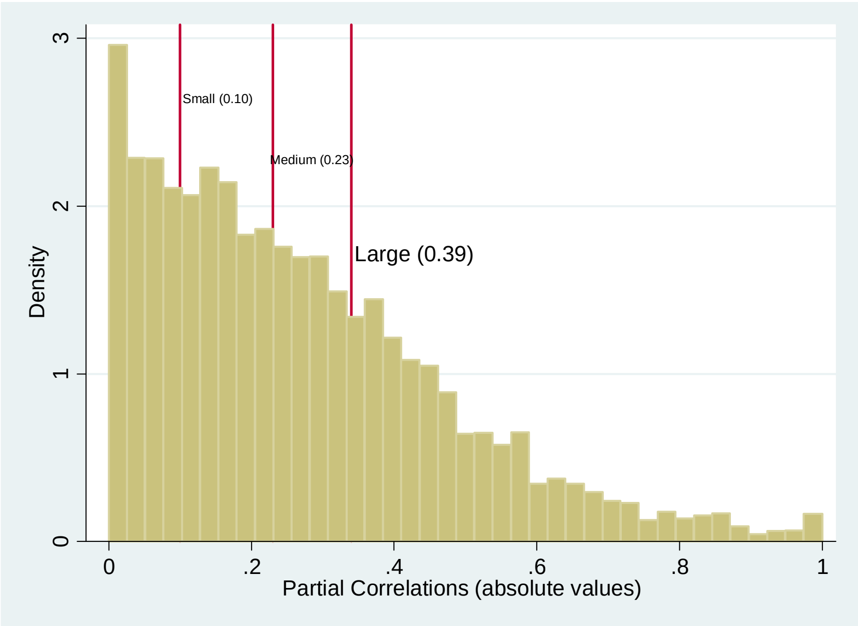
Table 4 reports field specific centiles. That is, instead of using all observations to develop guidelines for all empirical economics, the guidelines are developed for each of the 41 fields separately. This enables researchers in those fields to assess their results relative to what others have found in their field, as opposed to what has been found in all fields.

**Figure 2: Histogram of Absolute Values of Partial Correlations, All Data (n = 22,141)**




*Note*: Vertical lines show the location of the first, second and third quartile.

**Figure 3: Histogram of Absolute Values of Partial Correlations, Empirical Growth Studies, (n =9,934)**




*Note*: Vertical lines show the location of the first, second and third quartile.

