

Workers in the crowd: the labor market impact of the online platform economy

Michele Cantarella ^{1,*} and Chiara Strozzi²

¹University of Helsinki and University of Modena and Reggio Emilia, Modena, Italy.

e-mail: michele.cantarella@helsinki.fi and ²University of Modena and Reggio Emilia and IZA, Modena, Italy. e-mail: chiara.strozzi@unimore.it

*Main author for correspondence.

Abstract

We compare individuals engaged in online crowdwork against workers in traditional occupations from the United States and Europe, investigating the determinants of access into crowd employment and the nature of the deterioration of salary conditions and job quality characterizing these markets. We do so by matching responses from comparable working conditions surveys, and controlling for most individual and socio-economic characteristics affecting pay. We find that most differences in earnings are largely unexplained by the ability of individuals, and that labor in crowdwork is vastly under-utilized.

JEL classification: J31, J42, F66

1. Introduction

Among the “mega-trends” which characterize the future of work, the growth of the online platform economy—the economic activity which enables “innovative forms of production, consumption, collaboration and sharing through digital interactions” (OECD, 2018)—has been steady and fast in recent years and has been contributing to the changing nature of work (Harris and Krueger, 2015; OECD, 2016). Technological progress and digitalization are the basis of its development. Due to the overall exponential growth of internet facilities, recent years have seen the rise of what is known as the gig economy, intended as the “collection of markets that match providers to consumers on a gig (or job) basis in support of on-demand commerce” (Donovan *et al.*, 2016).¹

A result of the growth of the online platform economy is the rise of new platform-mediated forms of work which can be distinguished in remote and in-person mediated service provision (Kenney *et al.*, 2019; Kenney and Zysman, 2019). While the former enables offshorable services to be provided globally, the latter includes typical work-on-demand platforms (such as Uber, Foodora, or Taskrabbit) offering services which are characterized by their local and tangible nature.

Crowdwork (or crowd employment) is one of these remote platform-mediated forms of work and is identified as an “employment form that uses an online platform to enable organizations or individuals to access an indefinite and

1 Katz and Krueger (2018), estimate that 0.5% of all occupations in United States in 2015 involved the direct selling of activities and services mediated by online intermediaries—a figure that proxies the size of the gig-economy. While Farrell and Greig (2017) have supported the claim that these markets have, overall, reached their peak in 2016 (Collins *et al.*, 2019) estimate that gig-economy jobs have grown by 1.9% from 2000 to 2016 in the United States.

unknown group of other organizations or individuals to solve specific problems or to provide specific services or products in exchange for payment” (Eurofound, 2015). Crowdwork transcends traditional work arrangements by *de facto* requiring a tripartite relationship in which an online intermediary agent (the platform) matches independent contractors with clients (either individuals or companies) who offer online tasks through an open call, setting pay levels and rating the contractor’s performance afterwards. Crowd employment can then be identified as a phenomenon that entails a new, and cheaper, way of outsourcing services by breaking them down into smaller tasks which are distributed to a large anonymous pool of workers through IT-based platforms (Prassl and Risak, 2015). Because of this, it has also been defined as “crowdsourcing” (cf. J. Howe, “The Rise of Crowdsourcing”, Wired Magazine, 14, 2006).²

Crowdwork arrangements, and the skill requirement for outsourced tasks, may vary greatly. However, while professionals who are contracted for their services over the Internet are known as online freelancers,³ workers engaged in microtasks are more properly identified as crowdworkers, in line with the evidence that they constitute the bulk of online platform workers (Berg *et al.*, 2018). The microtasks performed by crowdworkers generally include a series of simple and repetitive tasks which together comprise a large unified project and that can be performed independently over the Internet in a short period of time (e.g. image categorization, translation of a sentence, writing product reviews, tagging). Amazon Mechanical Turk (AMT) stands as a prime example of a microtask platform, being widely recognized as one of the most popular ones (Harris and Krueger, 2015). Clickworker and Crowdflower are two additional examples.

The economic conditions of crowdworkers have been analyzed in a number of recent studies, leading to the conclusion that these workers suffer from the erosion of fundamental labor rights, the loss of social protections and difficulties in exercising collective action (De Stefano, 2016). The available evidence is based on analysis of information acquired either through online questionnaires (e.g. Berg, 2015; Huws *et al.*, 2017; Berg *et al.*, 2018; Difallah *et al.*, 2018; Pesole *et al.*, 2018; Rani and Furrer, 2019), panel surveys (e.g. CIPD, 2017; Lepanjuuri *et al.*, 2018) or web plug-ins (e.g. Hara *et al.*, 2018).⁴

Until recently, the body of research on the economics of crowdsourcing has been remarkably thin. As suggested by Hara *et al.* (2018), this scarcity of literature could be due to the absence of publicly available data on crowdwork platforms and their workers. Nonetheless, as discussed by Horton *et al.* (2011), Paolacci *et al.* (2010), and Berinsky *et al.* (2012), crowdwork platforms present themselves as an ideal environment for empirical studies, particularly those based on experimental research.

Horton and Chilton (2010) studied reservation wages in crowd employment, finding evidence of “target earning” behavior. The relationship between unemployment and micro-task labor markets was further explored by Borchert *et al.* (2018) and Jackson (2019), who found that labor demand shocks affect temporary participation in online labor markets. Dube *et al.* (2020) address monopsony in online labor markets, finding that their peculiar structure allows platforms to impose a considerable mark-up on workers’ productivity. Others have suggested that crowd employment could be equated to a form of digital migration (Degryse, 2016).

All these studies confirm that crowd employment is characterized by poor working conditions. However, based on available evidence, it would be a mistake to assume that crowdworkers’ working conditions are to be fully attributed to a platform effect, as it could be argued that the characteristics of crowdworkers are intrinsically different from the characteristics of workers engaged in traditional professions outside online platforms. Assessing how much ability plays a part in differentiating these markets will allow regulators to understand whether such online service marketplaces should be treated as traditional jobs or not.

In light of this question, in this article, we analyze the earnings and working conditions of micro-tasks crowdworkers from the supply side by comparing them with traditional workers with similar skills, focusing on both the United States and Europe, analyzing cross-sectional data collected from three surveys on labor market conditions.

