DATA WOMBLING: WHAT RE-ANALYSIS OF NATURALLY OCCURRING STUDENT DATA CAN TELL US ABOUT COURSES, STUDENT PERFORMANCE AND ACCESS TO THE LEGAL PROFESSION

Cath Sylvester*, Rory O’Boyle** and Emma Hall***

Abstract: Education sites are data-rich environments and such is the embarrassment of riches that very little use is made of these data beyond their initial purpose. Naturally occurring data about students’ backgrounds, their previous attainment and their module by module progression are, in a sense, “lying all over the ground”, but the combination of data sets or longitudinal analysis rarely takes place. This article considers two projects which take inspiration from the Wombles and re-purpose these discarded data: using large-scale data sets to reveal the potential for more nuanced understanding of students’ trajectories through legal education and towards the legal profession in England and Wales and in Ireland. Simple linear assumptions about who can and should enter the profession are challenged by our findings, and the potential to identify important developmental shifts and critical experiences on the way to qualification poses questions for curriculum design and improvements to legal education and access in both jurisdictions.

Keywords: universities’ use of student data; quantitative analysis; key characteristics; student trajectories; access to the legal profession

I. Secondary Data

A. Introduction and definition

The importance of secondary data analysis in social science generally has increased hugely;¹ however, secondary analysis is not used that widely in educational

---

* Associate Head of Law and Associate Professor, Northumbria Law School, Northumbria University, United Kingdom.
** Diploma Manager, Law Society of Ireland, Dublin, Ireland.
*** Research Assistant and Undergraduate Sociology and Anthropology Student, University of Cambridge, United Kingdom.

research.² Educational researchers are much more likely to simply make reference to findings from previous studies, rather than going back and reanalysing the primary data. However, the potential for secondary analysis in educational research would appear to be enormous, with large-scale high-quality data sets freely available. For example, in the United Kingdom, this includes data from the General Household Survey, the Labour Force Survey, the British Household Panel, the Youth Cohort Study and the National Child Development Study.³ Apart from such national statistics, researchers working within the third-level sector often have access to a vast amount of longitudinal data relating to student performance and progression which, again, have huge potential to reveal relevant trends and patterns which if understood could influence educational policy and the targeted deployment of limited resources. As Gorard states, “the range of sources for secondary data is now phenomenal”.⁴ However, Papasolomontos and Christie argued that “it appears that analysis of publicly available datasets is not well established in educational research methodology” and that the “potential for studies of educational progression and lifelong learning has not yet been exhausted”.⁵

Glass defined “secondary analysis” as “the re-analysis of data for the purpose of answering the original research question with better statistical techniques, or answering new questions with old data”.⁶ Rather than “secondary analysis” or “secondary data analysis”, somewhat confusingly Tashakkori and Teddlie preferred to term such data analysis simply as an “unobtrusive measure” on the basis that the term is more inclusive. We will continue to use the term “secondary data analysis”; however, following Tashakkori and Teddlie’s logic, it is certainly “unobtrusive” in that it allows the researcher to examine the social phenomenon “without interfering or changing it”.⁸ Jary and Jary offered a broad definition of secondary data analysis as “an enquiry based on the re-analysis of previously analysed research data”.⁹ Such broad definition is welcome in that it encompasses the possibility of the data being re-analysed by the original researcher or by someone entirely different. It also encompasses the possibility of the data being re-analysed either for a new purpose or for the same purpose for which it was originally collected but using new statistical techniques.

---

⁶ See Glass, “Primary, Secondary, and Meta-Analysis of Research” (n.1).
⁸ See Glass, “Primary, Secondary, and Meta-Analysis of Research” (n.1) p.3.
combination of all these independent variables was likely to identify particularly “at risk” students.

V. Conclusion

We have reviewed two different case studies using secondary data analysis in two extremely data-rich environments, namely the Northumbria University and the LSI. In both cases, the data reveal significant relationships between students’ characteristics and background and their performance. Indeed, it is entirely fair to say that in both institutions we are only scratching the surface of the trends and patterns that this type of secondary data can potentially reveal with regards to student progression.

The Northumbria University study focussed on the early stages of the route to qualification, the QLD and LPC in the form of an M Law programme. The study gave rise to a number of hypotheses about how students progress through a four-year exempting programme. This study established that student achievement at the end of years 1 and 2 is not predictive of performance in the final year of the M Law programme. This brings into question many important judgments that are made on the basis of year 1 and 2 marks by employers, educators and quite possibly students, which can have a substantial impact on students’ access to key requirements necessary for qualification as a solicitor. These findings have possible implications for the Law School. We need a closer understanding of these early years of study. Are students using them to make progress towards their final classification or should we be reconsidering the content — is there anything we can learn from year 3 and 4 modules that might improve achievement at an earlier stage?