## Table 4: Field Specific Guidelines

| Field | Number of partial correlations | Small<br>25th centile | Moderate<br>50th centile | Large<br>75th centile |
|---|---|---|---|---|
| Aid and growth | 1,243 | 0.047 | 0.107 | 0.188 |
| Beta convergence | 610 | 0.191 | 0.327 | 0.499 |
| Board composition and performance | 67 | 0.050 | 0.100 | 0.160 |
| Board duality and performance | 30 | 0.040 | 0.055 | 0.132 |
| Business cycle correlations | 460 | 0.106 | 0.225 | 0.403 |
| Capital and growth | 1,671 | 0.225 | 0.364 | 0.530 |
| CEO pay and firm performance (UK) | 511 | 0.0371 | 0.084 | 0.158 |
| CEO pay and firm size (UK) | 265 | 0.110 | 0.293 | 0.498 |
| Commerce and aid allocation | 747 | 0.060 | 0.149 | 0.283 |
| Demand for water | 110 | 0.157 | 0.284 | 0.459 |
| Democracy and growth | 483 | 0.079 | 0.159 | 0.278 |
| Education and growth | 2,513 | 0.087 | 0.201 | 0.351 |
| Education and inequality | 847 | 0.076 | 0.162 | 0.292 |
| Exchange rate volatility and trade | 1,255 | 0.057 | 0.140 | 0.272 |
| FDI and economic growth | 876 | 0.103 | 0.214 | 0.338 |
| FDI spillovers | 24 | 0.024 | 0.154 | 0.245 |
| Government and growth | 799 | 0.107 | 0.262 | 0.425 |
| Human rights and aid | 493 | 0.050 | 0.090 | 0.200 |
| Income and aid allocations | 1,030 | 0.120 | 0.242 | 0.396 |
| Inertia and aid allocations | 204 | 0.251 | 0.501 | 0.690 |
| Inequality and growth | 677 | 0.146 | 0.248 | 0.355 |
| Inflation and central bank independence | 384 | 0.103 | 0.156 | 0.212 |
| Institutions and growth | 112 | 0.255 | 0.371 | 0.593 |
| Inter-government competition | 622 | 0.020 | 0.086 | 0.177 |
| Market orientation and performance | 47 | 0.193 | 0.295 | 0.350 |
| Military interests and aid | 1,143 | 0.072 | 0.168 | 0.293 |
| Minimum wage employment effects | 1,528 | 0.045 | 0.100 | 0.191 |
| Participation and job satisfaction | 41 | 0.182 | 0.334 | 0.482 |
| Pensions and savings | 583 | 0.033 | 0.112 | 0.312 |
| Pharmaceutical demand | 60 | 0.036 | 0.079 | 0.351 |
| Politics and taxes | 410 | 0.015 | 0.037 | 0.076 |
| Population and aid allocations | 738 | 0.096 | 0.214 | 0.412 |
| Population and growth | 486 | 0.103 | 0.216 | 0.368 |
| Social responsibility | 82 | 0.059 | 0.146 | 0.275 |
| Technical efficiency and gender | 16 | 0.032 | 0.059 | 0.141 |
| Transport noise | 31 | 0.094 | 0.165 | 0.216 |
| Unions and intangible capital | 71 | 0.064 | 0.126 | 0.225 |
| Unions and productivity | 77 | 0.037 | 0.108 | 0.230 |
| Unions and profits | 532 | 0.043 | 0.123 | 0.206 |
| Value of a statistical life | 39 | 0.050 | 0.078 | 0.111 |
| Wage curve | 208 | 0.023 | 0.056 | 0.097 |
| **Total** | **22,141** | | | |

## 4. Applications and Limitations

This section of the paper considers the usefulness of the guidelines, as well as several limitations.

### 4.1 Practical Applications

Comparing effect sizes is fraught with numerous problems. Is an own price elasticity of 1.4 for alcohol consumption greater than the elasticity of bilateral trade with respect to distance of 1? In one sense, many elasticities are not comparable, as they relate to disparate fields. Is it valid to compare them? There is an argument to be made that each field (literature) should be interpreted on its own and analysts should form their own opinion on the size and significance of an association, although this does naturally introduce a degree of subjectivity into the assessment.

Where the aim of the meta-analysis is to test a specific hypothesis, such as whether an income elasticity equals 1, then the guidelines presented here are not useful. Similarly, where a specific value is needed for Cost-Benefit analysis, partial correlations should not (indeed cannot) be used.

Nonetheless, there are at least two scenarios where the guidelines are useful. First, as Cohen (1988) argued, guidelines are useful for new research areas, where it is unclear how large is a large effect. Second, many empirical studies look at only statistical significance and not practical significance. McCloskey has been complaining about this practice for several decades (see, for example, McCloskey, 1995, 1999). The practice does not look like it is going away any day soon.[18] Partial correlations offer a convenient way of deriving comparable estimates for such studies. Guidelines are useful when a large proportion of

[18] Indeed, many meta-analyses also just focus on the statistical significance of the meta-average, often ignoring the practical significance of the meta-average.

studies focus on statistical significance and they can be combined effectively into a meta-analysis using partial correlations. While the guidelines are necessarily arbitrary, they are useful in identifying the size of a partial correlation found in a new study, relative to all other studies in that body of literature. *It is important to note that the intention here is not to promote a focus on statistical significance. Rather, the point is how to assess the size of an effect from a literature where so many studies focus on statistical significance (both primary studies and meta-analyses). The partial correlation enables us to use this information and switch the focus away from statistical significance into practical significance.*

The guidelines are useful to both meta-analysis as well as individual researchers. An individual researcher may well want to compare her findings with the rest of the literature. Where other studies have not identified the size of their empirical effect, the researcher can convert some of the regression output from prior studies into partial correlations and then compare the partial correlation of her study with the partial correlations of some of the (selected) prior studies. Meta-analysis takes this process one large step further, by assessing all the prior evidence.

