# Worker Welfare in the Gig Economy

Jack Fisher *

November 8, 2022

Latest version

**Abstract**

Around the world, the last decade has seen rapid growth in the prevalence of individuals earning income from digital platforms that mediate work for the solo self-employed—gig work. Concerns that these work arrangements undermine labor protections have motivated regulatory legislation. But policymakers suffer from a lack of evidence that quantifies the benefits of gig work and how prospective policies affect worker welfare. I use unique administrative data, which spans the UK's food delivery market, to estimate worker surplus in this typical gig labor market. Evidence that workers learn about their own value of gig work over time, which a new survey corroborates, allows for the identification of the joint distribution of gig work valuations and outside options. Structural estimates imply a median monthly surplus for a gig worker equal to one third of the median employee's monthly income, and an aggregate annual welfare gain of £15bn from a labor market that was nascent a decade ago. Policymakers face a steep trade-off between ensuring benefits for full-time gig workers and maintaining gig work's appeal to low-hours participants, who enjoy most of the aggregate surplus. A counterfactual policy evaluation, which is calibrated to match aspects of California's Proposition 22, supports this conclusion.

*J.W.Fisher@lse.ac.uk. I am deeply grateful to Johannes Spinnewijn for his time, patience, and invaluable advice. I am also thankful to Alan Manning, Alessandro Gavazza, and John Van Reenen for their encouragement and support. For helpful comments, I thank Richard Blundell, Arnaud Dyèvre, Rachel Griffith, Camille Landais, Lu Liu, William Matcham, Canishk Naik, William Parker, Christopher Pissarides, Daniel Reck, Nina Roussille, Mark Schankerman, Benjamin Schoefer, Kate Smith, and Catherine Thomas. I acknowledge financial support from the Economic and Social Research Council, and the Royal Economic Society. I have no relevant or material financial interests that relate to the research described in this paper.

# 1 Introduction

In both the US and the UK, the last decade has seen rapid growth in the number of solo self-employed workers who connect with customers via digital intermediary platforms (Bertolini et al., 2021; Collins et al., 2019; Katz and Krueger, 2019). The vast take up of this type of work—gig work—suggests that individuals are enjoying a significant surplus, but this casual observation is at odds with qualitative evidence that many workers have negative experiences of gig work (Broughton et al., 2018; Dubal, 2019; Ravenelle, 2019).

In light of a desire to better protect workers in the gig economy with new regulation (Adams et al., 2018; Dubal, 2021; Goldman and Weil, 2021; Harris and Krueger, 2015; Kolsrud, 2018; Prassl, 2018), a comprehensive understanding of the gig work surplus is of first order importance for policymakers. However, little is known about the size of this surplus and there is no framework to assess the impact of prospective changes to the gig work environment. For example, what are the welfare effects of rebalancing hourly earnings and fixed benefits, or of reducing individuals' fixed costs to gig work?

There are three reasons for our limited understanding of the gig work surplus. Firstly, gig work differs from traditional employment in many ways (Boeri et al., 2020), which makes it hard to estimate workers' welfare in the gig economy relative to what they would enjoy in its absence. Gig work involves flexible hours (Chen et al., 2019; Katsnelson and Oberholzer-Gee, 2021; Mas and Pallais, 2017), uncertain pay (Angrist et al., 2021; Cook et al., 2021; Parrott and Reich, 2018), jobs of varying difficulty (Athey et al., 2021), and often takes place at specific times in the worker lifecycle (Jackson, 2019; Koustas, 2018, 2019; Katz and Krueger, 2017). Although researchers have investigated the implications of these amenities individually, there has not yet been a holistic analysis that subsumes these factors.

Secondly, data limitations impose an obstacle; the best data on the gig economy is siloed within different platforms. If a researcher is able to access data from one platform, then observed workers may be selected based on the platform's characteristics (*e.g.*, the structure of financial incentives and the user interface) and they may switch between platforms so that only a partial picture of labor supply is revealed (Caldwell and Oehlsen, 2021). Lastly, gig workers face frictions and behavioral biases (Camerer et al., 1997; Chen et al., 2020; Fisher, 2022; Thakral and Tô, 2021), which pose problems to revealed preference approaches. Indeed, the unusual traits of gig work likely make it hard for workers to judge their own surplus and, in turn, reveal their preferences.

This paper tackles these problems head on to estimate the size and structure of the gig work surplus through the joint distribution of gig work valuations and outside options. I introduce new data sources, evince novel patterns in gig work participation, and bring these contributions together with a structural model of gig work, which can evaluate worker welfare in counterfactual scenarios. The framework uses a new identification strategy, which leverages misperceptions and learning to infer workers' outside options from their endogenous exit decisions.

The results imply a large surplus with the median gig worker enjoying a monthly surplus equal to one third of the median employee's monthly income. At the aggregate level, this equates to a £15bn annual surplus for workers, which mainly accrues to less than full-time participants. The distribution of the surplus causes a trade-off for policymakers between ensuring benefits for full-time gig workers and maintaining gig work's appeal to most participants. Moreover, the analysis reflects some workers' gloomy experiences through misperceptions, which depress the gig work surplus by 17% through an allocative inefficiency.

Table 1: Gig Work Surplus Matrix

| | | Outside Option $\nu$: High | Outside Option $\nu$: Low |
|---|---|---|---|
| Valuation $\theta$ | High | $\theta_H - \nu_H \approx 0$ | $\theta_H - \nu_L > 0$ |
| | Low | $\theta_L - \nu_H < 0$ | $\theta_L - \nu_L \approx 0$ |

The ideal experiment to identify an individual's total surplus from gig work would offer workers increasing amounts of money which, if accepted, would prevent them from working in the gig economy. The lowest amount at which a worker accepts is equal to the worker's surplus. Absent this experiment, it is desirable to infer the size and structure of surpluses from gig workers' observable actions.

The crux of the problem is that the difference between workers' valuations of gig work, conditional on working in the gig economy, and their outside options determines their gig work surplus. There is significant unobservable variation in both dimensions. For example, in terms of valuations, some individuals engage in less than 40 hours of gig work a month and others regularly work more than 40 hours a week; these people clearly extract very different value from the gig economy. Further, they likely face contrasting outside options. Part-time workers may engage in leisure activities absent the gig economy, while full-time workers may substitute to another full-time job or a patchwork of part-time jobs.

To make this point clear, table 1 considers a gig economy with four types of workers who are defined by a combination of two characteristics: whether their valuation of gig work is high $\theta_H$ or low $\theta_L$, and whether their outside option is high $\nu_H$ or low $\nu_L$. Individuals with low outside options and high valuations benefit the most from gig work, and those with high outside options and low valuations would be better off outside of the gig economy. For the remaining individuals, the surplus from gig work is ambiguous but small. In this rudimentary setting, identification of workers' valuations $\{\theta_H, \theta_L\}$ and outside options $\{\nu_H, \nu_L\}$, and the proportion of types in the population allows for a description of the gig work surplus $\mathbb{E}[\theta - \nu | \theta > \nu]$.

In a more flexible setting, the task is to identify the joint distribution of gig work valuations and outside options. To do so, I use a new data source and focus on a typical part of the gig economy in the UK: Food delivery by motorcycle. This industry makes up around one fifth of the UK's gig economy (Cornick et al., 2018) and exhibits features that are characteristic of gig work more generally, such as payment on a per job basis and flexible hours (from hereon, "gig economy" will generally refer to this part of the gig economy).

The new data source is administrative data from a vehicle insurer (the "firm"), which provides mandated insurance to a sizeable share of gig workers in this market. The insurance primarily provides cover for damage to a third party that occurs while working, but is best seen from the worker's perspective as a necessary cost. The six largest food delivery platforms send information on each delivery to the firm to facilitate its insurance policy offering. As a result, this data avoids the pitfalls of working with an individual platform.

The firm offers insurance policies with either a variable or fixed premium. The fixed policy is preferable when workers expect to work many hours, while the variable policy minimizes costs for workers who intend to work few hours. Therefore, a cost-minimizing worker's policy choice contains information about their expectation of the hours that they will work in the gig economy. I contrast policy choice with realized hours using the firm's data and show that one in five workers make non-cost-minimizing decisions. Broadly, the lack of cost-minimization could be due to either *ex ante* misperceptions of future hours, or *ex post* shocks that affect hours after the policy choice has been made.

*Ex post* shocks would suggest that subsequent behavior is not related to cost-minimizing policy choice. Instead of this—and consistent with workers suffering from misperceptions—optimistic individuals (*i.e.*, those who do not work enough to explain their policy choice) reduce their hours after they enter, and exit faster. Conversely, pessimistic workers (*i.e.*, those who work too much to account for their policy choice) increase their hours initially, and exit at a slower rate. Different factors could drive misperceptions. For example, lower than expected customer demand that leads to low earnings, or unexpectedly high running costs of workers' vehicles—both find support in a survey of gig workers that I present in this paper.

Intuitively, if an individual's valuation of gig work determines the hours that they work, then misperceptions of hours reveal misperceptions of valuations. In this context, the dynamic patterns of participation and engagement in the gig economy also provide evidence of learning. Optimistic individuals enter the gig economy in anticipation of high value but, in learning that this is not the case, reduce their hours and exit faster, while the opposite is true for pessimists. In line with the hypothesis that learning ameliorates misperceptions, half of survey respondents subscribe to learning "a lot" about the "costs versus benefits" of gig work.

Misperceptions and learning lead to a mirror image of the ideal experiment outlined above. Consider an individual who is overly optimistic about their valuation of gig work. They will enter the gig economy and soon discover that their optimism was misplaced. Over time, they will reduce their hours as they learn about their true valuation and, if their perceived valuation drops below their outside option, they will exit the gig economy. In essence, the value of gig work is incrementally reduced, whereas the ideal experiment incrementally increases the outside option. Therefore, outside options can be identified as workers exit at the point where the trajectory of their perceived valuation crosses their outside option.

Equipped with this logic, the task is to estimate money metric valuations of gig work, workers' perceptions of valuations, and the learning process. Identification requires more than data on intensive and extensive margin participation in the gig economy because it does not permit a way to classify individuals based on their misperceptions. Observing workers' insurance policy choices is important to partition them into groups which are informative of their misperceptions.[1] This additional margin of variation is sufficient to identify workers' surpluses from gig work provided some structure, which is necessary to model the complicated environment that workers face. It is also attractive given the ability of the data to speak simultaneously to different features of the economic environment and the usefulness of non-marginal counterfactual evaluations (Mahoney, 2022).

I develop a model of individuals' participation, insurance policy choice, and hours of work in the gig economy to reflect the patterns in the data. Importantly, it allows for heterogeneity in three key areas: both outside options and valuations can vary across workers, and workers differ in their misperceptions of their valuations. The model explicitly connects the intensity of engagement in the gig economy, measured by hours, to valuations of gig work, and describes the learning process.

Workers decide to participate in the gig economy, if they perceive that its value exceeds their outside option. Upon entering, they select either a fixed or variable insurance policy based on their expected hours,

[1] A worker who enters the gig economy and reduces their hours, before eventually exiting, would only indicate over optimism if we take learning as given, which is itself something that needs to be evinced.

which are driven by their perceived valuation. Then, workers learn about their true valuation as they partake in gig work and adjust their hours accordingly since workers' perceptions always drive their engagement. Finally, workers will leave the gig economy, if their perceived value falls below their outside option, or if they receive an exogenous shock. Throughout, individuals' valuations are subject to *ex post* shocks to account for their influence when bringing the model to the data.

Simulated method of moments (SMM) provides estimates of workers' valuations, outside options, and misperceptions. I specify individual-level heterogeneity to follow a joint log-normal distribution, which helps to capture the skewed distribution of hours in the data and allows for economical computation with a convenient but flexible pattern of correlations between worker characteristics. Empirical moments from the administrative data, coupled with external moments on the employment share of the gig economy and labor supply elasticities, identify the model's parameters in conjunction with structural and stationarity assumptions.

The model and estimates of its structural parameters imply a large gig work surplus. The typical worker enjoys a monthly surplus of £1,066, over one third of mean employee's monthly earnings in the UK. But this masks huge dispersion (SD £1,775) with some workers extracting thousands of pounds of surplus from gig work, while others are on the margin of participating. At an hourly rate, the analysis implies workers enjoy 70 to 80% of their wage as a surplus. The bulk of the gig work surplus—55%—is received by workers who work less than 60 hours per month in the gig economy. This result is driven by the fact that the vast majority of individuals only dabble in gig work, and despite larger average surpluses for high-hours individuals.

I highlight several factors that could explain the large gig work surplus. Most gig workers are in the bottom half of the income distribution and often receive negative income shocks prior to entering (Bernhardt et al., 2022; Cornick et al., 2018; Koustas, 2018), which points towards a high marginal utility of income. Workers who especially value gig work amenities are more likely to select into participation. Further, these individuals' outside options may be particularly low. Alternative employment opportunities that can flexibly adapt to changing schedules are rare, and many gig workers are non-nationals who speak English as a second language, which may reduce other earning options.

The concentration of the gig work surplus amongst workers at the low end of the hours distribution poses a difficulty for policymakers. Broadly, regulators want to compel platforms to offer benefits for regular gig workers.[2] This comes at a cost to platforms, however, statutory incidence does not equal economic incidence so workers will bear some of this cost too. Moreover, this cost will likely be borne by all gig participants because platforms cannot *a priori* distinguish who will qualify for benefits, and multi-homing across platforms undermines any targeted incidence. Therefore, if platforms pass on some of the cost through, for example, an hourly wage penalty, the gig work surplus falls sharply because this severely hurts low-hours workers, who generate the majority of the surplus. In other words, policymakers face a steep trade-off between ensuring benefits for full-time gig workers and maintaining gig work's appeal to the majority of participants.

A counterfactual policy evaluation that is calibrated to match aspects of California's Proposition 22 crystallizes this point. In particular, I model the introduction of mandatory benefits for workers who reach certain hours thresholds. I find that such a policy reduces worker welfare if even half of the cost is born by workers through an hourly wage penalty. Yet, this intervention can increase the gig work surplus by up to

[2]This may be through new legislation or clarifications around the legal tests that are used to define employment status. See efforts along these lines in, for example, the US, Europe, and India.

11% if there is minimal incidence on workers. These impacts rationalize a complementary minimum wage requirement, which could obstruct firms from passing on the cost of mandated benefits to workers.

In the same vein, I use the rich heterogeneity in the model to consider an innovation that reduces the fixed costs of gig work. As a concrete example, I use the introduction of the variable policy, which affects welfare through four channels: (i) it can help existing gig workers save money on the variable policy, (ii) it allows individuals who wish to work only a few hours in the gig economy to participate, (iii) it can attract pessimistic workers who discover that they value gig work more than expected, and (iv) it can attract optimistic workers who, with hindsight, should stay out of the gig economy. Overall, the introduction of the variable policy increases welfare by 4.7% primarily through increased participation. These numbers correspond to an annual aggregate welfare gain of £709mn for workers, which represents a significant return to innovation and policy in this direction.

Next, I turn to the effect of misperceptions on the gig work surplus. Although the population correctly perceives the value of gig work on average, variation in misperceptions gives rise to an allocative inefficiency that stifles the gig work surplus. Namely, pessimistic workers who would receive a positive surplus do not participate, and overly optimistic participants may lose out compared to their outside option. Absent misperceptions the gig work surplus would be 21% higher, which stems equally from eradicating optimistic and pessimistic misperceptions. Notably, even halving the standard deviation of misperceptions attains three quarters of the first-best surplus, which suggests increasing transparency around gig work conditions could be a fruitful area for policy.

Relatedly, the results in this paper are congruent with the negative qualitative evidence surrounding gig work. Quantitatively, 45% of each new entering cohort is overly optimistic about gig work and would not enter but for misperceptions. Over the course of these individuals' tenure in the gig economy, they lose on average £1,183 relative to their next best alternative. Fortunately, learning ameliorates these losses such that misperceptions are not severe on average amongst participants.

**Literature review.** The rise of the digital gig economy has spurred a renewed interest in alternative work arrangements (Boeri et al., 2020; Mas and Pallais, 2020). This paper most closely relates to studies by Chen et al. (2019) and Chen et al. (2020) that use data on Uber rideshare drivers. Both papers aim to estimate the additional surplus that workers receive due to flexible work. To do so, they estimate workers' reservation wages and the market's wage hour-by-hour. The sum of differences between the market and reservation wages over a given schedule corresponds to their measure of a surplus. Then, the value of flexible work is calculated as the difference between the surplus under the flexible schedule and an alternative. They find flexibility almost doubles the gig work surplus on average. Notably, these estimates are much higher than those from discrete choice experiments (Datta, 2019; Mas and Pallais, 2017).

This paper's relative contribution is threefold. Firstly, the welfare gains here correspond to the gig work surplus rather than the additional surplus gig workers receive due to flexible hours. The welfare estimates in this paper plausibly constitute a catch-all surplus that accounts for many of the different features associated with gig work, including but not limited to the value of flexible work. Secondly, the approach in this paper lends itself to the evaluation of prospective policies that aim to improve worker welfare. In particular, the paper takes participation in the gig economy seriously by estimating workers' outside options through a structural model. Thirdly, the data and methodology used in this paper provide a novel way to identify misperceptions about the value of gig work, which leads to the possibility that some workers may be made worse off by entering the gig economy. This mechanism is consistent with qualitative evidence that many individuals are disappointed by what the gig economy offers, but is missing in the economics literature on gig work (Broughton et al., 2018; Dubal, 2019; Ravenelle, 2019).

Naturally, this work builds on many other studies of the gig economy. For example, numerous papers have documented the rise of the gig economy (Bernhardt et al., 2022; Collins et al., 2019; Katz and Krueger, 2019) and carefully assessed its underlying forces (Abraham et al., 2019; Cullen and Farronato, 2021; Ganserer et al., 2022; Garin et al., 2022). There has also been a great deal of descriptive work on the motives and demographics of gig workers (Garin and Koustas, 2021; Hall and Krueger, 2018; Chen et al., 2022), which can explain the wide range of surpluses that gig workers enjoy.

Relative to these studies, I establish novel patterns in gig work hours and survival by leveraging a new cost-structure choice. I argue these results are strongly suggestive of misperceptions and learning. Moreover, I translate heterogeneity in worker choices over hours, exit, and policy into a distribution of surpluses via a structural model that captures workers' value of gig work and their next best alternative.

Work mediated by digital platforms inherently involves many different parties and this paper only considers the worker's perspective. There is a body of work that considers the welfare effects for customers (Cohen et al., 2016), the incentives of platforms to manage either side of the market (Akbarpour et al., 2021; Castillo, 2020; Rochet and Tirole, 2006; Weyl, 2010), and how surpluses are shared between consumers and workers in the context of the knowledge gig economy (Stanton and Thomas, 2021).

There are also strands of the literature from labor economics and industrial organization (IO) that connect with this research. From labor economics, the long history of estimating labor supply elasticities (Blundell and MaCurdy, 1999), the estimates from which serve as a useful moment in my estimation, as well as more recent work assessing workers' outside options (Caldwell and Harmon, 2019; Caldwell and Danieli, 2020) and their perceptions of these alternatives (Jäger et al., 2022). Lastly, IO work on usage-based pricing (or two-part tariffs, or second-degree price discrimination) provides a useful foundation for the policy choice modeling in this setting (DellaVigna and Malmendier, 2006; Economides et al., 2008; Goettler and Clay, 2011; Grubb and Osborne, 2015; Hoffman and Burks, 2020; Lambrecht and Skiera, 2006; Nevo et al., 2016).

I contribute to this latter literature by demonstrating that it is possible to identify heterogeneous outside options with misperceptions, learning, and endogenous exit. The logic is similar to how Bresnahan and Reiss (1990, 1991) exploit variation in the number of entrants under different market condition to estimate entry costs.

The paper proceeds as follows. Section 2 discusses the institutional setting and the data for this study. Section 3 presents a series of reduced form facts, which motivate the model presented in section 4. Section 5 explains how this model is brought to the data, while section 6 discusses the estimates' implications for gig worker welfare in the *status quo* and in counterfactuals. Finally, section 7 concludes and presents complementary areas of future research.

## 2 Empirical Setting

This section describes the institutional environment in which gig workers operate, the data available, and the sample that I use for the analysis. To summarize, I will study motorcycle food delivery carried out by solo self-employed workers in the UK. These workers' experiences are emblematic of the broader gig economy in that platforms mediate their work, they are free to enter and exit, they have flexible hours, and they face uncertain wages. The data source is a firm that provides insurance to a sizeable share of this market, and collects administrative data on these individuals from many different platforms. I complement this data with a new survey of workers' experiences in the gig economy.

### 2.1 Institutional Details

Some of the most visible forms of gig work involve moving passengers and goods on the road; Toyota Prii[3] with smart phones fixed on dashboards and motorcycles adorned with insulated food delivery boxes are now quintessential sights for many cities around the world. Indeed, Cornick et al. (2018) find that this makes up over half of gig work in the UK, which is the setting for this study.

In this sense, I will focus on an exemplary part of the gig economy: food delivery by motorcycle. Specifically, solo self-employed workers who carry out this job for intermediary, digital platforms will be the subject of this paper. Like many in the gig economy, these individuals are free to onboard, and pick their own hours and location of work. They are generally paid a set fee per job but, when this is combined with fluctuating demand and supply, as well as other shocks (*e.g.*, traffic and waits at restaurants), they receive an uncertain wage.[4] Further, gig workers are self-employed so they are entitled to very few employment rights beyond health and safety and discrimination protections. For example, they do not receive sick pay and are not guaranteed a minimum wage.

A particular job, if accepted, requires the worker to drive to a restaurant, pickup a meal, and then deliver the meal to the customer. Platforms differ in the ways that they provide information and offer compensation. For example, some platforms tell workers where the customer is located prior to the acceptance of a job, while others only disclose the location of the restaurant. Compensation often adjusts to the distance of a job but this is only done in a coarse fashion.

Importantly, individuals working on UK roads must have an enhanced level of vehicle insurance. This additional insurance is called Hire and Reward (H&R) insurance and is a necessity for many gig workers, including motorcycle food delivery workers. This insurance covers damage to third parties while working and further coverage can be purchased to protect one's own vehicle under certain circumstances (*e.g.*, fire and theft). This insurance does not cover the food being delivered, which is dealt with by the intermediary platform.

From the perspective of a gig worker, this imposition can be seen as an unavoidable cost to be minimized. The H&R market offers insurance in two forms: variable and fixed. Variable policies are paid approximately by the hour, while the fixed policies insure workers for a 30 day period and are paid for upfront. From a cost-minimization perspective, if one expects to work few hours over the next 30 days, then they should prefer

[3] The plural of Prius by popular choice, see https://pressroom.toyota.com/toyota-announces-the-plural-of-prius/.

[4] Some platforms also provide other financial incentives, such as a bonus for making a set amount of deliveries within a month, but these are uncommon and information on their details is lacking.

the variable policy since it would not be economical to pay for a full 30 days of coverage. Conversely, the fixed policy is preferable when individuals expect to work many hours. Both policies are easy to use; the fixed policy is paid as a direct debit and the variable policy is paid for via a digital wallet, which can be auto-topped up from workers' bank accounts. When workers choose between either the fixed or variable policy online, both options appear equally prominent, side-by-side with their premiums listed.[5]

## 2.2 Data

The data for this paper comes from a H&R insurer (the "firm") that offers both the variable and fixed policy to prospective gig workers. The firm receives data from many different intermediary platforms in order to facilitate its insurance policies and, therefore, does not suffer from individuals selecting into or switching between work providers. Further, the firm provides insurance to a significant share of gig workers in the food delivery market.

The data contains information on jobs completed by workers and their fixed or variable policy choice, along with the premiums they faced and some worker covariates. In particular, the data contains information about the length of jobs, when they took place, a unique worker identifier, the age and gender of the worker, the type of insurance policy, and the premium. The main omission from the data is worker compensation.

Given the choice environment that workers face *vis-á-vis* a 30 day policy, and the aim of estimating a longer-term and broader surplus from gig work, I aggregate the data from the worker-job-level to the worker-month-level to construct, for example, a monthly hours worked variable.[6] The monthly data also has the advantage that it reduces the influence of high-frequency shocks that affect workers' labor supply decisions.[7] I treat a worker's first appearance in the dataset as their first entrance into the gig economy. While it is possible that workers may have already undertaken gig work, the rarity of policy switchers within the firm's data suggests that switching across firms is not a significant problem. Similarly, I define exit from the gig economy as a worker not reappearing in the data.

The data spans January 2018 to October 2021. I restrict to worker-months observed from the start of 2019 and onward because the insurer was growing rapidly in 2018 and did not offer a consistent menu of policies. This period of time includes the Covid-19 pandemic, which was a period of continuity and even growth for the food delivery market.[8] Some workers have multiple spells in the gig economy; for these workers I keep their first spell, where a spell is defined as working consecutive months with a break of no longer than three months. The fixed policy can offer additional forms of coverage to a worker's own vehicle for a higher premium, while the variable policy provides only third party coverage. In order to adjust for this, I use reports of willingness to pay (WTP) for additional coverage from the survey (discussed below) to correct

[5]Appendix A presents additional evidence which supports the claim that workers are primarily concerned with cost-minimization. In particular, it shows that individuals who change policy tend to switch to more economical policies. This appendix also contains a more detailed discussion of the potential differences between the fixed and variable policies, and ulterior influences that may affect policy choice, such as variance in income.

[6]Thus, the measure of labor supply is the sum of time spent on food deliveries over the course of a month. Workers often spend 20 to 30% of their time idle between jobs, which I account for in my structural estimation.

[7]It is possible that this aggregation also leaves more time for shocks to workers' valuations to realize. Therefore, in order to provide further insight into workers' dynamic behavior, I augment the main monthly analysis in section 3 with weekly data. Further, I allow for shocks to workers' valuations in the structural model in section 4.

[8]In appendix C, I show that the reduced form evidence is broadly consistent before and during the pandemic.

workers' premiums.[9] At present, I remove switchers from the analysis although the model is being developed to integrate these individuals.[10] Workers can also opt for an annual policy; these policies are taken up by less than a fifth of individuals who seem to be engaged in permanent, full-time work and, as such, they are qualitatively different from the vast majority of workers who are the focus of this study, so I do not include these people in the main analysis.[11]

The firm also has quote data, which contains the menu of prices that workers face when they make their participation and policy decision. This is useful for two reasons. Firstly, it reveals the distribution of fixed policy premiums faced by the population without any selection. I leverage the observed selection into policies based on premiums in the estimation. Secondly, it allows for the construction of individual-level "break-even" points. That is, the number of hours at which both the fixed and variable policy entail the same cost. For illustrative purposes, I often calculate an average break-even point as equal to the average observed fixed premium divided by the average hourly premium, which equals 110 hours (=£103/£0.94ph). Sometimes workers receive more than one quote; if these differ, I use the average of a worker's quotes.

Table 2 presents summary statistics for the analysis sample broken down by the type of policy, where the observations have been collapsed to the worker-level to ensure representativeness across workers. In total, I observe 86,024 (=16,575×5.19) worker-months. 64% of workers select the variable policy and these workers tend to work less both in terms of hours and the number of jobs that they complete in a month, but they stay in the gig economy longer than their peers on the fixed policy. Hourly premiums for variable policyholders are considerable at £0.94 per hour.

I complement this administrative data with a survey conducted in collaboration with the firm. The survey was sent out in June 2022 to the firm's active customer base who had subscribed to receiving promotional material.[12] The survey contained questions regarding workers' experiences of the gig economy, especially relative to their expectations, and their policy choice. The survey received over 500 responses in total though not all questions were answered by all respondents.

