The Importance of Client Size in the Estimation of the Big 4 Effect: A Comment on DeFond, Erkens, and Zhang (2016)

Alastair Lawrence
lawrence@haas.berkeley.edu
Haas School of Business
University of California at Berkeley
2220 Piedmont Avenue
Berkeley, CA 94720

Miguel Minutti-Meza
mminutti@bus.miami.edu
School of Business Administration
University of Miami
5250 University Drive, Coral Gables
FL 33146

Ping Zhang
pzhang@rotman.utoronto.ca
Rotman School of Management
University of Toronto
105 St. George Street
Toronto, Ontario M5X 3E6

April 2016

The Importance of Client Size in the Estimation of the Big 4 Effect:
A Comment on DeFond, Erkens, and Zhang (2016)

DeFond, Erkens, and Zhang (2016, hereafter DEZ) provide comprehensive analyses highlighting how random variations in propensity score matching (PSM) design choices affect inferences concerning the existence of the Big 4 auditor effect. The conclusion of DEZ is that Lawrence, Minutti-Meza, and Zhang (2011, hereafter LMZ) fail to find a Big 4 effect because of PSM's sensitivity to design choices. We believe that DEZ emphasizes the need to think carefully when implementing PSM.

We do not interpret our findings in LMZ, as characterized by DEZ, as a challenge to the general understanding of audit quality, or as proof that there is no Big 4 effect. Our main message is that differences in audit quality proxies "between Big 4 and non-Big 4 auditors largely reflect client characteristics and, more specifically, client size." (LMZ, p. 259) Early versions of LMZ used attribute matching (i.e., matching variable-by-variable) and find a statistically significant Big 4 effect only in matched samples in which size was not the main matching attribute (LMZ, Tables 3, 6, and 9). In later versions, thanks to the advice of an astute reviewer, LMZ also included PSM matching (LMZ, Tables 2, 5, and 7).[1] Ensuring a tight match on client size is of paramount importance in determining the Big 4 effect for two interrelated reasons. First, client size is the primary variable explaining auditor choice. In Big 4 choice models, using client size alone it is possible to correctly classify 86 percent of the client-year observations as Big 4 and non-Big 4 clients. Second, proxies for audit and financial reporting quality typically increase in client size. The plot in Figure 1 illustrates how absolute discretionary accruals (the main measure employed in LMZ and DEZ) are nonlinearly decreasing in client size. Consequently, client size is arguably a prime confounding variable when comparing audit quality between Big 4 and non-Big 4 auditors.

In our opinion, the findings in DEZ generally reiterate the main message in LMZ that the estimated Big 4 effect is largely influenced by client characteristics, and in particular by client size. First, The economic magnitude of the estimated Big 4 effect substantially decreases for the PSM samples. For example, in the full sample regression analyses in LMZ (p. 267) using absolute

[1] Arguably, PSM allows reaching more general conclusions by matching simultaneously on multiple attributes. This methodology has been widely implemented in a number of disciplines. For example, the seminal PSM paper by Rosenbaum and Rubin (1983) has been cited over 15,000 times. The pros and cons of PSM design choices have been extensively examined in the statistical literature. For example, see Morgan and Winship (2015, p. 175-180) that shows simulation results comparing the performance of PSM under various design choices versus other forms of matching, including coarsened exact matching (CEM). CEM does not perform significantly better than PSM.

discretionary accruals, the coefficient on the Big 4 variable is -0.0179, while in DEZ the mean coefficient on the Big 4 variable among the randomized PSM samples ranges between -0.003 and -0.007 (DEZ, Table 2 and Table 3, Panels A and B).[2] Hence, there is between a 61 to 83 percent reduction in the economic magnitude of the estimated Big 4 effect between the full sample and the mean of the PSM samples. We believe that researchers examining the Big 4 effect need to ponder not only its statistical significance, but also its economic magnitude, and whether statistical differences if any, are meaningful.

Second, the statistical significance of the Big 4 coefficients in DEZ's randomized PSM samples is increasing in relative imbalances in client size between Big 4 and non-Big 4 auditors. Figure 2, which illustrates a replication of DEZ's random PSM variation analyses for absolute discretionary accruals using our data, indicates that the *t*-statistics on the Big 4 coefficients of PSM models are decreasing in the Big 4 to non-Big 4 client size ratio (i.e., ratio of the mean total assets of Big 4 to non-Big 4 clients in each sample). Specifically, starting from 1,000 samples with random PSM variations, only 626 (62.6 percent) have a client size ratio between 0.95 and 1.05, indicating a close balance on client size, and among these samples only 101 (10.1 percent) have Big 4 coefficients that are negative and statistically significant at the ten percent level or lower (*t*-statistic < -1.68).[3]

Third, DEZ generally find substantially weaker results supporting their inferences for restatements (e.g., DEZ Table 3, shows that only 54.4 percent and 15.6 percent of the PSM samples with and without replacement, respectively, have a statistically significant Big 4 effect at the five percent level). Restatements are a proxy increasingly used in the literature, readily available in Audit Analytics for recent years, that constitute an arguably more direct measure of audit quality. We agree with DEZ's critique that some of the proxies in LMZ are crude and indirect measures of audit quality; however, they were among the proxies employed in extant research published in top journals at the time demonstrating the existence of the Big 4 effect, and hence, the reason for their use.

