Correspondence

Response to comments on: "Does your surname affect the citability of your publications?"

1. Introduction

In this response we will show that Weber's (2017) critical comments on our article (Abramo & D'Angelo, 2017) are theoretically and empirically ill-founded. Before going into the details, we wish to take away three main mystifications.

First, Weber (2017) would have the reader believe that the aim of our article is to answer "the question whether the position of a researcher's last name in the alphabet matters for his or her scientific career". From the title of the paper, our question is remarkably different: "whether the position of a researcher's last name in the alphabet matters for the citability of his or her publications."

Second, given our true research question, it is untrue that "our findings stand in contrast to almost all literature on the topic". To our knowledge, and seemingly also to the best of Weber's (2016), "all literature" consists of only three other studies, by Shevlin and Davies (1997), Huang (2015), and Ong, Chan, Torgler, and Yang (2015). The first two focus on the sciences, like ours, and one of these aligns with our findings. The third manuscript (not peer-reviewed) concerns economics and management.

Third, in contrast to Weber's criticisms, our findings are fully substantiated, as we will show next.

2. Correlation legitimacy

Weber's main criticism is that citability is affected by many important factors, and because the surname initial is a minor factor, it cannot be detected by a statistical test if the other factors are omitted. Here is our reply.

First, to avoid mixing different aspects, we first assume that the surname initial is uncorrelated with the other factors. Using the notation Y citability, X surname initial, and Z all other factors, we represent the issue by a linear regression model: $Y = \alpha + \beta X + \gamma Z + \varepsilon$. In this framework, the "mere correlations" in Weber's words, refer to a model where Y is regressed on X only, so that the effect of Z is absorbed by the error ε. As is well known, if X and Z are uncorrelated then the least squares estimate of the coefficient β is unchanged when Z is omitted (e.g. Greene, 2003, §8.2.). Then Z is irrelevant to the point estimate but remains relevant for tests, because the omission of Z causes an increase in the residual variance: if X and Z are uncorrelated, the increase in the residual variance raises the standard error of β, so that the t-test β/S.E. decreases and may become non-significant (technically, omitting Z causes a reduction in the power of the statistical test for X). Therefore, the issue is real but only pertains to statistical significance, whereas point estimation is not affected. The core objective of our work, namely estimating the size of the correlations between citability and surname initial, is not compromised. The analysis found tiny correlations, and if the surname initial and the other factors are uncorrelated, these estimates are unbiased. So the conclusion is that the surname initial is essentially irrelevant, beyond any considerations of the power of statistical tests.

3. Ethnicity and other factors

The question becomes whether the surname initial is correlated with some kind of relevant factor, so that omitting such factors causes a bias in the estimation of the correlation between citability and surname initial. This is the classical risk of analyzing observational data. Returning to our regression framework, we know that if Z is correlated with X, then omitting Z will produce a bias in the estimation of the effect of X (the strength of the bias depends on both the effect of Z and its correlation with X).

Weber (2017) suggests the example of ethnicity as an omitted variable. In his words: "It is not unreasonable to think that Asian scholars in Italy perform better than Italian scholars (why would Italian universities otherwise hire them given that

there may be troubles with missing Italian language skills) while their last names come on average later in the alphabet. It could well be the case that alphabetical discrimination exists, but that it is not detected by the correlations or that the correlations point to the opposite, because Asian scholars in Italy perform better than other scholars." Engaging Weber's argument, we object that it could also be reasonable to think that Asian professors in Italy do not perform better than Italian ones. For example, they may have been hired precisely on the basis of their mother-language skills or knowledge of own-country issues, for the advantages this would represent in arts and humanities programs, but without other greater qualifications. Also, if we assume that Italian academic recruitment is meritocratic, then Asian scholars would indeed have performed better than other candidates for their particular position, opened to competition, but not necessarily better than other candidates that were hired for other positions. We remind our readers that the field of observation is limited to all professors in Italian faculty, not extending to "all scholars". Numerous studies have then shown that the Italian academic recruitment process is infested by widespread favoritism, so that the "hire the best" syllogism is not at all a constant (Abramo & D'Angelo, 2015; Abramo, D'Angelo, & Rosati, 2014; Abramo, D'Angelo, & Rosati, 2015; Abramo, D'Angelo, & Rosati, 2016; Abramo, D'Angelo, & Soldatenkova, 2017; Aiuti, Bruni, & Leopardi, 1994; Amadori, Bernasconi, Boccadoro, Glustolisi, & Gobbi, 1992; Biggin, 1994; Fabbri, 1987; Gaetani & Ferraris, 1991; Garattini, 2001; Gerosa, 2001; Perotti, 2008; Zagaria, 2007).

