What Do Big Data Tell Us About Why People Take Gig Economy Jobs?

By Dmitri K. Koustas*

Why do households take gig economy jobs? There are now several studies examining labor supply of individuals of a particular gig economy company, but little is known about the economic activity of these individuals outside of the gig economy, or even on other gig economy platforms. New surveys have been designed with the intent to capture the alternative workforce and the gig economy, but these surveys have thus far faced challenges in implementation and by their nature only provide point-in-time estimates.[1] "Big" data from financial accounts provide a unique opportunity to examine the complete economic activity of gig economy workers. I employ one such dataset in this study, and another paper in this same issue uses related data from a large financial provider (Farrell and Greig, 2019).[2]

After identifying gig workers in accounts data, I document how key components of the household balance sheets of these workers evolve around starting a gig job. I show that households have outside income and liquid assets that are deteriorating rapidly in the quarter before starting a gig economy job and partially recover in the quarter afterwards. There are two main hypotheses consistent with these findings: 1) a voluntary reduction in outside labor supply and running-down of assets while waiting to gear up for gig work, or 2) financial distress due to outside shocks. The latter explanation can have important implications for studies focusing on gig economy activity only. I briefly discuss two examples: omitted variable bias from not observing outside shocks, and biases from multijob-holding and credit constraints when estimating structural labor supply elasticities.

I. The Gig Economy in Big Data

This paper employs a unique, transaction-level dataset from a large financial aggregator and bill-paying application.[3] A strength of these particular data is the comprehensive coverage of accounts across different financial providers: users of the app can choose to link almost any financial account, including bank accounts, credit card accounts, and utility bills. Each day, the app automatically logs into web portals for a user's accounts and obtains account balances and daily transactions. The app had approximately 2.1 million active users over the period 2012-2016 available for this study.

Baker (2018) provides an overview of benefits and caveats of data like these in detail, and so I will only briefly address issues specific to the data and context that they are used here. Because these data require households to voluntary select in, a potential concern is non-random selection into

* Harris Public Policy University of Chicago, 1307 E. 60th St. Room 3055, Chicago, IL 60637, dkoustas@uchicago.edu. This research uses data from a project at the University of Michigan and the University of California, Berkeley, supported by a grant from the Alfred P. Sloan Foundation. It uses anonymized data from a financial aggregation and bill-paying computer and smartphone application (the "app"). I am grateful to Michael Gelman, Yuriy Gorodnichenko, Shachar Kariv, Matthew D. Shapiro, Dan Silverman, and Steven Tadelis for collaboration on this project and comments on this paper.

[1]Most notably, the 2017 Contingent Worker Supplement to the Current Population Survey, had a supplement for electronically mediated work.

[2]IRS tax returns data provide a complete picture of outside income, see for instance Abraham et al. (2018); Jackson, Looney and Ramnath (2017); Collins et al. (2019). However, tax data are lower frequency and do not contain information on savings or spending.

[3]These same data have previously been used to study the high frequency responses of households to shocks such as the government shutdown (Gelman et al., Forthcoming), anticipated income (Gelman et al., 2014), and the 2014 fall in gasoline prices (Gelman et al., 2016).

the app; for instance, if users of the app are more financially responsible in that they are more likely to find sources of extra income to smooth income shocks, this could be a threat to external validity. Reassuringly, the propensity to do gig work within these data seems in line with external datasets (1-2%). While it might be desirable to reweight the sample to be representative of the population, the particular data available for this study only have demographic information attached for a small subset of accounts. These demographic data have been validated with outside sources in previous work, and the data are found to be broadly representative of the U.S. population with bank accounts (Gelman et al., 2014, 2016).

For the purposes of this paper, I focus on the "on-demand" gig economy. Using published lists and anecdotal sources, I compile a list of gig firms that are well known for having an easy sign up process, and where households have a reasonable expectation of earning money on any given day. I exclude companies from my sample that require specialized knowledge (the so-called "expert economy") or have uncertain demand. I identify these sources of gig economy in the app data by searching for income into bank accounts from these firms.[4] I find approximately 25,000 earners on 10 popular gig platforms meeting these criteria, with the two ridesharing platforms Uber and Lyft comprising approximately 90 percent of the sample.