At a high-level, the PSM models employed in both LMZ and DEZ aim to compare two groups of auditors that serve remarkably different audit markets and clientele. For instance, the

---

[2] The coefficient on the Big 4 variable in the matched sample in LMZ (p. 267) is -0.0018.

[3] Increasing the acceptable client size ratio to the range from 0.9 to 1.1 results in 179 (17.9 percent) samples with Big 4 coefficients that are negative and statistically significant at the ten percent level or lower. This effect is more pronounced in the full sample where there is a large size imbalance.

mean market capitalization of Big 4 and non-Big 4 audit clients is $4.1 billion and $147 million, respectively.[4] In recent years, non-Big 4 firms only audit less than two percent of the largest 500 U.S. companies, in terms of market capitalization, whereas they audit approximately 70 percent of companies with market capitalizations under $500 million. Thus, researchers lack non-Big 4 counterfactuals for large clients, such as Apple, General Motors, and Goldman Sachs, and using PSM researchers are forced to discard many valid observations, which is a nontrivial concern. From an athletic standpoint, an equivalent example would be comparing the quality of two track coaches, one whom specializes in long distance runners to another whom primarily specializes in sprinters, by examining their respective athletes' performances. To make this comparison, PSM would likely match observations for the coaches' middle distance runners, a group neither coach specializes in but their athletes overlap. Several design choices, such as ensuring that specific athletes with overlapping characteristics are selected for the comparison, should be made carefully and with purpose, rather than randomly. The distribution of estimates in the DEZ's analyses highlights the importance of not just randomizing PSM design choices, but instead ensuring that they reflect careful thought of the underlying economics and potential biases of each setting.

Looking forward, we encourage future research to establish comparatively more direct proxies for audit quality. For example, (a) there is an increasing interest in the drivers of audit quality at the office, engagement, and partner level; (b) there is relevant information in SEC fillings yet unused by audit researchers, such as the 10-K Schedule II that reports material valuation allowances and reserve accounts that are particularly subject to client and auditor judgment; (c) there can be new insights in enhanced audit report disclosures outside the U.S.; and, (d) there are opportunities to collaborate with the PCAOB to obtain interesting new data. Moreover, regarding competition in the audit market, it is unclear whether there is a single national level market, or instead the market is segmented by industry, location, client size, etc., factors that may give non-Big 4 auditors an advantage in certain settings. We know little about how non-Big 4 auditors persist in a highly competitive market and how the unaudited client quality influences efficient auditor choices.

[4] The descriptive statistics are measured using data from 2000 to 2014 from Compustat and Audit Analytics.

References

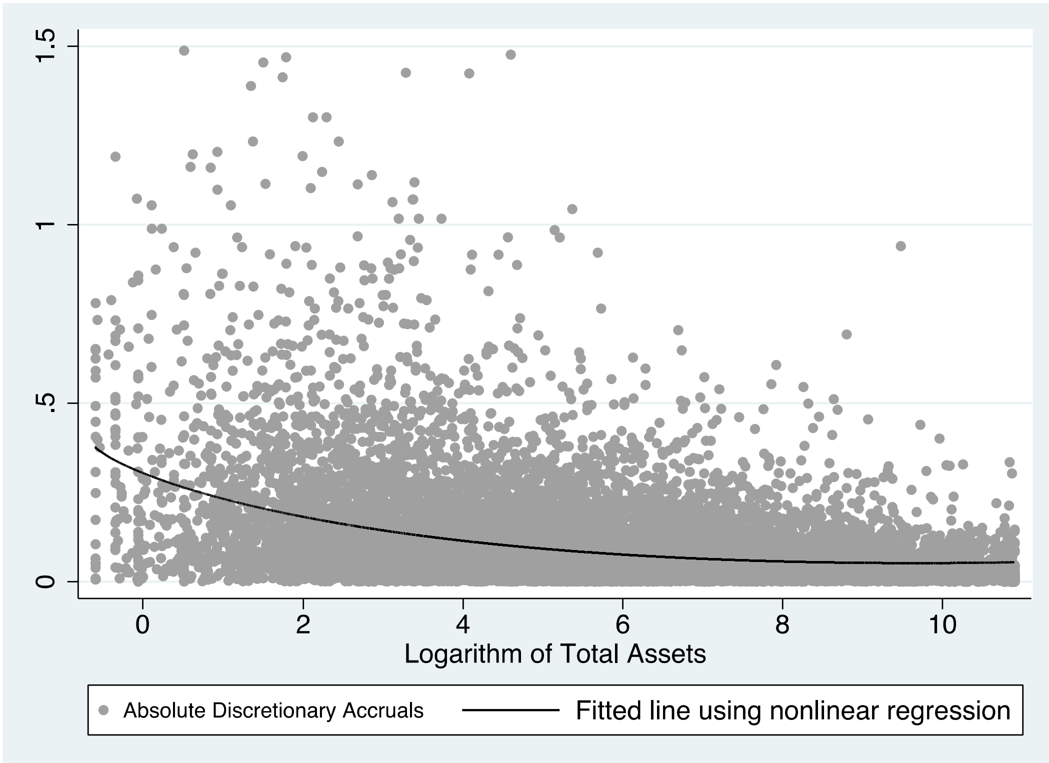
DeFond, M., D. Erkens, and J. Zhang. 2016. Does PSM really eliminate the Big N audit quality effect? *Forthcoming Management Science*.

