**STATA Regression Data Description:**

Cross-disciplinary convergence in Human Brain Science

# **Author(BoldBlue) / Publication -Level Data**

**nAuthorID:** scholar’s unique numerical ID

**scopus\_id:** scholar’s unique Scopus profile ID

**h\_index:** scholar’s h-index in the citation count census year (2019)

**num\_coauthors:** scholar’s total number of scopus coauthors

**num\_publications:** scholar’s total number of scopus publications

**citations:** scholar’s total number of scopus citations

**min\_pub\_year:** scholar’s first publication year

**max\_citation:** scholar’s highest-cited article in the citation count census year (2019)

**mean\_coauthors:** scholar’s average number of coauthors per publication

**Tauip:** scholar’s age as published academic: Tauip = Yp – min\_pub\_year+1

**Eidsp**: article’s unique Scopus ID

**DOIp**: article’s unique digital object identifier (DOI)

**JournalIDp**: article’s unique numerical ID corresponding to the publishing journal

**Cp**: article’s total citation count up to the census year (2019)

**Yp**: article’s publication year

**Kp**: article’s coauthor count based upon author list in PubMed record

**Zp**: article’s log-normalized citation impact z-score: zp = (log[1+Cp] – mu\_t)/sigma\_t , where mu\_t) and sigma\_t are the mean and standard deviation of log[1+Cp] for all other articles pubished in the same year; this variable is well fit by the normal N(0,1) distribution.

**[KR1-KR4]p**: article’s counts of the number of Scopus coauthors from regions 1-4: 1= North America; 2= Europe; 3 = Australasia; 4= rest of world

**[SA1-SA6]p**: article’s counts of the number of Major MeSH terms (keywords identified as major theme – indicated with \* in PubMed MeSH list for a given article) belonging to a given subject area (top-level MeSH classification). 9 of the MeSH top-level branches are combined into 6 representative SA clusters. Note: Article-level MeSH keywords are assigned by expert annotators by the US National Library of Medicine.

**[CIP1-CIP9]p**: article’s counts of the number of Scopus authors belonging to each of nine total CIP (Classification of Instructional Programs) groups. Note: CIP are designated by the US Dept. of Education.

**IRegionRefinedp**: categorical variable (8 types) indicating the regional location of the scopus coauthors: 0 = World only; 1 = NA; 2 = EU; 3 = AA; 4 = NA + EU; 5 = NA + AA; 6 = EU + AA; 7 = NA+EU+AA (independent of World = other); Types [4-7] are cross-region whereas types [0-3] are mono-region.

**Regionp**: categorical variable (4 types): = 1 if IRegionRefinedp = 1; = 2 if IRegionRefinedp = 2; = 3 if IRegionRefinedp = 3; and 0 otherwise

**XRegionp**: binary indicator variable = 1 if 2+ scholars from different regions are coauthors of the article, and = 0 otherwise

**MeanZJp**: Journal’s mean Zp value calculated across all its articles, independent of year.

**XSAp**: binary indicator variable = 1 if any 2+ SA are present, and 0 otherwise

**XCIPp**: binary indicator variable = 1 if authors belonging to any 2+ CIP are present, and 0 otherwise

**NEUROSHORTXSAp**: binary indicator variable = 1 if at least one MeSH belongs to SA=1 and at least one MeSH belongs to any SA = {2,3,4}

**NEUROLONGXSAp**: binary indicator variable = 1 if at least one MeSH belongs to any SA={1-4} and at least one MeSH belongs to any SA = {5-6}

**NEUROSHORTXCIPp**: binary indicator variable = 1 if at least one coauthor affiliation belongs to any CIP={1,3} and at least one other coauthor affiliation belongs to any CIP = {2,4-7}

**NEUROLONGXCIPp**: binary indicator variable = 1 if at least one coauthor affiliation belongs to any CIP={1-7} and at least one other coauthor affiliation belongs to any CIP = {8-9}

**IYBRAINPROJECTp**: binary indicator variable = 1 if Yp >=2014 and 0 otherwise

**NRegp**: article’s count variable indicating the total number of regions represented (independent of concentrations), ie min = 1 and max = 4

**NSAp**: article’s count variable indicating the total number of SA represented (independent of concentrations), ie min = 1 and max = 5 (for broad SA) and max = 4 (for narrow Neuro SA)

**NCIPp**: article’s count variable indicating the total number of CIP represented (independent of concentrations), ie min = 1 and max = 9

**Y05yeari**: scholar’s period of first publication, based upon non-overlapping 5-year windows splitting decades (e.g. 1970-1974, 1975-1979, etc.)

**CIPi**: scholar’s home department CIP code

**NPubsnAuthorIDi**: author’s total number of articles after cleaning (e.g. excluding articles with no keywords, and excluding articles from journals with less than 20 articles in entire sample).

**NPubsnAuthorID2000i**: author’s total number of articles after cleaning and in the period [2000-2017]

**NPubsnAuthorID2009i**: author’s total number of articles after cleaning and in the period [2009-2017]

**XSACIPp**: binary indicator variable =1 if XSAp = XCIPp = 1 and 0 otherwise

**NEUROLONGXSACIPp (**and similarly for **NEUROSHORTXSACIPp)**: binary indicator variable =1 if NEUROLONGXSAp = NEUROLONGXCIPp = 1 and 0 otherwise