Pay insecurity study- Technical report

Silvia Avram

University of Essex

December 2020

1 Objectives

The objective of this study was to test labour supply responses to uncertainty about work availability and pay. Traditionally, the desirability of unstable jobs has been evaluated either by comparing the job satisfaction of workers with different types of contracts or through more in-depth interviews (of Personnel and Development, 2013; Henly and Lambert, 2014; Pennycook et al., 2013). Unfortunately, both types of studies are vulnerable to biases common to observational studies. Jobs with variable hours and pay tend to be on average lower paid, provide less training, offer fewer opportunities for promotion and tend to be concentrated in particular occupations and industries (Farina et al., 2019; Koumenta and Williams, 2019). Workers in these types of jobs tend to disproportionately be younger, female, immigrant and have lower educational qualifications. Both worker and job characteristics can confound the relationship between uncertainty/ instability and labour supply behaviour. Observational studies, be they quantitative or qualitative, have difficulties in separating out these effects.

To overcome these difficulties, this study used an experimental approach. It sought to replicate the labour supply decision and was carried out with low-income, working age participants. Experiments have the advantage of providing a controlled environment where the various aspects of the labour supply decision can not only be observed but also manipulated. In this case, the insecurity aspect can be clearly separated and varied across the treatment and control groups. The confounding effects of worker characteristics can similarly be addressed by randomly assigning participants to treatment and control groups.

To make the experimental setting as realistic as possible, high monetary incentives and a real effort task (transcribing Latin text) were used. When giving instructions to participants, language associated with real-world decision making was deliberately used. Completing the experimental task is referred to as 'working' and associated payments as 'earnings'. The payment associated with the non-work alternative is referred to as a 'benefit'. Withdrawal of this alternative is referred to as a 'benefit sanction'. The language is designed to encourage participants to identify the experimental context with a real-world work decision situation.

2 Experimental sample

301 participants were recruited from the working age low income population in the UK to participate in an experiment carried out through a purpose built online

platform. To be eligible, participants had to be aged between 18 and 60 (inclusive), have a family income of less than £20,000 per year and not be undergraduate students. Graduate students who satisfied the age and income conditions and who were working were accepted. All participants were UK residents at the time they participated.

Two data collection methods were used. 68 participants tool part in nine face to face (f2f) sessions between July 2019 and February 2020. The sessions took place either at a location in Colchester Town or at the EssexLab centre hosted by the University of Essex. Participants were provided with iPads connected to the internet for purposes of participating in the experiment. Subsequently, data was collected through the online recruitment platform Prolific (www.prolific.co). Prolific is a unique recruitment platform in that it is specifically designed for research purposes. It employs strong ethical safeguards and emphasizes data quality including by vetting its participant pool and providing screening tools. 233 individuals took part in four sessions organized between May and June 2020.

There are some differences between the Prolific and face-to- face (F2F) samples. F2F participants were more likely to be female, more educated, more likely to be in work and less likely to have ever received out of work benefits. More information is available here: https://www.iser.essex.ac.uk/research/publications/working-papers/iser/2020-10.
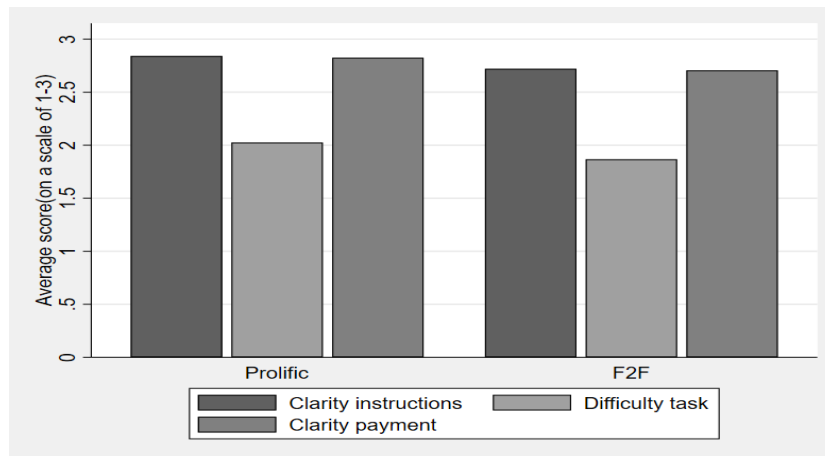
3 Experimental design

Participants completed the experiment on a custom made web application built using oTree (Chen et al., 2016) and temporarily hosted by a cloud service platform. Full experimental instructions are available. Face to face participants completed the experiment using iPads connected to the internet provided by the research team. Prolific participants used their own devices. They were advised that the experiment is designed to be completed using either a computer or a tablet but they were free to use whatever device they wished. Participants were advised that the experiment could take up to 90 minutes. However, no participant took longer than 70 minutes and on average they spent around 35-40 minutes.