We match workers by controlling for most observable characteristics affecting the hourly rate of pay and other elements of job quality. This allows us to capture most observable determinants of skills, and also to identify which

- 2 While crowdwork enables offshorable services to be provided globally, typical work-on-demand platforms (such as Uber, Foodora, or Taskrabbit) offer local services which are characterized by their physical and tangible nature.
- 3 Their tasks are characterized by high abstract content and low routine intensity (e.g. business analysis, management consulting, editing/post production). Upwork is a typical marketplace of online freelancers.
- 4 For an overview of recent surveys on the online platform economy, see Peetz (2019).

factors affect access to crowdwork markets. We later test for the presence of self-selection through unobservable characteristics using an instrumental variable approach, dealing with potential endogeneity issues.

Our findings indicate that earnings in crowdwork are mostly indifferent to skill and that crowdworkers earn less than half (on an hourly basis) than traditional workers with comparable ability, while working only a few hours less per week. In addition, the human capital on these platforms appears highly under-utilized, as crowdworkers are more likely to prefer to work more than they currently do.

The rest of the article is organized as follows. Section 2 introduces our research question and frames the policy relevance of our research. Section 3 describes the data used in our analysis, while the subsequent two sections present our results: Section 4 studies self-selection in crowdwork and its effect on earnings by controlling for all observable determinants of skill, while Section 5 offers further controls for endogenous self-selection into online crowd employment. Sections 6 and 7 offer our policy considerations and conclusions. The Appendix is dedicated to a theoretical discussion of labor supply in crowdwork ([Appendix A](#)), a discussion on the instrumental variable approach ([Appendix B](#)), and an analysis of the role of institutions on the development of crowdwork markets ([Appendix C](#)). Additional descriptive statistics and results, as well as robustness checks are offered in the [Supplementary Appendix](#).

2. Crowdwork, earnings, and skill-biased self-selection

At the time of the writing, online micro-task labor markets such as AMT fill a relatively small niche of the labor markets of the United States and Europe. However, their organization mirrors the one of the other online platforms mediating work (such as more skill-intensive platforms like Upwork) and, more generally, of the platform economy (as discussed by [Kenney and Zysman, 2016](#)). It is then important to assess the extent to which crowdwork markets, and platform-mediated work in general, should be treated as though they are like other, more traditional, forms of work.

There are several reasons why this issue is relevant. Indeed, the almost laissez-faire stance that many policy makers have adopted toward these markets has raised concerns not only about the quality and fairness of labor condition in online markets, but has also led to the worry that the growth of these markets would open up opportunities for social dumping vis-à-vis traditional labor markets, as more and more offshorable tasks (using the definition of [Autor and Dorn, 2013](#)) end up being performed in online platforms ([Bogliacino et al., 2020](#)).

To better assess these issues, a thorough discussion on the nature of crowd employment is needed. Such a discussion should be informed by how much differences in working conditions between crowd employment and traditional work are related to the characteristics of platform work, and how much do they relate to individual productivity instead.

Many of these platform-specific factors have already been highlighted in the literature. The large power platform owners can wield over their ecosystems ([Kenney and Zysman, 2019](#)) suggests how informational asymmetries might be ingrained in platform-mediated labor markets ([Drahokoupil and Piasna, 2017](#)). Others have suggested that deterioration of salaries in online labor markets might be related to excess supply linked to unemployment shocks ([Borchert et al., 2018](#)), or higher competition from foreign workers ([Degryse, 2016](#)).

More importantly, evidence of monopsony power in crowdwork ([Dube et al., 2020](#)) suggest heterogeneity in labor supply curves between crowdwork and traditional employment, the former being characterized by low elasticity. Similarly, this heterogeneity explains the target earning behavior studied in [Horton and Chilton \(2010\)](#).

It can be inferred from these studies that this heterogeneity might derive from endogenous characteristics of platform workers related to their individual productivity. For this reason, we intend to test the two following hypotheses:

*H*₀: Participation in crowdwork is prevalent among individuals who are unable to enter the traditional labor market.

*H*₁: The salary penalty in crowdwork is also determined by the lower productivity of individuals participating in this market.

Both hypotheses are strictly tied together. Nonetheless, the way these hypotheses stand up to our testing will have deep policy implications for crowd employment. Regulation of these online labor market platforms depends on the how much these conditions are tied to factors of labor supply.

Understanding how much individual ability influences salaries can allow for a better comparison of forms of work. Answering these research questions would allow policy-makers to understand how comparable the online

workforce is to the traditional workforce and to assess how much the wage rate per unit of labor differs across forms of work, regardless of individual productivity.

Which factors contribute to self-selection is also important. To workers unable to access traditional markets, crowd employment might stand as the only source of income, no matter the salary penalty, and offer them the opportunity to re-enter employment at a lower rate of pay. At the same time, workers in traditional employment might want to seek auxiliary earnings through crowdwork, and the lower earnings might already be determined by endogenous factors relating to experience and ability that already affect their potential earnings. In these instances, regulation of platforms should focus only on curbing the monopsonistic tendencies of these platforms, leaving most concerns on social dumping aside.

Also, especially if participation in platform markets is tied to other external processes, such as unemployment shocks (Borchert *et al.*, 2018), or job polarization (Autor and Dorn, 2013), it is clear that regulatory action should focus on the idiosyncrasies of traditional labor markets rather than digital platforms. Different countries will also have different labor market institutions, and it is worth investigating if online labor platforms develop faster in contexts where employment protections are lower.

In general, understanding the role of productivity-biased self-selection into online labor markets can help bridge a gap within the aforementioned literature: while we know that these platforms can foster monopsony and are characterized by low elasticity of labor supply (Horton and Chilton, 2010; Borchert *et al.*, 2018; Bogliacino *et al.*, 2020; Dube *et al.*, 2020), how individual productivity is tied to online labor markets is still unclear. Dube *et al.* (2020) have suggested that, along with structural factors such as the lack of bargaining power, the presence of informational asymmetries and opportunities for wage discrimination, the monopsony power of online labor platforms might also arise from the fact that the real wage rate of a task posted on the platform is often unknown to crowdworkers. As the real wage rate includes the time spent performing a task, however, only separating the self-selection effects from the platform effects can allow for a proper comparison with traditional forms of work. By identifying these effects, regulators can understand how the need to protect traditional labor markets from task-offshoring can be balanced with the new opportunities for entering employment that platform markets offer to workers unable or unwilling to enter the traditional workforce.