As an example of this, consider the recent meta-study by Doucouliagos, Stanley and Haman (2010) on the links between CEO pay and firm performance in the UK. The key *theoretical* variable of interest here is the elasticity of executive pay with respect to performance. However, the authors were able to collect only 187 elasticities and 217 semi-elasticities from 44 studies. In contrast, they were able to collect 511 partial correlations. Meta-analysis of the elasticities indicated no link between pay and performance. In contrast, the larger dataset of partial correlations suggested a small and statistically significant association (partial correlation = +0.08).

## 4.2 Limitations

*Selection bias*

Selection bias typically takes the form of a strong preference for reporting estimates that are statistically significant. The result of this preference will be to make estimates appear to be *larger* than they are, potentially inflating the guidelines presented here. This is a valid criticism that should be borne in mind, especially for some fields where there is a strong bias towards finding empirical support for certain theoretical propositions (e.g. downward slopping demand curves). However, as figures 1 and 2 illustrate, there is no absence of zero effects in the dataset used to construct the guidelines. Moreover, since the principal aim here is to identify the size of an effect relative to what others have found, then the issue of selection bias is not as important as it is for individual meta-analyses.

Conversely, in some areas, the selection bias will be in favour of the null hypothesis (e.g. tests of market efficiency), so the bias will be towards finding small effects, rather than large ones. None of the studies included in the dataset appear to have a theoretical bias towards accepting the null. Hence, the distribution illustrated in figures 1 and 2 is not likely to be biased towards small effects.

*Guidelines are subjective*

The guidelines presented here are based on the distribution of partial correlations observed in the field. That is, they are driven by the observed data. There is, obviously, an element of subjectivity in identifying the centile cut-offs for the qualitative categories. Thus, while the $25^{th}$ centile is chosen to represent small effects, others might prefer a larger or smaller cut-off.

What one person regards as reasonable, someone else might not. The real test of course will be whether these – or any other – guidelines are found to be useful in the field.

It should be recalled that Cohen acknowledges that his own guidelines “were made subjectively” (Cohen, 1992, p. 156).[19] However, they have subsequently been confirmed to be broadly consistent with evidence in the field. The approach taken here is in reverse. The actual distribution of partial correlations is observed first and guidelines constructed from this. That is, the guidelines presented here are more empirically orientated and, hence, less subjective. Furthermore, Cohen constructed his guidelines so that the difference between small and medium was the same as between medium and large. While this might seem a reasonable approach, the guidelines presented here are driven by the data and no such symmetry is imposed on the guidelines.

*Guidelines are likely to be time variant*

As already noted, the guidelines are preliminary. They are drawn from the data available at the time they were constructed. Hence, it is possible that with more observations at hand (both for the 41 fields listed in Table 1 and for other fields), they can be revised upwards or downwards. For example, what might be regarded as a medium effect today might subsequently be found to be a small effect. This is also a valid criticism. The guidelines are inherently dynamic, as opposed to the static ones developed by Cohen. Perhaps, however, there is some merit in having time varying effect sizes. Will what was considered to be large in 1988 still apply in 2028?

[19] The widely used 1, 5 and 10 percent levels of statistical significance (size of the test) are also arbitrary.

