



April 1, 2022

Dear Editors:

Please accept the attached manuscript, “The REDI package: Random Empirical Distribution Imputation of continuous from categorical incomes” as a submission to *Stata Journal*. This manuscript discusses a new community-contributed Stata program that I hope will be of use to a large segment of Stata users. The **redi** package implements a novel method for imputing continuous values from binned data using an unlinked dataset. I present the syntax for the package and a brief overview of the method’s advantages and limitations.

The paper goes on to present an example of the package using a small, publicly available General Social Survey dataset that has been used for other Stata demonstrations. The example can be replicated by placing all files in the same directory and running the “redi\_gss\_example.do” file. Additionally, my code and replication instructions for the project are available at GitHub ([https://github.com/mollymking/redi\\_code](https://github.com/mollymking/redi_code)).

Although a related theoretical paper that presents the underlying foundations of this method has been accepted and is forthcoming at *Sociological Methodology* (and can be viewed at <https://osf.io/preprints/socarxiv/eswm8/>), the present paper is unpublished and not under consideration elsewhere.

In summary, this paper:

- Develops a Stata package (**redi**) that implements a novel method, Random Empirical Distribution Imputation, which enables researchers to impute discrete individual income observations from categorical data.
- Exhibits the usefulness of this package in transforming categorical to continuous income using a small (n=254) sample GSS dataset that anyone can manage with their version of Stata.

I can best be reached by email at [mmking@scu.edu](mailto:mmking@scu.edu) or at the following mailing address:

1365 Market Street  
Santa Clara, CA 95050

Thank you for your consideration.

Sincerely,

Molly King, Ph.D.