3. Data

The identification of crowdworkers in existing general working conditions surveys is not trivial. The European Working Conditions Survey (EWCS) (Eurofound, 2017) and the American Working Conditions Survey (AWCS) (Maestas *et al.*, 2017) both contain comparable information on wages, job quality and skills but, in both instances it is often not possible to disentangle platform workers from any other freelancer working from home (186 observations in total). As micro-task crowdsourcers tend to perform specific, routine intensive activities, we expect that equating them to freelancers working from home will likely present a serious source of bias. Also, due to the current size of the platform economy, platform workers, even if correctly identified, are necessarily under-represented in general surveys.

Dedicated surveys on crowdworkers can be used to bridge this gap. To help achieve the aim of this paper, we construct a sample composed of both crowdworkers and traditional workers by extracting responses from comparable surveys.

3.1 Crowdworkers vs. traditional workers

Our sample of crowdworkers includes information on European and US crowdworkers from two rounds of the ILO Survey on Crowdworkers (Berg, 2015; Berg *et al.*, 2018) conducted on major online micro-task platforms⁵ in 2015 and 2017, including crowdworkers from the United States, the Europe, and other countries. The two rounds of the survey provides cross-sectional data on earnings, demographics and working quality indicators for 1167 (2015) and

5 The platforms are: Amazon Mechanical Turk (US, EU), Crowdfunder (EU), Clickworker (EU), Microworkers (EU), and Prolific Academic (EU).

2350 (2017) crowdworkers from around the world. Using information from both rounds of the survey, we extracted a group of 1393 US crowdworkers and 1000 European crowdworkers.⁶

We used information from the AWCS and the EWCS to construct a baseline group of traditional workers. The AWCS comprises 1946 working age individuals from the United States out of a total 3109 respondents, while a sample of 32,429 employed working-age individuals from the EU28 area was extracted from the EWCS. Both surveys include post-stratification weights.

The contents of the ILO Survey have been harmonized with data from the 2015 rounds of the AWCS and the EWCS into a single cross-section, providing a shared set of common variables and adjusting earnings for inflation and purchasing power parity.

While post-stratification weights in the AWCS and EWCS ensure survey representativeness for the traditional work group, concerns might arise with the crowdworker subsample. The procedures used to collect this information by the ILO attempt to minimize the incidence of selection bias by posting the survey as a paid task on the platform at random intervals during the fieldwork period. However, we should recognize that selection bias in the sample might still be present. Crowdworkers who are less motivated to respond to a survey, or who prefer tasks with a higher rate of pay, might be underrepresented in the sample.

It should be noted that [Horton and Chilton \(2010\)](#) and [Berinsky *et al.* \(2012\)](#) have argued that the external validity of estimates based on experiments conducted on the crowdwork markets might be limited, but in the literature no concerns have been raised so far regarding self-selection within the online platform on which the experiment was conducted. However, if sample selection issues persist, using instrumental variable estimation techniques, along with being able to control for sample selection through an adequate set of observables, can reasonably help mitigate these issues.

Weighted mean comparison *t*-tests for a number of key dimensions across the crowdwork and traditional work groups, are shown in [Table 1](#) (United States: $n = 3339$ and Europe: $n = 33,281$), revealing differences in earnings, age, education and marital status across forms of work.⁷ In logarithmic terms, summary statistics reveal the earnings differential between crowdwork and traditional work to be about -64.5% in the United States and -67.5% in Europe (-66.0% overall). This differential moves to -69.1% and -71.4% (-70.3% overall) when accounting for time spent undertaking unpaid activities. [Figure 1](#) shows participation in crowdsourcing conditional on age for both forms of work, showing how platform workers tend to be from younger age cohorts in which participation in traditional forms of employment is lower.

3.2 Selected labor market indicators and controls

To compare crowdworkers and traditional salaried workers, we focus on a number of key labor market indicators. With our data being extracted from a range sources, some variables have been subjected to recoding, for the sake of harmonization. Further details on the recoding process are available in [Supplementary Appendix B](#).

Hourly nominal earnings have been selected as our main variable of interest. Given that crowdworkers do not have fixed work hours, and that crowdwork is an auxiliary source of income to many individuals who have another occupation, the hourly rate of pay will not suffer from distortions caused by the number of hours worked per week. As platform workers are paid by task, hourly earnings are determined first by the demand for those skills and characteristics over which clients can discriminate (factors which can be observed in a survey) and, on the supply side, by the ability of each individual worker to complete these tasks efficiently (which is mostly unobserved).

Furthermore, the design of the ILO survey allows for an estimation of the time crowdworkers spend performing unpaid activities on the platform, while the AWCS and EWCS allow for measuring time spent commuting. This permitted the investigation of other variations in the earnings differential between crowdworkers and traditional workers, after accounting for unpaid activities such as commuting for traditional workers, and performing unpaid task, looking for tasks, earning qualifications, or leaving reviews for crowdworkers.

Additionally, along with indicators of work/life balance and job satisfaction, the EWCS, AWCS and ILO surveys contain items for identifying if respondents would like to work more than what they currently do, either in

⁶ Eight hundred fifty-two observations are extracted from European Member States, and 148 observations from EWCS guest countries.

⁷ Summary statistics are reported in the [Supplementary Appendix, Tables A.1–A.3](#).