II. Evolution of Household Balance Sheets Around Starting a Gig Job

In this section I investigate how household balance sheets evolve around starting a gig economy job. I use an event-study framework, which provides a non-parametric way of exploring the evolution of key variables around starting a gig job, controlling for individual heterogeneity in baseline levels of an outcome, as well as seasonality and trends. The point of this exercise is to examine the pretrends, not to estimate causal effects. The event-study specification I use is standard and given as follows:

$$y_{it} = \sum_{k \in K} \beta_k D_{it}^k + \alpha_i + \alpha_t + \epsilon_{it} \quad (1)$$

where y_{it} is an outcome variable of interest, a_i is an individual fixed effect, and α_t is a time period fixed effect. $D_{it}^k = I\{t = E_i + k\}$ is a indicator for time to first gig pay, E_i, with negative k indicating a future event date, and positive k indicating the event occurred k periods in the past.

I run my specifications at two different frequencies, depending on the dependent variable. For income outside of the gig economy, I aggregate the data to biweekly frequency to account for the fact that most payroll income is paid at biweekly frequency and divide by two to convert biweekly income to weekly frequency. In this case, I consider the week before first gig payment and the week of the gig payment as the event date. (Households will start working the week before receiving their first gig payment.) For all other variables, I run the specification at weekly frequency, dating the event as the week of first payment, and omitting the indicator for the period two weeks before first gig earnings. The β_k coefficients are thus relative to the period before the household first started working in the gig economy. All dependent variables are winsorized at the 1% level to account for outliers, and the sample is restricted to be balanced 4 weeks pre and post the event.[5]

A. Results

I present results graphically for my key variables of interest in Figures 1 and 2, plotting coefficients up to one quarter pre and post starting a gig job.[6]

[4] See the online data appendix for more detail.

[5] The final sample is approximately 17,000 after this restriction. Outside of this window, the sample will be unbalanced and the composition of the sample may change, requiring caution in interpreting coefficients. I have experimented with longer balanced panels and find results to be nearly identical over the one-quarter window (albeit less precise).

[6] While households can be followed for longer, as the time period extends, concerns rise about account attrition and non-syncing accounts.

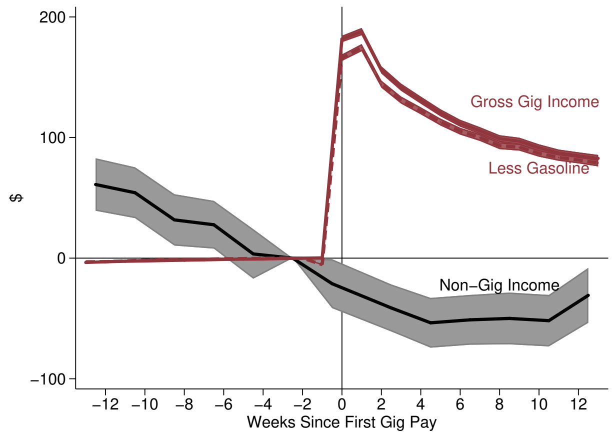

FIGURE 1. EVENT STUDIES AROUND STARTING A GIG JOB: INCOME IN AND OUTSIDE OF GIG ECONOMY

Note: Figure plots the event-study coefficients from Equation 1, for different outcomes of interest, see text. Effects are relative to the period before starting a gig economy job. The x-axis shows "Weeks Since First Gig Pay." "0" indicates the first week any gig pay is observed. Negative values indicate weeks before first gig pay is received, and positive values indicate weeks after first gig income is received. Standard errors are clustered at the individual level. 95% confidence intervals are shaded around the estimates.

As shown in Figure 1, gross gig income reaches about $200 in the first weeks after starting a gig job, before declining by about $100 one quarter later. Note that gig income here includes weeks with $0 from not working. Considerable debate has focused on the costs associated with gig work; gig income will be net of fees paid to the gig platform, but not of taxes, expenses and other depreciation. The transactions data allow us to to identify gasoline spending and remove it.[7] The dashed line shows the coefficients for gig income net of gasoline spending. The increase in income is about $13 less per week.

The figure also shows that non-gig income falls over the quarter before the household starts a gig job, bottoms out about one month after the household starts a gig job, and partially recovers over the next two months. The total earnings loss over the window can be calculated by comparing peak earnings one quarter before to the trough one month after, implying a total drop in outside income of approximately $115 per week, on average. Doing a similar calculation for every period in the window and summing them up implies total earnings losses of $900. The total rideshare earnings less gasoline over these 13 weeks is $1,550, implying that the earnings losses are more than accounted for by gig earnings (before taxes/depreciation), on average.

