$9^{th}$ International Conference on Computational Social Science $IC^2S^2$
July 17-20, 2023, Copenhagen, Denmark

# Pubs, schools and places of worship: scaling laws in rural England

*rural studies, geography, facilities, villages, scaling*

## Extended Abstract

Computational methods have been used extensively in order to understand how macroeconomic indicators and public amenities (road surface, housing) scale with population in urban environments[1]. More recently, social media data has been used to extend that study to gain an even deeper understanding of urban communities [2]. However, given the number of rural communities, the distance between them, and given the problems associated with studying smaller communities using traditional ethnographic methods [3], we suggest that there are opportunities to better understand rural communities using analytical methods.

Given no formal definition of what constitutes a village, we choose to focus on the parish (as the smallest historic administrative unit in England) and to solely focus on parishes that are classed as wholly rural as per the Office for National Statistics' assessment in 2011. This leaves us with 5,066 wholly rural parishes out of 6,587 in England overall - the fact that more than 75% of all parishes in England are wholly rural (and yet those parishes only contain around 11% of England's population) further underlines the challenges of studying rural populations.

In order to study the prevalence of facilities, we introduce a data set containing the amenities present in the output areas (OAs) of England, taken by combining OpenStreetMap data [4] and geographic bounds published by the Office for National Statistics. Data is collected (and published) on 616,216 amenities in 181,881 output areas (with populations of around 300) and is combined with information pertaining to an area's indices of multiple deprivation, area, and population. Additional higher-level geographical groupings (county and region) are joined in to enable comparison of statistics across larger areas.

Comparing the facilities present across rural and non-rural parishes we find that bicycle parking, parking spaces and fast food are the most over-represented amenities in non-rural parishes, and that post boxes, grave yards and benches are the most over-represented amenities in rural parishes. Further credence is given to the idea of the village church, pub or school being at the heart of village life with all of these facilities being over-represented in rural communities. Given their over-representation in rural communities and their importance to the traditional view of "idyllic English village life", we choose to focus on pubs, schools and places of worship.

Based on a hypothesis that there is a relationship between the number of inhabitants a village has and the likely number of facilities in that village, we plot the average number of each type of facility against the population size of a parish and show a clear relationship (see figure 2). As per Bettencourt et. al [1] we then fit a power law to determine the typical numerical relationship between population size and the number of each type of facility.

While there is a clear relationship between population size and number of each type of amenity in rural parishes, there is also a large degree of variance and many parishes are over or under-provisioned. We hypothesize that the area of a parish (and hence its population density) may explain some of that variance. We also investigate whether the wealth/prosperity of a parish is contributory factor. We measure this using a population-weighted calculation of the

average indices of multiple deprivation of a parish (IMD - a way of representing the relative deprivation of an area across 7 key indicators[5]).

We first plot a correlation matrix to demonstrate relationships between our variables of interest (shown in figure 1) and then build a logistic model to estimate the probability of a given amenity in a parish. Using this we find that there is clear statistical evidence that population, area and indices of multiple deprivation all contribute to the likelihood of a parish having a pub, school or place of worship. We find that pubs and schools behave in a much more similar fashion than places of worship do, and we find that they are much more predictable in their nature.

Next, we study whether there are regional variations in the probability of a parish having an amenity. Again, we built a logistic model with the aforementioned features but this time we included a one-hot encoded variable representing the region of England in which the parish was in. Controlling for the size and relative deprivation of a parish, we show how growing population alters the probability of a parish having a facility, by region. An example of this is shown in figure 3.

We find that there are statistically significant differences between some regions for some facilities. In general, the regional differences between pubs and schools are much smaller in magnitude, perhaps reflecting the fact that they are much more susceptible to the traditional laws of supply and demand than places of worship are.

We conclude that there are a range of possible extensions and interesting questions to be answered by studying rural communities and perhaps even the differences between urban and rural communities. Especially with the increased availability of remote working, developers and house-hunters alike may find it interesting to be able to choose villages which are either under or over-provisioned with respect to local facilities. A clustering exercise finding similar rural communities in terms of size, population, deprivation and facilities could yield an interesting base for comparative ethnographies to understand differences and similarities between seemingly similar rural communities. Finally, we also suggest that we have, in essence, built a formula for amenities in rural villages. When considering planning permission applications to allow the building of new houses, local councils should consider whether a parish is over or under-served with respect to facilities, and using this data they now have a way to do so.

# References

[1] L. M. Bettencourt, J. Lobo, D. Helbing, C. Kühnert, and G. B. West. Growth, innovation, scaling, and the pace of life in cities. *Proceedings of the national academy of sciences*, 104(17):7301–7306, 2007.

[2] A. J. Gross, D. Murthy, and L. R. Varshney. Pace of life in cities and the emergence of town tweeters. *Sage Open*, 7(4):2158244017745113, 2017.