## 3 Patterns in Gig Work Participation

This section presents four empirical facts about gig work participation. Firstly, there is dramatic variation in the number of hours worked per month across individuals. Secondly, hours worked do not predict survival in the gig economy. Thirdly, workers do not always make cost-minimizing policy choices. Fourthly, cost-minimization is correlated with trends in hours worked and survival. These patterns are consistent with workers having misperceptions about the value of gig work and learning, and new survey evidence corroborates this interpretation. Namely, workers report that the realities of gig work frequently deviate from their expectations and that they learn about these differences over time.

[9] Precisely, for those fixed policy workers who purchase additional coverage, I deduct the average WTP conditional on the WTP being greater than the extra premium for additional coverage. Further detail on these filters and adjustments are provided in appendix B and I present reduced form evidence and structural estimates, where I focus strictly on third party only policies, in appendix C and F, respectively.

[10] This is unlikely to significantly affect the results because for every 100 gig workers who exit, only seven switch policies. Further, many policy switches take place at the start of a second stint in the gig economy, in which case the first spell is kept in the analysis sample.

[11] Naturally, it is also more difficult to evaluate whether policy choices are cost-minimizing since it requires a minimum of a year's worth of data. In appendix F, I compare annual and 30 day policyholders behavior, and adjust the empirical moments in the estimation to inform how their exclusion affects the results.

[12] Given selection based on this subscription decision and voluntarily responding, I use the survey to corroborate my interpretation of the reduced form evidence and to benchmark results from the structural model.

Table 2: Worker-Level Summary Statistics

| Statistic | Variable | Fixed | Both |
|---|---|---|---|
| Number of workers | 10,589 | 5,986 | 16,575 |
| Mean number of jobs | 114.19 | 274.95 | 172.25 |
| Mean duration (months) | 5.42 | 4.76 | 5.19 |
| Mean monthly hours | 43.76 | 87.51 | 59.56 |
| SD monthly hours | 38.24 | 59.98 | 51.72 |
| Mean monthly premium (£) | – | 93.20 | – |
| SD monthly premium (£) | – | 41.85 | – |
| Mean hourly premium (£) | 0.94 | – | – |
| SD hourly premium (£) | 0.31 | – | – |

**Notes:** This table shows summary statistics at the worker level from the analysis sample. The worker-month-level data is collapsed to the worker-level. Then, for example, the mean hours row displays the mean number of hours worked by workers during an average month, and standard deviations are computed across workers. Mean duration is constructed as the average number of months workers spend in the gig economy. Monthly premiums are constructed as total premiums paid in a 30 day period and hourly premiums are constructed as monthly premiums divided by hours worked in the corresponding 30 day period.

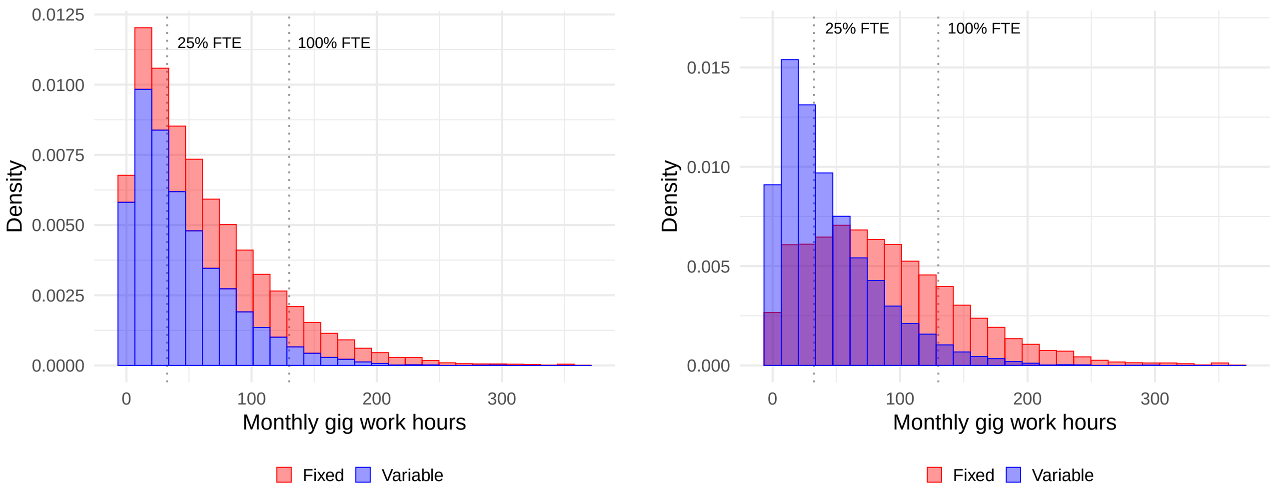
These facts motivate four features of the model in section 4: (i) workers' have different valuations of gig work, which manifest as a distribution of hours worked in the gig economy; (ii) outside options vary across individuals to justify those working few hours—or, equivalently, those with low valuations—remaining in the gig economy; (iii) workers may misperceive their valuations and this can lead to non-cost-minimizing decisions; and (iv) individuals learn about their true valuations over time, which leads to the observed evolution of hours and survival for (non-)cost-minimizing workers.

## 3.1 Hours Worked

There is enormous dispersion in the number of hours worked in a month by different workers, which suggests that workers are extracting very different value from gig work. Figure 1 illustrates this dispersion and how it relates to policy choice. Panel 1a presents the empirical distribution of hours worked and the share of policies that make up each hours bin, as reflected by the coloring of the bar. There are two key takeaways from this graph. Firstly, a majority of workers (*circa* two thirds) are on the variable policy; the blue area is greater than the red area. Secondly, most individuals do not work many hours in the gig economy but there is a strong right skew. The modal number of hours worked is approximately 20 hours per month while the mean is 60 hours per month with a standard deviation of 52 hours. Panel 1b reveals a third fact: the distribution of hours looks very different conditional on policy choice, as one would expect given selection into policies. For the variable policy, most of the mass is compact around its mean of 44 hours per month. Conversely, hours are more dispersed under the fixed policy with a mean and standard deviation of 88 and 60 hours per month, respectively.

The firm's quotes data contains information on individuals' age and gender, which can be linked to work

Figure 1: Distribution of Hours Worked




(a) Stacked density by policy choice

(b) Overlapping density by policy choice

**Notes:** This figure plots the probability mass of binned hours worked per month for different samples. The sample in panel 1a is all workers in the analysis sample and the coloring of the bar is determined by the share of fixed versus variable policy holders. The proportion of the bar that is blue represents the share of workers who are on the variable policy in that hours bin. Two samples are used in panel 1b; the distribution of hours worked by fixed and variable policy holders is shown by the red and blue bars, respectively. The dotted grey lines show where different shares of full-time equivalent (FTE) work fall in the distribution of gig work hours. Each observation in a bin is a worker, where repeated worker observations have been averaged over. For this figure, I have removed individuals with less than three monthly observations in order to reduce the impact of noise, which leaves 9,575 workers.

hours data. Over 90% of workers in the sample are male, which limits statistical power to discern differences in gig labor supply between genders. Still, regression analysis suggests that women on the variable policy tend to work approximately 4 hours less per month than their male counterparts (p-value 2.4%), while there is no statistically significant difference in hours worked between genders for the fixed policy. Interacting gender with age yields imprecise estimates. In general, hours tend to increase with age regardless of policy although this is more pronounced on the fixed policy where, on average, individuals under 30 work 12 hours less per month than those over 40.

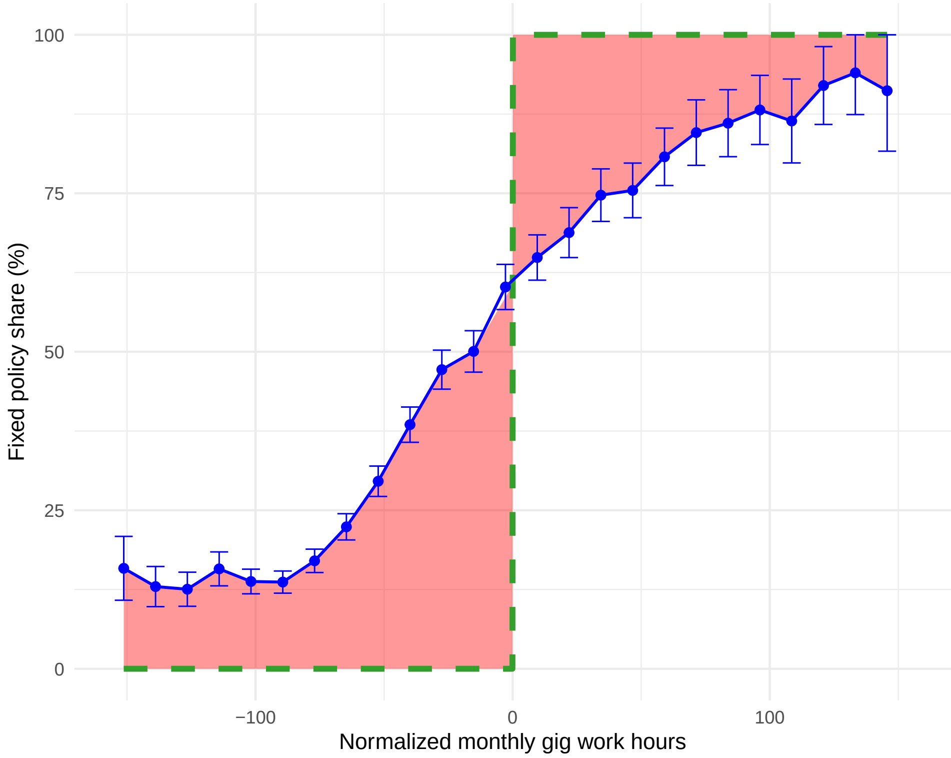
## 3.2 Policy Choice

Given information on hours worked and quoted premiums it is possible to assess the quality of policy choice from a cost-minimization perspective. To do so, I construct worker-level break-even points, which describe the number of hours above which the fixed policy is most economical. While the data display patterns firmly consistent with an intention to cost-minimize, many individuals would be better off on the alternative policy and there is some *a priori* evidence of optimism from gig participants.

In section 4, this motivates misperceptions over workers' valuations of gig work, which can lead to misperceptions of the number of hours they will work and, in turn, non-cost-minimizing policy choices. Moreover, individuals with higher perceived valuations select into gig work, which can cause the appearance of optimism.

Figure 2 provides a convenient lens through which to view policy choice quality (Handel et al., 2020). The graph shows the share of individuals on the fixed policy for different normalized hours bins relative to their

Figure 2: Fixed Policy Share by Normalized Hours Worked




**Notes:** This figure plots the share of workers who are on the fixed policy by normalized hours bins. Normalized hours are hours minus an individual's break even point, which is constructed from the quote data as a worker's quoted or actual monthly premium divided by the analagous hourly premium. Each observation in a bin is a worker, so the hourly bin that an individual falls into is determined by their average monthly hours. The green dashed line indicates the perfect cost-minimizer's policy choice, which is vertical at the break-even point. Standard errors are constructed by applying the law of large numbers to the average of Bernoulli random variables (*i.e.*, $\sqrt{p \cdot (1-p)/N}$ where $p$ is the share of policies on the fixed policy in a bin and $N$ is the number of observations in that bin).

break-even point. Normalized hours are constructed as hours of work in the gig economy minus an individual's break-even point. A perfect cost-minimizer would exhibit a step function so that when they work below the break-even point, they are always on the variable policy, and when they are above the break-even point, they are always on the fixed policy. This is illustrated by the dashed green line. Of course, at the break-even point a perfect cost-minimizer would be indifferent between policies and any fixed policy share is compatible with optimization. If one were to take this fictional cost-minimizer and introduce imperfect foresight to their predictions of hours worked, then this would smooth the step function and lead to a monotonically increasing line that crosses the break-even point at 50%.

The data reveals a pattern similar to this, as shown by the blue line in figure 2. Workers far from the break-even point (*e.g.*, those on -150 and 150 normalized hours) all but minimize their costs and, moving between these extremes, workers have an increasing tendency to opt for the fixed policy. Therefore, the data is strongly supportive of workers minimizing costs in their policy choice. Yet there are still a significant portion of workers who make non-cost-minimizing choices. This is illustrated by the red shaded regions in figure 2, which highlight deviations from the perfect step function. Further, the blue line crosses the break-even point above the 50% level, which could be indicative of optimism from gig workers.

Table 3: Worker Categories

| | | Policy Choice | |
|---|---|---|---|
| | | Fixed | Variable |
| Cost-Minimizing | Yes | Minimizers 13.1% | Minimizers 59.1% |
| | No | Optimistic 20.1% | Pessimistic 7.7% |

**Notes:** This table shows the constructions of different worker categories, where the color denote how the categories are shown in subsequent figures. The percentages reflect the proportion of worker-months that fall into each category.

**Categorizing workers.** For the remainder of this section, I will categorize workers based on a combination of their policy choice and whether their choice was cost-minimizing. Practically, I consider a worker to be cost-minimizing if their policy choice minimized their costs for the majority of months during their tenure in the gig economy.[13] The categories are summarized in table 3 alongside their unconditional share of the population. The colors in the matrix correspond to how these groups are depicted in the figures below. Workers who make cost-minimizing choices are grouped together and referred to as "minimizers", and fixed and variable policy holders who make non-cost-minimizing decisions will be called "optimistic" and "pessimistic", respectively.

Broadly, deviations from cost-minimization could be driven by two factors: firstly, *ex post* shocks that affect workers' hours after they have entered the gig economy and, secondly, *ex ante* misperceptions about how much they will work. If *ex post* shocks are responsible for non-cost-minimizing behavior, then these categories should not be predictive of subsequent behavior in the gig economy. Conversely, if *ex ante* misperceptions cause non-cost-minimization, then this partitioning of the data should be correlated with subsequent gig work engagement. Of course, minimizers are also subject to both phenomena, but not such that they have revealed it through their policy choice.

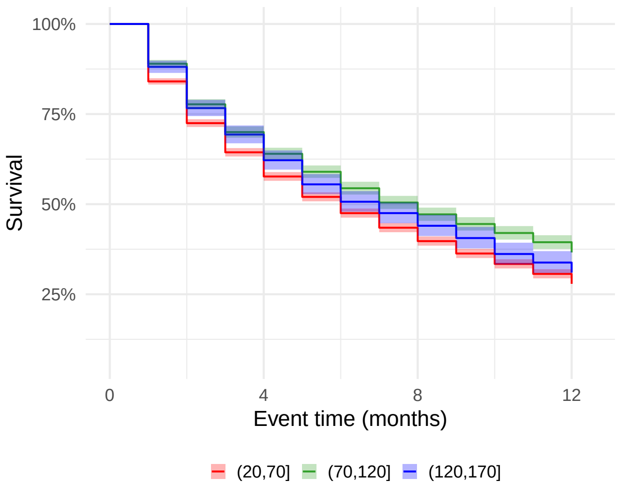
## 3.3 Dynamics of Survival and Hours

In this subsection, I leverage repeated observations of gig workers over time to examine dynamic aspects of worker behavior and how this correlates with their categories. Figure 3 depicts individuals' survival probabilities over time. Panel 3a shows that hours do not predict survival; grouping workers by the average number of hours they work does not lead to noticeably different survival trajectories and, moreover, any differences are not monotonic in hours. Intuitively, this implies that workers have different outside options since low-hours workers would not remain in the gig economy if they had to sacrifice the same outside option as a full-time gig worker. Further, policy choice is not predictive of survival. Panel 3b shows that fixed and variable policyholders have almost indistinguishable survival paths. Thus, hours and policy choice alone are not informative of tenure in the gig economy.

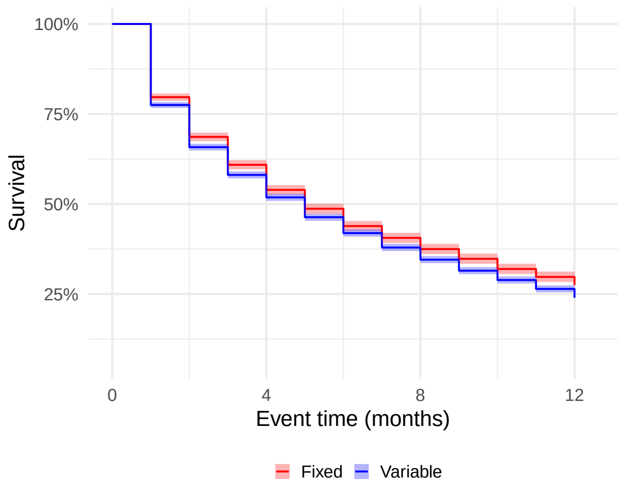
In contrast, categories are a strong predictor of survival, as shown by panel 3c. The optimistic group (*i.e.*, those who select the fixed policy but do not work enough to make it worthwhile) initially drop out of the gig

[13] The results below are robust to other ways of classifying choice quality. Appendix C replicates the analysis below with alternative categories, where non-cost-minimization is calculated by whether either an individual minimized their total insurance premiums over the course of their spell in the gig economy, or whether they minimized their costs for a typical month.

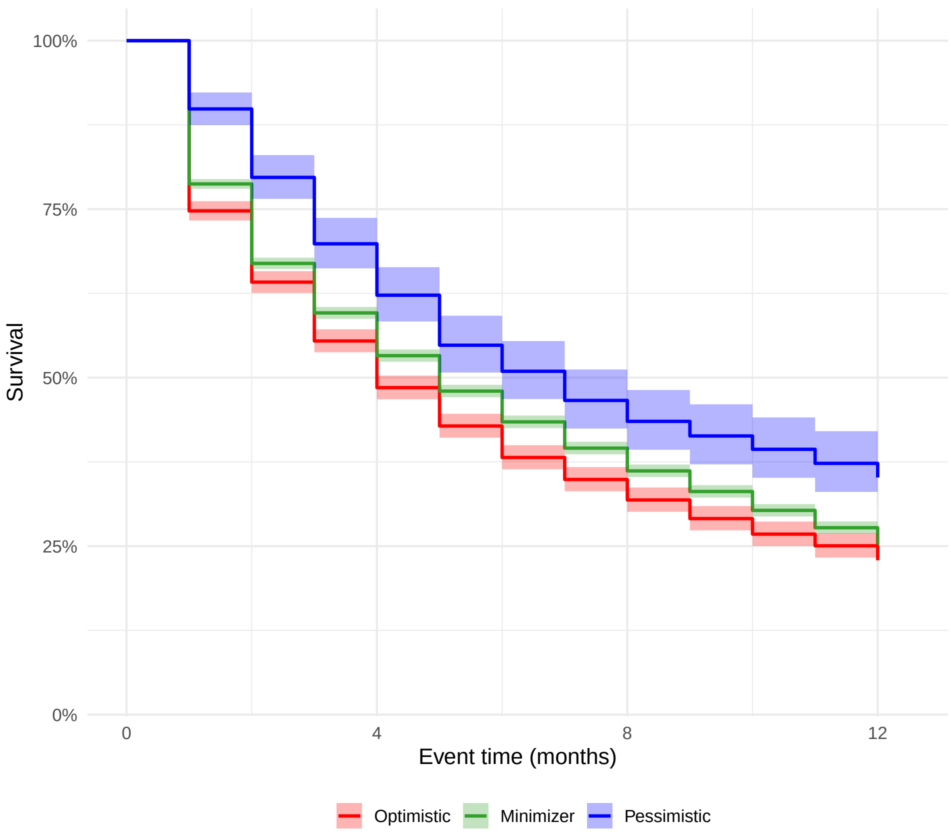
Figure 3: Survival




(a) By hours




(b) By policy




(c) By category

**Notes:** This figure plots Kaplan-Meier survival curves for different groups in the gig economy. In panel 3a, the green, red, and blue lines denote the hours bins [20, 70), [70, 120), and [120, 170), respectively. Panel 3b shows fixed and variably policyholders in red and blue, respectively. In panel 3c, the green, red, and blue lines denote the minimizers, optimistic, and pessimistic categories, respectively. Event time is tenure month in the gig economy (*i.e.*, $t = 1$ is workers' first month in the gig economy so if an individual does not have a second month in the gig economy, then they exit in the first period).

economy faster than the other groups; the red line falls below the other two lines in the first period. Conversely, the pessimistic group (*i.e.*, those who select the variable policy but would have saved money on the fixed policy) exit slowest at first, as evinced by the fact the blue line starts above the others. Minimizers' survival probabilities are somewhere in between those of the optimistic and pessimistic groups, which reflects their less severe exposure to forces that could push them off the cost-minimizing policy.

This evidence suggests that *ex ante* misperceptions play an important role in non-cost-minimizing policy choices because, if the categories only reflected *ex post* shocks, then these dynamic patterns should not be evident. Further, the precise path of survival points towards learning. That is, the categories influence survival most at the beginning of a worker's tenure in the gig economy, which implies that workers enter with misperceptions and learn about these over time such that some individuals exit. After sufficient time has passed, misperceptions have all but gone and the categories do not affect survival further.

To evince these patterns, and to confirm that categories are highly predictive of survival relative to hours, I estimate a Cox proportional hazards model with time-varying coefficients.[14] Table 4 displays the estimates. Monthly hours, although statistically significant, have little meaningful impact on survival. The top row of estimates suggest that increasing a workers average number of hours per month by ten marginally decreases the baseline hazard rate by 1%. Meanwhile, the preferred estimates in column (1) suggest the optimistic and minimizer categories increase the baseline hazard rate by 48% and 22%, respectively, relative to the pessimistic category over the first two months of a worker's spell in the gig economy. Thereafter, their effects wane; the minimizer category's impact is no longer discernible from zero and the optimistic category's effect falls by two thirds—consistent with learning about misperceptions.

Lastly, I show that categories are also correlated with the evolution of hours worked. Figure 4 displays how workers' hours evolve over time. The figure shows average hours worked in each week of tenure relative to workers' second week in the gig economy to avoid the fact that workers may not begin working at the start of their first week.[15] Again, the different categories display contrasting behavior. Optimistic workers see their hours initially fall while pessimistic workers see their hours increase at the start of their tenure. Cost-minimizing workers also see their hours fall, though less so than optimistic workers, which is consistent with the survival evidence for minimizers.

This is also supportive of learning and misperceptions. Some workers have severe misperceptions such that, when they become aware of them, they leave the gig economy. For others it is still worthwhile participating but they adjust their intensive margin accordingly.

### 3.4 Survey Evidence

Survey responses from over 300 of the firm's customers provide strong evidence that expectations of gig work often deviate from reality, and that the true value of gig work is learned over time. Figure 5 presents responses to four questions contained in the survey. Panel 5a shows how individuals found gross earnings (*i.e.*, earnings before costs) relative to their expectations. Earnings expectations appear to be accurate on average, but with significant dispersion such that the majority of workers are left either pleasantly surprised or disappointed in almost equal proportion. Panels 5b and 5c show that workers report costs and the difficulty of the job,

[14]A parametric model also allows for the inclusion of controls although, in appendix C, I show that the visual patterns remain when splitting the sample by age, time periods, and policy coverage. Appendix C also contains a linear probability regression that reveals the same patterns as the Cox proportional hazards model with time varying coefficients.

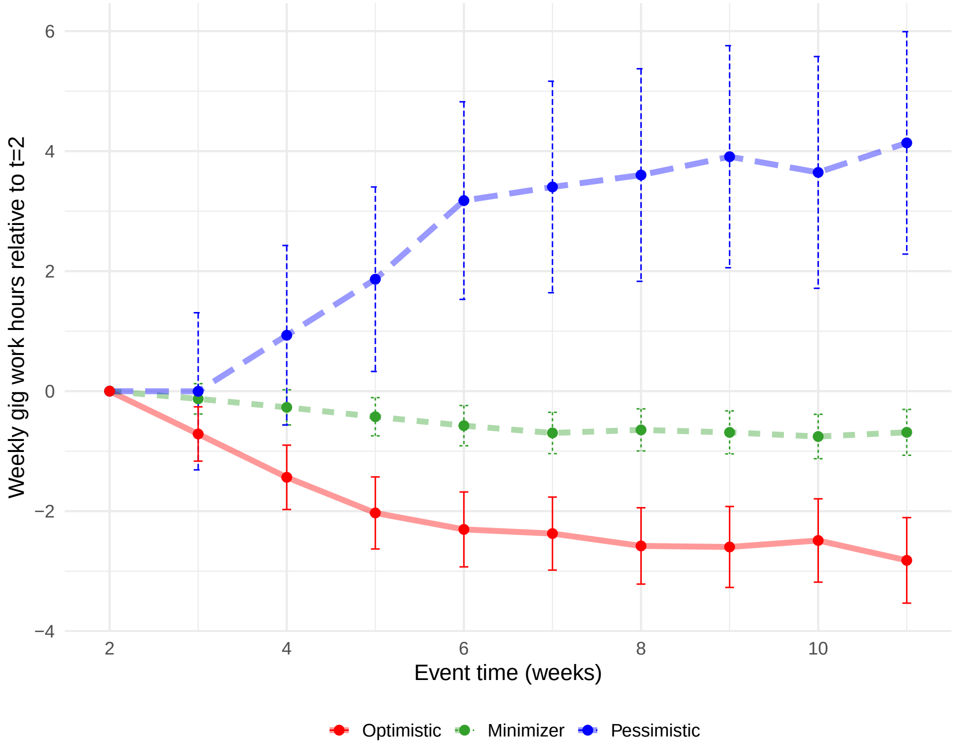
[15]As with the survival evidence, appendix C shows that the patterns in hours dynamics are robust across different cuts of the data.

Table 4: Cox Proportional Hazards Model with Time-Varying Coefficients

| | *Dependent variable:* | | |
|---|---|---|---|
| | Tenure in the gig economy (months) | | |
| | All controls | Time controls | No controls |
| | (1) | (2) | (3) |
| Mean hours | $-0.001^{**}$ | $-0.001^{***}$ | $-0.001^{***}$ |
| | (0.0003) | (0.0003) | (0.0003) |
| Minimizer (<= 2 months) | $0.221^{**}$ | $0.181^{*}$ | $0.179^{*}$ |
| | (0.100) | (0.095) | (0.095) |
| Optimistic (<= 2 months) | $0.481^{***}$ | $0.378^{***}$ | $0.353^{***}$ |
| | (0.104) | (0.098) | (0.098) |
| Minimizer (> 2 months) | $-0.053$ | $-0.067$ | $-0.072$ |
| | (0.070) | (0.068) | (0.068) |
| Optimistic (> 2 months) | $0.173^{**}$ | 0.083 | 0.074 |
| | (0.078) | (0.073) | (0.073) |
| Low hours | Yes | Yes | Yes |
| Time controls | Yes | Yes | No |
| Age | Yes | No | No |
| Gender | Yes | No | No |
| Cover | Yes | No | No |
| Observations | 23,969 | 25,729 | 25,729 |
| $R^2$ | 0.076 | 0.071 | 0.066 |

**Notes:** $^{*}$p<0.1; $^{**}$p<0.05; $^{***}$p<0.01. This table shows estimates from a Cox proportional hazards model with time varying coefficients on the categories variable; the effect of this factor variable is allowed to differ between the first two months of a workers spell and any remaining months. Coefficients reflect percentage changes in the base hazard rate associated with the corresponding variable. The main panel of the table shows estimates for these coefficients and estimates of the coefficient on a workers average number of hours per month. The table displays three specifications: column (3) includes no controls, column (2) includes only time controls, and column (1) includes time controls and additional covariates from the quotes data. Observations are censored at October 2021. All specifications also include a dummy for low hours because panel 3a suggests survival may not be monotonic in hours and to proxy for optimistic misperceptions amongst variable policyholder minimizers. Standard errors are shown in parentheses.

