**Supporting information**

For the dual purposes of correction for heteroskedasticity as well as a systematic exploration of the inversely U-shaped variation in regression residuals, we use one integral estimation method. In case of *a priori* information that can specify heteroskedasticity as some function  of the independent variable of the linear regression such that ,

one resorts to weighted least squares estimation where observation *i* is weighted by  for all *i*.

Nevertheless, since the inversely U-shaped heteroskedasticity we observe here is a new phenomenon, we have no such prior information except for what we can observe empirically. The skewed, bell-shaped pattern observed in Figure 1 is representative of all disciplines. Hence, we use one and the same functional form in all regressions.

The functional form that can best capture this pattern is the probability distribution function of the Gamma distribution, where *θ* and *γ* are parameters. Accordingly we assume and estimate the parameters,  as well as  (which is yet another parameter that scales the functional form) and by ordinary least squares estimation on the basis of the residuals of initial (non-weighted) linear regression estimation. Note that the superscript *j* denotes individual disciplines.

Figure S1 depicts the estimated variance functions (i.e., the continuous red curves) along with the scatter plot of absolute values of the respective error terms against (log) city size for the whole dataset (all disciplines together) and three selected disciplines. These examples also show that, despite the similarity in the general form, scientific disciplines are heterogeneous in terms of the variation not explainable by scaling laws. For example, while such variance is highest at the city size of 73,466 (i.e., *e11.205*) at the aggregate level, it is 252,760 for arts and humanities and 165,703 for engineering.