Table 1. Differences across crowdworkers and traditional workers in the US and EU

	US			EU		
	Traditional	Crowdwork	Diff.	Traditional	Crowdwork	Diff.
Hourly nominal earnings (USD)	30.774 (207.851)	7.208 (7.483)	−23.566***	17.058 (91.886)	6.585 (28.970)	−10.473***
Hourly nominal earnings (USD) [†]	26.947 (195.506)	5.371 (5.023)	−21.580***	15.448 (81.167)	4.694 (20.919)	−10.750***
Weekly work hours	39.056 (11.655)	21.180 (20.511)	−17.876***	37.176 (11.901)	14.697 (24.137)	−22.479***
Weekly work hours [†]	39.056 (11.655)	28.266 (26.422)	−10.789***	37.176 (11.901)	19.903 (32.601)	−17.273***
Age	41.024 (12.615)	35.027 (10.934)	−5.997***	42.207 (11.390)	35.543 (11.137)	−6.663***
Female	0.463 (0.499)	0.476 (0.500)	0.013	0.478 (0.500)	0.426 (0.495)	−0.051***
Married or living with a partner	0.516 (0.500)	0.434 (0.496)	−0.082***	0.697 (0.459)	0.493 (0.500)	−0.204***
No. of people in household	3.063 (1.672)	2.665 (1.429)	−0.398***	2.882 (1.268)	2.819 (1.260)	−0.063
Main earner in household	0.603 (0.489)	0.789 (0.408)	0.186***	0.595 (0.491)	0.815 (0.389)	0.220***
Educ.: no high school diploma	0.064 (0.244)	0.009 (0.092)	−0.055***	0.161 (0.367)	0.052 (0.222)	−0.109***
Educ.: high school diploma	0.502 (0.500)	0.374 (0.484)	−0.128***	0.448 (0.497)	0.309 (0.462)	−0.139***
Educ.: technical/associate	0.097 (0.296)	0.157 (0.364)	0.061***	0.147 (0.354)	0.102 (0.303)	−0.045***
Educ.: bachelor's degree	0.208 (0.406)	0.348 (0.477)	0.141***	0.127 (0.333)	0.322 (0.468)	0.195***
Educ.: master's degree	0.094 (0.292)	0.097 (0.296)	0.003	0.108 (0.311)	0.165 (0.371)	0.056***
Educ.: higher	0.036 (0.185)	0.015 (0.122)	−0.021***	0.009 (0.092)	0.051 (0.219)	0.042***
Health: very good	0.132 (0.338)	0.244 (0.429)	0.112***	0.261 (0.439)	0.257 (0.437)	−0.003
Health: good	0.407 (0.491)	0.534 (0.499)	0.128***	0.532 (0.499)	0.523 (0.500)	−0.008
Health: fair	0.345 (0.475)	0.180 (0.384)	−0.165***	0.185 (0.389)	0.178 (0.383)	−0.007
Health: poor	0.099 (0.299)	0.037 (0.190)	−0.062***	0.020 (0.140)	0.033 (0.178)	0.013**
Health: very poor	0.018 (0.132)	0.005 (0.071)	−0.013	0.002 (0.048)	0.008 (0.090)	0.006*
Caregiving (15 h/week)	0.149 (0.356)	0.207 (0.405)	0.058***	0.170 (0.375)	0.112 (0.315)	−0.058***
Caregiving (40 h/week)	0.082 (0.275)	0.207 (0.405)	0.124***	0.020 (0.139)	0.112 (0.315)	0.092***

Notes: Mean-comparison *t*-tests across crowdworkers (ILO data) and traditional workers (AWCS and EWCS data) from the US and EU. Standard errors in parentheses. Summary statistics and *t*-test are calculated from weighted US and EU reference samples. Sample restricted to employed and self-employed individuals in working age. [†]Adjusted for time spent in unpaid activities.

P* < 0.05; *P* < 0.01; ****P* < 0.001.

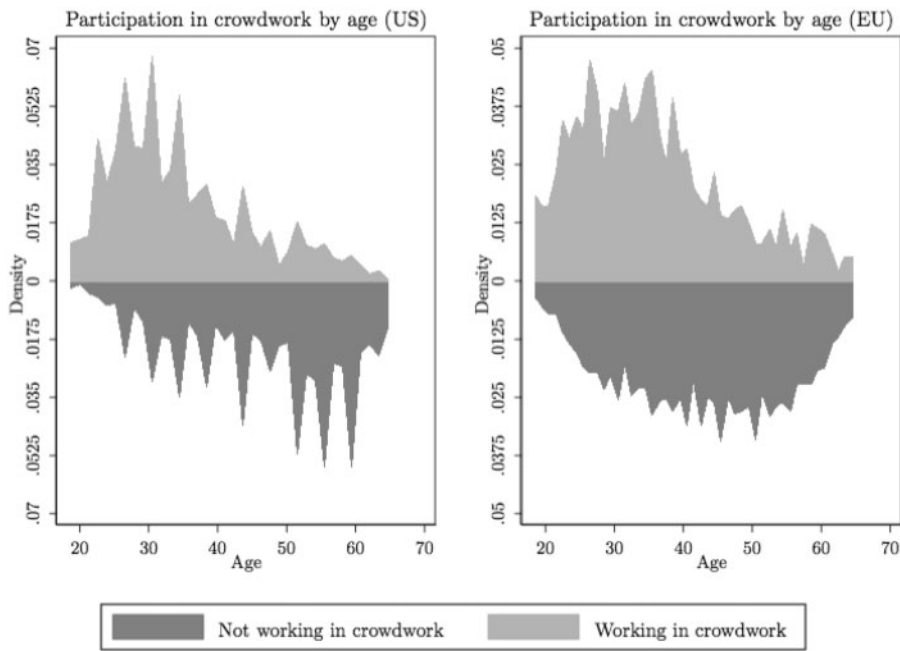


Figure 1. Participation in crowdsourcing versus traditional occupations by age.

Notes: The figure shows the probability density functions of age by type of work across the US and European samples. Comparison sample is restricted to employed and self-employed individuals in working age, excluding freelancers working from home.

crowdwork or in traditional work. These aspects of crowdwork, along with weekly work hours, are also studied in our analysis, serving as proxies for labor utilization in the platform.

In our analysis, a number of controls were considered. We first control for age—and its squared term—gender, and include constants for health status and educational attainment. In the literature, returns to education on earnings have been widely documented (Angrist and Krueger, 1991; Card and Krueger, 1992), while gender pay gaps have also been studied thoroughly (Altonji and Blank, 1999).

We can also expect marital status and the number of people living in the household to affect earnings and working conditions in general, as it is a well-known fact that married men tend to earn more than similar unmarried men (Hersch and Stratton, 2000). For this reason, the interaction between gender and marital status is also included in the specification. We then include a number of controls for other household characteristics, starting with a proxy for household income (i.e. whether the household is able to make ends meet each month), and then adding controls which specify whether the respondent is the main earner in his/her household and whether the occupation consists of the main source of income of the respondent.