The finding as to whether performance in modules with significant practice orientation such as year 3 SLO and Property Law and Practice has any relationship to high achievement in year 4 is more tenuous. There does appear to be a correlation

![Figure 8: Age Profile of PPC Students](image-url)
between these subject marks and year 4 achievement although causation may be a result of other factors. These initial findings assumed module content and in particular the mix between practice/skills/theory/black letter law content based largely on our understanding of the M law curriculum; however, no systematic breakdown of module content has been carried out. Similarly, no detailed consideration has been given to teaching and learning and delivery of these modules. Methods of assessment will also have an impact on marks that there is evidence to suggest that modules assessed by coursework normally attract higher marks. In addition, the focus on marks is only one variable among many that may impact on performance in year 4. In year 4, students have close supervision (one supervisor to six students in SLO year 4) and individual supervisors for their dissertation, and there is evidence in student feedback of high levels of engagement in these modules. It may be that high marks in these modules have been affected by the way that these modules are delivered during the course of year 4. In addition, student motivation and ability to self-study more effectively may have an impact on year 4 marks.

The correlation between year 3 marks and the year 4 recalculated class are harder to interpret, but they are not inconsistent with some of what we have suspected; that clinical or practice focussed modules may encourage higher achievement by students with previously lower marks and that modules with a higher practice content may improve marks in both practice and academic modules and the biggest question of all that clinical modules impact on every aspect of legal education not just the improvement of legal skills and processes.

With regards to the Irish case study, the research objectives were to identify attributes or characteristics of students that are predictive of success, or, conversely, attributes or characteristics of students that are predictive of failure at both the FE-1s and PPC level using secondary data analysis. Such secondary analysis of available data is invaluable, because if properly understood, it will help to identify “at-risk” students for both the FE-1 exams and the PPC. This information could then be used to build a predictive model to inform intervention strategies designed to assist such students.

One major issue was the extent to which the FE-1s are predictive of PPC performance. This is not only merely important for the perspective of building a predictive model but also has importance with regards to the validity of the assessment process. This is because the PPC acts as preparation for practice, completion of which permits entry into the profession. On the premise therefore that the PPC is assessing the aptitudes necessary to be a solicitor, if the FE-1s are intended to select those most suitable to be solicitors, then you would expect a reasonably strong correlation between students’ FE 1 and PPC performance. By way of comparison, from the United States, there is good evidence that the

---

LSAT score is highly effective as a predictor of first-year law school grades.\textsuperscript{75} It is therefore welcome that our results show a reasonably strong correlation between students’ FE-1 and PPC results, indicating that the FE-1 is acting as an appropriate filter. However, given the importance of this issue, a wider scale study with regards to such correlations is warranted.

Another area focused on was the performance of law degree holders against that of non-law degree holders. From the American perspective, Kissam states, “that a student’s experience in taking twenty-five to thirty … exams in the course of a law school career will constitute effective, if not the most appropriate, preparation for taking bar exams”.\textsuperscript{76} Applying the same logic to the Irish context, we found that somewhat surprisingly the holding of a law degree does not appear to confer any particular advantage for students undertaking the FE-1 exams, although a certain advantage was evident at the PPC stage. This raises interesting issues with regards to the value of an undergraduate law degree as a means of preparation for professional assessments. Again, given the importance of this issue, a wider scale study is warranted.

There were also other specific categories of students that warrant further investigation. For example, gender was seen to be statistically significant at the PPC level, with females outperforming males. Age was also statistically significant, with older students performing less well at the PPC level. The performance of those attending the FE-1s via the Preliminary Examination (ie, non-degree holders) was substantially below average. The Preliminary Examination has the potential to liberalise entry routes into the profession, so the performance of this cohort of students is important. However, the sample contained a very small set of such students (five in total). A fuller more wide-scale study of the performance of such students over a number of intakes would be helpful.

Finally, this type of secondary data analysis is invaluable in describing what is happening with regards to student attainment and progression. However, there is also a strong case to be made for potential follow-up qualitative research to ascertain why it is happening. For example, in the Irish context, we have seen that more than anything, the FE-1s acts as the main filter into the profession. Of particular interest, therefore, are those individuals who initially attempt the FE-1s but do not gain sufficient credit to progress to the PPC. However, with no clear defining characteristics of such students, it would be important to understand more about their motivation in the first place to attend the FE-1s and why and at what point they decide not to progress the process further. Motivation is also relevant to student performance in the Northumbria M Law programme and further qualitative research considering that the importance of delivery, module content, assessment and student choice would further explore the issues emerging from these initial findings from analysis of student grades.

\textsuperscript{75} Marjorie Shultz and Sheldon Zedeck, “Predicting Lawyer Effectiveness: Broadening the Basis for Law School Admission Decisions” (2011) 36(3) Law & Social Inquiry 620.