As a test for the degree of time variation in the data, the guidelines were constructed recursively for 10 year time periods for the economic growth studies (the largest group of studies). Table 5 shows that as more data has become available, effect sizes in the empirical economic growth literature have declined.[20]

**Table 5: Guidelines at 10 Year Intervals**

| Centile | Prior to 1980 (1) | Prior to 1990 (2) | Prior to 2000 (3) | All (4) |
|---|---|---|---|---|
| 25 (small) | 0.161 | 0.138 | 0.144 | 0.104 |
| 50 (moderate) | 0.322 | 0.270 | 0.269 | 0.226 |
| 75 (large) | 0.432 | 0.384 | 0.411 | 0.386 |
| N | 102 | 372 | 3,537 | 9,934 |

*Notes*: The table reports absolute values. N denotes the number of partial correlations

*Guidelines ignore variance explained*

Instead of using partial correlations, guidelines can be constructed using the squared partial correlation.[21] The squared partial correlation measures the amount of the variance in Y that is uniquely explained by one variable (say X1) as a proportion of the variance not explained by the other explanatory variables (say X2 and X3). Once again, however, some guideline/cut-off has to be established for the essentially qualitative dimension of small, medium and large. For example, one can adopt guidelines that state that an effect is: (a) small if $r^2$ is 0.10, so that the variable is explaining only 10% of the remaining variance; (b) medium if $r^2$ is 0.25, so that the variable is explaining one quarter of the remaining variance; and (c) large if $r^2$ is

[20] Indeed, there is much anecdotal evidence that effect sizes have been declining throughout most empirical economics fields. This is a fascinating issue that warrants further investigation.
[21] I thank Tom Stanley for this insight.

0.50, so that the variable is explaining one half of the remaining variance.[22] Such guidelines are, of course, also arbitrary and subjective.

Hunter and Schmidt (2004, p. 190) caution against the use of such variance based measures, noting that: "The problem with all percentage variance accounted for indexes of effect size is that variables that account for small percentages of the variance often have very important effects on the dependent variable. Variance-based indexes of effect size make these important effects appear much less important that they actually are, misleading both researchers and consumers of research".

Column 4 of Table 1 presents the $r^2$ associated with the median partial correlation value for each of the fields included in the dataset. Table 6 lists the squared partial correlations associated with the partial correlations for the small, medium and large categories based on the new guidelines. The threshold for a large effect ($r$ = 0.327) is associated with a 10.7% reduction in unexplained variation, which is not a trivial outcome.

**Table 6: Partial Correlations and Squared Partial Correlations**

| Size category | $r$ | $r^2$ | *Proportionate reduction in unexplained variation* |
|---|---|---|---|
| Small | 0.070 | 0.005 | 0.5% |
| Medium | 0.173 | 0.030 | 3.0% |
| Large | 0.327 | 0.107 | 10.7% |

[22] These guidelines were suggested by T.D. Stanley in personal correspondence.

*Guidelines ignore costs*

Effective policy requires reliable cost-benefit analysis. That is, it is not just the size of an outcome that is important. The cost of implementing an intervention is also important. For example, consider two interventions (A and B), both of which have an identical policy outcome deemed to be large (by whatever criteria is adopted) but intervention A is half as costly as intervention B. In this case, intervention A is a more cost effective intervention than B and effect size categories do not distinguish between them.

This is, of course, another valid point. Indeed, all of the fields included in the dataset essentially ignore this issue. For example, studies on the effects of inequality on growth explore whether inequality increases or decreases economic growth, as they wish to test rival theoretical predictions (de Dominicis, Florax and de Groot, 2008). There is little consideration of the costs arising as a result of the process of changing inequality in a country (other than what is revealed in terms of growth).

*Effect is not effectiveness*

Are field comparisons useful? For example, while many economists would regard a 1% increase in GDP per capita as a large effect, many would also argue that a 1% reduction in the inflation rate is not a large effect.[23] At the end of the day, perhaps what are needed are field specific guidelines for economic significance. While Table 4 presents these for partial correlations, does a large statistical effect translate into a large economic effect?