We always control for state-level effects which can mitigate regional variability in outcomes by including a dichotomous variable for each US federal and EU member state (79 in total). These dummies should control for cross-country and cross-state variability in wage-setting institutions across national and federal entities, and also capture wider differences in earnings between the Europe and the United States. This means that, while our individual level regressions will not reveal the direct effect of labor market institutions, the confounding effect of these factors will nonetheless be held as fixed. For a formal discussion on how our model controls for cross-country heterogeneity, and for an analysis of the institutional determinants for participation and earnings in crowdwork, see Appendix C.

Also, the ILO survey presents items indicating the last occupation of the respondent before he/she started their involvement in crowdwork, enabling us to match crowdworkers with traditional workers through 44 different ISCO-08 occupations. These occupational fixed effects are then included in our specification. As these intercepts should reflect the respondent's condition prior to entering any crowdwork occupation, this group of predictors will not act as a confounder, but will rather enhance our estimates by providing a further control for individual ability. Indeed, as

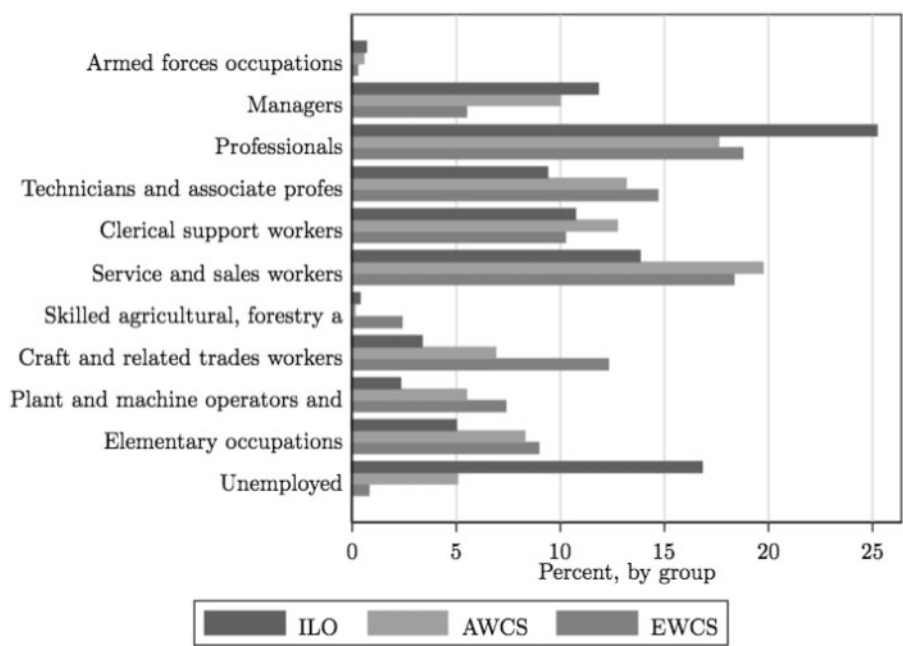


Figure 2. Occupational breakdown by survey.

Notes: The figure shows the last non-crowdwork occupation held by respondents from each survey used in the final sample. Occupations are aggregated by ISCO-08 major groups, and figures are presented as group-specific percentages. ISCO-08 groups: (0) Armed forces occupations, (1) Managers, (2) Professionals, (3) Technicians and associate professionals, (4) Clerical support workers, (5) Service and sales workers, (6) Skilled agricultural, forestry and fishery workers, (7) Craft and related trades workers, (8) Plant and machine operators and assemblers, (9) Elementary occupations.

participation in crowdwork markets would stand as a source of endogenous variation for the respondent’s current occupation, this inter-temporal quality allows for comparing occupations and individual ability before an individual’s decision to enter crowd employment is taken. As the motivation for entering these markets might be endogenous, the occupational vector will also include a dummy (“Kept last non-crowdwork occupation”) specifying whether this individual is still engaged in this profession at the time of the interview.

Finally, we include a set of controls for how many years the respondent has been working in their last non-crowdwork occupation. These controls provide a much better proxy for experience in traditional labor markets than age alone, given the likelihood of longer labor force interruptions by women than men.

The occupational structure of each survey is shown in Figure 2, in which last non-crowdwork occupations are aggregated into ISCO-08 major groups, and characteristics within each work-type are shown as a proportion of the total.

4. Analyzing differences between crowdworkers and traditional workers

4.1 Empirical specification

We estimate differences in earnings and working conditions between platform and traditional workers by treating them as two distinct groups. The primary group comprises crowdworkers from the ILO survey, while the comparison group includes workers from the AWCS and EWCS. From this point of view, our approach has drawn inspiration from Dehejia and Wahba (1999), who took individuals from different observational groups, matched them, and compared them with the aid of econometric tools.

A simple comparison of average earnings between the two groups might produce biased results. To account for selection bias and offer a more appropriate comparison between types of workers, we compare outcomes controlling for observable characteristics through ordinary least squares regressions. The baseline OLS specification we adopt is the following, where *i* refers to each individual:

$$Y_{is} = \text{Crowdwork}_{is}\beta + X'_{is}\gamma + \psi_s + \epsilon_{is} \quad (1)$$

The dummy Crowdwork_i (alternatively, “Working in crowdwork”) designates individuals in crowdwork, while Y_i is the labor market outcome of interest. Hourly earnings is the main focus, but the relationship between crowdwork and other facets of working conditions (such as working hours and factor utilization) will also be investigated.

The vector X'_i is our vector of controls and as mentioned earlier, includes gender, educational level, marital status, health status, household characteristics, and occupational controls, whereas ψ_s controls for regional (state-level) fixed effects.

To evaluate how much individual and socio-economic characteristics affect self-selection into crowdwork, we follow this baseline empirical specification, where the variables are characterized in the same way as in [equation \(1\)](#):

$$\text{Crowdwork}_{is} = X'_{is}\gamma + \psi_s + \epsilon_{is} \quad (2)$$

This model allows us to analyze the relationship between participation in crowdwork and observable skills, enabling us to define how crowdworkers set themselves apart from the rest of the workforce.