Figure 4: Hours Worked Over Time by Category



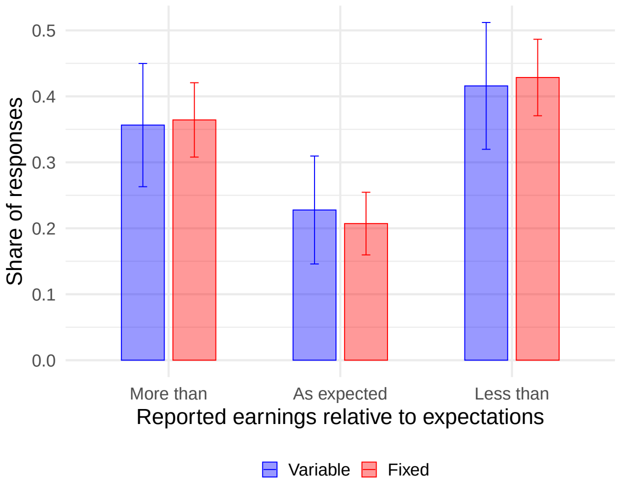

**Notes:** This figure plots three sets of coefficients from three separate regressions, which are run on a balanced panel of each category of worker. Weekly hours are regressed on fixed effects and event time dummies, where $t = 2$ corresponds to their second month in the gig economy (*i.e.*, event time is tenure in the gig economy), as well as calendar time controls. Intuitively, each coefficient represents the difference in hours at a given point in time from their hours in their first month in the gig economy. SEs are clustered at the worker level with a HC3 weighting scheme.

respectively, to be much higher than expected. Approximately three quarters of workers found costs to be more than expected, while half of them find the work more difficult than anticipated.

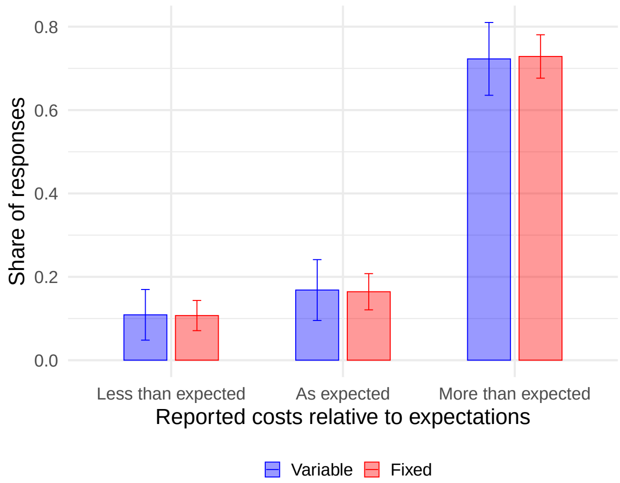
Survey responses are generally consistent across policies but, in figure C8, I show the share of optimists and pessimists' responses relative to minimizers are in line with the hypothesis of misperceptions and learning. Pessimists are more likely to find gig work better than expected relative to optimists and minimizers. Conversely, optimists more frequently report gig work to be worse than expected, while minimizers report their experiences as expected most often. These differences are not statistically significant because of low sample-size, and noise from self-reports of hours and premiums, but are supportive nonetheless.

Inaccurate initial perceptions suggest room for learning. This is confirmed by panel 5d, which presents workers' responses to the question of whether they have learned about the "costs versus benefits of this job since [they] started". Around 90% of workers report experiencing learning over the course of their tenure in the gig economy. Interestingly, some differences between policy holders emerge in this figure. Fixed policy holders are almost 20 percentage points more likely to say that they learned "a lot" relative to variable policy holders.

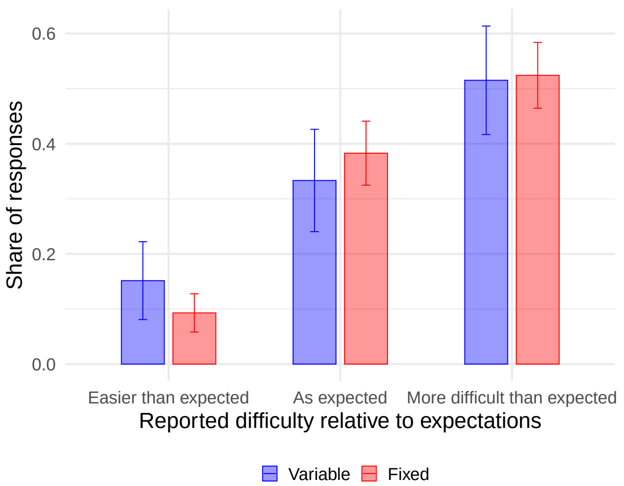
Figure 5: Experiences and Learning



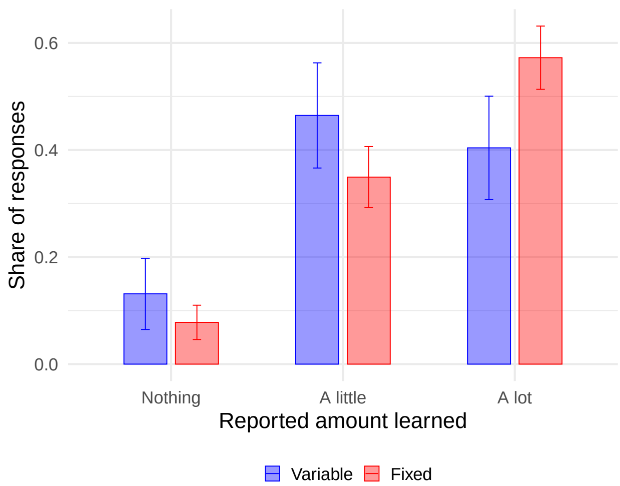

(a) "Are you earning (before costs) more or less than you expected in this job?"




(b) "Are the costs in this job (*e.g.*, fuel, insurance) more or less than you expected?"




(c) "Is the difficulty of this job more or less than you expected when you started?"




(d) "Have you learned much about the costs vs benefits of this job since you started?"

**Notes:** This figure plots the share of responses to four questions from the survey. The sample contains 85 variable users and 251 fixed users. Standard errors are calculated by applying the law of large numbers to the average of a random variable that follows the multinomial distribution with one trial (*i.e.*, $\sqrt{p_j \cdot (1-p_j)/N}$ where $p_j$ is the share of responses for a given category $j$ and $N$ is the number of observations).

### 3.5 Discussion

To recap, the distribution of hours worked in the gig economy exhibits are strong right-skew so that, while the majority of individuals work a small number of hours, some workers participate full-time and more. Gig workers suffer from *ex ante* misperceptions about their hours, which cause non-cost-minimizing policy choices. A combination of policy choice and (non-)cost-minimization is predictive of survival and the evolution of hours. In particular, optimists tend to reduce their hours of work in the gig economy and exit faster than other workers.

These pieces of evidence are consistent with a story whereby workers have misperceptions about their valuations of gig work relative to their outside option, but they learn over time about their true valuations. Take an individual who thinks they value gig work more than they in fact do—an optimist—this individual will be more prone to enter and more likely to select the fixed contract. After entering they will learn that their optimism was misplaced and they will reduce their hours, and potentially exit. In an opposite fashion, a worker who initially undervalues gig work—a pessimist—would be more likely to select the variable policy and subsequently increase their hours with a much lower propensity to exit.

Survey evidence lends further support to this narrative. Individuals' responses confirm that misperceptions of the realities of gig work and learning are common phenomena for gig workers. A model that encapsulates this thesis follows in section 4.

## 4 A Theory of Gig Work

In this section, I develop a model of workers' participation, choice of insurance policy, and hours in the gig economy. The model captures the key features of the economic environment as evinced by the reduced form empirics: Workers have different valuations, which lead to a distribution of hours worked, but they may misperceive these valuations such that many individuals end up on an unnecessarily expensive policy, while some would be better of outside of the gig economy altogether. In addition, workers learn about their true valuation of gig work over time, which manifests in an evolution of hours worked and survival that is correlated with policy choice and (non-)cost-minimization. The model also allows for workers to experience *ex post* shocks to their valuations.

### 4.1 The Model

A worker $i$ is endowed with an individual-specific quadruple $\{\theta_i, \nu_i, \phi_i, P_i\} \in \mathbb{R}^4_+$ that contains their true valuation of gig work $\theta_i$, their outside option $\nu_i$, their initial misperception of their valuation $\phi_i$, and their fixed policy premium $P_i$. If the worker enters the gig economy, upon entering they decide between the fixed and variable policy $\omega \in \Omega = \{\omega_F, \omega_V\}$ and then, each period, they pick how many hours to work in the gig economy $h \in \mathbb{R}_+$. These choices entail a normative flow utility for worker $i$ of

$$u(h, \omega; \theta_i) = \theta_i \cdot \frac{h^{1-\frac{1}{\varepsilon}}}{1-\frac{1}{\varepsilon}} - \big(p(\omega) + \kappa\big) \cdot h - P_i(\omega), ^{16} \tag{1}$$

[16] The following analysis goes through if this utility function also had an individual-specific intercept, however, in practice this would not be separately identified from the outside option.

where $\varepsilon \in (1, +\infty)$ governs the response of hours to the value of gig work and the variable cost of hours. The variable cost is made up of an exogenous linear cost to working $\kappa \in \mathbb{R}_+$ and the hourly insurance premium $p(\omega)$, which equals $p \in \mathbb{R}_+$ if the worker opts for the variable policy $\omega = \omega_V$ and is zero otherwise. The worker also faces a fixed premium $P_i(\omega)$, which equals $P_i$ if the worker chooses the fixed policy $\omega = \omega_F$ and is zero otherwise.[17] If the worker decides not to enter the gig economy, then they receive their outside option $\nu_i$ every period. Utility and surplus are always measured with the normative utility function described by equation (1).

When the worker is making their decision about gig work participation, they misperceive the value of gig work. That is, before they enter the gig economy, they perceive their value of gig work to be $\hat{\theta}_{i,0} = \phi_i \cdot \theta_i$. If the worker decides to participate in the gig economy, they will learn about their misperception over time. Concretely, their misperception will erode such that after $t$ periods it is equal to

$$\Phi(t, \phi_i) = \frac{t}{t+\lambda} + \frac{\lambda}{t+\lambda} \cdot \phi_i, \tag{2}$$

where $\lambda \in \mathbb{R}_+$ determines the speed of learning (an increase in $\lambda$ implies slower learning).[18] This functional form is microfounded by a model of individual level Bayesian learning (see appendix D.2). Equation (2) implies that $\lim_{t\to\infty} \Phi(t, \phi_i) = 1$ so the worker will all but perceive their true valuation after sufficient time has passed.

At any point in time, the worker will behave in accordance with their perceived value of gig work $\hat{\theta}_{i,t} = \Phi(t, \phi_i) \cdot \theta_i$, which generates a perceived flow utility for worker $i$ at time $t$ given by

$$u_i(h, \omega; \hat{\theta}_{i,t}) = \hat{\theta}_{i,t} \cdot \frac{h^{1-\frac{1}{\varepsilon}}}{1-\frac{1}{\varepsilon}} - \big(p(\omega) + \kappa\big) \cdot d - P_i(\omega).^{19} \tag{3}$$

Misperceptions mean that individuals can find themselves on non-cost-minimizing policies, and that the opportunity to work in the gig economy can lead to welfare losses when individuals enter because of an inflated valuation.

The worker may exit the gig economy at any period. This could be due to either endogenous exit, if a worker's perceived value falls below the threshold that makes gig work worthwhile relative to their outside option, or because of an exogenous shock that removes them from the gig economy with probability $\eta$. This latter feature captures persistent shocks to valuations or outside options that render the surplus from gig work negative.

Given this series of decisions—participation, policy choice, hours, and exit—the model is solved via backward induction. Exit is either exogenous or equivalent to participation, so I start with hours, then policy choice, and then participation and exit.

**Hours worked.** Hours are chosen each period by workers conditional on their participation and policy in order to maximize perceived flow utility. Therefore, worker $i$ in period $t$ will pick hours to maximize equation

[17] Heterogeneity in the fixed premium $P_i$ is motivated by the fact that the firm personalizes prices for the fixed policy but not the variable policy.

[18] As well as learning about misperceptions, workers may also learn to improve their productivity on the job, which is not explicitly modelled here. Intuitively, this would be an alternative force that raises the true and perceived value of gig work over time. Since systematic dynamics in the model are captured by learning about misperceptions, I expect this mechanism to influence estimated misperceptions and the rate of learning. The direction of the effect is ambiguous *ex ante*.

[19] Note that the subscript $i$ for the utility function $u_i(\bullet)$ captures the role of individual specific fixed policy premiums $P_i$, which are suppressed as an argument of the function (this is true also for the value functions $V_i(\bullet)$ and $\tilde{V}_i(\bullet)$ that are defined later).

(3)

$$h^*_{i,t}(\omega) = \left(\frac{\hat{\theta}_{i,t}}{p(\omega)+\kappa}\right)^{\varepsilon}. \text{[20]} \tag{4}$$

**Policy choice.** Workers make their policy decision before they have entered into the gig economy, and so suffer from their initial misperception. They believe that their flow utility during their tenure in the gig economy will remain constant (*i.e.*, they are naïve about their learning) so they pick whichever contract yields a higher flow utility. Denote the value function

$$V_i\big(\omega;\hat{\theta}_{i,0}\big) = u\big(h^*_i(\omega),\omega;\hat{\theta}_{i,0}\big). \tag{5}$$

Therefore, individuals pick the policy which maximizes their perceived flow utility

$$\omega^*_i = \arg\max_{\omega}\big\{V_i\big(\omega_F;\hat{\theta}_{i,0}\big), V_i\big(\omega_V;\hat{\theta}_{i,0}\big)\big\}. \tag{6}$$

**Participation and exit.** Since workers believe they will remain in the gig economy until they exogenously exit, at which point they receive their outside option, they will decide to enter the gig economy if and only if

$$V_i\big(\omega^*_i;\hat{\theta}_{i,0}\big) > \nu_i. \tag{7}$$

Similarly, worker $i$ will exit at time $t$ if their perceptions evolve such that

$$V_i\big(\omega^*_i;\hat{\theta}_{i,t}\big) \leq \nu_i. \tag{8}$$

Panel 6a in figure 6 shows the mechanics of the model when there are no misperceptions and outside options are constant across individuals. If a worker's valuation exceeds an initial threshold, which is shown by the dashed red line, then they will decide to enter the gig economy and select the variable policy. Under this policy, their hours increase with their valuation, however, if their valuation exceeds a further threshold shown by the dashed green line, then they will opt for the fixed policy. Workers' hourly wage rate jumps up when they cross this threshold because they no longer face the hourly premium.

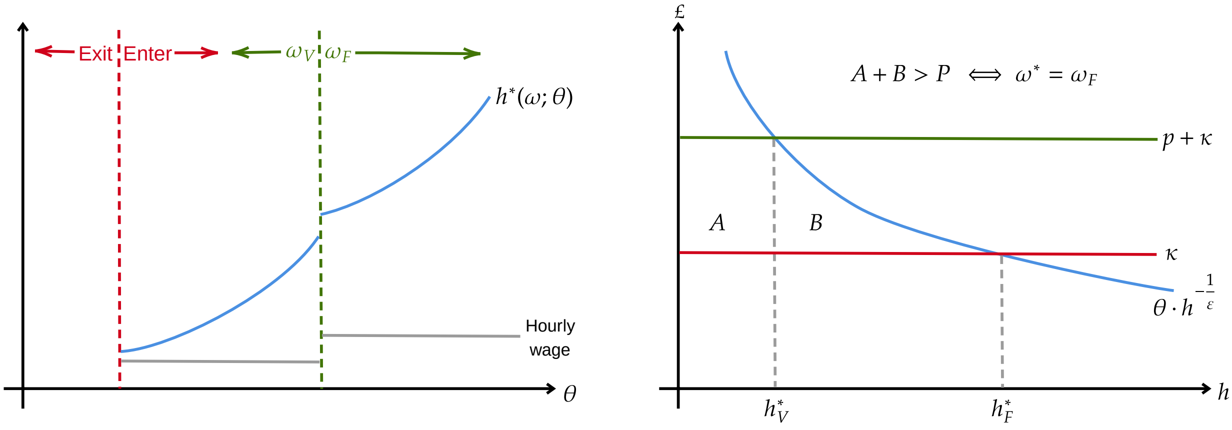
## 4.2 Introducing Shocks to Valuations

To help bring the model to the data, I introduce transitory shocks to workers' valuations. This is important because, although the reduced form evidence supports the hypothesis of misperceptions and learning, *ex post* shocks still affect the cost-minimizing nature of workers' policy choices. Consequently, any patterns in the data are the result of a confluence of these factors that the model should reflect. It is also evident that hours vary within workers across months, which further motivates shocks to the valuations that drive hours.

I consider that the workers' valuations may also be subject to independently and identically distributed shocks $\rho_{i,t} \in \mathbb{R}_+$ each period. The distribution of shocks is known to workers and they leave the worker's valuation unchanged in expectation $\mathbb{E}[\rho_{i,t}] = 1$. Thus, on any given period, the worker's true valuation is $\theta^{\rho}_{i,t} = \rho_{i,t}\cdot\theta_i$, although it will be perceived to be $\hat{\theta}^{\rho}_{i,t} = \rho_{i,t}\cdot\Phi(t,\phi_i)\cdot\theta_i$. That is, the worker's normative flow

[20] In the empirical implementation of the model, a time period is set to one month and hours are truncated from above at 480 (=16×30) to reflect time constraints, and workers fully account for this, although it does not bite much in practice.

Figure 6: Model Illustration




(a) The Worker's Problem

(b) Money-metric Identification

**Notes:** These figures illustrate different aspects of the model. Panel 6a shows the workers problem for a fixed outside option $\nu$ with no misperceptions $\phi = 1$ and a range of valuations $\theta$. The dashed red line denotes the threshold valuation required for a worker to enter the gig economy, and the dashed green line is the valuation threshold at which a worker would prefer to be on the fixed policy. The blue curve denotes the hours worked per month, while the grey line denotes the hourly wage rate. Panel 6b illustrates how the model identifies money-metric valuations. The blue curve denotes the marginal benefit of an additional hour in the gig economy, and the red and green lines show the marginal cost under the variable and fixed policy, respectively.

utility is given by

$$u_i(h, \omega; \theta_{i,t}) = \theta_{i,t}^{\rho} \cdot \frac{h^{1-\frac{1}{\varepsilon}}}{1 - \frac{1}{\varepsilon}} - \big(p(\omega) + \kappa\big) \cdot h - P_i(\omega), \tag{9}$$

although it is perceived to be

$$u_i(h, \omega; \hat{\theta}_{\rho,i,t}) = \hat{\theta}_{i,t}^{\rho} \cdot \frac{h^{1-\frac{1}{\varepsilon}}}{1 - \frac{1}{\varepsilon}} - \big(p(\omega) + \kappa\big) \cdot h - P_i(\omega). \tag{10}$$

If a shock is sufficiently low, then the worker may not want to work in the gig economy. In this case, they can access their outside option at a discount $\nu_i \cdot (1 - \psi)$, where $\psi \in (0, 1)$, and work zero hours in the gig economy. I refer to this as temporary exit. Temporary exit reflects two features of reality. Firstly, if an individual participates regularly in the gig economy, they have less time to invest in and raise the value of their outside option. Secondly, it captures any refundable fixed costs to enter into the gig economy. Moreover, temporary exit reflects intermittent short breaks in gig work that are observed in the data.

Conditional on participation and policy choice, hours are chosen according to equation (4) with $\hat{\theta}_{i,t}^{\rho}$ replacing $\hat{\theta}_{i,t}$, if the worker does not temporarily exit. Workers will temporarily exit if their shock $\rho_{i,t}$ is sufficiently low. In this case, they will receive their outside option at a discount $\nu_i \cdot (1 - \psi)$ so, formally, they will make a temporary exit and work zero hours if and only if

$$\begin{aligned} u(h_{i,t}^*(\omega), \omega; \hat{\theta}_{\rho,i,t}) &= \frac{1}{\varepsilon - 1} \cdot \frac{\big(\hat{\theta}_{i,t}^{\rho}\big)^{\varepsilon}}{\big(p(\omega) + \kappa\big)^{\varepsilon - 1}} - P_i(\omega) \leq \nu_i \cdot (1 - \psi) \\ \iff \quad \rho_{i,t} &\leq \left( \Big(\nu_i \cdot (1 - \psi) + P_i(\omega)\Big) \Big/ \frac{1}{\varepsilon - 1} \cdot \frac{\big(\hat{\theta}_{i,t}\big)^{\varepsilon}}{\big(p(\omega) + \kappa\big)^{\varepsilon - 1}} \right)^{\frac{1}{\varepsilon}} = \Gamma(\nu_i, \hat{\theta}_{i,t}, P_i). \end{aligned} \tag{11}$$

In summary, hours worked are determined by

$$h_{i,t}^*(\omega) = \begin{cases} \left(\frac{\hat{\theta}_{i,t}^{\rho}}{p(\omega)+\kappa}\right)^{\varepsilon} & \text{if } \rho_{i,t} > \Gamma(\nu_i, \hat{\theta}_{i,t}, P_i), \\ 0 & \text{otherwise}. \end{cases} \tag{12}$$

Workers' participation decisions must incorporate the possibility of temporary exit and engage more fully with the dynamic nature of the problem they face. They discount the future with discount factor $\beta \in (0,1)$ such that they perceive their discounted sum of utility under policy $\omega$ for worker $i$ at time $t$ to be

$$\begin{aligned} \tilde{V}_i(\omega; \hat{\theta}_{i,t}) =& \mathbb{E}\left[\sum_{t=0}^{\infty}(\eta \cdot \beta)^t \cdot \max\left\{\frac{1}{\varepsilon-1} \cdot \frac{\left(\hat{\theta}_{i,t}^{\rho}\right)^{\varepsilon}}{\left(p(\omega)+\kappa\right)^{\varepsilon-1}} - P(\omega), \nu_i \cdot (1-\psi)\right\}\right] + \sum_{t=0}^{\infty}(1-\eta^t) \cdot \beta^t \cdot \nu_i \\ =& \frac{1}{1-\eta \cdot \beta} \cdot \Bigg(\mathbb{P}\left(\rho_{i,t} \leq \Gamma(\nu_i, \hat{\theta}_{i,t}, P_i)\right) \cdot \nu_i \cdot (1-\psi) \dots \\ & \dots + \mathbb{P}\left(\rho_{i,t} > \Gamma(\nu_i, \hat{\theta}_{i,t}, P_i)\right) \cdot \left(\frac{1}{\varepsilon-1} \cdot \frac{\left(\hat{\theta}_{i,t}\right)^{\varepsilon}}{\left(p(\omega)+\kappa\right)^{\varepsilon-1}} \cdot \mathbb{E}\left[\rho_{i,t}^{\varepsilon} | \rho_{i,t} > \Gamma(\nu_i, \hat{\theta}_{i,t}, P_i)\right] - P_i(\omega)\right)\Bigg) \dots \\ & \dots + \frac{\nu_i \cdot \beta \cdot (1-\eta)}{(1-\beta) \cdot (1-\eta \cdot \beta)}, \end{aligned} \tag{13}$$

where the expression makes use of the fact that exogenous exit is an absorbing state. Equation (13) breaks down the value function into three lines. The first line contains the utility derived from temporary exit multiplied by the probability of temporary exit on any given period. This line is also multiplied by $1/(1-\eta \cdot \beta)$ to reflect the fact that these pay-offs form a geometric sequence with common ratio $\eta \cdot \beta$. The second line contains the expected flow utility conditional on working in the gig economy scaled by the probability of receiving a sufficiently large shock to work in the gig economy. The last line reflects the expected utility derived from the full value of a worker's outside option, which is received upon exit.

Therefore, worker $i$ selects policy

$$\omega_i^* = \arg\max_{\omega} \left\{\tilde{V}_i(\omega_F; \hat{\theta}_{i,0}), \tilde{V}_i(\omega_V; \hat{\theta}_{i,0})\right\}, \tag{14}$$

and enters the gig economy if and only if

$$\tilde{V}_i(\omega_i^*; \hat{\theta}_{i,0}) > \frac{\nu_i}{1-\beta}. \tag{15}$$

Exit at a given time $t$ occurs endogenously if and only if

$$\tilde{V}_i(\omega_i^*; \hat{\theta}_{i,t}) \leq \frac{\nu_i}{1-\beta}. \tag{16}$$

## 4.3 Discussion

The model sets out a sequence of decisions which are designed to capture the way workers engage with the gig economy. Here I discuss some aspects of the model.

**Heterogeneity.** The model allows for heterogeneity both across and within workers. Individuals differ in three key dimensions: gig work valuations, outside options, and misperceptions. A desire to capture the reduced form evidence in the most simple way motivates introducing heterogeneity in these specific areas. Firstly, there is significant variation in the average number of hours worked between workers, which motivates the flexibility of individual specific valuations of gig work. Secondly, workers make different policy choices; some make cost-minimizing decisions and, for those who do not, there is significant dispersion in how far their hours are from their break-even point. Differences in workers perceptions of their valuations are necessary to capture this feature of the data. Finally, it is intuitive that workers face different outside options and that this is critical to their participation in the gig economy. For example, one worker's participation may be precarious because they compare gig work with an alternative job, which is only marginally worse. For this individual, slight changes in their valuation could push them out of gig work. Conversely, an individual who has recently been made unemployed and has few employment prospects is much more attached to the gig economy.

The model also incorporates heterogeneity within a worker's spell in the gig economy. Workers are subject to transitory shocks that affect their valuation of gig work, as well as permanent shocks that remove them from gig work. These permanent shocks capture persistent negative shocks to the surplus that workers derive from gig work. Learning also generates dynamics in within-worker labor supply to the gig economy.

**Gig work valuations.** Workers' valuations of gig work encompass a broad range of factors that determine how gig work affects their utility. Perhaps most intuitively, one can think of the wage as being an important determinant of valuations, but even this is mediated by other factors such as workers' marginal utility of income. Amenities also likely play an important role, for example, flexible hours, absence of a boss, and varying demands of the job. All these characteristics of gig work, and more, make up valuations. The model allows workers to value these aspects differently but it does impose that, conditional on working the same number of hours, workers valuations are equivalent. In other words, workers who work the same number of hours value gig work equally, although the value may be derived from different factors. This does not mean workers with the same number of hours receive the same surplus since they will have different misperceptions and outside options.

**Outside options.** Outside options constitute workers' welfare in a world where they do not work in the gig economy. Broadly speaking, workers have two margins of adjustment. Firstly, they can replace the hours worked in the gig economy with another activity (*e.g.*, leisure or some alternative work). Secondly, they can adjust the bundle of activities that they undertake in a day. That is, workers can reorganize all the activities in their day and not just the time they would spend in the gig economy. For example, a worker who complements a full-time job with gig work may instead take up two part-time jobs, if gig work is not available. The model matches these features of reality in two ways: it has a linear cost to hours, which reflects the opportunity cost of time, and a fixed outside option that captures the re-bundling effect.[21]

**Misperceptions.** Workers can have disappointing experiences of gig work if their misperceptions cause them to be overly optimistic about the value of gig work. This can manifest visibly as workers being on the non-cost-minimizing policy, and as workers promptly reducing their hours and leaving the gig economy. In the model, workers are assumed to be ignorant of the possibility that they misperceive the value gig work.

[21] Individual-level outside options $\nu_i$ may also pick up some opportunity cost of time because this is assumed to be homogeneous across workers through $\kappa$, but in reality it may vary.

In other words, they act as if they have perfect knowledge of their valuation. This assumption allows the model to abstract from several potential influences. In particular, workers neither shade their valuations (Capen et al., 1971; Smith and Winkler, 2006; Thaler, 1988), nor stay in the gig economy to learn about their valuations, nor pick their hours to maximize an uncertain return.