[3] S. Neal and S. Walters. Strangers asking strange questions? a methodological narrative of researching belonging and identity in english rural communities. *Journal of Rural Studies*, 22(2):177–189, 2006.

[4] OpenStreetMap contributors. Planet dump retrieved from https://planet.osm.org . https://www.openstreetmap.org , 2017.

[5] R. A. Payne and G. A. Abel. Uk indices of multiple deprivation-a way to make comparisons across constituent countries easier. *Health Stat Q*, 53(22):2015–2016, 2012.

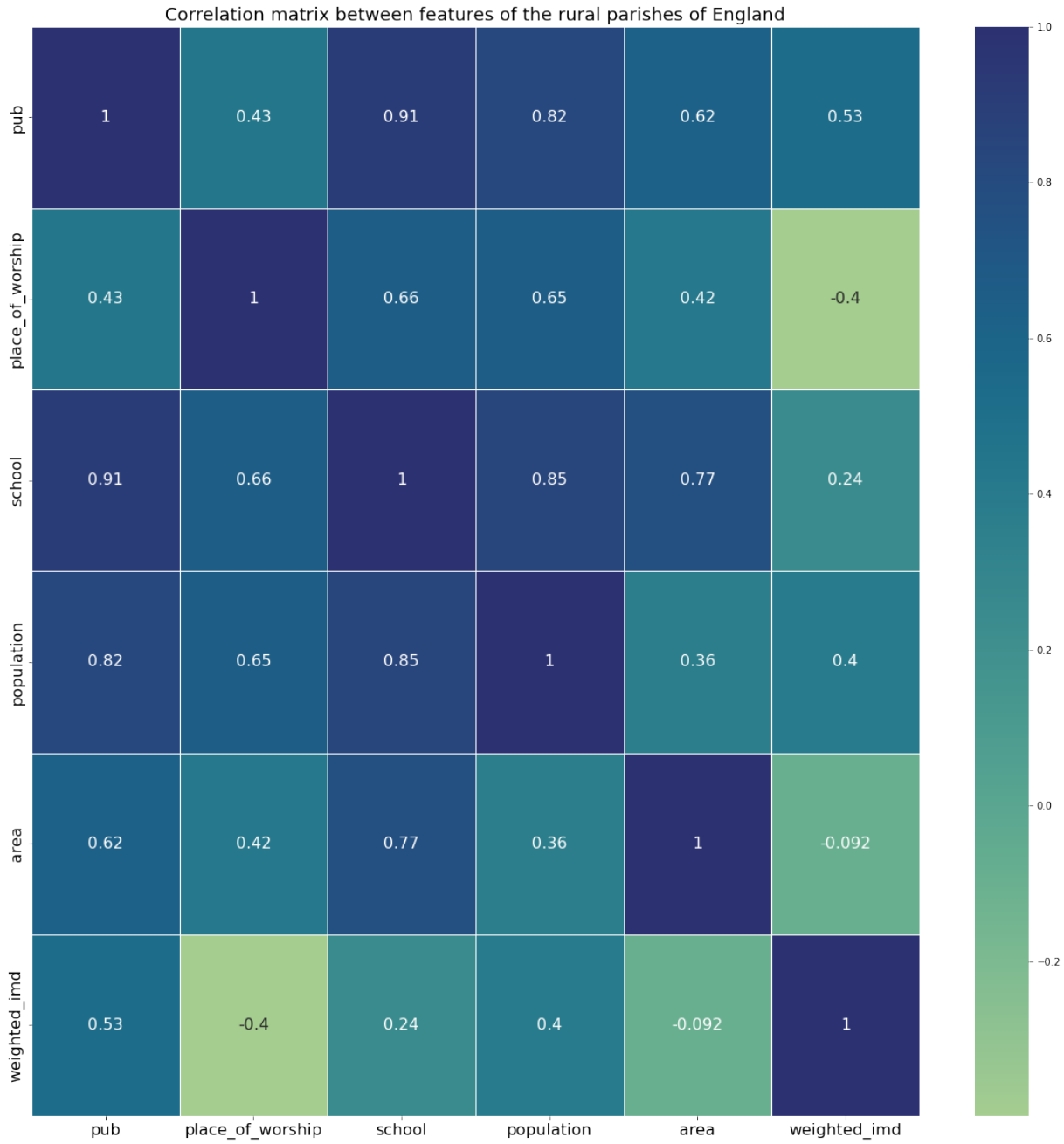


Figure 1: Correlations between the various metrics of interest per parish

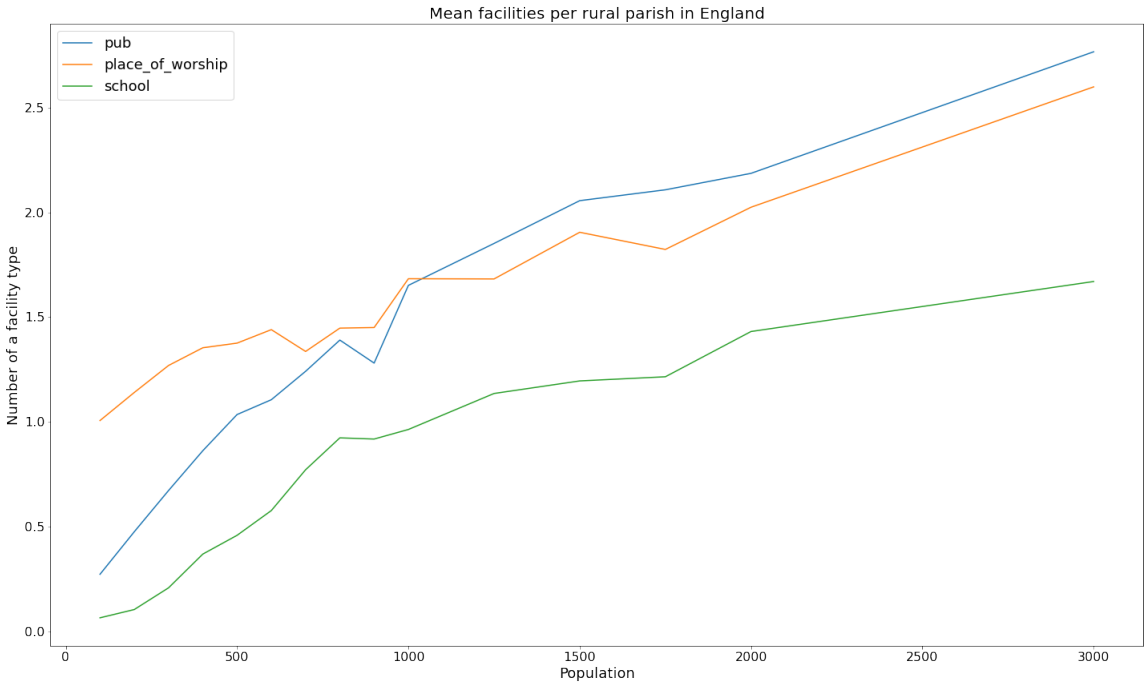


Figure 2: The relationship between the population of a rural parish and the average number of facilities within its boundaries

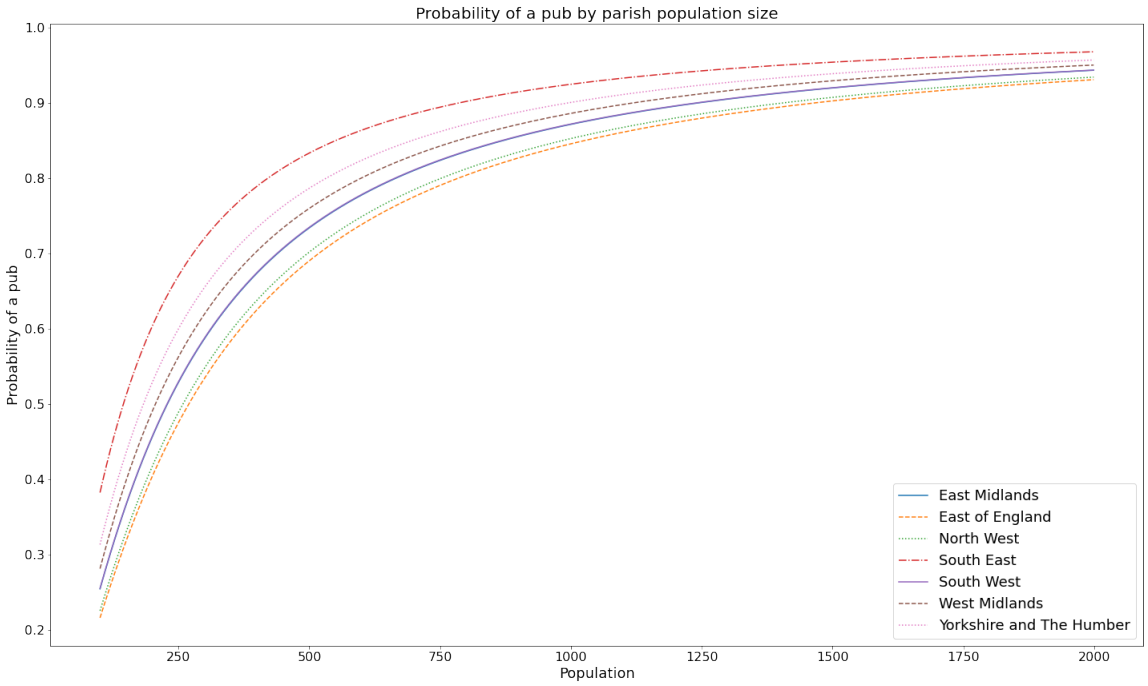


Figure 3: The relationship between the probability of a rural parish having a pub and its population, by region