4.2 Results

In the following sections, we present our regression results. While in [Section 4.2.1](#), we compare labor market outcomes across crowdworkers and traditional workers (cfr. [equation \(1\)](#)), in [Section 4.2.2](#), we analyze the determinants of the participation in crowdwork (cfr. [equation \(2\)](#)).

4.2.1 Working conditions in crowdwork

Columns (1) to (3) from [Table 2](#) present initial OLS results for hourly earnings, using a sample of US crowdworkers from the ILO survey and traditional workers from the AWCS. Additional controls are added with each specification, with an initial sample including a total of 3217 workers.⁸

Initial key controls from column (1) are meant to capture individual characteristics such as gender and age. Intercepts for the US and EU state of residence, health and education are also included. Standard errors are robustly clustered at the level of the federal and member state.

Household characteristics are controlled for in column (2), where we add regressors for number of people in the household, marital status, and three dummies indicating whether the respondent is the main contributor to the household's income, whether crowdwork is his/her main source of income, and whether there is enough household income to cover basic expenses.

As shown in the table, the effect of crowdwork on earnings is always negative and significant, pointing in both cases to a 64% lower earnings. Hourly earnings are also generally higher (18.0% higher) across crowdworkers whose main occupation is crowdwork. The third column, shows the full specification, which adds the set of occupational controls, along with a dummy denoting whether the respondent has kept his/her last non-crowdwork occupation. The regression shows that crowdwork retains a negative and significant association with earnings (pointing to 60.9% lower earnings), yet its magnitude is lowered from the introduction of the occupational controls.

Columns (5) to (7) present estimates for the effect of crowdwork on hourly earnings in the European sample (27,741 observations). The sign and magnitude of the crowdwork coefficient is always negative and significant and, after controlling for all covariates and occupational effects in column (7), we obtain similar estimates, but larger than the ones obtained for the US sample, equaling to a 71.1% reduction in hourly earnings. This time, the inclusion of the occupational controls leads to a moderate increase in the coefficient's magnitude, from 68.4% in column (6).

A full sample (US + EU) specification is presented in column (9). The difference in general region-specific gender effects is isolated by the coefficient of the $EU \times \text{Female}$ interaction term. The effect of crowdwork on PPP-adjusted net hourly earnings is estimated to be 64.1% lower. The omission of occupational controls increases this coefficient by a negligible margin (63.9%, not shown in the table), suggesting that prior occupational skills play little role in explaining differences in earning outcomes between crowd employment and traditional forms of work. Also, in all instances, the negative effect of working in digital labor market is slightly reduced when crowdwork poses as the main source of income for the respondent. Furthermore, earnings always appear unconnected with whether crowdworkers have retained their last occupation prior to entering crowd employment.

⁸ Observations with missing values are excluded from the estimation.

Table 2. Estimates of the effect of online crowdwork on earnings in the US and EU (OLS)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	US			EU			US+EU			
	H. earnings	H. earnings	H. earnings [†]	H. earnings	H. earnings	H. earnings [†]	H. earnings	H. earnings [†]	H. earnings	H. earnings [†]
Working in crowdwork	-1.042*** (0.038)	-1.037*** (0.041)	-0.939*** (0.067)	-1.135*** (0.063)	-1.187*** (0.073)	-1.153*** (0.030)	-1.243*** (0.060)	-1.399*** (0.048)	-1.027*** (0.057)	-1.206*** (0.045)
Female	-0.240*** (0.046)	-0.161** (0.064)	-0.136** (0.055)	-0.144** (0.056)	-0.126*** (0.010)	-0.045** (0.018)	-0.048** (0.019)	-0.051** (0.019)	-0.181*** (0.047)	-0.187*** (0.046)
EU × female										0.140*** (0.044)
Age	0.051*** (0.010)	0.036*** (0.013)	0.024** (0.011)	0.022* (0.011)	0.023*** (0.004)	0.015*** (0.004)	0.009** (0.004)	0.009** (0.004)	0.011*** (0.004)	0.011*** (0.004)
Age squared	-0.001*** (0.000)	-0.000** (0.000)	-0.000** (0.000)	-0.000** (0.000)	-0.000*** (0.000)	-0.000** (0.000)	-0.000	-0.000	-0.000** (0.000)	-0.000** (0.000)
No. of people in household		-0.019 (0.012)	-0.025** (0.011)	-0.020* (0.011)		0.006 (0.007)	0.005 (0.006)	0.006 (0.006)	0.000 (0.006)	0.002 (0.005)
Married or living with a partner		0.246*** (0.069)	0.201*** (0.063)	0.195*** (0.064)		0.090*** (0.017)	0.085*** (0.014)	0.084*** (0.015)	0.092*** (0.015)	0.090*** (0.016)
Married/living with partner × Female		0.030 (0.096)	0.048 (0.093)	0.054 (0.086)		-0.025 (0.021)	-0.031 (0.022)	-0.026 (0.021)	-0.020 (0.025)	-0.014 (0.024)
Main earner in household		0.408*** (0.049)	0.368*** (0.048)	0.375*** (0.045)		0.150*** (0.013)	0.127*** (0.014)	0.129*** (0.014)	0.151*** (0.016)	0.154*** (0.015)
Main source of income		0.176*** (0.046)	0.254*** (0.062)	0.199*** (0.063)		0.131* (0.068)	0.094** (0.040)	0.053 (0.063)	0.195*** (0.042)	0.141*** (0.044)
Monthly hh. income covers expenses		0.366*** (0.041)	0.324*** (0.036)	0.326*** (0.034)		0.200*** (0.010)	0.167*** (0.013)	0.169*** (0.012)	0.179*** (0.014)	0.181*** (0.013)
Kept last non-crowdwork occupation			0.015 (0.042)	0.006 (0.038)			-0.055 (0.065)	-0.084 (0.083)	-0.001 (0.032)	-0.016 (0.033)
Observations	3217	3085	3085	3085	27,741	27,562	27,562	27,562	30,647	30,647
Adjusted R-squared	0.363	0.415	0.441	0.499	0.373	0.400	0.422	0.438	0.418	0.445
Federal/member state controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Education controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Health controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Occupation controls	No	No	Yes	Yes	No	No	Yes	Yes	Yes	Yes
Experience controls	No	No	Yes	Yes	No	No	Yes	Yes	Yes	Yes