Figure 2 examines liquid assets and liquidity. The left-hand panel shows net balances (total liquid assets in bank account and checking accounts, net of credit card debt). Net balances decline by over $400 in the period before starting a gig job, stabilize when starting a gig economy job, and recover slightly over the post period. The right-hand panel focuses on two measures of credit constraints: the share of the sample with less than $100 in bank balances during the week, and the share with less than $100 in available credit (credit limit minus the credit card balance). Both measures are rising in the period before starting in the gig economy. The share with less than $100 in balances falls by over 2 percentage points around the time of starting in the gig economy, before rising again, while the growth in the share with little available credit only slows. These figures suggest that credit constraints are likely to be present both before and after entering gig work.

[7]See Gelman et al. (2016) for a discussion of how gasoline spending is identified in the data.

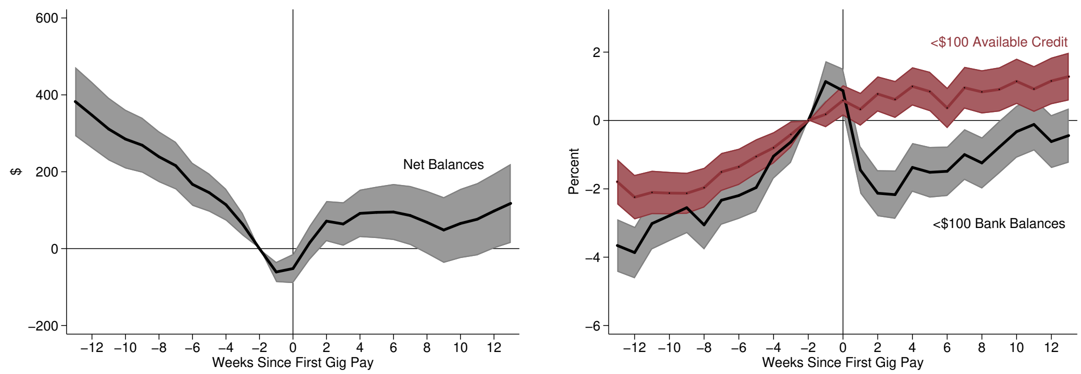

FIGURE 2. EVENT STUDIES AROUND STARTING A GIG JOB: ASSETS AND LIQUIDITY

Note: See notes for Figure 1.

III. Discussion

The analysis above reveals striking pretrends in income and assets. There are a number of potential explanations that are consistent with these findings. One interpretation is that the decline comes from gearing up for gig work. A second interpretation is that a gig worker is facing outside shocks, such as unemployment or wage cuts. This latter explanation has potentially important implications for the validity of previous studies focusing on gig economy activity only. Below, I discuss two examples.

A. *Bias from Not Observing Outside Shocks*

Quite simply, large, persistent outside shocks at the same time households start in the gig economy are likely to confound most analyses of the treatment effects of gig economy participation. For example, suppose one were to run the naive estimator:

$$Y_{it} = \beta AnyGig_{it} + \alpha_i + \alpha_t + \epsilon_{it}$$

where $AnyGig_{it}$ is an indicator for having a gig economy job. It is well known that shocks to the marginal utility of wealth increase labor supply: if households select in after a shock, $\mathrm{Cov}(AnyGig, Shock) > 0$. Suppose the shock would lower Y_{it}; then $\mathrm{Cov}(Y, Shock) < 0$, biasing β downwards.

In many other studies, only the "post" period is observed. Studies using proprietary company data will only be able to focus on the period with gig earnings. Figure 1 shows that outside income starts to recover, which could help explain high rates of attrition from the gig economy.[8] In surveys that identify current participation in the gig economy, a naive comparison with other workers in the cross-section could show that gig work is correlated with economic distress; the pretrends suggest that this is not causal, but is the result of events occurring many months earlier. To deal with these issues, in Koustas (2018), I examine consumption-smoothing behavior before and after Uber enters into the market, the timing of which is exogenous from the perspective of rideshare drivers. In addition, I use coworkers at a main job who face similar income processes as a control group.

