APPENDIX I

NOTES ON TECHNICAL METHODOLOGY

The database used as the tool of consolidation and reference in this study differs in a number of ways from many of those currently available for academic and public use online.[1] Rather than being simply a holding-place for transcribed, interpreted data that is comprehensively taken as the authority on its subject properties, the model that has been built to study Jacobite prosopography, this iteration of which is called *The Jacobite Database of 1745* (JDB1745), is based upon a distributed authority source-persona premise. In such a model, data from all sources within the scope of the study are entered, even if they are conflicting sources, without authoritative judgments made upon the properties themselves as they are entered, within reason. In this sense, the distributed authority model is a method of studying Jacobitism in context as opposed to simply being a mechanised tool used to compile and crunch altered data. What emerges from this model is an objective meta-resource that allows contrast and comparison of information relative to each particular researcher's query methods.[2] In this case, these research methods focus on the topics and themes explored by this thesis.

In order to populate the database with enough property information to present a cohesive and topical study, a three-tiered methodology was devised to gather and enter the data chosen for inclusion.[3] First, numerous printed lists of persons connected with Jacobitism were digitally catalogued, beginning with the most relevant and referenced sources used by scholars of the subject. These first sources are the most efficient for building up a core of data to work with, as they contain large numbers of alleged Jacobites with numerous personal factoids spread over a wide breadth of categories. Second, both manuscript and published sources

---

1 For some prominent examples of historical databases currently in use, see the following: The Medieval and Renaissance Italy Prosopographical Data Project <http://www.slu.edu/department-of-history/departmental-projects>; Prosopography of the Byzantine World <http://blog.pbw.cch.kcl.ac.uk>; People of Medieval Scotland 1093-1314 <http://www.poms.ac.uk]>; Open Domesday <http://domesdaymap.co.uk>; Scotland, Scandinavia and Northern European Biographical Database <http://www.st-andrews.ac.uk/history/ssne/>; Irish in Europe Project <http://www.irishineurope.ie>, Bomb Sight: Mapping the WW2 Bomb Census <http://bombsight.org>.

2 As with every well-defined taxonomy, authority records can be provided based upon the consensus of the sources, but not at the expense of any of the sources' informational integrity.

3 For the purposes of this database methodology, the terms *properties*, *elements*, and *fields* refer to the record's characteristic metadata while the *values* are the characteristics and information relative to a particular record. For more on this classification scheme, see Miller, *Metadata for Digital Collections* (New York, 2011), p. 28-29.

known to contain a significant number of named Jacobite participants were entered, regardless of overlap with the previous sources or with each other. Third, a vast selection of diverse primary sources were scanned for Jacobite-related names not only to add further persons to the corpus of data, but also as a sort of control for the first two tiers of entry. This tertiary entry scheme includes, for example, trial records, prison receipts, burgh and council minute books, personal correspondence, and other items from various archives.[4]

A few words should be stated about the normalisation of the data contained within this study, as the methodology of organisation and analysis are central to the nature of this particular database and, therefore, of the results that it offers. It would be most efficient to describe this normalisation in terms of structure/taxonomy, interpretation, and data functionality. A much more detailed evaluation of this could be described, but a brief encapsulation must suffice for the purposes of this thesis. The taxonomical organisation of the included data was created to feature all elements of information that are given in any of the sources that were used. Rather than submit to a 'square-peg/round-hole' methodology that pre-determines the type of information included from the sources, JDB1745 is built upon a survey of extant categories of information present across all sources. For instance, because at least one manuscript list used in the study approximates estate values of prominent Jacobite personalities, a field for measuring estate values will be present for all all entries within the database, even if that information is currently unknown.[5] This ensures the most comprehensive survey possible of the included data with the broadest number of categories of information, regardless of whether it is being used at this stage of the project.

As well, the specific data drawn from the sources in this study are done so with as little modification as possible, preserving the context for analysis rather than skewing the data through interpretation during the entry process. This concept of 'clean' data is central to the creation of JDB1745 and reinforces a need for the historical database to present information in a contextually-appropriate manner. To this end, data fields are entered without correcting the misspellings and informational errors made by the original recorder. Marginalia are preserved and

---

4 All three tiers are outlined in detail in Chapter I: Introduction, pp. 20-26.

5 For example, Estimates of Yearly Rents of Persons Concerned in the Rebellion (28 February, 1746), SPS.54/28/47d.

abbreviations are retained throughout.[6] Fields for research notes and interpretation are present within the database's interface, but not at the cost of the integrity of the original sources. A fundamental part of the normalisation process, then, is coding authority structures for names, place-names, and occupations, just to name a few operable categories. As an example for our purposes, all occurrences of *shoemaker* must be able to be accessed, regardless of spelling or abbreviation, and also must include all instances and variations of *cobbler*, *cordwainer*, and *bootmaker*. To address this, a customised variation of the Booth-Armstrong occupational classification matrix was adopted, with relevant additions and subtractions due to our period of study being a century earlier than that of the model.[7] Name thesauri were created for equating the dozens of forms of similar surnames as they have been written, and demographic analysis would be biased or inaccurate without doing the same using standardised onomastic conventions, when possible, to determine locale.[8] In the case of place-names, granularity is limited by comprehension of the places in question. Therefore, while there are known quantities of counties and parishes, smaller steadings like estates and farms might not be able to be positively identified.

Lastly with regard to methodology, the target functionality of the data becomes vital as part of the normalisation schema. The first and most important part of developing a historical database is having a firm grasp on its scope and eventual aims.[9] Ensuring that the included data is functional to these aims by providing relevant information in a manner that is useful to the researcher was paramount even as the sources were being gathered and entered. Very little information present in these particular sources was left out from entry, as virtually all of it is usable in some manner for historical study and has found some way to be included in this analysis of the Jacobite constituency. It is important to note, however, that the project's intention and its implementation of data functionality is not limited by this particular thesis. JDB1745 is designed to be an objective database instead of an interpretive one; rather than being created with a particular thesis in mind, it is a device available to discover and challenge limitless theses.

----

6 The exception to this is the practice of notating 'dittos' within lists of information, usually written as *do*. As the database entries do not necessarily follow the same hierarchical positioning as found in the original source, this would make searching such entries impossible. Instead, an exact copy of the data that was repeated is used in its stead.

7 W.A. Armstrong, 'The Use of Information about Occupation' in E.A. Wrigley, ed., *Nineteenth-Century Society: Essays in the Use of Quantitative Methods for the Study of Social Data*, (Cambridge, 1972), pp. 191–310.

8 See sections on coding occupational and nominal data in Charles Harvey and Jon Press, *Databases in Historical Research* (Basingstoke, 1996) pp. 226-231.

9 Ibid., p. 105.