Lawrence, A., M. Minutti-Meza, and P. Zhang. 2011. Can Big 4 versus Non-Big 4 differences in audit-quality proxies be attributed to client characteristics? *The Accounting Review* 86 (1): 259-286.

Morgan S., and C. Winship. 2015. Counterfactuals and Causal Inference Methods and Principles for Social Research. Second Edition, Analytical Methods for Social Research, New York: Cambridge University Press.

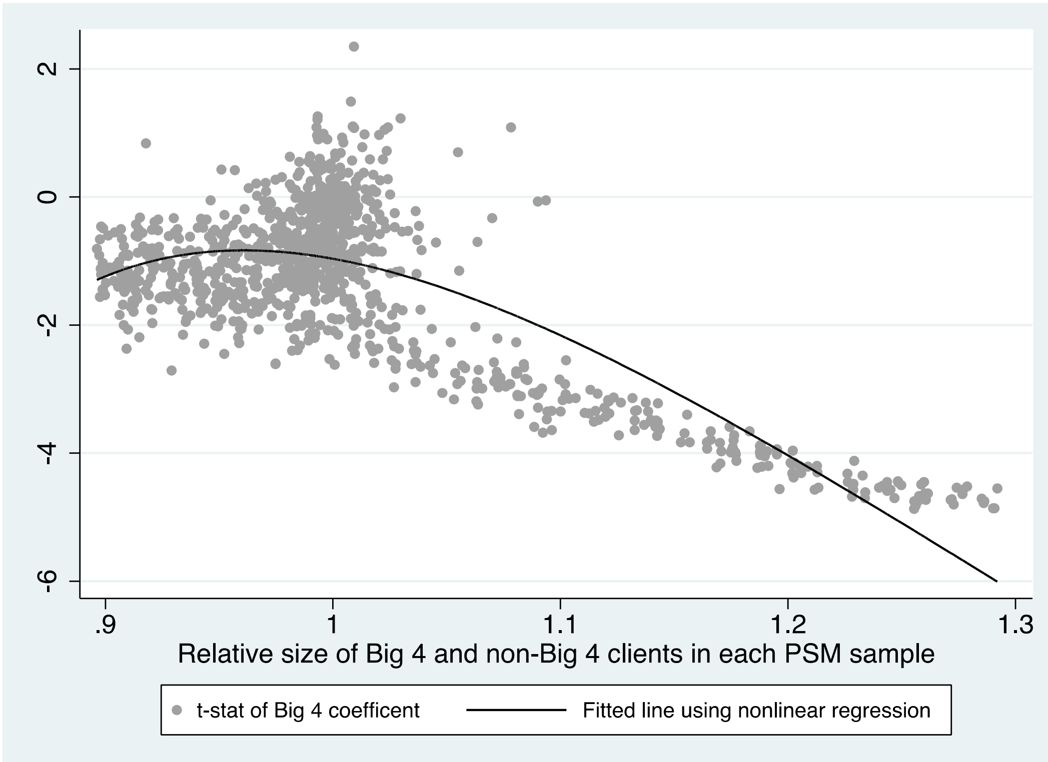
Rosenbaum R., and D. Rubin. 1983. The central role in the propensity score in observational studies for causal effects. *Biometrika* 70 (1): 41-55.

Figure 1–Association between Absolute Discretionary Accruals and Client Size



This figure shows a scatterplot of absolute discretionary accruals and client size, measured as the logarithm of total assets, as well as a fitted nonlinear regression line showing the association between these two variables. The plot uses data from years 2003 to 2006, the sample period in LMZ which overlaps with that in DEZ. The data and variables calculations are as described in LMZ.

Figure 2–Effect of Imbalance in Client Size on the Statistical Significance of the Big 4 Coefficient in 1,000 Variations to PSM Design Choices



This figure shows a scatterplot of the *t*-statistic of the Big 4 coefficient and the ratio of the mean total assets of Big 4 to non-Big 4 clients for 1,000 variations to PSM design choices, as well as a fitted nonlinear regression line showing the association between these two variables. The plot uses data from years 2003 to 2006, the sample period in LMZ which overlaps with that in DEZ. The data and variables calculations are as described in LMZ. Each PSM sample was generated matching 1:1 non-Big 4 to Big 4, using propensity scores with the first-stage functional form in LMZ, without replacement, including random variations in the nonlinear terms, and pruning a random percentage between 1 and 99 of the worst matched pairs.