B. *Estimating Labor Supply Elasticities*

My empirical results imply that people who become gig workers are affected by income losses and credit constraints. This can have important implications for estimating key structural parameters. One example is the intertemporal labor supply elasticity, commonly estimated by regressing changes in log hours on changes in log wages. On one hand, the gig economy appears to provide a perfect opportunity to estimate la-

[8] Alternative explanation include better juggling of gig and non-gig work, or simply learning that gig work is not as appealing as anticipated.

bor supply elasticities, given that hours are flexible, there is considerable variation in wages, and experiments can be designed to provide exogenous variation in wages. As a result, a number of papers have attempted to estimate labor supply elasticities using variation in wages on a popular gig economy platform.[9]

The empirical specification regressing log hours on log wages implicitly has a number of key assumptions: a single market wage, transitory wage shocks and perfect capital markets. Switching across other gigs or jobs outside the gig economy will bias naive estimates of the labor supply elasticity estimated on a single job *upwards*, as workers adjust the allocation of their hours across jobs (Caldwell and Oehlsen, 2018). On the other hand, credit constraints will bias estimates of labor supply elasticites *downward*; intuitively, if constrained households cannot borrow across periods, they will want to work more even when wages are temporarily low (Domeij and Floden, 2006). In sum, multi-job holding and credit constraints present challenges for estimating labor supply elasticities among gig workers.

IV. Conclusion

While it was already well-known that the gig economy serves as a source of secondary income for many households, there has been little evidence to date on the evolution of outside income and assets of gig economy households. The personal finance data used in this study shows that participating households are facing declines in income and a significant running down of assets before entering the gig economy. Implications of financial distress from outside the gig economy have largely been ignored in the new literature on the gig economy, but are likely to matter given the large magnitudes.

REFERENCES

Abraham, K., J. Haltiwanger, K. Sandusky, and J. Spletzer. 2018. "Measuring the Gig Economy: Current Knowledge and Open Issues." *NBER Working Paper No. 24950.*

Angrist, J., S. Caldwell, and J. Hall. 2017. "Uber vs. Taxi: A Driver's Eye View." *NBER Working Paper No. 23891.*

Baker, S. 2018. "Debt and the Response to Household Income Shocks: Validation and Application of Linked Financial Account Data." *Journal of Political Economy*, 126(4): 1504–1557.

Caldwell, S., and E. Oehlsen. 2018. "Monopsony and the Gender Wage Gap: Experimental Evidence from the Gig Economy." *Mimeo.*

Chen, K., J. Chevalier, P. Rossi, and E. Oehlsen. 2017. "The Value of Flexible Work: Evidence from Uber Drivers." *NBER Working Paper No. 23296.*

Collins, B., A. Garin, E. Jackson, D. Koustas, and M. Payne. 2019. "Understanding the Trend in U.S. Alternative Work Arrangements: Evidence from Tax Data." *Mimeo.*

Domeij, D., and M. Floden. 2006. "The labor-supply elasticity and borrowing constraints: Why estimates are biased." *Review of Economic Dynamics*, 9(2): 242 – 262.

Farrell, D., and F. Greig. 2019. "The Evolution of the Online Platform Economy: Evidence from Five Years of Banking Data." *American Economic Review, Papers and Proceedings.*

Gelman, M., S. Kariv, M. Shapiro, D., and S. Tadelis. Forthcoming. "How Individuals Smooth Spending: Evidence from the 2013 Government Shutdown Using Account Data." *Journal of Public Economics.*

Gelman, M., S. Kariv, M. Shapiro, D. Silverman, and S. Tadelis. 2014. "Harnessing naturally occurring data to measure the response of spending to income." *Science*, 345(6193): 212–215.

[9]See, e.g.: Hall, Horton and Knoepfle (2018); Angrist, Caldwell and Hall (2017); Chen et al. (2017)

Gelman, M., Y. Gorodnichenko, S. Kariv, D. Koustas, M. Shapiro, D. Silverman, and S. Tadelis. 2016. "The Response of Consumer Spending to Changes in Gasoline Prices." *NBER Working Paper No. 22969.*

Hall, J., J. Horton, and D. Knoepfle. 2018. "Pricing efficiently in designed markets: Evidence from ride-sharing." *Mimeo.*

Jackson, E., A. Looney, and S. Ramnath. 2017. "The Rise of Alternative Work Arrangements: Evidence and Implications for Tax Filing and Benefit Coverage." *Office of Tax Analysis Working Paper 114.*

Koustas, D. 2018. "Consumption Insurance and Multiple Jobs: Evidence from Rideshare Drivers." *Mimeo.*