Another important assumption is that workers misperceive the value of gig work rather than another parameter in the model. In particular, misperceptions could be about outside options, where optimism about gig work maps to pessimism about outside options in a similar way to that found by Jäger et al. (2022). However, the dynamics of hours and survival in the data, as well as the survey evidence, motivates the choice to embed misperceptions in the value of gig work. In particular, it is hard to conceive how misperceptions in outside options could generate dynamics within the gig economy.

**The gig work surplus.** The gig work surplus differs across and within workers because of the model's rich heterogeneity. Differences in surpluses across workers are driven primarily by variation in gig work valuations and outside options, but they can also differ because of suboptimal labor supply decisions driven by misperceptions. Within workers, shocks to valuations cause the surplus to vary over time and to eventually go to zero because of exogenous exit shocks.

The model's gig work surplus measure does not capture the possibility that outside options could change with, for example, a mass exodus from the gig economy. Plausibly, in the short-run, outside options may fall in such a situation as the supply of labor to other markets increases while demand remains unchanged leading to a fall in wages. Thus, the gig work surplus identified here serves as a lower bound. In the long run, the prospect for outside options is unknown in such a scenario.

# 5 Estimation

This section describes how I estimate the structural parameters in the model and provides a guide to how variation in the data helps identification. Further, I present the estimates and model fit before section 6 analyses their implications for welfare and counterfactual scenarios in the gig economy.

## 5.1 Simulated Method of Moments Estimator

The variable premium $p$ is set equal to its average of £0.94 in the data and I feed the model the empirical distribution of quoted fixed premiums $P_i$, which can be seen in figure B2. These are the premiums offered by the firm—not necessarily taken by customers—and so do not suffer from selection. Moreover, they are all but uncorrelated with hours worked, which implies they are unrelated to worker valuations $\theta_i$ from the perspective of the model. Through conversations with the firm it is also apparent that customers' policies are not priced based on any proxy for misperceptions $\phi_i$. Therefore, I assume individuals' fixed premiums are independent of their other characteristics.

To reduce the burden of estimation, I also fix the exogenous exit rate $\eta$ equal to the exit rate of pessimistic workers. In the model, these workers have no reason to leave aside from exogenous shocks. Further, the elasticity parameter $\varepsilon$ is mechanically adjusted in order to ensure an intensive margin labor supply elasticity compatible with empirical evidence. Lastly, $\beta$ is set to a standard value.

The remaining parameters are estimated using SMM. This requires an assumption about the distribution of heterogeneity in the population. I assume that individuals' valuations $\theta_i$ and outside options $\nu_i$ follow a joint log-normal distribution. This has three main advantages. Firstly, it helps to capture the skewed distribution of hours that is evident in the data. Secondly, it allows for an easily specified but rich pattern of correlations between these individual characteristics. Thirdly, the log-normal distribution permits convenient closed for expressions that are helpful with computation. For similar reasons, I specify *ex ante* misperceptions to follow an iid log-normal distribution with a mean equal to one.[22] Thus, in summary, individual heterogeneity is distributed according to

$$\begin{pmatrix} \theta_i \\ \phi_i \\ \nu_i \end{pmatrix} \sim \log\mathcal{N}\left(\begin{bmatrix} \mu_\theta \\ -\sigma_\phi^2/2 \\ \mu_\nu \end{bmatrix}, \underbrace{\begin{bmatrix} \sigma_\theta^2 & 0 & \sigma_{\theta,\nu} \\ 0 & \sigma_\phi^2 & 0 \\ \sigma_{\nu,\theta} & 0 & \sigma_\nu^2 \end{bmatrix}}_{=\Sigma}\right).$$

In practice, I estimate elements of the Cholesky decomposition $L$ of the covariance matrix $\Sigma = LL^T$ in order to ensure the latter is positive semidefinite, where I fix two elements of the lower triangular matrix $L$ to reflect the constraint imposed on the covariance matrix. Further, outside options are truncated at £10,000 since the model's parameters are identified from observed participants, thus, the data cannot speak to individuals who are far from entering the gig economy. The idiosyncratic shocks to worker valuations are assumed to be log-normal iid distributed $\rho_{i,t} \sim \log\mathcal{N}(\mu_\rho, \sigma_\rho^2)$ with $\mu_\rho = -\sigma_\rho^2/2$ so that $\mathbb{E}[\rho_{i,t}] = 1$.

This leaves ten parameters to estimate in the model; six parameters describing the joint log-normal distribution, the linear cost to hours $\kappa$, the rate at which misperceptions correct $\lambda$, the variance of shocks that affect workers valuations $\sigma_\rho^2$, and the sunk portion of gig workers' outside options $\psi$. Let $\boldsymbol{\zeta}$ denote this vector of parameters

$$\boldsymbol{\zeta} = \{\mu_\theta, \sigma_\theta^2, \sigma_\phi^2, \mu_\nu, \sigma_\nu^2, \sigma_{\theta,\nu}, \sigma_\rho^2, \psi, \kappa, \lambda\}.$$

I construct the difference between the $j$th model moment $\hat{m}_j(\bullet)$ and the $j$th data moment $m_j(\bullet)$ to be $e_j(\bullet)$ so that

$$e_j(\tilde{X}, X|\boldsymbol{\zeta}) = \hat{m}_j(\tilde{X}|\boldsymbol{\zeta}) - m_j(X),$$

where $X$ denotes the observed data and $\tilde{X}$ denotes the simulated data. The estimated parameters are those that minimize the weighted sum of errors

$$\hat{\boldsymbol{\zeta}} = \arg\min_{\boldsymbol{\zeta}} \ \boldsymbol{e}(\tilde{X}, X|\boldsymbol{\zeta})^T W \boldsymbol{e}(\tilde{X}, X|\boldsymbol{\zeta}),$$

where $\boldsymbol{e}(\bullet)$ is the stacked deviations of the moments $(J \times 1)$ and $W$ is a weight matrix $(J \times J)$.

[22] Previous iterations of the model, which allow for misperceptions to be systematically different from one and for a correlation with valuations, do not improve the model fit.

Standard errors for the parameters are computed according to

$$\hat{\mathbb{V}}(\hat{\zeta}) = \frac{N+1}{N} \cdot \left(\hat{\mathcal{J}}^T W \hat{\mathcal{J}}\right)^{-1} \left(\hat{\mathcal{J}}^T \hat{\Sigma} W \hat{\Sigma}^T \hat{\mathcal{J}}\right) \left(\hat{\mathcal{J}}^T W \hat{\mathcal{J}}\right)^{-1},$$

where $N$ is the number of simulations, $\mathcal{J}$ is the Jacobian matrix of the moment conditions ($J \times K$), and $\Sigma$ is the variance-covariance matrix of the moment conditions ($J \times J$) (Hansen, 1982). The hat notation $\hat{\bullet}$ reflects the fact that these components are functions of the data. In particular, the variance-covariance matrix of the moment conditions is estimated by block bootstrapping at the worker level and recomputing the moments. For moments that do not come from the main dataset, I construct their variances and assume zero correlation with the other moments.

The ten parameters of the model are identified with 23 empirical moments:

1. The labor market share of this section of the gig economy.
2. The proportion of workers on the variable policy.
3. The mean number of hours worked in a month by workers on the variable policy.
4. The standard deviation of the number of hours worked in a month by workers on the variable policy.
5. The mean number of hours worked in a month by workers on the fixed policy.
6. The standard deviation of the number of hours worked in a month by workers on the fixed policy.
7. The share of variable policy workers who are not on the cost-minimizing policy.
8. The share of fixed policy workers who are not on the cost-minimizing policy.
9. The mean distance between the break-even point and the number of hours worked in a month for non-cost-minimizing variably policy workers.
10. The standard deviation of the distance between the break-even point and the number of hours worked in a month for non-cost-minimizing variably policy workers.
11. The mean distance between the break-even point and the number of hours worked in a month for non-cost-minimizing fixed policy workers.
12. The standard deviation of the distance between the break-even point and the number of hours worked in a month for non-cost-minimizing variably policy workers.
13. The initial hazard rate of cost-minimizing variable policy holders.
14. The initial hazard rate of cost-minimizing fixed policy holders.
15. The initial hazard rate of non-cost-minimizing fixed policy holders.
16. The initial decline in hours per month of non-cost-minimizing variable policy holders.
17. The initial decline in hours per month of non-cost-minimizing fixed policy holders.
18. The mean quoted fixed premium for fixed policy holders.
19. The mean quoted fixed premium for variable policy holders.
20. The standard deviation of quoted fixed premiums for fixed policy holders.
21. The standard deviation of quoted fixed premiums for variable policy holders.

22. The frequency of zero hour months.
23. The average within worker standard deviation in hours.

The labor market share is derived from two sources separate to the firm's administrative data. Bertolini et al. (2021) report that 17.1% of the UK workforce work for digital platforms on at least a monthly frequency from a survey of 2,201 workers, and Cornick et al. (2018) state that 21% of the 95 gig workers in their sample work in food delivery. Hence, a labor market share of 3.59% (=17.1%×21%).[23] I reconcile the labor market share with the model through a stationarity assumption (*i.e.*, that the employment share of the gig economy is in steady state).

Lastly, I lean on the labor supply elasticity literature to pin down $\varepsilon$. Given the short term nature of gig work for most workers (the standard duration is less than six months), a Frisch elasticity seems the most appropriate. Moreover, given these are self-employed workers with total control over their hours one would expect this group of individuals to exhibit a high Frisch elasticity. I take the Frisch elasticity estimate of 0.80 (SE 0.10) from Fisher (2022), which is estimated on a sample of self-employed taxi drivers who are subject to exogenous variation in their wage rates due to London tube strikes. I combine this with online earnings data for delivery riders and scale this to account for idle time.[24]

In practice, I find the minimum of the objective function using a multi-start simplex search method, the weight matrix is set to normalize the error function to percentage deviations, and I simulate six million workers.

## 5.2 Sources of Identification

While the parameters are identified from a combination of structural assumptions and moments of the data, it is helpful to consider the particularly close linkages between some parameters and moments. Workers' hours are determined by an unobserved distribution of valuations of gig work $\{\mu_\theta, \sigma_\theta^2\}$. Valuations can be made up of many factors (*e.g.*, productivity, the wage rate, an individual's marginal utility of income, and their preference for the type of work), and the higher an individual's valuation the more they want to work in the gig economy. The model implicitly assumes that, aside from misperceptions, participants who work the same hours have the same valuations—and, thus, the distribution of valuations is implied by the hours distribution.

The money-metric nature of this valuation comes primarily from the policy choice, which can be framed as the opportunity to buy a higher wage rate at some upfront cost. Panel 6b in figure 6 shows the logic behind this. The additional benefit from switching to the fixed policy is the hourly earnings that are saved rather than paid to the firm (*i.e.*, area $A$), plus the sum of marginal benefits net of marginal costs for the additional hours that are worked (*i.e.*, area $B$). Therefore, the linear cost to work $\kappa$ must imply a solution to this problem that matches the data. To the extent that policy choice pins down a money-metric, variation in fixed premiums $P_i$ also helps identification.

Naturally, the prevalence of non-cost-minimizing choices and a measure of how far these choices were from justifying their policy choice implies the distribution of misperceptions $\sigma_\phi^2$. For example, an individual

[23] I assume independence between these two statistics and use that fact that for two independent random variables $X$ and $Y$ with finite first and second moments $\mathbb{V}(X \cdot Y) = \mathbb{V}(X) \cdot \mathbb{V}(Y) + \mathbb{V}(X) \cdot \mathbb{E}(Y)^2 + \mathbb{E}(X)^2 \cdot \mathbb{V}(Y)$ in order to construct the variance-covariance matrix of the empirical moments.

[24] For example, see Glassdoor or Indeed.

on the fixed policy who would save money on the variable policy must have overestimated their valuation (*i.e.*, $\phi_i > 1$) but this alone does not provide information about the extent of the individual's over optimism. However, if the individual was only one hour [100 hours] away from the break-even point, then their misperception must have been small [large]. Selection into participation plays an influential role in determining the distribution of misperceptions amongst the pool of gig workers.

The speed of learning $\lambda$ is primarily identified from the change in workers' hours upon entering the gig economy. A greater initial adjustment in hours, for example, indicates a quick rate of learning.

Outside options $\{\mu_\nu, \sigma_\nu^2\}$ are most connected to the employment share of the gig economy and survival rates. Intuitively, the employment share can identify a constant outside option; with a fixed distribution of valuations, a homogeneous outside option can be adjusted to ensure the correct employment share. The estimation leverages endogenous exit in the model to identify heterogeneity in outside options. With knowledge of the learning process and valuations, it is possible to infer outside options as the perceived valuations at which optimistic workers decide to leave the gig economy.

The covariance between valuations and outside options $\sigma_{\theta,\nu}$ comes from conditioning the moments on a combination of policy choice and whether the policy choice was cost-minimizing. For example, given knowledge of misperceptions and the speed of learning, differential exit rates amongst variable minimizers and optimistic workers is indicative of outside options. Again, selection into participation plays a key role in mediating the covariance amongst gig workers compared to the population as a whole.

The variance of shocks $\sigma_\rho^2$ is identified from the mean within worker standard deviation of hours worked, and the fraction of the outside option that is sunk due to regularly participating in the gig economy $\psi$ is inferred from the prevalence of interruptions to workers' spells in the gig economy (*i.e.*, the fraction of months where workers work zero hours but reappear in the data subsequently).

## 5.3 Parameters and Model Fit

Table 5 presents the structural estimates and their associated standard errors in the top panel. The lower panel of the table shows the fixed model parameters. Namely, the variable premium $p$, the mean and standard deviation of the fixed premium distribution $\{\mu_P, \sigma_P\}$, the discount factor $\beta$, exogenous exit rate $\eta$, and the elasticity parameter $\varepsilon$. The estimates map to a joint distribution of characteristics in the simulated population, which are shown in table 6.

Across the whole population, the mean and standard deviation of valuations $\theta_i$ is equal to 47 and 28, respectively. The distribution of valuations exhibits a right skew so that the median valuation equals 41. Workers with higher valuations are more likely to participate in the gig economy, all else equal, so participants exhibit higher valuations equal to 211 on average. This valuation would translate to a variable policy flow utility (*i.e.*, $u(h_i^*(\omega_V), \omega_V; \theta_i)$) of £1,398. I discuss the magnitude of these flows net of outside options in the next section.

The mean outside option for participants equals £674 with a standard deviation and median of £491 and £608, respectively. Since the average participant works approximately 50% FTE in the gig economy, it is useful to compare this with the median earnings of a part-time worker in the UK.[25] The average gig workers' outside option is just one quarter of this amount, which implies low quality and/or low pay alternative opportunities

[25]The Office for National Statistics classifies this as somebody who works less than 30 hours per week.

Table 5: Parameter Estimates

| $\zeta_{1:5}$ | $\hat{\zeta}_{1:5}$ | $\zeta_{6:10}$ | $\hat{\zeta}_{6:10}$ |
|---|---|---|---|
| $\mu_\theta$ | 3.70 | $\mu_\nu$ | 12.66 |
| | ( 0.07) | | ( 0.03) |
| $\sigma^2_\theta$ | 0.30 | $\sigma^2_\nu$ | 4.54 |
| | ( 0.03) | | ( 0.03) |
| $\sigma_{\theta,\nu}$ | -1.08 | $\sigma^2_\phi$ | 0.14 |
| | ( 0.05) | | ( 0.01) |
| $\lambda$ | 1.00 | $\kappa$ | 39.41 |
| | ( 0.13) | | ( 1.13) |
| $\sigma^2_\rho$ | 0.03 | $\psi$ | 0.27 |
| | ( 0.00) | | ( 0.00) |
| $\mu_P$ | 101.82 | $\sigma_P$ | 24.08 |
| $p$ | 0.94 | $\beta$ | $0.95^{1/12}$ |
| $\eta$ | 0.07 | $\varepsilon$ | $\Delta \hat{\log}(h) \cdot \log(1 + p/\hat{\kappa})$ |

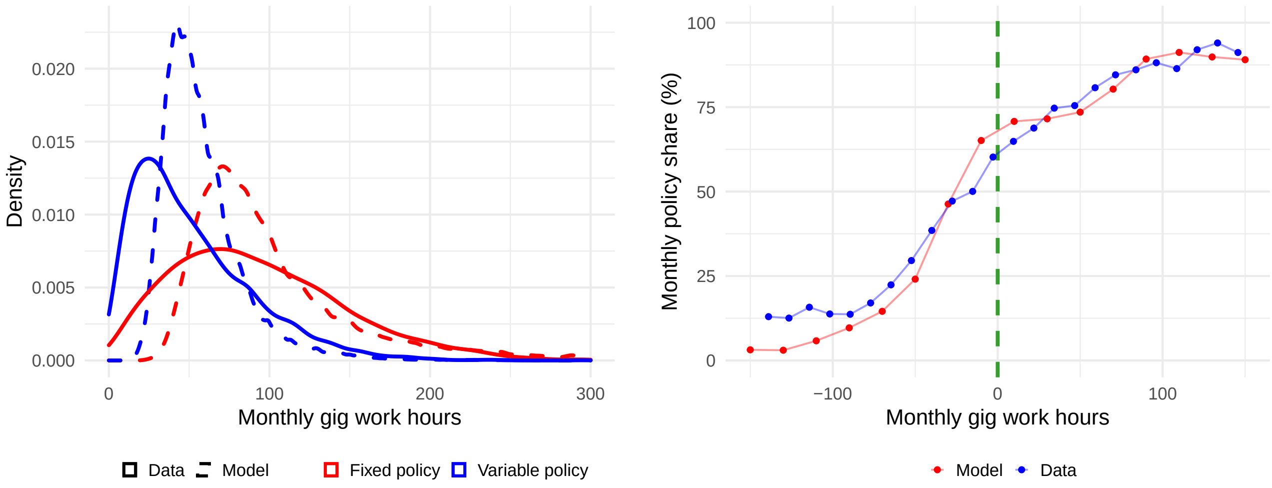
**Notes:** The top panel of this table presents estimates of the structural parameters from the model. Standard errors are contained in the parentheses, which are estimated as described in subsection 5.1. The second panel shows the variable policy premium and the mean of the quoted fixed premium distribution used in the estimation, which correspond to the empirical averages of these parameters, as well as the set discount factor.

Table 6: Simulated Population Characteristics

| Statistic | Population | Participants |
|---|---|---|
| Mean valuation $\theta$ | 47.24 | 211.17 |
| SD valuation $\theta$ | 28.07 | 48.81 |
| Mean misperception $\phi$ | 1.00 | 1.04 |
| SD misperception $\phi$ | 0.39 | 0.11 |
| Mean outside option $\nu$ | 9,744 | 674 |
| SD outside option $\nu$ | 1,258 | 491 |
| Correlation $\rho_{\theta,\nu}$ | -0.62 | -0.32 |
| Correlation $\rho_{\theta,\phi}$ | 0.00 | -0.27 |
| Correlation $\rho_{\nu,\phi}$ | 0.00 | 0.53 |

**Notes:** This table presents statistics that describe the simulated population. The first column shows these statistics for the entire population, while the second column conditions on participating in the gig economy.

Figure 7: Model Fit




(a) Hours Distribution

(b) Policy Choice

**Notes:** These figures plot comparisons of the model output with the data. Panel 7a shows the density of hours. The dashed lines reflect the model while the solid lines illustrate the data. The fixed policy holders are shown in red and the variable policy holders are shown in blue. Panel 7b is analogous to figure 2 but with the models predictions overlaid in red.

for these individuals. The mean outside option for the whole population is large—£9,744—because the vast majority of workers will not be drawn into the gig economy given any plausible variation in their economic environment.

*Ex ante*, the population level correlation between gig work valuations and outside options is ambiguous. Wealthy and high income individuals, who conceivably have high outside options, also likely have low valuations of gig work since they have low marginal utilities of income. Conversely, those with high valuations would work more in the gig economy which entails a greater opportunity cost of time that could be captured by a greater outside option.[26] The estimation suggests the former mechanism dominates and finds a moderate negative correlation between gig work valuations and outside options of -0.62. For gig participants, the negative correlation is largely undone by selection into the gig economy—workers with high outside options must have high valuations to justify their engagement—such that for those in the gig economy the correlation is much smaller at -0.32.

Misperceptions amongst gig workers are not severe because of learning. Nonetheless, the average gig worker is slightly optimistic because those with high misperceptions are more prone to entering. The average participant overestimates the value of gig work ($\phi_i - 1$) by 4%. Further, gig workers' misperceptions $\phi_i$ exhibit a significant standard deviation of 11%.

The participation decision drives a negative correlation of -0.27 between misperceptions and valuations for gig workers. Intuitively, this reflects two forces. Firstly, regression to the mean; individuals with a high valuation are likely to have a less extreme misperception given the variables are uncorrelated. Secondly, individuals with low valuations are more likely to have optimistic misperceptions, if they are participating in the gig economy.

Aside from individual specific characteristics, the model implies a reasonable degree of concavity of util-

[26]The linear cost to working $\kappa$ can also capture the opportunity cost of time but this parameter is restricted to be homogeneous so outside options can still pick up the opportunity cost of time associated with high valuations.

ity with respect to hours worked in the gig economy (*i.e.*, $1 - 1/\varepsilon$) equal to 0.59. The speed of learning parameter implies that misperceptions erode (*i.e.*, $1/(1+\lambda)$) by 50% at the end of the first period of work in the gig economy.

Concerning the stochastic element of the model, the standard deviation of valuations shocks $\sigma_\rho$ is estimated to be 0.18. This implies that constant individual level valuations are responsible for 79% of the estimated variation in valuations (*i.e.* $\sqrt{\mathbb{V}(\theta_i)/\mathbb{V}(\theta^\rho_{i,t})}$). These shocks can cause individuals to temporarily exit the gig economy and receive their outside option at a discount $\psi$ estimated to be 27%.

Table 7 compares the empirical moments with those from the model when it is evaluated at the parameter values from table 5. Overall, the model fits the 23 empirical moments well; the model's predictions are close to the data. Figure 7 provides visual confirmation by contrasting the data with model predictions of policy and hours choices. The quality of the fit supports the view that the estimation captures structural elements of workers engagement with the gig economy.

# 6 Welfare and Counterfactuals

This section describes the gig work surplus implied by the model and estimates of its structural parameters, and considers worker welfare in counterfactual scenarios. Specifically, I analyse the impact of mandatory benefits that workers qualify for by working a sufficient number of hours in the gig economy, and how fixed costs influence the gig work surplus. For the former, I construct a counterfactual experiment to reflect elements of California's Proposition 22, which involves hours thresholds to qualify for a health insurance stipend and, for the latter, I study the introduction of the variable policy. Lastly, I examine how misperceptions stymie the gig work surplus.

Broadly, the analysis reveals a large gig work surplus, which is concentrated amongst low-hours workers. The distribution of the gig work surplus across hours worked suggests that policymakers face a tricky trade-off between guaranteeing costly worker protections and maintaining the appeal of gig work to the majority of participants. Counterfactual analysis of mandatory benefits for workers who reach an hours threshold confirms this; worker welfare falls if they bear even half of the economic incidence associated with the costs of this policy. Fixed costs to gig work, which make dabbling in the gig economy unattractive, impose significant losses. The allocative inefficiency that arises from misperceptions is considerable. Eradicating misperceptions increases the gig work surplus by 21%, which stems almost equally from correcting optimistic and pessimistic perceptions.

## 6.1 The Gig Work Surplus

In this subsection, I measure the size and distribution of the gig work surplus by subtracting workers' outside options from their utility flows to construct a monthly measure of the surplus from gig work. This means, for example, a worker $i$ who temporarily exits receives a negative surplus equal to $-\psi \cdot \nu_i$ and that, naturally, non-participants receive no surplus.

Figure 8 presents the estimated distribution of the monthly gig work surplus. The mean monthly surplus for a gig worker equals £1,066, but this masks significant heterogeneity with a standard deviation of £1,775. Moreover, the ratio of the $30^{\text{th}}$ to $70^{\text{th}}$ percentile is equal to 6.5. This median monthly surplus equals £673 or,

Table 7: Model Fit

| Moment | Data | Model |
|---|---|---|
| Labor market share (%) | 3.6 | 3.8 |
| Variable policy share (%) | 66.8 | 68.9 |
| Mean hours per month, variable policy | 53.4 | 58.1 |
| SD hours per month, variable policy | 48.4 | 37.5 |
| Mean hours per month, fixed policy | 95.3 | 104.5 |
| SD hours per month, fixed policy | 68.6 | 74.8 |
| Share non-cost-minimizing (%), variable policy | 11.6 | 9.4 |
| Share non-cost-minimizing (%), fixed policy | 60.6 | 62.3 |
| Mean hours per month from cost-minimizing, variable policy | 40.1 | 40.9 |
| SD hours per month from cost-minimizing, variable policy | 36.6 | 47.3 |
| Mean hours per month from cost-minimizing, fixed policy | 55.6 | 40.8 |
| SD hours per month from cost-minimizing, fixed policy | 41.7 | 31.0 |
| Hazard rate for cost-minimizers (%), variable policy | 28.9 | 23.8 |
| Hazard rate for cost-minimizers (%), fixed policy | 20.7 | 17.6 |
| Hazard rate for non-cost-minimizers (%), fixed policy | 29.2 | 38.3 |
| Decline in hours per month for non-cost-minimizers, variable policy | -10.2 | -9.7 |
| Decline in hours per month for non-cost-minimizers, fixed policy | 2.6 | 2.9 |
| Mean quoted fixed premium (£), variable policy | 97.0 | 95.6 |
| Mean quoted fixed premium (£), fixed policy | 106.5 | 104.4 |
| SD quoted fixed premium (£), variable policy | 25.1 | 19.9 |
| SD quoted fixed premium (£), fixed policy | 24.3 | 26.5 |
| Share of zero hours months (%) | 5.5 | 5.9 |
| Mean within worker SD hours per month | 31.0 | 37.7 |

**Notes:** This table presents the targeted empirical moments alongside their model implied counterparts. The first column contains the empirical moments and the second column contains the model analogue.

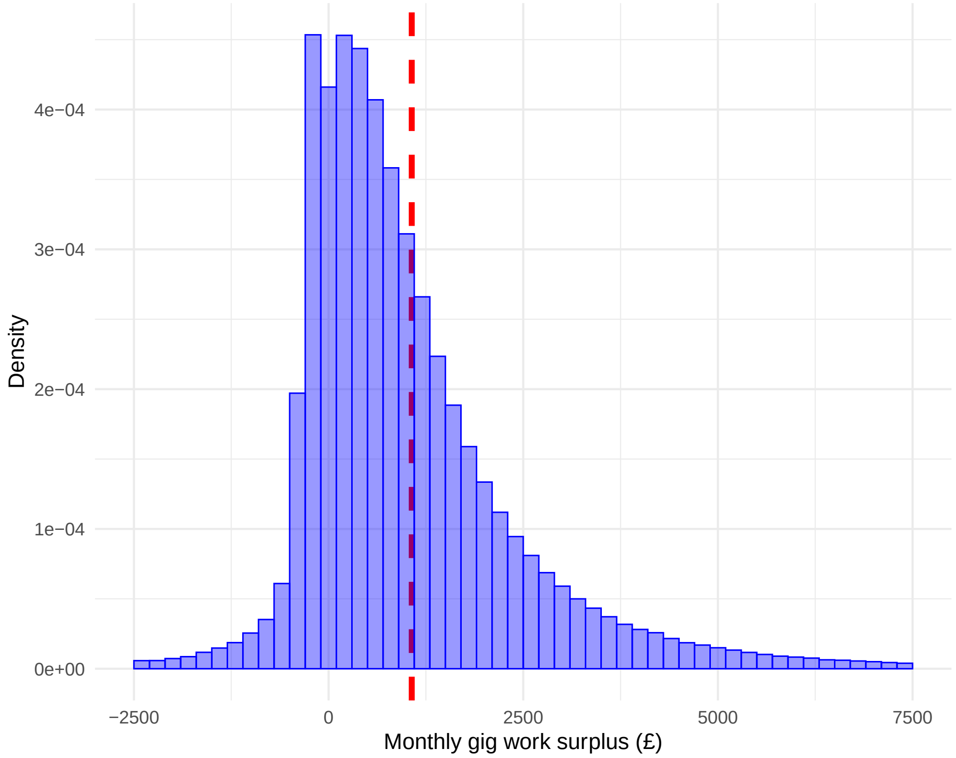
equivalently, 34% of the median employee's monthly income in the UK.