We can get a sense of the links between partial correlations (a statistical measure) and elasticities (an economic measure), by running the following regression:

[23] I thank Tom Stanley for this observation.

$$Elasticity_i = \beta_0 + \beta_1 r_i + \varepsilon_i \qquad (1)$$

Table 7 reports the results of this regression for two of the fields in the dataset.[24] Estimation was through robust regression, because of the existence of some very large elasticities. In both cases, there is robust relationship between partial correlations and elasticities, as there should be. The parameter estimates can be used to predict the value of the elasticity associated with the guidelines. For example, in the case of the effects of the minimum wage on employment literature (column 1), a large partial correlation (defined here as 0.327) implies a minimum wage elasticity of 0.169. That is, a 10% increase in the minimum wage would reduce employment by 1.69%. This is an economically significant effect. It is important to note that Table 7 does not attempt to assess the effects of the US minimum wage on employment. Rather, the aim is to assess the links between partial correlations and elasticities. In their meta-analysis of this data, Doucouliagos and Stanley (2009) found that the selection bias corrected weighted average minimum wage employment effect in the US was effectively zero.

In the case of CEO pay and firm performance literature (column 2), the large partial correlation threshold implies an elasticity of 0.257. That is, a 10% increase in firm performance (say an increase in the return on equity from 0.10 to 0.11) would increase CEO pay by 2.57%. Again, this is a large effect. As already noted, in their meta-analysis of this dataset, Doucouliagos, Stanley and Haman (2010) found a precision weighted average effect of only 0.08. That is, they found a small effect. The implied elasticity associated with this overall (small) effect is only 0.06. That is, a 10% increase in firm performance (say an increase in the return on equity from 0.10 to 0.11) would increase CEO pay by 0.60%. This is

[24] These fields were chosen *purely* because of the availability of a relatively large number of elasticities and, naturally, partial correlations.

a small economic effect. Hence, at least for these two cases, the effect size categories for partial correlations correspond broadly with elasticity sizes.

One difference between these two examples is that for the minimum wage literature, the implied elasticity associated with small partial correlations (0.10) is not that different from the large effect (0.17). In contrast, for CEO pay, the difference between the implied elasticity for small partial correlations (0.06) is much larger than for large effects (0.26).

**Table 7: Elasticities and Partial Correlations**

| | Minimum wages (US studies) (1) | CEO Pay-performance (UK studies) (2) |
|---|---|---|
| Constant | 0.080 | 0.002 |
| | (17.85) *** | (0.54) |
| Partial correlation | 0.273 | 0.781 |
| | (13.26)*** | (29.59)*** |
| $R^2$ | 0.13 | 0.21 |
| Number of observations | 1458 | 178 |
| | - *Implied elasticities* - | |
| Small ($r$ = 0.07) | 0.10 | 0.06 |
| Medium ($r$ = 0.17) | 0.13 | 0.13 |
| Large ($r$ = 0.33) | 0.17 | 0.26 |

*Note:* All estimates relate to equation 1. Robust regression used to derive parameter estimates. $R^2$ is derived from the corresponding OLS regression. Figures in brackets are t-statistics.

Table 7 seems to suggest that there is a link guidelines constructed for partial correlations and the size of elasticities. However, only 2 fields are analysed there and it is rather premature to conclude that this is a universal pattern. It is important to recall that the aim of this paper is not to present guidelines for economic measures of effect. Rather, it is to develop guidelines for the size of partial correlations. The guidelines are useful in assisting an individual researcher to assess how her results compare to those found by others, and they also assist meta-analysts when they are forced to use statistically based measures of effect.

They do not replace the need for an economic measure of effect, such as elasticities. If a literature has a clear and established notion of what is large, than that should be used instead.

## 5. Summary

Partial correlations enable a comprehensive set of empirical effect sizes to be included in a meta-analysis. They are important also to an individual researcher who needs a statistical measure of an effect size. They are especially important for meta-analysis where a comprehensive analysis of effect sizes is needed. There are currently no guidelines for what might be considered to be small/medium/large partial correlations. The aim of this paper is to develop such guidelines for the practical significance of partial correlations. These guidelines are based on the actual distribution of reported partial correlations in empirical economics. It is, of course, possible to focus on sub-fields (e.g. economic growth studies as opposed to all studies, see Table 3), or even individual fields themselves (e.g. inequality on growth rather than all economic growth studies, see Table 4). That is, they enable researchers to assess the size of the effects they have discovered relative to those by researchers in the same field, or economics in general.

