Journal of Accounting Research – Data Description Sheet

Paper: *Gaming the IRS's Third-Party Reporting System: Evidence from Pari-Mutuel Wagering*

Author: Victor (Duke) Ferguson II

1. This paper is based on my solo-authored dissertation at Purdue University. I (Victor "Duke" Ferguson) was the only person who worked with the data for the purposes of this project.

2. The data for the main analysis of this project was obtained from The Jockey Club and Equibase Inc. on August 15, 2019. The data for Table 4 was downloaded from the IRS's website, the download for the final version of this table occurred on September 8, 2022.

3. While I do not have a non-disclosure agreement regarding the data or limits on the publication of results, Equibase Inc. does generally charge a fee for this type of data. The organization shared this data with me based on the understanding that it would be used only for my academic research. With this in mind, making the data public would be inappropriate. My contacts at Equibase have been shared privately.

   Data for Table 4 is publicly available from the IRS at https://www.irs.gov/statistics/soi-tax-stats-county-data.

4. The data used in the primary analysis of this paper originated from a single dataset obtained from Equibase Inc., no collection was necessary. The raw data used for Table 4 was downloaded from the IRS's website and the needed variables were copied and pasted into a usable dataset. A complete description of the steps needed to process the data is included in the paper.

5. The Stata code used to reach the results tabulated in the paper will be provided. Because only one dataset was used in the primary analysis no mergers were necessary, racetracks are identified by three-letter codes. For Table 4 state names serve as identifiers. I have included an excel sheets listing these identifiers.

6. Data will be maintained my me (Victor "Duke" Ferguson) for at least six years.