A small number of workers suffer a negative surplus because of their misperceptions and temporary exit. In an average month, only 16% of participants receive a negative surplus due to misperceptions.[27] However, 45% of each new entering cohort would be better off with their outside option and they lose £1,183 on average. Further, in a typical month, a worker who temporarily exits loses £291 relative to the full value of their outside option. These results reconcile the tension between the huge take up of gig work, which implies a surplus for workers, and the prevalence of negative stories surrounding gig work.

It is also informative to consider where the gig work surplus falls along the hours distribution. Figure 9 presents the average monthly gig work surplus by hours bin and the share of the total gig surplus that each bin accounts for. Two clear patterns emerge. Firstly, the surplus increases on average as hours increase. Fur-

[27] Jäger et al. (2022) offers a useful comparison to this result. This paper finds that "if workers had correct beliefs, at least 10% of jobs... would not be viable", which is close to the 16% number found here.

Figure 8: The Gig Work Surplus Distribution




**Notes:** This figure shows the histogram of the monthly gig work surplus from the estimated model. The red dashed line shows the mean of this distribution.

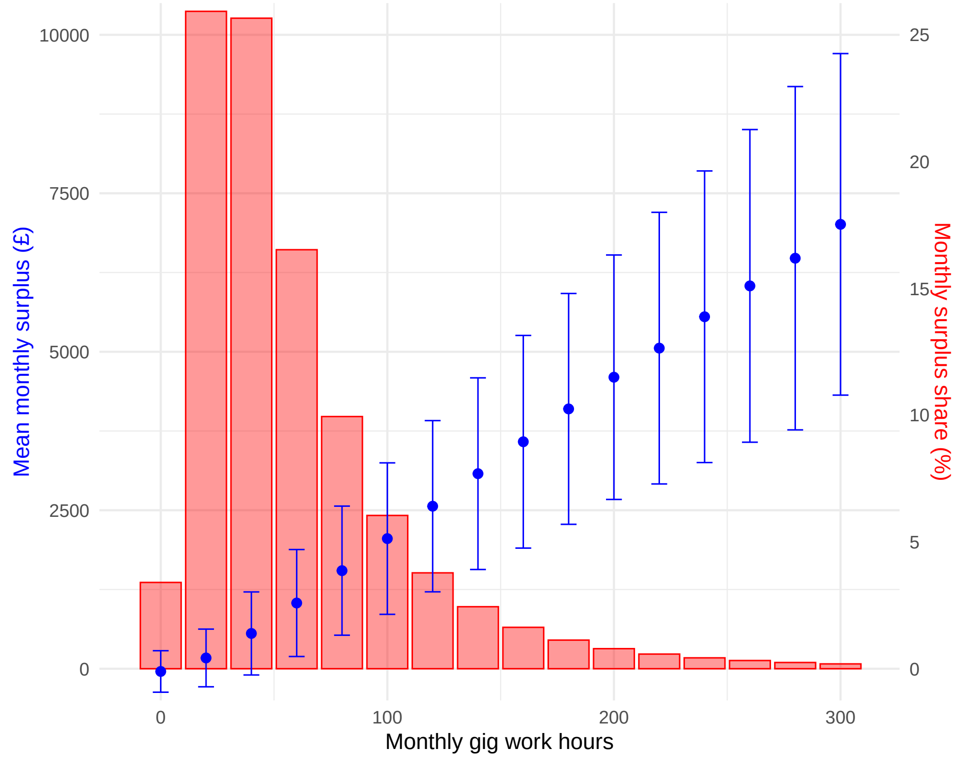
ther, the relationship is roughly linear so that the hourly surplus is more or less constant across the hours distribution. Secondly, the total worker surplus generated by the gig economy is concentrated amongst workers who work less the 50% full-time, which has important consequences for the counterfactual experiments that follow. In particular, it makes it difficult for regulators to enshrine protections for those engaged full-time in the gig economy without damaging the surplus of low hours workers, who comprise the majority of the gig work surplus.

I provide three benchmarks to compare these results against. Firstly, the mean hourly surplus is £12.04, which represents between 70 to 80% of the likely wage in the market after accounting for idle time. Secondly, I asked survey respondents "How much would your earnings have to drop per month for you to stop doing this type of work?". With significant caveats, the response to this question should reflect the difference between an individual's value of gig work and their outside option. Comfortingly, the median response is £462, which is roughly in line with the estimates presented here.[28] Thirdly, Chen et al. (2019) estimate a median "base" surplus of £571 using an exchange rate of dollars to pounds of 0.8. Thus, the results presented here are consistent with the scant evidence on the gig work surplus.

**Discussion.** The gig work surplus is unambiguously large, both in terms of its level (*i.e.*, the monthly surplus) and the rate at which it accrues (*i.e.*, the average hourly surplus). Taking the numbers above seriously, at an aggregate level, this part of the gig economy alone generates £15bn (= 32.8 million $\times$ 3.59% $\times$ £1,066 $\times$

[28] Survey responses were censored from above at £1,000 so it is not possible to compare means.

Figure 9: The Gig Work Surplus by Hours Worked




**Notes:** This figure plots the mean monthly gig work surplus for participants by different hours bins. The error bars show a one standard deviation in the monthly gig work surplus. The red bars show the share of the total gig work surplus accounted for by each hours bin.

12) in worker surplus annually.

Broadly, workers can derive a large surplus through either a high valuation of gig work or a low outside option, or both. In terms of valuations, workers can derive value from the gig economy through many channels. Income earned from gig work provides utility via consumption. These additional earnings may be particularly valuable if workers have a high marginal utility of income. There are reasons to believe this is the case; most of these individuals are in the bottom half of the income distribution and are likely to receive negative income shocks prior to entry into the gig economy (Cornick et al., 2018; Koustas, 2018). Moreover, a desire from workers to top-up their income with gig work implies a higher marginal utility of income than if this were not the case, *ceteris paribus*. Given institutional arrangements around self-employment, there is also greater scope for tax avoidance, as well as evasion, which can further inflate the value of earnings in the gig economy.

In addition to income, gig work entails a unique combination of amenities. To name a few, workers can pick their hours, do not have a boss, are paid weekly, and receive some income insurance because they can earn more or less depending on their circumstances. Individuals who especially value these amenities are more likely to select into gig work, which elevates the gig work surplus.

A large gig work surplus could also come from low outside options. This could be leisure, which may deliver little utility due to, for example, underemployment, complementarity with income, or other poor employment opportunities. Since gig work hours often fit around traditional work hours, the alternative employment offering may be particularly bad. In the UK, many gig workers are low-skilled migrants with

English as a second language, which could make traditional employment difficult to find. At an extreme, although food delivery platforms implement right to work checks, there are anecdotal stories that these checks are possible to circumvent (*e.g.*, by working under a different identity) such that gig participants may not have the opportunity to work elsewhere.

### 6.2 Hours-Based Benefits: Proposition 22's Health Insurance Stipend

In 2020, California passed a ballot initiative—Proposition 22—that provided gig workers with a range of new protections, while denying these workers the broader benefits received by employees.[29] One of the benefits under the Proposition includes a health insurance stipend for workers who meet an hours threshold. Precisely, workers who average more than 25 hours per week over a quarter are entitled to 100% of the average premium for a specified health insurance policy. Workers who average between 15 and 25 hours per week over the corresponding period are entitled to half that amount.[30]

While the setting for this paper is the UK, the Proposition provides a useful benchmark to think about the scale and structure of protections that may become available to gig workers in the UK and elsewhere. For this counterfactual experiment, I consider that workers who exceed 100 hours per month receive a fixed pecuniary benefit of £400, and that workers who work between 60 and 100 hours per month receive £200. These numbers are approximately in line with the aforementioned hours threshold and health insurance stipend.

The legislation places the statutory incidence on platforms to pay workers who qualify for benefits but the economic incidence will fall on a combination of the platforms, customers, and workers themselves. Importantly, it is the economic incidence, which depends on a number of factors outside the scope of this paper, that will determine the welfare effects of this policy (Besley and Case, 2000; Gruber and Krueger, 1991; Gruber, 1994).

To this end, I assume that the incidence on workers will manifest as a reduction in all their hourly earnings. An hourly earnings penalty is the most transparent way to model how platforms may pass on any costs. The important aspect of this assumption is that it imposes a variable cost on workers. Moreover, it is reasonable to assume that all workers will face this penalty because *a priori* platforms cannot determine which workers will qualify for benefits, and multi-homing across platforms will undermine any targeted incidence.

I consider a range of incidence $s$ from zero to full in order to study the welfare impacts of the counterfactual policy. The penalty on hourly earnings $c(\bullet)$ associated with incidence $s$ is found by numerically solving

$$c(s) \cdot H\big(c(s)\big) = s \cdot \Big(200 \cdot N_{200}\big(c(s)\big) + 400 \cdot N_{400}\big(c(s)\big)\Big),$$

where $H(\bullet)$ is the aggregate number of hours worked, and $N_{200}(\bullet)$ and $N_{400}(\bullet)$ denote the number of workers who qualify for the respective benefits.

This policy is non-marginal; it introduces sizeable non-convexities into workers' economic environment and, thus, lends itself to evaluation in a structural model. The presence of hours-based benefits induces several labor supply responses from gig workers that depend on the extent to which the economic incidence

[29] Proposition 22 exempted digital rideshare and delivery platforms from Assembly Bill 5.

[30] The text of the law can be found here https://vig.cdn.sos.ca.gov/2020/general/pdf/topl-prop22.pdf, and the section relating to the health insurance stipend is at the bottom of page 32.

falls on workers. Intuitively, the greater the incidence on gig workers, the higher the hourly wage penalty that individuals face. Workers can exhibit four potential labor supply responses: (i) non-participants may join the gig economy as they are encouraged by the benefits, (ii) some gig workers may exit as the wage penalty reduces their surplus below zero, (iii) individuals close enough to the hour thresholds may discretely increase their hours to qualify for the benefits, and (iv) the remainder of participants will reduce their hours as their hourly earnings fall to cover some proportion of the cost of the benefits. Specifically, workers' labor supply in the gig economy will follow

$$h^*_{i,t}(\omega) = \begin{cases} \left(\frac{\hat{\theta}^{\rho}_{i,t}}{p(\omega)+\kappa+c}\right)^{\varepsilon} & \text{if } \rho_{i,t} > \rho_{25}(\nu_i, \hat{\theta}_{i,t}), \\ 25 \times 4 & \text{if } \bar{\bar{\rho}}(\nu_i, \hat{\theta}_{i,t}) < \rho_{i,t} < \rho_{25}(\nu_i, \hat{\theta}_{i,t}), \\ \left(\frac{\hat{\theta}^{\rho}_{i,t}}{p(\omega)+\kappa+c}\right)^{\varepsilon} & \text{if } \rho_{15}(\nu_i, \hat{\theta}_{i,t}) < \rho_{i,t} < \bar{\bar{\rho}}(\nu_i, \hat{\theta}_{i,t}), \\ 15 \times 4 & \text{if } \bar{\rho}(\nu_i, \hat{\theta}_{i,t}) < \rho_{i,t} < \rho_{15}(\nu_i, \hat{\theta}_{i,t}), \\ \left(\frac{\hat{\theta}^{\rho}_{i,t}}{p(\omega)+\kappa+c}\right)^{\varepsilon} & \text{if } \underline{\rho}(\nu_i, \hat{\theta}_{i,t}) < \rho_{i,t} < \bar{\rho}(\nu_i, \hat{\theta}_{i,t}), \\ 0 & \text{if } \rho < \underline{\rho}(\nu_i, \hat{\theta}_{i,t}), \end{cases} \tag{17}$$

where $\underline{\rho}(\bullet)$, $\bar{\rho}(\bullet)$, $\rho_{15}(\bullet)$, $\bar{\bar{\rho}}(\bullet)$, and $\rho_{25}(\bullet)$ are defined in appendix E. Workers' policy choices and participation from the gig economy follow as before via backward induction.

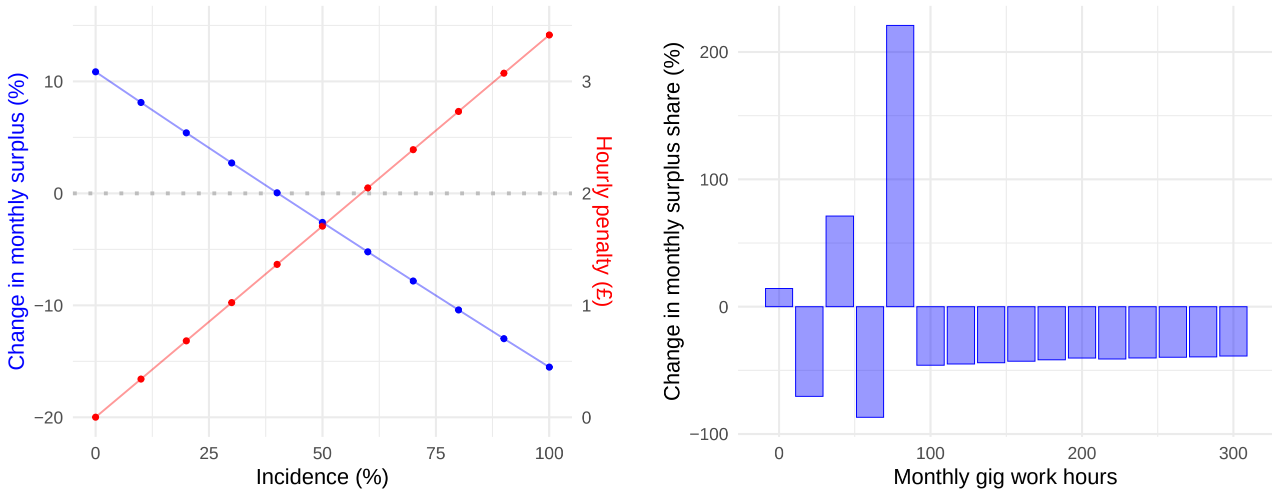
I quantitatively evaluate the welfare effects of this policy. From the policymaker's perspective there is a clear trade-off; higher mandated benefits help those workers who qualify to receive the transfer, but all workers bear the cost through lower hourly wages. This latter factor is exacerbated by the fact that a higher incidence on wages entails a higher wage penalty which discourages work further and necessitates a higher penalty again. Consequently, the efficacy of this policy in terms of raising gig worker welfare will depend critically on the degree of incidence and a complicated set of behavioral responses.

The welfare impacts of this policy are illustrated in figure 10. Panel 10a shows the gig work surplus under a range of different incidence levels. The analysis suggests that a moderate degree of incidence on workers—anything greater than 40%—would cause this hypothetical policy to decrease worker welfare. This level of incidence leads to a £1.37 drop in hourly earnings. That said, if there is minimal pass-through to workers, the policy could increase worker welfare by as much as 11%.

Panel 10b shows how the distribution of the gig work surplus changes across hours worked when incidence is evaluated at 40%. Note that this keeps the size of the gig work surplus the same. The share of the surplus that falls into the bins (40,60] and (80,100] grows substantially because workers bunch at 60 and 100 hours in order to received the mandated benefits. Everywhere else the share of the surplus falls since other hours bins see a decrease in the number of workers and the remaining workers face an hourly wage penalty, which makes them worse off.

The results of this analysis suggest that an appropriately structured minimum wage could complement mandatory benefits by legally limiting the degree of incidence on workers. Indeed, Proposition 22 included a minimum wage requirement that applies while workers are actively on jobs, but it is unclear whether the level set has any bite. It is also possible that platforms could find another margin through which to make workers pay but these are at least less obvious and could be tackled additionally.

Figure 10: Proposition 22 Counterfactual




(a) The Gig Work Surplus by Incidence

(b) % Change in the Gig Work Surplus Share by Hours

**Notes:** This figure plots outcomes under the Proposition 22 counterfactual policy. Panel 10a plots the percentage change in the total gig work surplus as a function of incidence in blue. The red line plots the hourly wage penalty associated with the different degrees of incidence. The dotted grey line shows the total surplus in the *status quo*. Panel 10b plots the percentage change in the share of the gig work surplus associated with each hours bin at an incidence level of 40%.

## 6.3 Higher Fixed Costs: Introduction of the Variable Policy

The introduction of the variable H&R policy constitutes a real life example of a reduction in the fixed costs associated with gig work. Therefore, in this counterfactual experiment I compare the gig work surplus in the *status quo* with a world without the variable policy. I also consider the removal of the fixed policy for completeness.

Without the variable policy it is possible that fixed policy premiums may adjust. Although there is little evidence of this in this market, I consider welfare absent the variable policy with and without adjustment in fixed premiums. To capture potential adjustment I make two assumptions. First, insurers' costs are linear in hours worked; discussions with the firm indicate that exposure to risk (*i.e.*, time on the road) is the greatest single driver of claims and that the relationship is linear up to a reasonable approximation. Second, the fixed policy market is competitive such that profits are zero (Einav and Finkelstein, 2011).

Under these assumptions I can back out the expected claims from an hour of driving as observed total fixed policy premiums collected divided by the sum of hours worked under the fixed policy. This implies an hourly expected claims cost $C$ of £0.95 per hour, which is very close to the variable policy premium. I assume that the firm will adjust all fixed policy premiums by a proportional amount $\alpha$, so that a given monthly premium $P_i$ will become $\alpha \cdot P_i$. I solve for $\alpha$ by requiring zero profits

$$H(\alpha) \cdot C = Q(\alpha),$$

where $H(\bullet)$ is aggregate hours worked and $Q(\bullet)$ denotes total premiums collected.

Again, this is a non-marginal counterfactual experiment that is well suited to a structural model. The introduction of the variable policy could spur two labor supply responses: (i) workers may switch policies

Table 8: % Change in the Gig Work Surplus under Different Scenarios

| Scenario | Agg. Surplus | Employment Share | Mean Surplus |
|---|---|---|---|
| *Status quo* | £40 | 3.8% | £1,066 |
| No variable policy (NVP) | -4.7 | -4.8 | 0.1 |
| NVP with endogenous $P_i$ | 0.8 | -2.4 | 3.2 |
| No fixed policy | 0.6 | 0.3 | 0.4 |
| No misperceptions | 20.6 | 23.0 | -2.0 |
| Half SD of misperceptions | 15.2 | 9.3 | 5.4 |
| No optimists | 8.7 | -8.3 | 18.6 |
| No pessimists | 11.7 | 31.5 | -15.0 |

**Notes:** This table shows the welfare affects of removing the variable policy. Column 1 shows the scenario considered and columns 2, 3, and 4 show the *per capita* surplus, the employment share of the gig economy, and the gig work surplus conditional on working, respectively. The top panel shows the base levels of these variables, and the bottom panel shows the percentage changes under the different scenarios.

and reduce the number of hours that they work, and (ii) some workers may enter the gig economy for the first time on the variable policy. These new entrants could benefit or be made worse off, with hindsight, depending on their perceptions.

Table 8 compares the welfare outcomes under the *status quo* and the counterfactuals. Without the variable policy and no adjustment in fixed policy premiums the gig work surplus is reduced by 4.7%, which stems mainly from a reduction in gig work participation. This means that, in aggregate, the introduction of the variable policy constituted a £709mn boon for workers. In contrast, welfare would be marginally higher without the fixed policy because, in the *status quo*, overly optimistic individuals select into the fixed policy but would be better off on the variable policy. The removal of the fixed policy benefits these workers more than it harms those who are correctly on the fixed policy.

Interestingly, if fixed policy premiums adjust to the removal of the variable policy, then worker welfare is slightly higher in a world without the variable policy. Participants in the gig economy exhibit a negative correlation between valuations and misperceptions, so that many low hours workers who select the variable policy are overly optimistic. Consequently, a small amount of fixed costs can prevent these individuals from making the mistake of entering. However, this scenario entails that fixed premiums should have risen by 30% since 2017—something that has not been observed. This counterfactual highlights that competition amongst firms that serve gig workers is crucial to the gig work surplus.

## 6.4 Reduced Misperceptions: Alleviating an Allocative Inefficiency

Misperceptions cause an allocative inefficiency in the gig economy: some optimistic workers participate in the gig economy when, with hindsight, they should not, meanwhile some pessimistic individuals would be better off inside the gig economy but are not. This inefficiency is material. The "No misperceptions" counterfactual in table 8 reveals that the gig work surplus would be 20.6% higher absent misperceptions.

Interestingly, both sides of this allocative inefficiency are roughly equally responsible for the welfare loss albeit in very different ways. Correcting all optimists' misperceptions causes a 8.3% reduction in participation but a 18.6% increase in the mean surplus. In contrast, if all pessimists are disabused of their misperceptions, there is a 31.5% increase in the employment share and a 15% fall in the typical surplus.

The model suggests that policies aimed at reducing misperceptions could be a fruitful pursuit. Halving the standard deviation of misperceptions obtains three quarters of the welfare gains from eradicating misperceptions. Think-tanks have touted policies that require platforms to increase transparency by, for example, providing predictions of hourly earnings to workers. It remains to be seen how these information treatments would translate to a reduction in misperceptions.

## 7 Conclusion

This paper evinces new empirical facts about workers' engagement with the gig economy: the majority of individuals work less than 50% FTE but a minority of workers work in excess of full-time; individuals select fixed and variable cost structures in line with their hours but many of them still make non-cost-minimizing choices; these decisions are correlated with survival and hours dynamics. Survey evidence supports the view that this results from misperceptions about the value of gig work combined with learning.

I develop a structural model of participation in the gig economy, which captures these core features and is amenable to the evaluation of prospective gig work policies and other counterfactuals. The model fits the data well and, together, they produce precise estimates of the model's structural parameters.

The estimates imply that participants enjoy a surplus of £1,066 per month from gig work. When aggregated, this implies an annual worker surplus of £15bn from a labor market that was nascent a decade ago. Despite this, the model can explain many workers' negative experiences of gig work. Misperceptions cause some individuals to enter the gig economy when they would be better off with their outside option. Learning and endogenous exit ameliorate these losses, although misperceptions still materially reduce the size of the gig work surplus.

Counterfactual experiments reveal that policymakers must be conscious of the incidence of any mandatory benefits that they impose. If the associated cost is passed onto workers in terms of lower hourly earnings, for example, then the gig work surplus will fall sharply as gig work becomes less attractive for the majority of participants. Further analysis suggests that convexifying workers' budget constraints through the gig economy can yield significant welfare gains by allowing individuals to fine-tune their bundle of leisure and consumption. For example, the introduction of the variable policy, which reduced fixed costs for gig work, is estimated to have increased the gig work surplus by 4.7%.

In my view, this study has two main limitations that can serve as a foundation for future research. Firstly, valuations of gig work are inferred from the hours individuals work; the model imposes that individuals who work more in the gig economy value each hour of work more. This does not rank individuals' surpluses since outside options are heterogeneous, but one can imagine other useful information (*e.g.*, wages and preferences over amenities) would enrich the model.

Secondly, individuals' outside options are fixed. In reality, workers choose a bundle of activities and the amount of time to devote to each activity subject to time constraints. Viewing gig work through this perspective and engaging with the IO literature on bundles and discrete-continuous choice problems could offer new insights and, in practice, make outside options endogenous to the opportunities available to workers.

# References


Abraham, K. G., Haltiwanger, J., Sandusky, K., and Spletzer, J. (2019). The rise of the gig economy: fact or fiction? In *AEA Papers and Proceedings*, volume 109, pages 357–61.

Adams, A., Freedman, J., and Prassl, J. (2018). Rethinking legal taxonomies for the gig economy. *Oxford Review of Economic Policy*, 34(3):475–494.

Akbarpour, M., Alimohammadi, Y., Li, S., and Saberi, A. (2021). The value of excess supply in spatial matching markets. *arXiv preprint arXiv:2104.03219*.

Angrist, J. D., Caldwell, S., and Hall, J. V. (2021). Uber versus taxi: A driver's eye view. *American Economic Journal: Applied Economics*, 13(3):272–308.

Athey, S., Castillo, J. C., and Chandar, B. (2021). Service quality on online platforms: Empirical evidence about driving quality at uber. *Available at SSRN 3499781*.

Bernhardt, A., Campos, C., Prohofsky, A., Ramesh, A., and Rothstein, J. (2022). Independent contracting, self-employment, and gig work: Evidence from california tax data. Technical report, National Bureau of Economic Research.

Bertolini, A., Bogg, A., Colclough, C., Cole, M., Farrar, J., Ford, M., Gutierrez, C., Huws, U., Powdrill, T., Rix, M., Spencer, N., Sharp, T., Williamson, J., and Wishart, R. (2021). *Seven Ways Platform Workers are Fighting Back*.

Besley, T. and Case, A. (2000). Unnatural experiments? estimating the incidence of endogenous policies. *The Economic Journal*, 110(467):672–694.

Blundell, R. and MaCurdy, T. (1999). Labor supply: A review of alternative approaches. *Handbook of labor economics*, 3:1559–1695.

Boeri, T., Giupponi, G., Krueger, A. B., and Machin, S. (2020). Solo self-employment and alternative work arrangements: A cross-country perspective on the changing composition of jobs. *Journal of Economic Perspectives*, 34(1):170–95.

Bresnahan, T. F. and Reiss, P. C. (1990). Entry in monopoly market. *The Review of Economic Studies*, 57(4):531–553.

Bresnahan, T. F. and Reiss, P. C. (1991). Entry and competition in concentrated markets. *Journal of political economy*, 99(5):977–1009.

Broughton, A., Gloster, R., Marvell, R., Green, M., Langley, J., and Martin, A. (2018). *The Experiences of Those in the Gig Economy*.

Caldwell, S. and Danieli, O. (2020). *Outside options in the labor market*. Pinhas Sapir Center for Development, Tel Aviv University.

Caldwell, S. and Harmon, N. (2019). Outside options, bargaining, and wages: Evidence from coworker networks. *Unpublished manuscript, Univ. Copenhagen*, pages 203–207.

Caldwell, S. and Oehlsen, E. (2021). Gender differences in labor supply: Experimental evidence from the gig economy. *Unpublished*.

Camerer, C., Babcock, L., Loewenstein, G., and Thaler, R. (1997). Labor supply of new york city cabdrivers: One day at a time. *The Quarterly Journal of Economics*, 112(2):407–441.

Capen, E. C., Clapp, R. V., and Campbell, W. M. (1971). Competitive bidding in high-risk situations. *Journal of petroleum technology*, 23(06):641–653.

Castillo, J. C. (2020). Who benefits from surge pricing? *Available at SSRN 3245533*.

Chen, K.-M., Ding, C., List, J. A., and Mogstad, M. (2020). Reservation wages and workers' valuation of job flexibility: Evidence from a natural field experiment. Technical report, National Bureau of Economic Research.

Chen, M. K., Currier, L., Rossi, P. E., and Chevalier, J. A. (2022). Suppliers and demanders of flexibility: The demographics of gig work. *Unpublished*.

Chen, M. K., Rossi, P. E., Chevalier, J. A., and Oehlsen, E. (2019). The value of flexible work: Evidence from uber drivers. *Journal of political economy*, 127(6):2735–2794.

Cohen, P., Hahn, R., Hall, J., Levitt, S., and Metcalfe, R. (2016). Using big data to estimate consumer surplus: The case of uber. Technical report, National Bureau of Economic Research.