Notes: Standard errors clustered by federal and member state in parentheses. Comparison sample restricted to employed and self-employed individuals in working age, excluding freelancers working from home. [†] Adjusted for time spent in unpaid activities.
* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

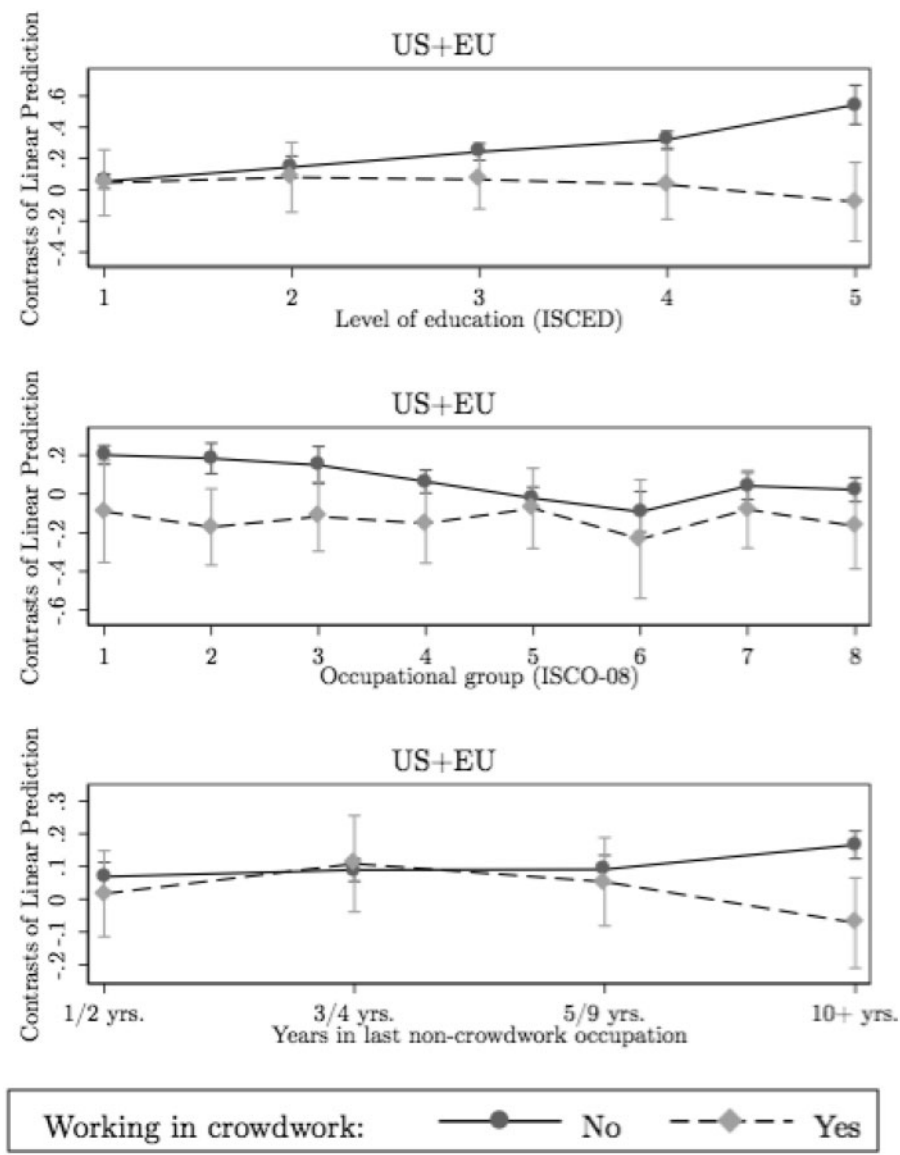


Figure 3. Heterogeneous returns from observable skills in crowdwork (US + EU).

Notes: Contrasts of linear predictions (Predictive Margins) of Education, Occupation and Years of experience with 95% CIs. Levels of education: (Baseline) no high school diploma, (1) high school diploma, (2) technical/associate degree, (3) bachelor’s degree, (4) master’s degree, (5) higher. Occupational groups: (Baseline) elementary occupations, (1) managers, (2) professionals, (3) technicians and associate professionals, (4) clerical support workers, (5) service and sales workers, (6) skilled agricultural, forestry and fishery workers, (7) craft and related trades workers, (8) plant and machine operators and assemblers.

Accounting for time spent in unpaid activities, hourly earnings—columns (4), (8) and (10)—fall well below our previous estimates, displaying an overall coefficient of -70.0% (United States: -67.8% ; Europe: -75.1%). The flexibility in working hours and the opportunity of working from home are not enough to offset the time lost performing unpaid chores.

We also test (Figure 3) for the presence of differential returns to observable skills in crowdwork by interacting crowdwork with each educational attainment and occupational group, maintaining the covariate specification from

Table 2, column (7).⁹ Figure 3 plots the predicted returns from education, occupational group and years of experience in the last non-crowdwork occupation, conditional on participation into crowdwork markets. The results indicate that disparities between traditional and platform workers persist and increase with the level of education, and that, similarly, earnings in crowdwork are generally indifferent from prior occupation. Most importantly, these figures suggest that crowdwork arrangements might offer almost no return to observable skills, as the returns to education and occupation from traditional occupations are canceled out in crowdwork (in line with the results from Rani and Furrer (2019)).

OLS estimates for work hours indicators are shown in Table 3. When investigating time spent on the platform, the estimates appear sensitive to the way work hours are computed. In particular, in columns (1), (4), and (7) we find that, on average, when only paid activities are considered, workers in crowdwork work 12.8 h less than traditional workers on a weekly basis, with a 11 h differential between US and the EU platform workers. When crowdwork is also the main source of income, these figures are further reduced, and all crowdworkers appear to be working circa 2 h less than traditional workers, all else being equal. However, wide differences between the US and the EU persist: when crowdworking is the respondent's main source on income, crowdworkers from the United States appear to be working even more (nearly 2 h) than their counterparts in traditional occupations, whereas crowdworkers from the Europe still work fewer hours than traditional workers.