The experiment had five stages.

Introductory stage

In the first stage, participants were given general information about the experiment: purpose and objectives, investigating team, use of data and right to withdraw, expected pay, how to make a complaint etc., and consent was obtained. They also

Source: Experimental data

Figure 1: Average rating of clarity and task difficulty

received information about the structure of the session and general instructions before starting.

The lottery stage

In the second stage participants took part in a lottery game designed to measure risk aversion. The lottery is a modification of the widely used Holt and Laury (2002) multiple price list proposed by Drichoutis and Lusk (2016). Participants were asked to choose between two lotteries, one considered 'safe' and one 'risky'. Each lottery had only two outcomes, each with a 50 percent chance of realization. In the 'safe' lottery, the two outcomes were relatively similar whereas in the 'risky' lottery one outcome represented a large gain and the other a very small one. The more favourable outcome in the 'risky' lottery represented a larger gain compared to both outcomes of the 'safe' lottery whereas the less favourable one represented a smaller gain than either outcome of the 'safe lottery'.

Participants were asked to repeatedly choose between the 'safe' and 'risky' lottery ten times. As they progressed, the difference between the maximum gain offered in the 'risky' lottery and the 'safe' lottery progressively got larger. A participant's risk aversion was measured by the point at which they switched from the 'safe' to the 'risky' lottery. Participants who were more risk averse switched later (or perhaps not at all), whereas participants who were more tolerant to risk switched sooner.

To encourage participants to pay attention to the various choices, they were told that at the end of the experiment, one of the ten pairs of lotteries will be

4

selected at random and the lottery they chose played out. The amount they won would be added to their total experimental earnings.

The main stage

The lottery stage was followed by the main stage. Participants were randomly allocated to three groups: two treatment groups and a control group. The purpose of the main stage was to simulate the decision to work under 'standard' and 'unstable' working conditions. All participants were asked to choose between receiving a fixed payment, 'a benefit', or completing a work task for higher pay. The task was designed so as to involve significant amounts of effort but be relatively simple and straightforward to complete. It consisted of transcribing short paragraphs in Latin from Tertullian, an early Christian author. Participants had 2 minutes to complete each transcription and were paid in full if they made three or fewer mistakes as measured by the Levenshtein distance. The purpose of this rule was to encourage participants to transcribe as accurately as possible while not penalizing them for small mistakes. Pilot trials showed that the two minutes allowed time was more than enough to allow even slow transcribers to complete the task. There were two unpaid trial rounds designed to familiarize participants with the task followed by 24 paid rounds.

The 24 rounds were grouped into two phases. In the first phase consisting of 12 rounds, the benefit option consisted of a £1 fixed payment for all participants that was always accessible irrespective of work decisions in previous rounds. Participants made the decision to work or not at the beginning of each round. If they chose 'benefit' they would receive the associated pay and move on to the next round. If they chose to work, they were presented with the work task if work was available. If work was available, they completed the task and received the associated payment subject to having three or fewer mistakes. If no work was available, they received no pay and moved on to the next round.

In the control group, work was always available and participants received £1.50 for each correctly transcribed text. In the treatment groups, the computer determined whether work was available using a coin toss, i.e. there was a 50 percent probability that work would be unavailable. The decision to work or not was made *before* work availability was revealed. A participant who chose to work but work was unavailable would not complete the transcription task but would also not receive any payments, similarly to a zero-hours contract.

In the first treatment group, participants were paid £1.50 per correctly transcribed paragraph. In the second treatment group, the payment was doubled to £3.00 per correctly transcribed paragraph. The two treatment groups were constructed so as to allow for a separation of income and uncertainty effects on labour

5

supply. In the first treatment group, participants could expect to earn half of what participants in the control group earned but also to work 50 percent less. This is the zero-hours setting in which the lack of work and pay is only compensated by increased leisure time. In the second treatment group, participants could expect to earn the same as participants in the control group but only work half of the time. Clearly, work should be more appealing in the second treatment group compared to the control group. However, the former group faced uncertainty about work availability whereas the latter group did not.

In the second phase consisting of the last 12 rounds, the availability and payments associated with work remained the same but the availability of the benefit changed. In the first phase, the benefit payment was only offered as an alternative to work. This setting was intended to mirror the time and administrative costs associated with accessing benefits and the fact that many out of work benefits do not easily allow combining income from employment and benefits. A temporary fall in earnings may not necessarily prompt a zero-hours worker to apply for benefits. Even when they are compensated as part of Universal Credit, this compensation normally takes time to be processed (Ball et al., 2017) .

