ECON 121, Applied Econometrics and Data Analysis **Prof. Tom Vogl**

Fall 2024 **University of California San Diego**

PROBLEM SET 1: NAMING TRENDS IN THE UNITED STATES

DUE BY 11:59 PM PDT ON THURSDAY 10/10

A number of cultural, political, and economic factors affect how parents name their babies. In this problem set, you will explore naming trends in the United States using the Social Security Administration Baby Names dataset (https://github.com/tomvogl/econ121/raw/main/data/ssa_names.csv). You should be able to do the entire problem set using `tidyverse` functions from the coding example and the course introduction slides.

To install R and *RStudio, follow this link. You are encouraged to work in a group of up to 4 members. You may write code together, but you must write verbal answers yourself. Please use a Markdown template for your code. Write verbal answers in the comments within the Markdown file, so that you produce a single PDF with code, results, and writing, which you will upload to Gradescope.*

1. List your group members.
2. Load `tidyverse`, read in the data, and summarize, with 1-3 sentences explaining the summary statistics.
3. What are the most popular boy and girl names over the entire period of the dataset? How many babies were given these names between 1940 and 2022? Explain in 1-2 sentences.
4. In cultures with more nonconformity, parents may be more likely to choose uncommon names, leading to a greater number of names. How has the number of unique boy names and unique girl names changed over time in the United States? For each sex separately, plot the number of unique names over time in a graph. Interpret the graph in 2-3 sentences.
5. Changes in the number of unique names may reflect changes in the number of babies rather than changes in conformity among parents. Redo the graph from question 3 to address this issue. For each sex and year, calculate the ratio of the number of unique names to the total number of babies. For each sex, plot the ratio over time in a graph. Did the number of unique names rise faster than the number of babies? Are the patterns consistent with increasing or decreasing conformity? Explain in 2-3 sentences.
6. Choose 1-3 names that interest you. State the names and explain why you chose them in 1-2 sentences.
7. Combining across all years, how common are your chosen names relative to the most popular names you found in question 3? Explain in 1-2 sentences.
8. Plot the time series of frequencies for all 1-3 names in one graph. Interpret your results in 3-4 sentences.