But when it comes to the crunch, the real fact here is that the body of foreign professors on Italian faculty is near zero. In spite of government initiatives to attract external academics, as of 2013 (ANVUR, 2013) there were still only 533 non-Italians out of a total of 53,459 professors. We observed publications from 2001 to 2004, so for our analysis the presence was still less. Of those who had arrived by 2013, 70 percent were European. Assuming (very doubtfully) similar arrivals of foreign and non-European faculty by 2001–2004, our observed population of 14,467 sciences professors would include 43 non-European professors. Weber (2017) would have us deal with foreign professors with late-alphabet names such as "Patel" or "Singh" in our dataset. It happens that there are no Patels or Singhs in Italian universities. The reader has the facts, and can judge whether the ethnicity variable of Asian, or any foreign professors, could determine bias in our analysis. Weber infers that the reason we exclude 66 professors (most likely foreigners) from the analysis, with surnames beginning H, J, K, W, X, Y, is to avoid an ethnicity problem. Not at all: we do this precisely to limit uncertainty in calculations for these letters, given the almost total lack of observations.

Weber (2017) puts forth ethnicity as just one example of the possible variables requiring control, neglecting to suggest any others. Let us specify them, because they are important and interesting.

Resources, both tangible and intangible, can certainly affect performance, and possibly the average quality of output. However in Italy, in the period under observation, public funding was not allocated to universities based on performance, and these did not allocate resources internally based on individual performance. Such non-competitive environments have not favored the emergence of world-class universities. Instead we observe a system of scarcely differentiated institutions, with internal dispersion of research performance remarkably higher than dispersion between universities (Abramo, Cicero, & D'Angelo, 2012; Abramo, D'Angelo, & Grilli, 2016; Bonaccorsi & Cicero, 2016). So professors in Italy can hardly benefit from differing resources or reputations of their home universities, to increase citability.

Gender is another variable related to performance in the sciences, with males showing higher performance than females (Abramo, D'Angelo, & Caprasecca, 2009a; Cole & Zuckerman, 1984; Fox, 2005; Long, 1987, 1992; Xie & Shauman, 1998), especially among top scientists (Abramo, D'Angelo, & Caprasecca, 2009b; Bordons, Morillo, Fernández, & Gómez, 2003). Men still predominate in the prestigious first and last author positions of the byline, and women are significantly underrepresented as authors of single-authored papers (Sugimoto, Larivière, Ni, Gingras, & Cronin, 2013; West, Jacquet, King, Correll, & Bergstrom, 2013). The results are less clear though when it comes to the average impact of publications: some studies indicate that women achieve higher average impact (Duch et al., 2012; Symonds, Gemmell, Braisher, Gorringe, & Elgar, 2006), while others encounter the opposite (Hunter & Leahey, 2010; Larivière, Vignola-Gagné, Villeneuve, Gelinas, & Gingras, 2011).

Other variables also related to research performance are: academic rank (Abramo, D'Angelo, & Di Costa, 2011; Blackburn, Behymer, & Hall, 1978; Ventura & Mombrù, 2006); age and seniority (Abramo, D'Angelo, & Murgia, 2016; Cole, 1979; Gingras, Larivière, Macaluso, & Robitaille, 2008; Kyvik & Olsen, 2008; Levin & Stephan, 1991; Over, 1988; Stroebe, 2010); and research collaboration, although the causal nexus between collaboration and research performance of the individual scientist has still not been fully clarified (Abramo, D'Angelo, & Murgia, 2017; He, Geng, & Campbell-Hunt, 2009; Lee & Bozeman, 2005; Ynalvez & Shrum, 2011). Empirical investigations have shown a relation between collaboration and both intensity of publication (Abramo, D'Angelo, & Di Costa, 2009; Ductor, 2015; Godin & Gingras, 2000; Katz & Martin, 1997; Lee & Bozeman, 2005), and impact (Bidault & Hildebrand, 2014; Laband & Tollison, 2000; Lawani, 1986; Li, Liao, & Yen, 2013; Wuchty, Jones, & Uzzi, 2007), even more so for international collaboration (Abramo, D'Angelo, & Solazzi, 2011; He, Geng, & Campbell-Hunt, 2009; Kato & Ando, 2013). Finally, individual talent must play a significant role on the impact of papers.

The reader can begin to understand that we considered many potential correlates in the run-up to preparing our method. We could continue the discussion, yet for each possible variable, we cannot think of any potential relation to surname initials, and cannot find any empirical evidence of such relations in the literature. Amongst all our countries, where is it that talent, age, or gender are related to surname initial, or resources are assigned by last name? Or are "weak scholars with last names late in the alphabet more likely to leave academia than weak scholars with last names early in the alphabet" (Weber, 2017)? This would be a curious variable, requiring inexplicable alphabetical behaviors – behavior unexpected for all in Italy, where a professor leaving any position is a startling exception.

Weber (2017) further criticizes: "the authors do not examine a single field by itself in which alphabetical name ordering is the norm", and "the authors exclude economics from their analysis". But empirical investigations must deal with reality as it