In the second stage, all participants across the three treatment groups were randomly assigned to two groups, creating a 3 by 2 factorial design. In the first group, the benefit payment was automatically awarded whenever a participant chose work but work was unavailable. In this setting, the benefit system provides an insurance mechanism against unavailability of work and associated falls in pay. In the second group, participants could face sanctions in the form of benefit loss whenever they chose to receive benefits twice in a row. More specifically, if a participant chose to receive benefits *both* in the current and previous rounds, the computer determined with a 50 percent probability whether a sanction would be applied. If a sanction was applied, the participant lost the benefit payment associated with the current round and moved on to the next round. Both benefit treatments were designed to increase the relative attractiveness of the work option, but did so in different ways. While the first treatment made the work option more attractive by reducing the income losses associated with work unavailability, the second treatment made the benefit option less attractive by imposing sanctions and introducing uncertainty on the benefit side, similarly to a workfare based approach.

The productivity stage

After the main stage, participants moved on to a productivity stage designed to measure their ability to complete the transcription task.

Whereas the task was simple and all participants could complete it if they chose to, the associated effort was likely to differ greatly from one individual to

another. More typing proficient individuals could complete the task more quickly and effortlessly and so may have been more inclined to choose work over receiving the benefit. Because participants were randomly allocated, typing proficiency is not expected to differ between treatment groups and will not impact on the estimate of average treatment effects. However, differences in typing proficiency could greatly increase variance and hence reduce the precision of the estimates. This issue can be addressed by collecting a measure of participant productivity and using it to adjust results.

To measure their productivity, participants took part in a challenge: they were asked to transcribe as many texts as they could in a 5 minute period. Participants were paid £1.5 for each correctly transcribed text (the same 3 mistakes or fewer rule applied) and were not able to progress to the final stage before the 5 minutes elapsed. On average, participants transcribed between 3 and 4 texts. However, there was considerable heterogeneity. The lowest number was two and the highest number was nine.

The wrap-up stage

In the last stage of the experiment, participants were asked a few demographic questions, rated the experimental instructions and the difficulty of the work task and played out one lottery randomly selected by the computer. Finally, they were given information about their total pay and how to receive it. Face to face participants were paid in cash immediately after they finished the experiment. They were free to leave and did not have to wait for other participants to finish as well. Prolific participants were paid through the Prolific platform within 48 hours of completing the experiment.

4 Differences between F2F and Prolific data collection

An effort has been made to keep the experimental design as constant as possible across the two data collection methods. F2F sessions took place before Prolific sessions. During F2F sessions, participants had an opportunity to ask questions which the research team then answered. On the Prolific platform, participants could asked questions using the messaging feature of the platform. These questions were then answered in real time by the research team.

Both F2F and Prolific participants completed the lottery stage. However, due to a technical error the lottery data of F2F participants was not saved correctly. As

a result, these data are missing from the final data file.

There are two additional differences.

Consent

In the F2F data collection, participants were initially given some a sheet of paper with information about the study. They were then asked to sign a paper consent form and could then proceed to start their session using a wi-fi connected tablet provided by the researchers.

Collection of paper based consent forms was obviously not possible in the case of the Prolific sample. As a result, the information text and consent questions were added to the on-line experimental platform before the start of the experiment. The same wording was used.Participants had to consent initially in order to proceed to the main part of the experiment. They could subsequently withdraw their consent. In that case, the entirety of their data was deleted.

Show-up fee

F2F participants were offered a £10 show-up fee in addition to any experimental earnings. The show-up fee was intended to motivate higher participation, as well as to cover some of the time and travel costs associated with taking part in an experimental session.

Due to financial considerations, Prolific participants were not offered the same £10 show-up fee. The pay structure and monetary incentives were otherwise identical.

References

Ball, M., Hampton, C., Kamerade, D., and Richardson, H. (2017). Agency workers and zero hours- the story of hidden exploitation. Report, Derbyshire Unemployed Workers' Centres.

Chen, D. L., Schonger, M., and Wickens, C. (2016). otree - an open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, 9:88–97.

Drichoutis, A. C. and Lusk, J. L. (2016). What can multiple lists really tell us about risk preferences? *Journal of Risk and Uncertainty*, 53(2/3):89–106.

Farina, E., Green, C. P., and McVicar, D. (2019). Zero hours contracts and their growth. *British Journal of Industrial Relations*, Online Version (28 Nov 2019).

Henly, J. R. and Lambert, S. J. (2014). Unpredictable work timing in retail jobs: Implications for employee work-life conflict. *Industrial & Labor Relations Review*, 67(7):986–1016.

Holt, C. A. and Laury, S. K. (2002). Risk aversion and incentive effects. *American Economic Review*, 92(5):1644–1655.

Koumenta, M. and Williams, M. (2019). An anatomy of zero-hour contracts in the uk. *Industrial Relations*, 50(1):20–40.

of Personnel, C. I. and Development (2013). Zero-hours contracts: Myth and reality. Report.

Pennycook, M., Cory, G., and Alakeson, V. (2013). A matter of time. the rise of zero-hours contracts. Report, The Resolution Foundation.