Collins, B., Garin, A., Jackson, E., Koustas, D., and Payne, M. (2019). Is gig work replacing traditional employment? evidence from two decades of tax returns. *Unpublished paper, IRS SOI Joint Statistical Research Program*.

Cook, C., Diamond, R., Hall, J. V., List, J. A., and Oyer, P. (2021). The gender earnings gap in the gig economy: Evidence from over a million rideshare drivers. *The Review of Economic Studies*, 88(5):2210–2238.

Cornick, P., Lepanjuuri, K., and Wishart, R. (2018). *The Characteristics of Those in the Gig Economy*.

Cullen, Z. and Farronato, C. (2021). Outsourcing tasks online: Matching supply and demand on peer-to-peer internet platforms. *Management Science*, 67(7):3985–4003.

Datta, N. (2019). Willing to pay for security: a discrete choice experiment to analyse labour supply preferences.

DellaVigna, S. and Malmendier, U. (2006). Paying not to go to the gym. *american economic Review*, 96(3):694–719.

Dubal, V. (2019). An uber ambivalence: Employee status, worker perspectives, & regulation in the gig economy. *UC Hastings Research Paper*, (381).

Dubal, V. B. (2021). Economic security & the regulation of gig work in california: From ab5 to proposition 22. *European Labour Law Journal*, page 20319525211063111.

Economides, N., Seim, K., and Viard, V. B. (2008). Quantifying the benefits of entry into local phone service. *the RAND Journal of Economics*, 39(3):699–730.

Einav, L. and Finkelstein, A. (2011). Selection in insurance markets: Theory and empirics in pictures. *Journal of Economic perspectives*, 25(1):115–38.

Fisher, J. (2022). The cost of labor supply biases. *Unpublished*.

Ganserer, A., Gregory, T., and Zierahn, U. (2022). Minimum wages and the rise in solo self-employment.

Garin, A., Jackson, E., and Koustas, D. (2022). New gig work or changes in reporting? understanding self-employment trends in tax data. *Understanding Self-Employment Trends in Tax Data (May 19, 2022). University of Chicago, Becker Friedman Institute for Economics Working Paper*, (2022-67).

Garin, A. and Koustas, D. (2021). The distribution of independent contractor activity in the united states: Evidence from tax filings. *Joint Statistical Research Program of the Statistics of Income Division of the IRS*.

Goettler, R. L. and Clay, K. (2011). Tariff choice with consumer learning and switching costs. *Journal of Marketing research*, 48(4):633–652.

Goldman, T. and Weil, D. (2021). Who's responsible here? establishing legal responsibility in the fissured workplace. *Berkeley J. Emp. & Lab. L.*, 42:55.

Grubb, M. D. and Osborne, M. (2015). Cellular service demand: Biased beliefs, learning, and bill shock. *American Economic Review*, 105(1):234–71.

Gruber, J. (1994). The incidence of mandated maternity benefits. *The American economic review*, pages 622–641.

Gruber, J. and Krueger, A. B. (1991). The incidence of mandated employer-provided insurance: Lessons from workers' compensation insurance. *Tax policy and the economy*, 5:111–143.

Hall, J. V. and Krueger, A. B. (2018). An analysis of the labor market for uber's driver-partners in the united states. *Ilr Review*, 71(3):705–732.

Handel, B. R., Kolstad, J. T., Minten, T., and Spinnewijn, J. (2020). The social determinants of choice quality: evidence from health insurance in the netherlands. Technical report, National Bureau of Economic Research.

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica: Journal of the econometric society*, pages 1029–1054.

Harris, S. D. and Krueger, A. B. (2015). *A Proposal for Modernizing Labor Laws for Twenty-First-Century Work: The" Independent Worker"*. Brookings Washington, DC.

Hoffman, M. and Burks, S. V. (2020). Worker overconfidence: Field evidence and implications for employee turnover and firm profits. *Quantitative Economics*, 11(1):315–348.

Jackson, E. (2019). Availability of the gig economy and long run labor supply effects for the unemployed. *Work. Pap., Stanford Univ., Stanford, CA Google Scholar Article Location*.

Jäger, S., Roth, C., Roussille, N., and Schoefer, B. (2022). Worker beliefs about outside options. Technical report, National Bureau of Economic Research.

Katsnelson, L. and Oberholzer-Gee, F. (2021). *Being the Boss: Gig Workers' Value of Flexible Work.* Harvard Business School.

Katz, L. F. and Krueger, A. B. (2017). The role of unemployment in the rise in alternative work arrangements. *American Economic Review*, 107(5):388–92.

Katz, L. F. and Krueger, A. B. (2019). The rise and nature of alternative work arrangements in the united states, 1995–2015. *ILR review*, 72(2):382–416.

Kolsrud, J. (2018). Sweden: Voluntary unemployment insurance.

Koustas, D. (2018). Consumption insurance and multiple jobs: Evidence from rideshare drivers. *Unpublished working paper.*

Koustas, D. K. (2019). What do big data tell us about why people take gig economy jobs? In *AEA Papers and Proceedings*, volume 109, pages 367–71.

Lambrecht, A. and Skiera, B. (2006). Paying too much and being happy about it: Existence, causes, and consequences of tariff-choice biases. *Journal of marketing Research*, 43(2):212–223.

Mahoney, N. (2022). Principles for combining descriptive and model-based analysis in applied microeconomics research. *Journal of Economic Perspectives*, 36(3):211–22.

Mas, A. and Pallais, A. (2017). Valuing alternative work arrangements. *American Economic Review*, 107(12):3722–59.

Mas, A. and Pallais, A. (2020). Alternative work arrangements. forthcoming. *Annual Review of Economics.*

Nevo, A., Turner, J. L., and Williams, J. W. (2016). Usage-based pricing and demand for residential broadband. *Econometrica*, 84(2):411–443.

Parrott, J. A. and Reich, M. (2018). An earnings standard for new york city's app-based drivers. *New York: The New School: Center for New York City Affairs.*

Prassl, J. (2018). *Humans as a service: The promise and perils of work in the gig economy.* Oxford University Press.

Ravenelle, A. J. (2019). *Hustle and gig: Struggling and surviving in the sharing economy.* Univ of California Press.

Rochet, J.-C. and Tirole, J. (2006). Two-sided markets: a progress report. *The RAND journal of economics*, 37(3):645–667.

Smith, J. E. and Winkler, R. L. (2006). The optimizer's curse: Skepticism and postdecision surprise in decision analysis. *Management Science*, 52(3):311–322.

Stanton, C. T. and Thomas, C. (2021). Who benefits from online gig economy platforms? Technical report, National Bureau of Economic Research.

Thakral, N. and Tô, L. T. (2021). Daily labor supply and adaptive reference points. *American Economic Review*, 111(8):2417–43.

Thaler, R. H. (1988). Anomalies: The winner's curse. *Journal of economic perspectives*, 2(1):191–202.

Weyl, E. G. (2010). A price theory of multi-sided platforms. *American Economic Review*, 100(4):1642–72.

# Appendices

## A Policy Choice

I argue that cost-minimization is the main motivation for workers when they chose their H&R insurance policy. In this appendix, I present supportive evidence and I discuss other potential influences.

### A.1 Cost-Minimization Motive

The primary difference between the fixed and variable policy is the tariff structure. The variable policy is paid for by the hour while the fixed policy comes with a fixed monthly cost. Therefore, gig workers will select the policy that minimizes their costs *ceteris paribus*. In practice, this means that workers who work less than 110 will generally be better off on the variable policy.

This logic is supported by figure 2, which shows the share of workers on the fixed policy across different hours bins. The diagram reveals that workers are increasingly likely to take the fixed policy as they work more hours. This pattern is consistent with a rational cost-minimizer who suffers from imperfect foresight. As described in the main text, a cost-minimizer with perfect foresight would behave in accordance with the dashed green line. Introducing noise into this individual's problem would distort the step function into an upward sloping line that crosses the break-even point at 50%. The exact shape of the line would depend on the distribution of the noise. Thus, the data is consistent with workers that try to cost minimize but who are subject to *ex ante* misperceptions and *ex post* shocks.

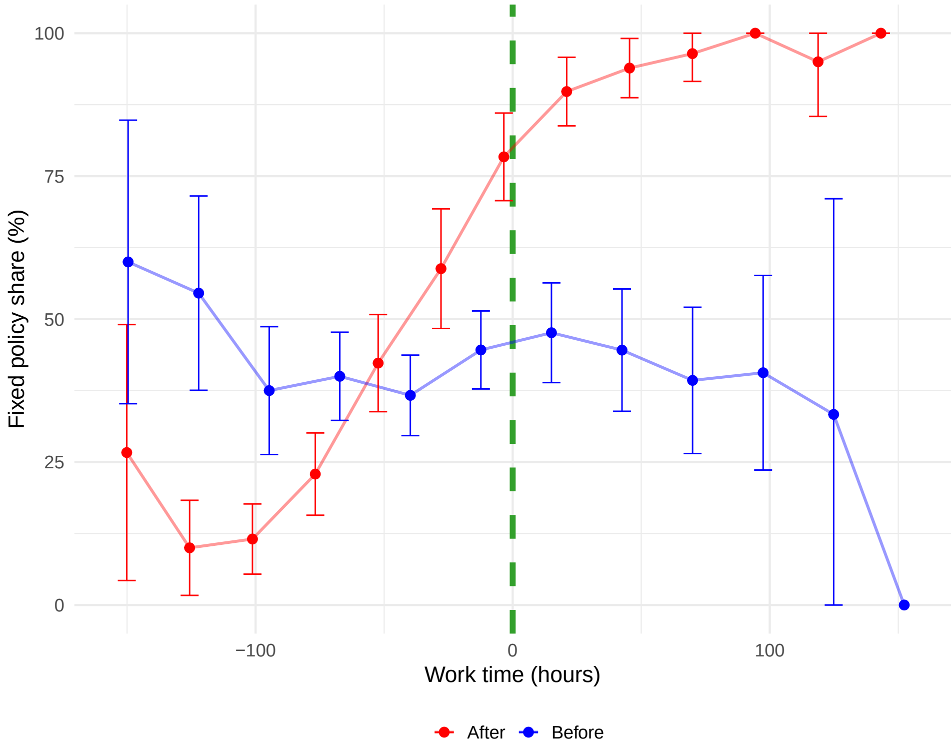
The few workers who switch, and who are excluded from the main analysis, can offer further support to the cost minimization hypothesis. Figure A1 shows that workers who switch policy initially make very poor decisions from a cost-minimization perspective. The blue line describes choices before switching; there is no discernible increase in the probability of opting for the fixed policy as hours worked increase. However, after switching, policy choices reflect a strong cost-minimizing tendency. The increasing red line indicates that the fixed policy share increases with hours. This evidence supports the view that switchers seek to correct non-cost-minimizing choices, which in turn supports the hypothesis that individuals want to minimize costs in their policy choice when they first enter the gig economy.

### A.2 Alternative Motives

Other factors besides cost minimization could influence the choice between the fixed and variable policies. In this subsection, I provide a taxonomy of these factors.

**Coverage.** The hourly policy only offers third party coverage, therefore, drivers who desire more comprehensive coverage may select a fixed policy despite higher costs. I deal with this by adjusting premiums for reported WTP for additional coverage from the survey, when individuals opt for greater coverage. Further, I check the robustness of reduced form and structural results to restricting attention to third party only policies in appendix C and F.

Figure A1: Fixed Policy Share by Hours Worked, Before and After Switching




**Notes:** This figure plots the share of workers who are on the fixed policy by hours bins for switchers before and after they switch. Each observation in a bin is a worker, so the hourly bin than an individual falls into is determined by their average monthly hours. The green dashed line indicates the perfect cost-minimizer's policy choice, which is vertical at the break-even point. Standard errors are constructed by applying the law of large numbers to a Bernoulli random variable (*i.e.*, $\sqrt{p \cdot (1-p)/N}$ where $p$ is the share of policies on the fixed policy in a bin and $N$ is the number of observations in that bin). In total, this figure contains 588 workers.

**Insurance value.** The different policy choices imply different variances in costs and overall income. In this case, the policy that minimizes variance in income would have some value aside from its implications for the expected level of costs. Although the fixed policy minimizes variance in costs, the variable policy minimizes variance in overall incomes since increases in costs from more hours worked are offset with higher earnings from work. The overall result is that income is more steady. As such, any insurance value would push in the direction of selecting the variable policy.

Quantitatively, reasonable degrees of risk aversion and the likely magnitude of fluctuations in income imply this story is not a significant driver of any patterns in the data. Consider an individual with CARA preferences and a degree of risk aversion equal to 0.0016 (Handel and Kolstad, 2015). In this case, reducing the variance of income by £1000 would increase money-metric utility by £0.80. Moreover, workers exhibit a slightly excessive tendency to opt for the fixed policy, which goes against this mechanism.

**Engine size.** Only scooters with an engine size of up to 125cc can opt for the variable policy. Fortunately, I can observe engine size with the quote data and exclude them from the analysis. Less than 1% of engine sizes exceed the threshold.

**Liquidity.** If agents are illiquid, then they may not be able to afford the up-front cost of the fixed insurance policy. This would push such individuals towards the variable policy regardless of their expected hours. Therefore, some of these individuals may find it more economical to select the fixed policy but are not able to do so. Given that few variable policyholders would reduce costs by being on the fixed policy, this does not seem to be a significant friction. Moreover, if illiquid workers are able to access credit, then liquidity issues should not affect their policy choice.

**Taxi meter effect.** The taxi meter effect would occur in this context if workers receive higher utility gross of insurance costs from the same number of hours on the fixed policy than on the variable policy (Lambrecht and Skiera, 2006). The motivating example, and origin for the name of the effect, is that taxi rides are less enjoyable simply because the customer can see the fare tick up on the taxi meter. Such an effect would push workers to choose the fixed policy over the variable policy *ceteris paribus*.

**Present bias.** Sophisticated present bias pushes workers towards selecting the fixed policy, while the effect of naïve present bias is ambiguous. An agent who is aware of their present bias (*i.e.*, sophisticated) may prefer to opt for the fixed policy to correct their inefficiently low level of labor supply (Lockwood, 2020). In the same vein, naïve present bias may push workers to choose the fixed policy since they overestimate how much they will work in the future (Augenblick and Rabin, 2019). On the other hand, naïve present bias may cause workers to favour the variable policy to avoid the upfront cost of the fixed policy, if workers suffer from present-bias over money or the timely consumption opportunities that it brings.

# B Data Cleaning and Filters

In this appendix, I describe the data cleaning procedures and the subsequent filters I apply to the data in order to arrive at my analysis sample. I also discuss the survey and quotes data in more detail.

## B.1 Data Cleaning and Filters

The raw data is in calendar month tranches, which were received from the firm, with observations at the job level (*i.e.*, each observation is a food delivery). I collapse each tranche down to the user-policy level, or just to the user level in the case of variable policy holders. Then, a variable policy holder's monthly observation is associated with a calendar month, while I merge fixed policy holders' policies that were divided over two calendar months such that their month corresponds to a 30 day policy duration. I then drop any user-policy-month-year duplicates, and removed any fixed policies that exceed 30 days. Some fixed policy holders' policies are further fragmented because of changes to their policy over the course of its duration (*e.g.*, a customer might have switched their coverage). I combine these policies by checking start and finish dates of the fragmented policies to see combinations of policies that consist of precisely 30 days. After this step, I drop any policies that are shorter than 14 days. Lastly, I trim observations according to monthly work time and premiums at the 0.1% percentile in order to remove outliers.

Hours of work may be understated for variable policy holders in the their first month because an observation is associated with a calendar month such that they may begin work halfway through a month. To deal with this, I *pro rata* work hours of these individuals according to what they did work while they were active in that month and, if there is less than two weeks left in the month, then I drop these partial observations. I have tested robustness to the two week threshold (*e.g.* using less than one week or less than three weeks) and the effect on the data is minimal. A similar problem arises for both fixed and variable policyholders in their final month before exiting; if workers leave after one week into their final month then their hours are not reflective of their engagement in the gig economy. I resolve this issue analogously.

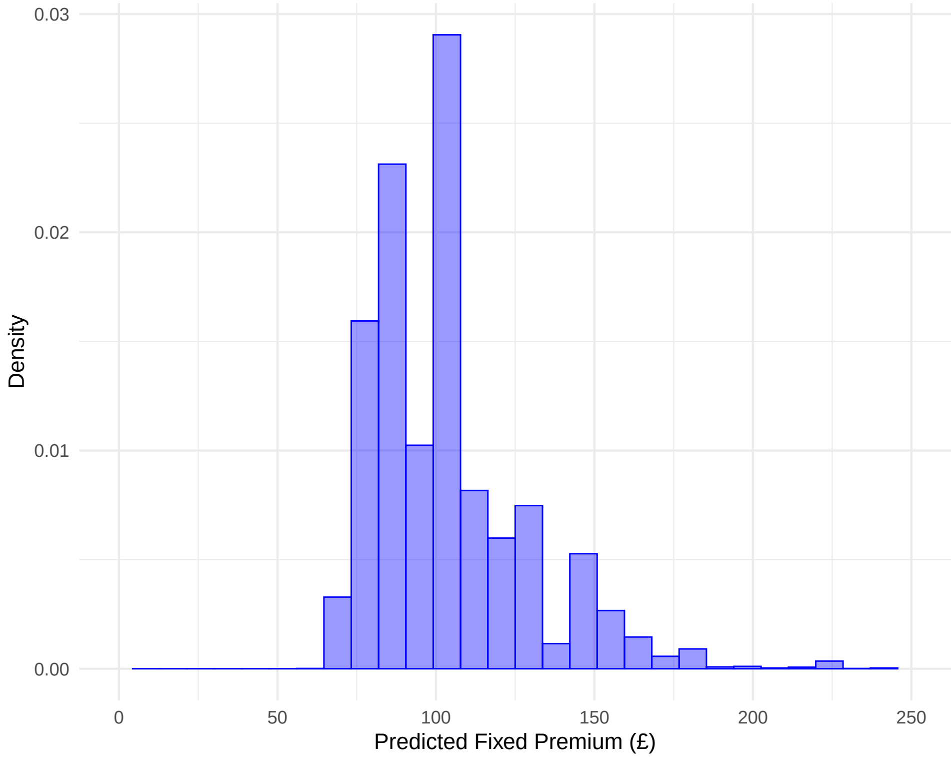
I apply four filters after data cleaning. Firstly, I remove annual policy holders, hence, their omission from the data cleaning discussion. Secondly, I keep only individuals who start after January 1 2019 since before this point the firm did not offer a consistent menu of policies. Thirdly, I identify workers who have more than one stint in the gig economy by flagging breaks of four months or longer. For these workers, I keep only their first stint in the gig economy. Finally, I remove individuals from the main analysis who switch policy during their initial spell.

## B.2 Survey Data

I complement the administrative data with survey data. The survey was conducted through the firm by emailing customers, if they had subscribed to receive promotional material. Fixed policyholders were oversampled because they have a greater tendency to subscribe affirmatively. 500 workers started the survey, of which 336 completed it. Of these, 251 are on the fixed policy.

The survey was sent out twice (the second time as a reminder) in June 2022. Therefore, the workers surveyed are not necessarily in my administrative data and, despite efforts, they cannot be merged. Therefore, to construct categories (*e.g.*, minimizers) I rely on self-reports of hours and the premiums that they face.

Figure B2: Empirical Distribution of Fixed Premiums




**Notes:** This figure plots the empirical distribution of offered fixed policy premiums from the firms quote data.

## B.3 Quotes Data

The firm sent data on the quotes that they have offered to all enquirers over the analysis period. This allows for the observation of an unselected distribution of quotes and, in particular, I can view the premiums offered to variable policyholders.

From the fixed policyholders, it is clear that quoted premiums typically fall shy of realized premiums. To resolve this issues, I non-parametrically predict realized premiums with quoted premiums for fixed policy holders by calculating the average realized premium within £5 quoted premium bins. Then, I use this model to construct counterfactual realized premiums for all other customers. Figure B2 shows the resulting distribution of premiums.

# C Robustness Checks: Reduced Form Evidence

In the main body of the paper, I present a number of reduced form facts. In this appendix, I present robustness checks which show that the patterns in the data persist across different subgroups of the population, with different definitions of categories, and over different time periods in the data. I also present additional empirical evidence.

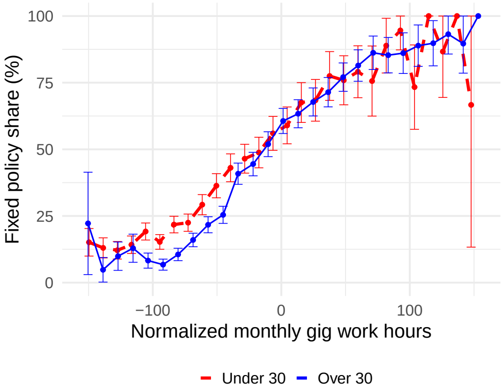
Figure C3 shows the analogue of figure 2 for various cuts of the data. Panel C3a shows the share of individuals on the fixed policy for those under and over 30, separately. Broadly, the two age groups show the same pattern, although over 30s are slightly more effective at minimizing costs by selecting the variable policy more frequently, when hours are below the break-even point. Panel C3b shows the equivalent data when including all types of coverage and when keeping soley third party only policies. Since only fixed policies can include levels of coverage other than third party only, the blue line is mechanically lower than the red line, where the latter reflects the fixed policy share for all types of coverage. Panel C3c shows the share of individuals on the fixed policy for the main analysis sample and for the sample where individuals whose predicted premiums exceed £175 are excluded. Panel C3d shows the fixed policy share for the period of time before the Covid pandemic, and for the course of the Covid pandemic during which there were several lock-downs in UK, where food delivery riders could still work. For transparency, figure C4 shows the fixed policy share by non-normalized hours where the dashed green line represents a perfect cost minimizer who faces the average fixed premium and the variable premium.

Figure C6 shows that the patterns of survival persist for the categories across various robustness checks. Panels C5a and C5b use alternative definitions to construct categories. The former categorizes workers based on whether they minimize their bill of the course of their tenure in the gig economy; and the latter compares average hours with break-even points to categorize workers. Panels C5c and C5d show the patterns of survival for different categories, using the baseline definition, for under and over 30s, respectively. Panels C5e and C5f illustrate the same data for the period before the Covid pandemic and for the course of the Covid pandemic, respectively. Panel C5g exhibits the pattern of survival for different categories, restricting to those on third party only policies. Panel C5h shows the survival trajectories for non-third party only policies by categories. These are necessarily fixed policies, so the diagram excludes pessimistic workers. Panel C5i shows the survival curves of minimizers broken down by fixed and variable policy holders. Panel C5j shows the survival patterns of workers who select different types of coverage. Panel C5k shows that survival function for workers preceding and during the Covid pandemic.

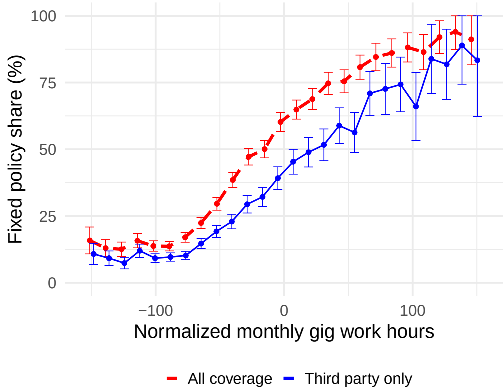
Table C1 reproduces table 4 with alternative controls. Columns (1) to (4) vary the controls included in the Cox proportional hazards model. In order, the columns exclude cover, gender, age, and low hours controls. Table C2 and C3 show the results from linear probability models, which show the same patterns as table 4 using the same controls. Table C2 shows an OLS regression where the outcome variable is the exiting the gig economy within the specified periods of time, while C3 shows the analogue conditional on reaching that tenure.

Figure C6 shows the trajectory of hours, like figure 4, for subgroups of the data. Panels C6a and C6b shows the trajectory of hours for categories using alternative definitions, which are analogous to those in figure C6. Panels C6c and C6d show the trajectory of hours over time for under and over 30s respectively. Panels C6e and C6f show the dynamics of hours before and during the Covid pandemic. Panel C6g shows the baseline

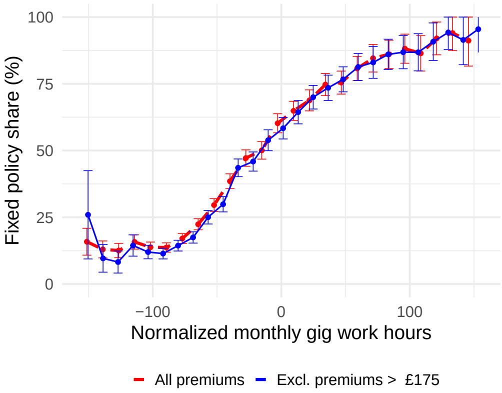
Figure C3: Fixed Policy Share by Normalized Hours for Different Samples



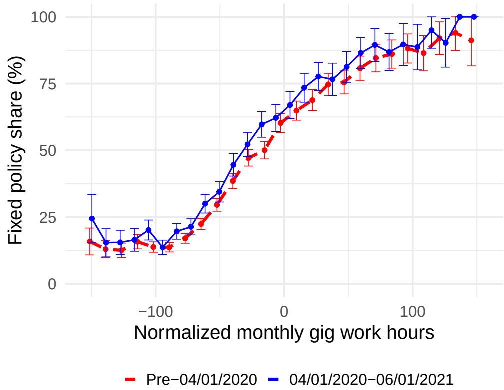

(a) By age




(b) By cover




(c) By premiums




(d) By time

**Notes:** These figures plot the share of individuals on the fixed policy by different groups. Standard errors are calculated by applying the law of large numbers to the average of a random variable that follows the binomial distribution with one trial (*i.e.*, $\sqrt{p_j \cdot (1-p_j)/N}$ where $p_j$ is the share of responses for a given category $j$ and $N$ is the number of observations).

Table C1: Cox Proportional Hazards Model with Time-Varying Coefficients

| | *Dependent variable:* | | | |
|---|---|---|---|---|
| | Tenure in the gig economy (months) | | | |
| | No cover control | No gender control | No age control | No low hours control |
| | (1) | (2) | (3) | (4) |
| Mean hours | $-0.001^{***}$ | $-0.001^{**}$ | $-0.001^{**}$ | $-0.006^{***}$ |
| | (0.0003) | (0.0003) | (0.0003) | (0.0003) |
| Minimizer (<= 2 months) | $0.179^{*}$ | $0.222^{**}$ | $0.262^{***}$ | $0.181^{*}$ |
| | (0.099) | (0.100) | (0.099) | (0.099) |
| Optimistic (<= 2 months) | $0.388^{***}$ | $0.480^{***}$ | $0.535^{***}$ | $0.322^{***}$ |
| | (0.103) | (0.104) | (0.104) | (0.104) |
| Minimizer (> 2 months) | $-0.095$ | $-0.053$ | $-0.007$ | $-0.289^{***}$ |
| | (0.070) | (0.070) | (0.070) | (0.070) |
| Optimistic (> 2 months) | 0.082 | $0.172^{**}$ | $0.230^{***}$ | $-0.122$ |
| | (0.076) | (0.078) | (0.078) | (0.078) |
| Low hours | Yes | Yes | Yes | No |
| Time controls | Yes | Yes | Yes | Yes |
| Age | Yes | Yes | No | Yes |
| Gender | Yes | No | Yes | Yes |
| Cover | No | Yes | Yes | Yes |
| Observations | 23,969 | 23,969 | 24,013 | 23,969 |
| $R^2$ | 0.075 | 0.075 | 0.073 | 0.038 |

**Notes:** $^{*}$p<0.1; $^{**}$p<0.05; $^{***}$p<0.01. This table shows estimates from a Cox proportional hazards model with time varying coefficients on the categories variable; the effect of this factor variable is allowed to differ between the first two months of a workers spell and any remaining months. Coefficients reflect percentage changes in the base hazard rate associated with the corresponding variable. The main panel of the table shows estimates for these coefficients and estimates of the coefficient on a workers average number of hours per month. Standard errors are shown in parentheses.

