

Online Appendix for “What Do Big Data Tell Us About Why People Take Gig Economy Jobs?”

by Dmitri K. Koustas

This appendix describes how gig income is identified in the app data and the construction of the final estimation sample.

1 How gig income is identified in the app data

I search transaction strings for credits into bank accounts from major “on-demand” gig economy platforms. The platforms I search for are chosen based on published lists and anecdotal sources. I consider a platform to be “on-demand” if the gig firm is widely known for having an easy sign up process, and where households have a reasonable expectation of earning money on any given day.

Refunds and reversals that are not gig income will also be picked up by this procedure. To deal with these, I exclude transaction strings that contain phrases commonly associated with a pending debit or a refund, such as “pre auth”, “pos adjustment,” “canceled,” “return,” “refund,” “reversal” or “adjustment.” I also search for corresponding debits from the same account up to one week earlier which would cancel out the income and indicate the income is a refund.

In the end, I am able to identify transaction strings associated with 10 firms which appear to be payments primarily to gig workers on these platforms based on the pay frequencies and amounts.

One concern is that I am also including non-gig workers at these firms. Gig companies tend to pay instantly or at weekly pay frequencies, whereas most traditional jobs pay on biweekly frequencies. To ensure I am not picking up traditional (non-gig) employees at these firms, I exclude individuals who have gaps between paychecks of 13-15 days, but never the 7 days indicative of a gig worker.

My final sample restricts to people with at least two payments from any gig company, with total earnings at that gig company greater than \$100. I also drop the top 1% of earners in either gig or non-gig jobs to ensure a small number of high earners are not driving results.

2 Construction of estimation sample

After merging the users I identify as gig economy workers to their transactions, I restrict to a sample of weeks with positive weekly spending.

All dependent variables are winsorized at the 1% level to account for outliers, and the sample in each event study regression is restricted to be balanced 4 weeks pre and post the event. The balance restriction before the event is particularly important to exclude app users that enter the data *after* starting a gig economy job.