

International Journal of Population Data Science

Journal Website: www.ijpds.org



AD|ARC (Administrative Data | Agricultural Research Collection): Linking Farms to Individual, Household and Business Data to create a Research-Ready Dataset for ADR UK

Nathan O'Connor¹, Beth Allen¹, Esther Lewis¹, Sarah Cummins², and Sian Morrison-Rees³

¹Office for National Statistics, London, United Kingdom

²Office for National Statistics, Titchfield, United Kingdom

³Swansea University, Swansea, United Kingdom

The aim of this project was to produce a complete record of farms in England to be linked to individual, household and business data, enabling the creation of a de-identified research-ready dataset (RRD) of farming individuals, households and businesses in England. Following lessons learned from the first phase of this project, a 'farm spine' was created using two farms data sources, 2021 June Survey of Agriculture and Horticulture and 2020–2022 Rural Payments Agency subsidy payment information. This spine utilised address and contact information at individual, farm holding, and business level, enhancing the matching potential when linking to census and business data. Deterministic linkage methods were used to match farms with associated households and businesses. The varied and nuanced nature of the data sources meant a non-standard approach was required when structuring the data for

linkage, as well as ongoing feedback and collaboration between stakeholders to best deliver an optimised, high-quality bespoke dataset. The resulting dataset links geographic and farm activity data, subsidy payments, business ownership and turnover information, and characteristics of farming households. This presentation includes details of the data preparation, challenges encountered, and solutions utilised at each stage of building this complex dataset. To our knowledge, the AD|ARC RRDs are the first datasets linking agricultural data to individual and household-level data at a population level. They provide an invaluable resource that will inform policymakers seeking to support agricultural and rural communities. Its multi-organisational approach and use of varied data sources highlight lessons learned for future data linkage projects.