Table C2: Linear Probability Model

| | *Dependent variable:* | | | |
|---|---|---|---|---|
| | Tenure <= 2 months | | 2 < Tenure <= 6 months | |
| | Controls | No Controls | Controls | No Controls |
| | (1) | (2) | (3) | (4) |
| Mean hours | 0.0002** | 0.0002* | 0.0001 | 0.0002** |
| | (0.0001) | (0.0001) | (0.0001) | (0.0001) |
| Optimistic | 0.083*** | 0.091*** | 0.010 | 0.025*** |
| | (0.010) | (0.009) | (0.010) | (0.009) |
| Pessimistic | −0.076*** | −0.066*** | 0.076*** | 0.067*** |
| | (0.020) | (0.021) | (0.020) | (0.020) |
| Low hours | Yes | Yes | Yes | Yes |
| Time controls | Yes | No | Yes | No |
| Age | Yes | No | Yes | No |
| Gender | Yes | No | Yes | No |
| Cover | Yes | No | Yes | No |
| Observations | 14,795 | 15,924 | 14,795 | 15,924 |
| $R^2$ | 0.163 | 0.077 | 0.135 | 0.009 |

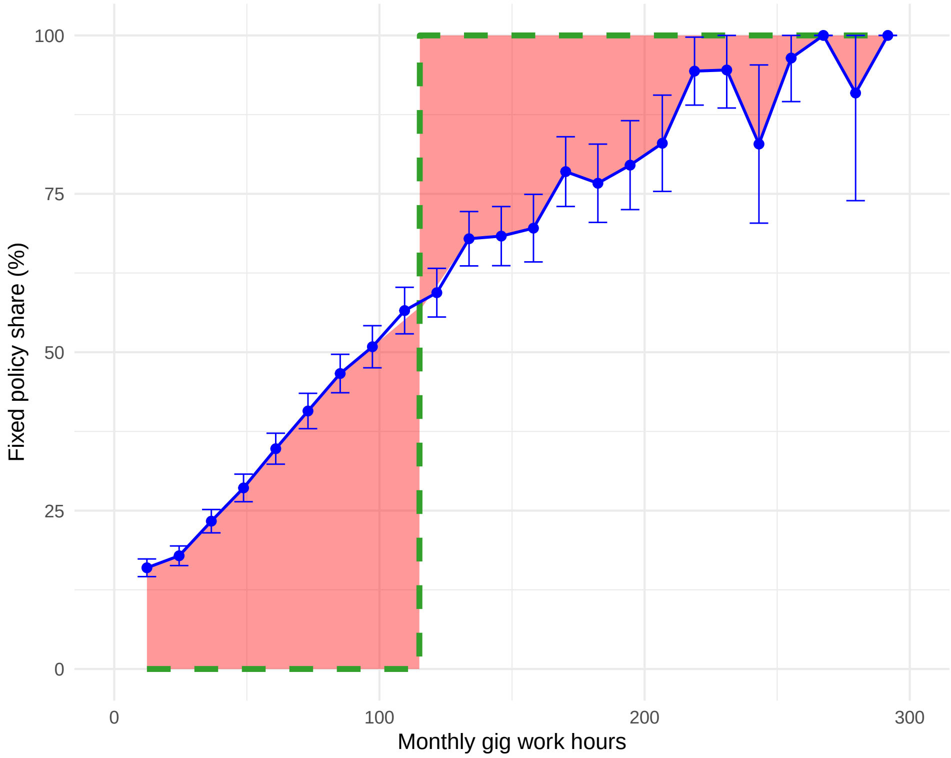
**Notes:** *p<0.1; **p<0.05; ***p<0.01. This figure plots the coefficients from an OLS regression, where the outcome is a binary variable that takes value one, if that worker survived for either less than two months—columns (1) and (2)—or between 2 and 6 months—columns (3) and (4). Columns (2) and (4) have no controls, similar to column (3) in figure 4. Columns (1) and (3) control for all available covariates, as in column (1) in figure 4

Table C3: Linear Conditional Probability Model

| | *Dependent variable:* | | | |
|---|---|---|---|---|
| | Tenure <= 2 months | | 2 < Tenure <= 6 months | |
| | Controls | No Controls | Controls | No Controls |
| | (1) | (2) | (3) | (4) |
| Mean hours | 0.0002** | 0.0002* | 0.0003** | 0.0005*** |
| | (0.0001) | (0.0001) | (0.0001) | (0.0001) |
| Optimistic | 0.083*** | 0.091*** | 0.105*** | 0.125*** |
| | (0.010) | (0.009) | (0.013) | (0.013) |
| Pessimistic | −0.076*** | −0.066*** | 0.038* | 0.049* |
| | (0.020) | (0.021) | (0.023) | (0.025) |
| Low hours | Yes | Yes | Yes | Yes |
| Time controls | Yes | No | Yes | No |
| Age | Yes | No | Yes | No |
| Gender | Yes | No | Yes | No |
| Cover | Yes | No | Yes | No |
| Observations | 14,795 | 15,924 | 9,174 | 9,805 |
| $R^2$ | 0.163 | 0.077 | 0.263 | 0.021 |

**Notes:** *p<0.1; **p<0.05; ***p<0.01. This figure plots the coefficients from an OLS regression, where the outcome is a binary variable that takes value one, if that worker survived for either less than two months—columns (1) and (2)—or between 2 and 6 months conditional on surviving beyond 2 months—columns (3) and (4). Columns (2) and (4) have no controls, similar to column (3) in figure 4. Columns (1) and (3) control for all available covariates, as in column (1) in figure 4.

Figure C4: Raw Fixed Policy Share




**Notes:** This figure plots the share of workers who are on the fixed policy by monthly hours bins. Each observation in a bin is a worker, so the hourly bin that an individual falls into is determined by their average monthly hours. The green dashed line indicates the perfect cost-minimizer's policy choice, which is vertical at the break-even point. Standard errors are constructed by applying the law of large numbers to the average of Bernoulli random variables (*i.e.*, $\sqrt{p \cdot (1-p)/N}$ where $p$ is the share of policies on the fixed policy in a bin and $N$ is the number of observations in that bin).

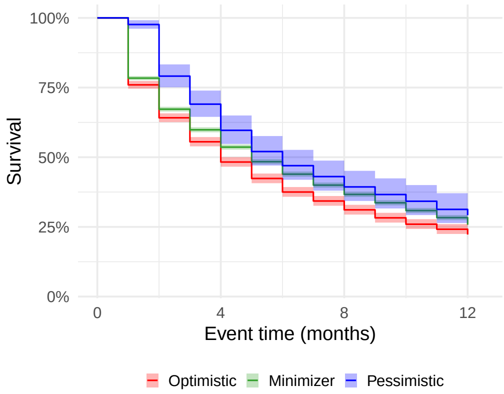
figure but without enforcing a balanced panel. Lastly, panel C6h shows the trajectory of hours for third party only policyholders.

Figure 4 shows the dynamics of hours at the weekly level. They can also be displayed at the monthly level—this is done in figure C7. I do this for all three definitions of the categories that I use.

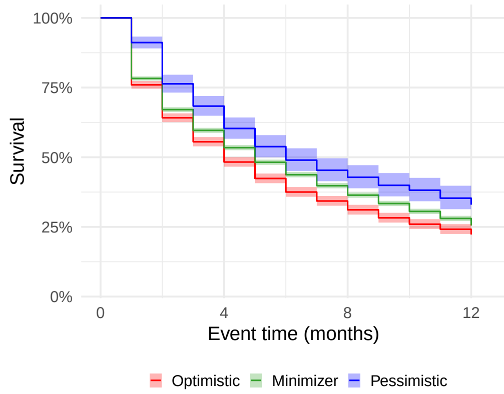
Figure C8 shows responses to the survey of gig workers' experiences for self-reported optimistic and pessimistic workers, respectively. To construct this figure, I subtract the share of responses by minimizers from those of the other categories in order to illustrate the relative prevalence of responses. Moreover, although the differences are not statistically significant, in order to get some reasonable precision I aggregated all the questions asked in the survey about workers experiences. These are questions about earnings, costs, and the difficulty of work. The figure shows that pessimistic workers are most likely to report aspects of gig work are better than they expected. Meanwhile, minimizers are most likely to report experiences as expected. Lastly, optimistic workers most frequently report gig work to be worse than expected.

Finally, figure C9 shows the aggregate responses to all experience questions at the finest level of response. There is a tendency to find gig work worse than expected, which is consistent with some evidence of optimism in fixed policy choice and results from the model. However, the majority of the mass falls in the "As expected" and "A little worse" bins, which indicates misperceptions are not too severe—again consistent with the model's results.

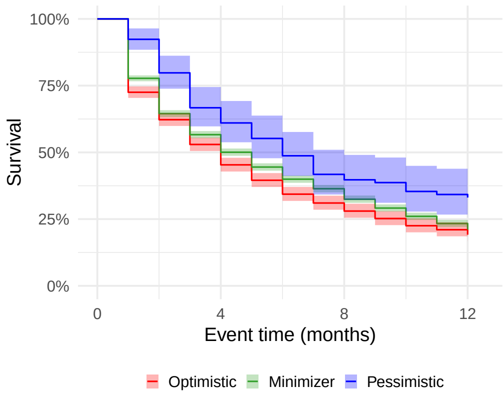
Figure C5: Survival by Categories for Different Definitions & Samples



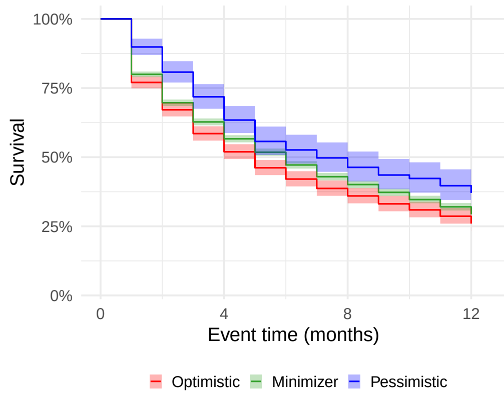

(a) Alternative categories: bill minimization



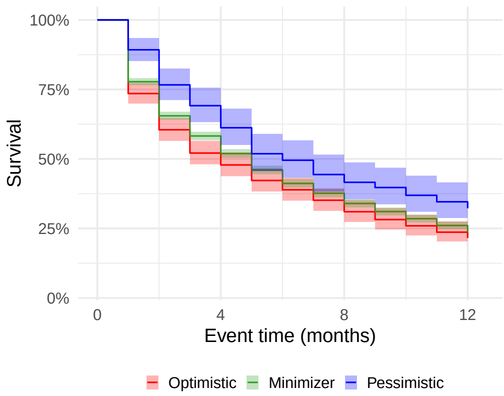

(b) Alternative categories: average hours



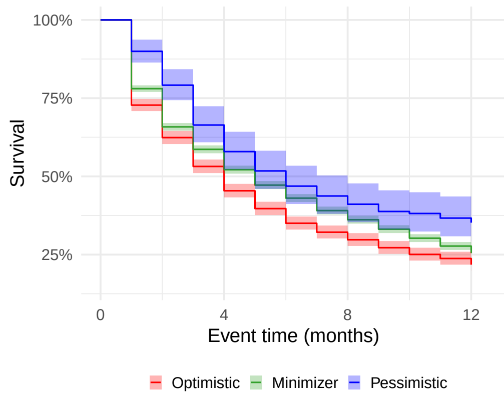

(c) Under 30s




(d) Over 30s



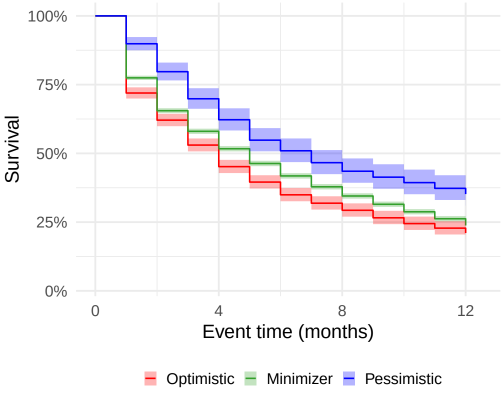

(e) Pre-04/01/202



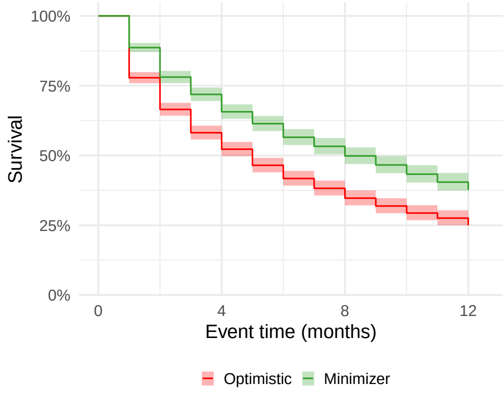

(f) 04/01/2020-06/01/2021

**Notes:** This figure plots Kaplan-Meier survival curves for different groups in the gig economy. The green, red, and blue lines denote the minimizers, optimistic, and pessimistic categories, respectively. Event time is tenure month in the gig economy (*i.e.*, $t = 1$ is workers' first month in the gig economy so if an individual does not have a second month in the gig economy, then they exit in the first period).

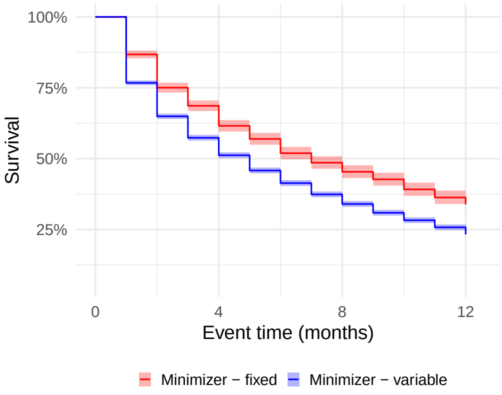
Figure C5: Survival by Categories for Different Definitions & Samples



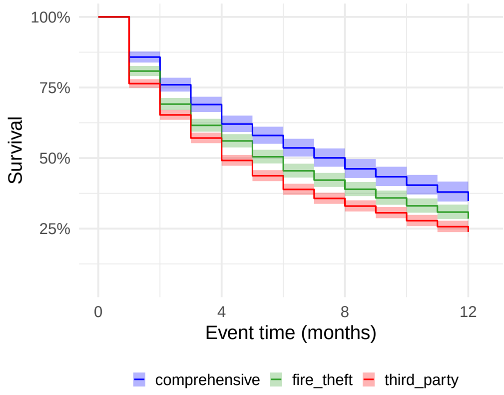

(g) Third party only



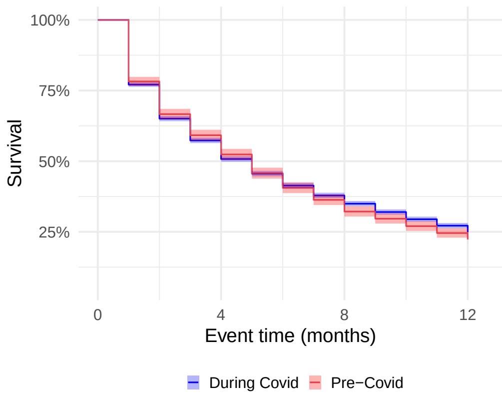

(h) Non-third party only




(i) By type of minimizer



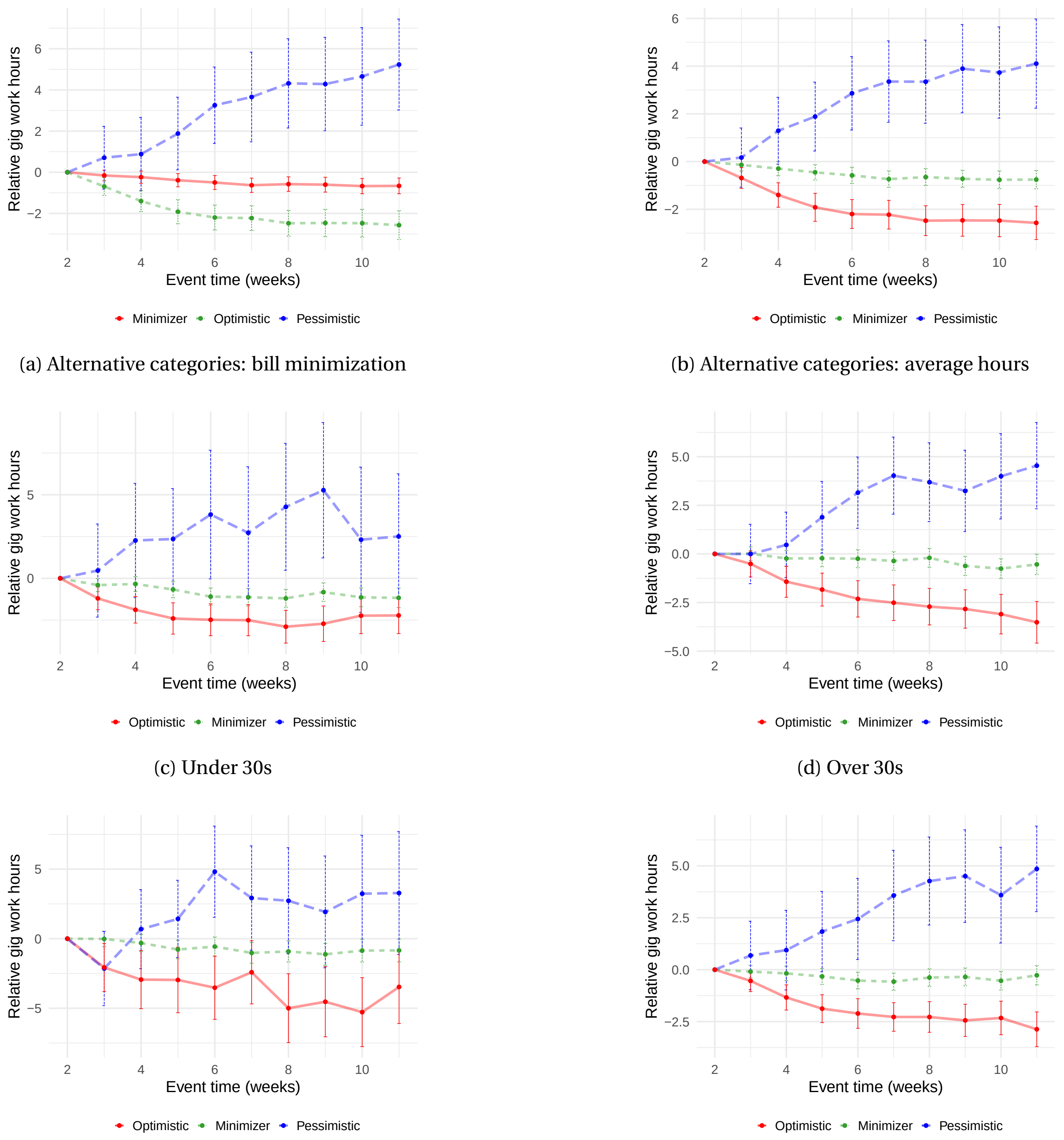

(j) By cover




(k) By time

**Notes:** This figure plots Kaplan-Meier survival curves for different groups in the gig economy. The green, red, and blue lines denote the minimizers, optimistic, and pessimistic categories, respectively. Event time is tenure month in the gig economy (*i.e.*, $t = 1$ is workers' first month in the gig economy so if an individual does not have a second month in the gig economy, then they exit in the first period).

Figure C6: Hours Worked Over Time by Categories for Different Definitions & Samples




(a) Alternative categories: bill minimization

(b) Alternative categories: average hours

(c) Under 30s

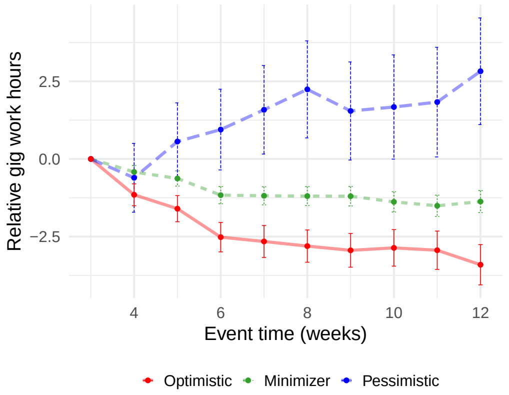
(d) Over 30s

(e) Pre-04/01/202

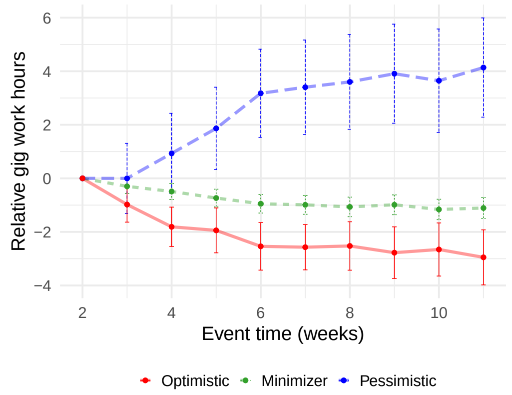
(f) 04/01/2020-06/01/2021

**Notes:** These figures plot three sets of coefficients from three separate regressions, which are run on a balanced panel (unless otherwise specified) of each category of worker. Weekly hours are regressed on fixed effects and event time dummies, where $t = 2$ corresponds to their second month in the gig economy (*i.e.*, event time is tenure in the gig economy), as well as calendar time controls. Intuitively, each coefficient represents the difference in hours at a given point in time from their hours in their first month in the gig economy. SEs are clustered at the worker level with a HC3 weighting scheme.

Figure C6: Hours Worked Over Time by Categories for Different Definitions & Samples




(g) Not balanced




(h) Third party only

**Notes:** These figures plot three sets of coefficients from three separate regressions, which are run on a balanced panel (unless otherwise specified) of each category of worker. Weekly hours are regressed on fixed effects and event time dummies, where $t = 2$ corresponds to their second month in the gig economy (*i.e.*, event time is tenure in the gig economy), as well as calendar time controls. Intuitively, each coefficient represents the difference in hours at a given point in time from their hours in their first month in the gig economy. SEs are clustered at the worker level with a HC3 weighting scheme.

Figure C10 shows the distribution of individual break-even points.

Table C4 presents summary statistics for the worker-level covariates from the quotes data.

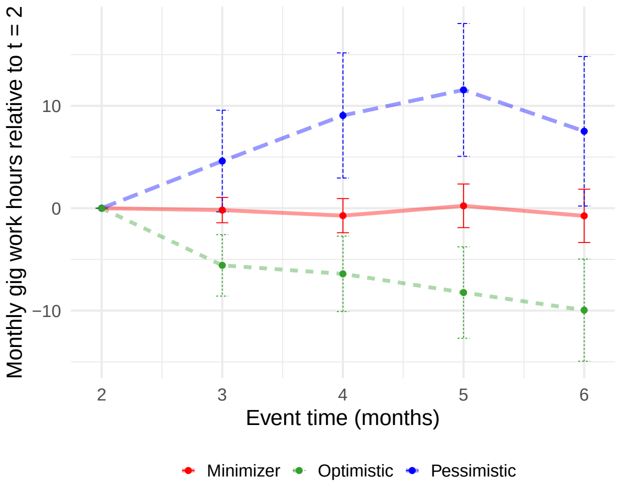
Table C4: Worker Covariate Summary Statistics

| Statistic | N | Mean | St. Dev. | Min | Pctl(25) | Median | Pctl(75) | Max |
|---|---|---|---|---|---|---|---|---|
| Age | 15,411 | 32.227 | 7.811 | 20 | 26 | 31 | 37 | 65 |
| Gender | 15,440 | 0.913 | 0.281 | 0 | 1 | 1 | 1 | 1 |
| Cover | 16,575 | 0.813 | 0.390 | 0 | 1 | 1 | 1 | 1 |
| Licence | 15,412 | 5.585 | 5.991 | 0 | 1 | 3 | 8 | 45 |

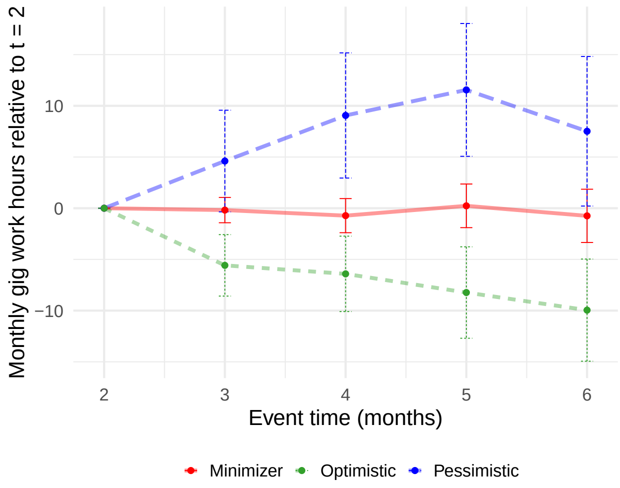
**Notes:** This figure shows summary statistics for worker-level covariates. The gender variable was received as binary and takes value one, if male, and zero otherwise. Cover is coded so that third party coverage takes value one, while fire and theft and comprehensive cover take value zero. The licence variable reports how long a worker has had their licence for.

Table C5 shows the average level of worker-level covariates within each category.

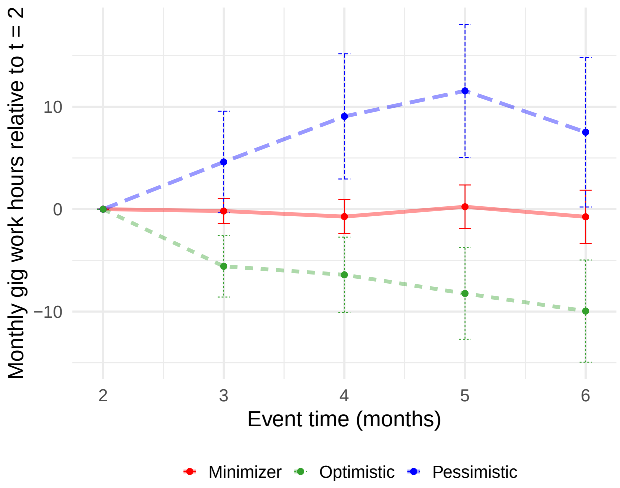
Figure C7: Hours Worked Over Time by Category at a Monthly Frequency




(a) By baseline categories




(b) By alternative categories: bill minimization




(c) By alternative categories: mean hours

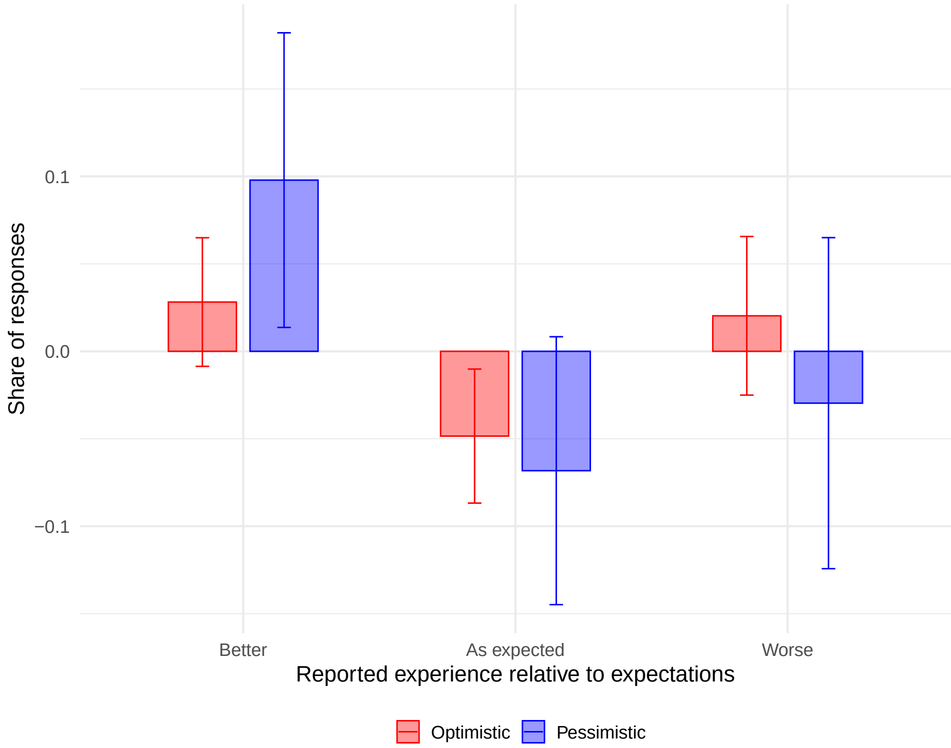
**Notes:** These figures plot three sets of coefficients from three separate regressions, which are run on a balanced panel of each category of worker. Monthly hours are regressed on fixed effects and event time dummies, where $t = 2$ corresponds to their second month in the gig economy (*i.e.*, event time is tenure in the gig economy), as well as calendar time controls. Intuitively, each coefficient represents the difference in hours at a given point in time from their hours in their first month in the gig economy. SEs are clustered at the worker level with a HC3 weighting scheme.