The new guidelines are significantly lower than those Cohen developed for zero order correlations. Researchers who apply Cohen's guidelines to partial correlations will thus tend to *understate* the economic significance of the underlying empirical effect.

## 6. References


Abreu, M., H.L.F. de Groot, and Florax, R.J.G.M. 2005. A meta-analysis of beta-convergence: The legendary two-percent. *Journal of Economic Surveys* 19: 389-420.

Bellavance, F. Dionne, G. And Lebeau, M. 2009. The value of a statistical life: A meta-analysis with a mixed effects regression model. *Journal of Health Economics* 28:444-464.

Cohen, J. 1988. *Statistical Power Analysis in the Behavioral Sciences*. 2nd ed. Hillsdale: Erlbaum.

Cohen, J. 1992. A power primer. *Psychological Bulletin* 112 (1): 155-159.

de Dominicis, L., R.J.G.M. Florax and de Groot, H.L.F. 2008. A meta-analysis of the relationship between inequality and economic growth. *Scottish Journal of Political Economy* 55(5): 654-682.

Djankov, S. and Murrell, P. 2002. Enterprise restructuring in transition: A quantitative survey. *Journal of Economic Literature* XL: 739-792.

Doucouliagos, C. 1995. Worker participation and productivity in labor-managed and participatory capitalist firms: A meta-analysis. *Industrial and Labor Relations Review* 49: 58-77.

Doucouliagos, C. and Paldam, M. 2008. Aid effectiveness on growth: A meta-study. *European Journal of Political Economy* 24: 1-24.

Doucouliagos, C. and Stanley, T.D. 2008. Theory competition and selectivity: Are all economic facts greatly exaggerated? Deakin University, Economics Working Paper No. 2008_14.

Doucouliagos, C., T.D. Stanley, and Haman, J. 2010. Pay for performance and corporate governance reform. Deakin University, Economics Working Paper No. 2010_04.

Fidrmuc, J. and Korhonen, I. 2006. Meta-analysis of the business cycle correlation between the Euro area and the CEECs. *Journal of Comparative Economics* 34: 518–537.

Fisher, R. 1954. *Statistical Methods for Research Workers*. Hafner Publishing Company, New York.

Greene, W. 2000. *Econometric Analysis*, London: Prentice Hall.

Hunter, J. and Schmidt, F. 2004. *Methods of Meta-Analysis: correcting error and bias in research finding*, London: Sage.

Lenth, R. V. 2001. Some practical guidelines for effective sample size determination. *The American Statistician* 55: 187-193.

Lipsey, M.W and Wilson, D.B. 1993. The efficacy of psychological, educational, and behavioural treatment: Confirmation from meta-analysis. *American Psychologist* 48: 1181-1209.

McCloskey, D.N. 1985. The loss function has been mislaid: The rhetoric of significance tests. *American Economic Review* 75: 201-05.

McCloskey, D.N. 1995. Insignificance of statistical significance. *Scientific American* 272: 32-33.

Roberts, C.J. and Stanley, T.D. (eds) 2005. Meta-regression analysis: Issues of publication bias in economics. Blackwell, Oxford.

Stanley, T.D. 2001. Wheat from chaff: Meta-analysis as quantitative literature review. *Journal of Economic Perspectives* 15(3): 131-150.

Stanley, T.D., 2008. Meta-regression methods for detecting and estimating empirical effects in the presence of publication selection. *Oxford Bulletin of Economics and Statistics* 70: 103-27.

Tosi, H.L., Werner, S., Katz, J.P., and Gomez-Mejia, L.R. 2000. How much does performance matter? *Journal of Management* 26: 301-339.

Wagenaar, A.C., M.J. Salois and Kormo, K.A. 2009. Effects of beverage alcohol price and tax levels on drinking: A meta-analysis of 1003 estimates from 112 studies. *Addiction*, 104: 179-190.

Ziliak, T. and McCloskey, D.N. 2008. *The Cult of Statistical Significance*, University of Michigan Press.