Table C5: Covariates by Categories

| Categories | Age | Licence | Cover | Gender |
|---|---|---|---|---|
| Minimizer | 33.00 | 5.95 | 0.87 | 0.93 |
| Optimistic | 32.12 | 5.32 | 0.50 | 0.93 |
| Pessimistic | 36.88 | 6.64 | 1 | 0.93 |

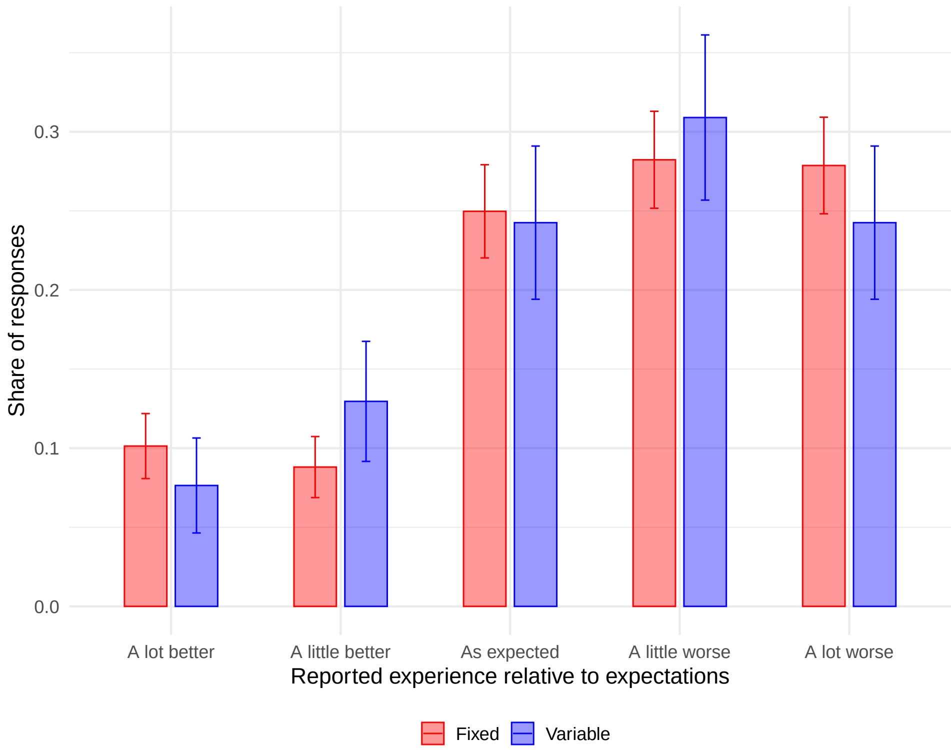
**Notes:** This table shows the average level of worker covariates in each category.

Figure C8: Survey Responses by Category Relative to Minimizers



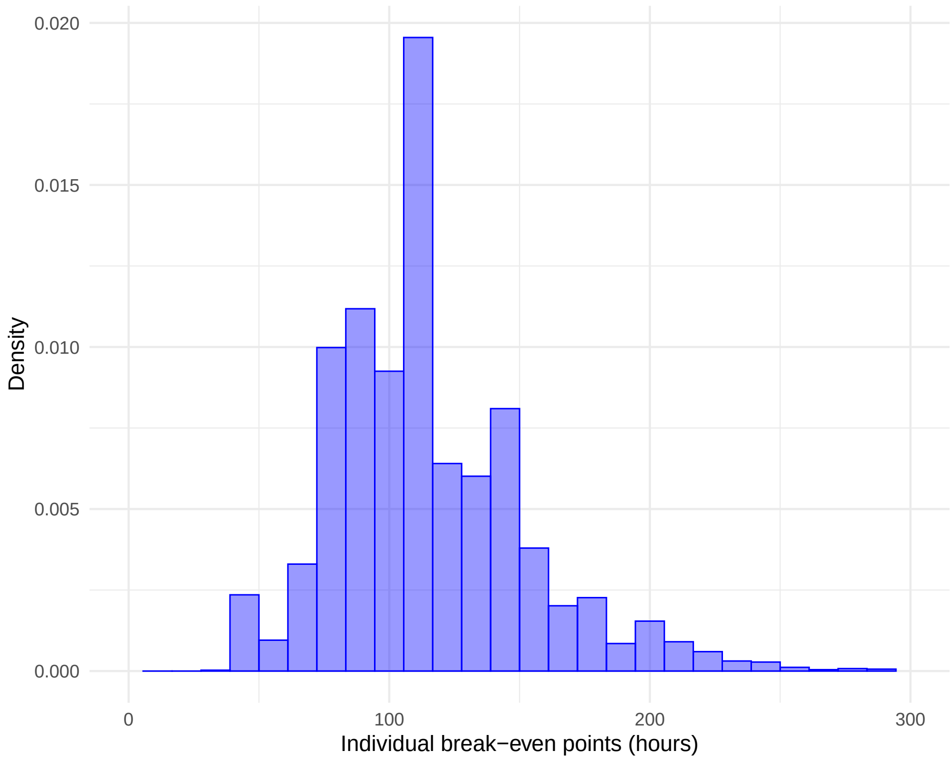

**Notes:** This figure shows the relative frequency of responses to an aggregate measure of workers' experiences in the gig economy.

Figure C9: Survey Responses by Category Relative to Minimizers




**Notes:** This figure shows the relative frequency of responses to an aggregate measure of workers' experiences in the gig economy.

Figure C10: Distribution of Break-Even Points




**Notes:** This figure shows the distribution of individual-level break-even points, which are constructed as a worker's quoted fixed premium divided by the variable premium.

# D Bayesian Learning

In this appendix, I show that the specified functional form for how gig worker's learn over time is derived from a micro-founded learning process, where individuals update their priors in a Bayesian fashion as they learn from normally distributed signals. As an example, I do this in the context of taxi drivers who learn about their wage rate.

## D.1 The Environment

Drivers receive a signal $\log(w_i) + \mu_{i,t}$ about their true wage rate $\log(w_i)$ after a weeks' driving is finished. So for week zero, drivers drive according to their prior $\log(w_{i,0})$. Note that this perceived wage rate also has an associated variance $\sigma_0^2$ which forms an exogenous, initial prior $\mathrm{N}(\log(w_{i,0}), \sigma_0^2)$ over the true wage rate $\log(w_i)$. Then, in week one, they update $\log(w_{i,0})$ to $\log(w_{i,1})$ using their prior, the signal, and the variance of the distribution from which the signal is drawn, where $\log(w_i) + \mu_{i,t} \sim \mathrm{N}\big(\log(w_i), \sigma_\mu^2\big)$.

Given homoskedastic variance across drivers' priors and signals, the perceived wage rate $\log(w_{i,t})$ and it's variance $\sigma_t^2$ acquires a convenient form

$$\log(w_{i,t}) = \frac{\sigma_\mu^2 + \sigma_0^2}{\sigma_\mu^2 + (t+1)\cdot\sigma_0^2}\cdot\log(w_{i,0}) + \frac{t\cdot\sigma_0^2}{\sigma_\mu^2 + (t+1)\cdot\sigma_0^2}\cdot\log(w_i) + \frac{\sigma_0^2}{\sigma_\mu^2 + (t+1)\cdot\sigma_0^2}\cdot\sum_{k=0}^{t}\mu_{i,t-k} \tag{18}$$

$$\sigma_t^2 = \frac{\sigma_0^2}{1 + t\cdot\frac{\sigma_0^2}{\sigma_\mu^2}}.$$

Note that the mean of beliefs is a homographic function in time, as is the learning process that I specify in section 4.

To see this, consider the variance of beliefs over time

$$\begin{aligned}
\sigma_1^2 &= \frac{\sigma_\mu^2\cdot\sigma_0^2}{\sigma_\mu^2 + \sigma_0^2}\\
\sigma_2^2 &= \frac{\sigma_\mu^2\cdot\sigma_1^2}{\sigma_\mu^2 + \sigma_1^2}\\
&= \frac{\sigma_\mu^2\cdot\big(\frac{\sigma_\mu^2\cdot\sigma_0^2}{\sigma_\mu^2+\sigma_0^2}\big)}{\sigma_\mu^2 + \big(\frac{\sigma_\mu^2\cdot\sigma_0^2}{\sigma_\mu^2+\sigma_0^2}\big)}\\
&= \frac{\sigma_0^2}{1 + 2\cdot\frac{\sigma_0^2}{\sigma_\mu^2}}\\
\sigma_3^2 &= \frac{\sigma_\mu^2\cdot\left(\frac{\sigma_0^2}{1+2\cdot\frac{\sigma_0^2}{\sigma_\mu^2}}\right)}{\sigma_\mu^2 + \left(\frac{\sigma_0^2}{1+2\cdot\frac{\sigma_0^2}{\sigma_\mu^2}}\right)}\\
&= \frac{\sigma_0^2}{1 + 3\cdot\frac{\sigma_0^2}{\sigma_\mu^2}}\\
&\ \vdots\\
\implies \sigma_t^2 &= \frac{\sigma_0^2}{1 + t\cdot\frac{\sigma_0^2}{\sigma_\mu^2}}.
\end{aligned}$$

The variance of beliefs are used to weight signals, and so the mean of beliefs over time look like

$$\begin{aligned}
\log(w_{i,t}) &= \frac{\sigma_\mu^2 + t\cdot\sigma_0^2}{\sigma_\mu^2 + (t+1)\cdot\sigma_0^2}\cdot\log(w_{i,t-1}) + \frac{\sigma_0^2}{\sigma_\mu^2 + (t+1)\cdot\sigma_0^2}\cdot\big(\log(w_i) + \mu_{i,t}\big) \\
&= \frac{\sigma_\mu^2 + t\cdot\sigma_0^2}{\sigma_\mu^2 + (t+1)\cdot\sigma_0^2}\cdot\left(\frac{\sigma_\mu^2 + (t-1)\cdot\sigma_0^2}{\sigma_\mu^2 + t\cdot\sigma_0^2}\cdot\log(w_{i,t-2}) + \frac{\sigma_0^2}{\sigma_\mu^2 + t\cdot\sigma_0^2}\cdot\big(\log(w_i) + \mu_{i,t-1}\big)\right)\dots \\
&\dots + \frac{\sigma_0^2}{\sigma_\mu^2 + (t+1)\cdot\sigma_0^2}\cdot\big(\log(w_i) + \mu_{i,t}\big) \\
&\vdots \\
&= \frac{\sigma_\mu^2 + \sigma_0^2}{\sigma_\mu^2 + (t+1)\cdot\sigma_0^2}\cdot\log(w_{i,0}) + \frac{t\cdot\sigma_0^2}{\sigma_\mu^2 + (t+1)\cdot\sigma_0^2}\cdot\log(w_i) + \frac{\sigma_0^2}{\sigma_\mu^2 + (t+1)\cdot\sigma_0^2}\cdot\sum_{k=0}^{t}\mu_{i,t-k}.
\end{aligned}$$

## D.2 Implications

Taking expectations of equation (18) with respect to signals yields

$$\mathbb{E}\left[\log(w_{i,t})\right] = \frac{\sigma_\mu^2 + \sigma_0^2}{\sigma_\mu^2 + (t+1)\cdot\sigma_0^2}\cdot\log(w_{i,0}) + \frac{t\cdot\sigma_0^2}{\sigma_\mu^2 + (t+1)\cdot\sigma_0^2}\cdot\log(w_i).$$

Note that this can be rewritten as

$$\mathbb{E}\left[\log(w_{i,t})\right] = \frac{\lambda}{\lambda + t}\cdot\log(w_{i,0}) + \frac{t}{\lambda + t}\cdot\log(w_i).$$

where $\lambda = \frac{\sigma_\mu^2 + \sigma_0^2}{\sigma_0^2}$. Thus, the speed of learning parameter $\lambda$ reflects how noisy the signal is relative to initial aggregate uncertainty. Further, this analysis reveals that $\lambda$ should be bounded from below by one. In practice, I find that $\lambda$ is estimated to be close to one, which implies that the signals gig workers receive are precise relative to the variance in initial misperceptions.

Intuitively, the way that learning is modelled in equation (2) is similar to estimating a full model of Bayesian learning, where the agents' signals are simulated. This is because, if a sufficient number of agents were simulated, then averaging over agents with the same initial misperception would lead to (2). However, the model does deviate from Bayesian learning because agents perceive that they have an infinitely precise signal the value of gig work, rather than a noisy posterior, which may affect their behavior.

# E Model Derivations and Extensions

This section provides some additional derivations for analysis conducted in the paper, and for extensions of the model presented in appendix F.

## E.1 Proposition 22 Hours Thresholds

The thresholds for workers' labor supply rule with hours-qualified benefits are given by

$$\rho_j = \frac{j^{1/\varepsilon} \cdot (p(\omega)+\kappa)}{\hat{\theta}_{i,t}} \text{ for } j = 15, 25,$$

$$\bar{\bar{\rho}}^{\varepsilon} \cdot \left( \frac{1}{\varepsilon - 1} \cdot \frac{\hat{\theta}^{\varepsilon}_{i,t}}{(p(\omega)+\kappa)^{\varepsilon-1}} \right) - \bar{\bar{\rho}} \cdot \left( \hat{\theta}_{i,t} \cdot \frac{25^{1-1/\varepsilon}}{1-1/\varepsilon} \right) + 200 - 400 + (p(\omega)+\kappa) \cdot 25 = 0,$$

$$\bar{\rho}^{\varepsilon} \cdot \left( \frac{1}{\varepsilon - 1} \cdot \frac{\hat{\theta}^{\varepsilon}_{i,t}}{(p(\omega)+\kappa)^{\varepsilon-1}} \right) - \bar{\rho} \cdot \left( \hat{\theta}_{i,t} \cdot \frac{15^{1-1/\varepsilon}}{1-1/\varepsilon} \right) - 200 + \big(p(\omega)+\kappa\big) \cdot 15 = 0.$$

And to find $\underline{\rho}$

$$\begin{aligned}
(1-\psi)\cdot\nu \le u(15,\omega;\rho_{15}) &\implies (1-\psi)\cdot\nu = u(h^*(\omega),\omega;\underline{\rho}) \\
u(15,\omega;\rho_{15}) < (1-\psi)\cdot\nu \le u(15,\omega;\rho_{15}) + 200 &\implies \underline{\rho} = \rho_{15} \\
u(15,\omega;\rho_{15}) + 200 < (1-\psi)\cdot\nu \le u(25,\omega;\rho_{25}) + 200 &\implies (1-\psi)\cdot\nu = u(h^*(\omega),\omega;\underline{\rho}) + 200 \\
u(25,\omega;\rho_{25}) + 200 < (1-\psi)\cdot\nu \le u(25,\omega;\rho_{25}) + 400 &\implies \underline{\rho} = \rho_{25} \\
u(25,\omega;\rho_{25}) + 400 < (1-\psi)\cdot\nu &\implies (1-\psi)\cdot\nu = u(h^*(\omega),\omega;\underline{\rho}) + 400.
\end{aligned}$$

## E.2 Switching

To introduce endogenous switching to the model, I allow workers to switch policy, as well as exiting, when their perceived valuation changes. The endogenous exit rule is now altered such that a gig worker exits if and only if

$$\tilde{V}_i\big(\omega^*_i;\hat{\theta}_{i,t}\big) \le \frac{\nu_i}{1-\beta} \text{ and } \tilde{V}_i\big(\Omega/\omega^*_i;\hat{\theta}_{i,t}\big) - \tau \le \frac{\nu_i}{1-\beta}, \tag{19}$$

where $\tau$ is a new parameter: the hassle cost of switching policy. Therefore, workers will switch policy if and only if

$$\tilde{V}_i\big(\Omega/\omega^*_i;\hat{\theta}_{i,t}\big) - \tau > \tilde{V}_i\big(\omega^*_i;\hat{\theta}_{i,t}\big) \text{ and } \tilde{V}_i\big(\Omega/\omega^*_i;\hat{\theta}_{i,t}\big) - \tau > \frac{\nu_i}{1-\beta}. \tag{20}$$

If neither of these conditions are satisfied, then policyholders will remain on their current policy.

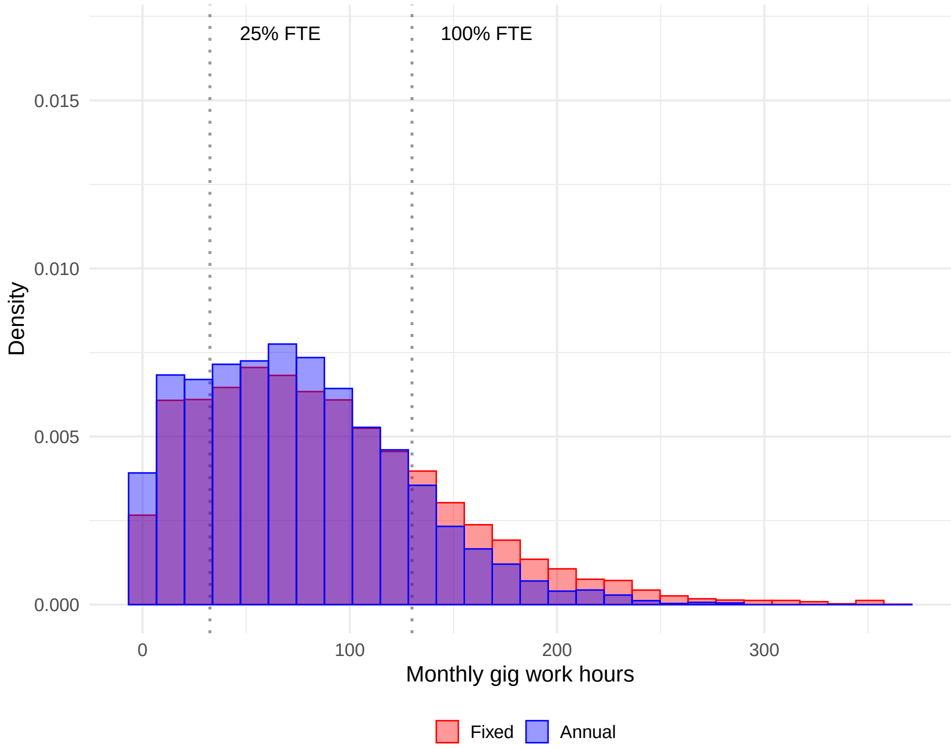
## E.3 Cholesky Decomposition of Variance-Covariance Matrix

$$\begin{pmatrix} \sigma_\theta^2 & \sigma_{\theta,\phi} & \sigma_{\theta,\nu} \\ \sigma_{\phi,\theta} & \sigma_\phi^2 & \sigma_{\phi,\nu} \\ \sigma_{\nu,\theta} & \sigma_{\nu,\phi} & \sigma_\nu^2 \end{pmatrix} = \begin{pmatrix} l_{1,1} & 0 & 0 \\ l_{2,1} & l_{2,2} & 0 \\ l_{3,1} & l_{3,2} & l_{3,3} \end{pmatrix} \times \begin{pmatrix} l_{1,1} & l_{2,1} & l_{3,1} \\ 0 & l_{2,2} & l_{3,2} \\ 0 & 0 & l_{3,3} \end{pmatrix}$$

$$\begin{aligned} \sigma_\theta^2 =& l_{1,1}^2 \\ \sigma_\phi^2 =& l_{2,1}^2 + l_{2,2}^2 \\ \sigma_\nu^2 =& l_{3,1}^2 + l_{3,2}^2 + l_{3,3}^2 \\ \sigma_{\theta,\phi} =& l_{1,1} \cdot l_{2,1} \\ \sigma_{\theta,\nu} =& l_{1,1} \cdot l_{3,1} \\ \sigma_{\phi,\nu} =& l_{2,1} \cdot l_{3,1} + l_{2,2} \cdot l_{3,2}, \end{aligned}$$

where $l_{2,1}$ and $l_{3,2}$ is set equal to zero in order to ensure $\sigma_{\phi,\theta} = 0$ and $\sigma_{\phi,\nu} = 0$.

Figure F11: Hours Worked Comparison with Annual Policyholders




**Notes:** This figure plots the histogram of monthly hours worked for fixed customers in blue, and the analogue for annual customers in red.

# F Robustness Checks: Structural Model Estimates (*In Progress*)

Figure F11 shows that, for the small minority of gig workers who opt for the annual policy, their monthly intensive margin labor supply is roughly the same as fixed policy workers.

## F.1 Variance in $\kappa$

In this subsection, I re-estimate the model allowing for independent variation in the linear cost parameter $\kappa$. Tablees F6, F7, and F8 show the resulting parameter estimates, population characteristics, and model fit. The model finds allowing for a standard deviation of £1.79 helps fit the empirical moments.

Table F6: Parameter Estimates

| $\zeta_{1:6}$ | $\hat{\zeta}_{1:6}$ | $\zeta_{6:12}$ | $\hat{\zeta}_{6:12}$ |
|---|---|---|---|
| $\mu_\theta$ | 3.69 | $\mu_\nu$ | 12.66 |
| | ( 0.05) | | ( 0.33) |
| $\sigma^2_\theta$ | 0.31 | $\sigma^2_\nu$ | 4.53 |
| | ( 0.02) | | ( 0.49) |
| $\sigma_{\theta,\nu}$ | -1.09 | $\sigma^2_\phi$ | 0.14 |
| | ( 0.05) | | ( 0.01) |
| $\lambda$ | 1.00 | $\kappa$ | – |
| | ( 0.09) | | (–) |
| $\sigma^2_\rho$ | 0.03 | $\psi$ | 0.27 |
| | ( 0.00) | | ( 0.00) |
| $\mu_\kappa$ | 3.67 | $\sigma^2_\kappa$ | 0.00 |
| | ( 0.02) | | ( 0.00) |
| $\mu_P$ | 101.82 | $\sigma_P$ | 24.08 |
| $p$ | 0.94 | $\beta$ | $0.95^{1/12}$ |
| $\eta$ | 0.07 | $\varepsilon$ | $\Delta \hat{\log}(h) \cdot \log(1 + p/\hat{\kappa})$ |

**Notes:** The top panel of this table presents estimates of the structural parameters from the model. Standard errors are contained in the parentheses, which are estimated as described in subsection 5.1. The second panel shows the variable policy premium and the mean of the quoted fixed premium distribution used in the estimation, which correspond to the empirical averages of these parameters, as well as the set discount factor.

Table F7: Simulated Population Characteristics

| Statistic | Population | Participants |
|---|---|---|
| Mean valuation $\theta$ | 46.50 | 210.86 |
| SD valuation $\theta$ | 27.93 | 48.65 |
| Mean misperception $\phi$ | 1.00 | 1.04 |
| SD misperception $\phi$ | 0.39 | 0.11 |
| Mean outside option $\nu$ | 9,745 | 672 |
| SD outside option $\nu$ | 1,255 | 487 |
| Mean linear cost $\kappa$ | 39 | 39 |
| SD linear cost $\kappa$ | 2 | 2 |
| Correlation $\rho_{\theta,\nu}$ | -0.62 | -0.31 |
| Correlation $\rho_{\theta,\phi}$ | 0.00 | -0.27 |
| Correlation $\rho_{\nu,\phi}$ | 0.00 | 0.52 |
| Correlation $\rho_{\theta,\kappa}$ | 0.00 | -0.01 |
| Correlation $\rho_{\kappa,\phi}$ | 0.00 | 0.01 |
| Correlation $\rho_{\nu,\kappa}$ | 0.00 | 0.04 |

**Notes:** This table presents statistics that describe the simulated population. The first column shows these statistics for the entire population, while the second column conditions on participating in the gig economy.

Table F8: Model Fit

| Moment | Data | Model |
|---|---|---|
| Labor market share (%) | 3.6 | 3.7 |
| Variable policy share (%) | 66.8 | 69.0 |
| Mean hours per month, variable policy | 53.4 | 57.7 |
| SD hours per month, variable policy | 48.4 | 37.5 |
| Mean hours per month, fixed policy | 95.3 | 105.5 |
| SD hours per month, fixed policy | 68.6 | 76.1 |
| Share non-cost-minimizing (%), variable policy | 11.6 | 9.2 |
| Share non-cost-minimizing (%), fixed policy | 60.6 | 61.7 |
| Mean hours per month from cost-minimizing, variable policy | 40.1 | 41.2 |
| SD hours per month from cost-minimizing, variable policy | 36.6 | 48.1 |
| Mean hours per month from cost-minimizing, fixed policy | 55.6 | 40.3 |
| SD hours per month from cost-minimizing, fixed policy | 41.7 | 30.8 |
| Hazard rate for cost-minimizers (%), variable policy | 28.9 | 23.5 |
| Hazard rate for cost-minimizers (%), fixed policy | 20.7 | 17.4 |
| Hazard rate for non-cost-minimizers (%), fixed policy | 29.2 | 38.0 |
| Decline in hours per month for non-cost-minimizers, variable policy | -10.2 | -9.3 |
| Decline in hours per month for non-cost-minimizers, fixed policy | 2.6 | 2.7 |
| Mean quoted fixed premium (£), variable policy | 97.0 | 95.3 |
| Mean quoted fixed premium (£), fixed policy | 106.5 | 104.4 |
| SD quoted fixed premium (£), variable policy | 25.1 | 20.1 |
| SD quoted fixed premium (£), fixed policy | 24.3 | 26.6 |
| Share of zero hours months (%) | 5.5 | 5.8 |
| Mean within worker SD hours per month | 31.0 | 37.6 |

**Notes:** This table presents the targeted empirical moments alongside their model implied counterparts. The first column contains the empirical moments and the second column contains the model analogue.

## Appendix References

Augenblick, N. and Rabin, M. (2019). An experiment on time preference and misprediction in unpleasant tasks. *Review of Economic Studies*, 86(3):941–975.

Handel, B. R. and Kolstad, J. T. (2015). Health insurance for" humans": Information frictions, plan choice, and consumer welfare. *American Economic Review*, 105(8):2449–2500.

Lambrecht, A. and Skiera, B. (2006). Paying too much and being happy about it: Existence, causes, and consequences of tariff-choice biases. *Journal of marketing Research*, 43(2):212–223.

Lockwood, B. B. (2020). Optimal income taxation with present bias. *American Economic Journal: Economic Policy*, 12(4):298–327.