However, if the indicator is adjusted for the time spent on unpaid tasks—as in columns (2), (5), and (8), Table 3—the magnitude of the coefficient changes again, showing an increase of 6.5 h in work hours across the United States and the Europe. In terms of weekly work hours, US crowdworkers are now indistinguishable from their counterparts in traditional occupations, whereas a 17.6 h gap persists between crowd and traditional workers in Europe. For individuals whose main occupation is crowdwork, the differential compared with traditional workers is even lower, to the point that, on average, US crowdworkers appear to be working 15 h more than comparable workers. Significant disparities with the European sample remain, indicating that for EU workers, these differences in work hours—including unpaid activities—between platform and traditional workers are reduced to 10 h when crowdwork provides the main source of an individual's income. These findings corroborate the results from Horton and Chilton (2010), and suggests that many workers are unable to meet their earnings targets.

Moving to factor utilization, in columns (3), (6), and (9) in Table 3, our OLS model suggest that a large proportion of platform workers would like to work more than they currently do in either crowdwork or in other forms of employment, regardless of crowdwork being the main source of income, suggesting a degree of factor under-utilization. This does not necessarily imply that crowdworkers would like to spend more time in crowdwork: in the Supplementary Appendix A (Table A.4, columns 1, 4, and 7), we show that, on average, crowdworkers are less willing to spend extra work hours on the platform rather than in any other occupation. The table also shows how job satisfaction tends to be lower when compared to traditional occupations, even though crowdworking seems to be related to a better work life balance.

In light of this, the crowdworker status presents some similarities to the condition of involuntary part-timers or inactive persons with a labor force attachment, where individuals would like to work more but are unable or too discouraged to look for other forms of employment. For that, crowdwork could be found to be related to slack in the labor market, possibly linked to a scarcity in demand. This is consistent with the interpretation of Katz and Krueger (2018), who found work in online platform work to be mostly involuntary and based on economic reasons.

These peculiarities in the crowdworker status justify the need to study crowdwork as an alternative to unemployment. We do so in Supplementary Appendix A, Table A.5, in which we re-adjust the AWCS/EWCS comparison groups to include unemployed individuals, and introduce an unemployment dummy to the group of occupation controls. Compared to all labor force, crowdwork appears to be linked to an overall 52.4% reduction in hourly earnings, on average (52.0% in the United States, 54.6% in Europe), moving to -60.2% (United States: -60.7%; Europe: -62.1%) after controlling for unpaid activities.

4.2.2 Determinants of participation in crowdwork

The OLS estimates from the previous section provide a first descriptive account of the different working conditions of crowdworkers and traditional workers, but do not allow us to identify how much observable individual and

9 For simplicity, the occupational controls have been grouped into major one-digit ISCO-08 occupational groups.

Table 3. Estimates of the effect of online platform work on working hours in the US and EU (OLS)

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Work hours	Work hours [†]	More work	Work hours	Work hours [†]	More work	Work hours	Work hours [†]	More work
Working in crowdwork	-9.939*** (1.675)	-2.572 (2.159)	0.598*** (0.037)	-21.421*** (2.245)	-17.649*** (2.128)	0.245*** (0.080)	-12.843*** (1.790)	-6.554*** (2.536)	0.495*** (0.060)
Female	-1.010 (0.948)	-0.971 (1.271)	0.023 (0.036)	-3.518*** (0.513)	-3.743*** (0.617)	0.004 (0.024)	-1.657*** (0.611)	-1.649* (0.850)	0.033 (0.036)
EU × female							-1.709** (0.819)	-1.870* (1.070)	-0.032 (0.033)
Age	0.884*** (0.221)	1.171*** (0.300)	-0.010* (0.006)	0.476*** (0.096)	0.526*** (0.104)	0.001 (0.004)	0.531*** (0.093)	0.617*** (0.104)	-0.001 (0.004)
Age squared	-0.011*** (0.003)	-0.014*** (0.003)	0.000 (0.000)	-0.006*** (0.001)	-0.007*** (0.001)	-0.000 (0.000)	-0.007*** (0.001)	-0.008*** (0.001)	0.000 (0.000)
No. of people in household	-0.399** (0.188)	-0.583** (0.229)	0.020* (0.010)	-0.275* (0.142)	-0.315* (0.158)	0.010* (0.005)	-0.329*** (0.119)	-0.406*** (0.138)	0.012** (0.005)
Married or living with a partner	1.890 (1.243)	1.460 (1.461)	-0.123*** (0.032)	1.143*** (0.260)	1.241*** (0.279)	-0.000 (0.019)	1.235*** (0.264)	1.311*** (0.308)	-0.009 (0.019)
Married/living with partner × female	-1.705 (1.495)	-1.688 (1.783)	-0.013 (0.051)	-0.272 (0.409)	-0.417 (0.400)	-0.013 (0.015)	-0.446 (0.405)	-0.631 (0.429)	-0.015 (0.015)
Main earner in household	5.085*** (1.016)	5.590*** (1.179)	-0.140*** (0.024)	3.295*** (0.592)	3.544*** (0.630)	0.023** (0.011)	3.428*** (0.531)	3.673*** (0.568)	0.011 (0.012)
Main source of income	12.191*** (1.331)	17.744*** (1.886)	-0.004 (0.022)	4.341* (2.330)	7.294*** (1.842)	0.035 (0.021)	10.592*** (1.301)	15.810*** (1.745)	0.106*** (0.024)
Monthly hh. income covers expenses	0.719 (1.033)	0.372 (1.208)	-0.120*** (0.022)	0.908*** (0.248)	0.807*** (0.280)	-0.060*** (0.007)	0.863*** (0.244)	0.738*** (0.278)	-0.063*** (0.007)
Kept last non-crowdwork occupation	1.554 (1.309)	0.944 (1.533)	0.012 (0.025)	-2.316 (1.445)	-1.126 (2.307)	-0.048** (0.023)	0.127 (1.029)	0.345 (1.338)	0.001 (0.016)
Observations	3085	3085	3084	27,562	27,562	27,017	30,647	30,647	30,101
Adjusted R-squared	0.317	0.227	0.419	0.272	0.241	0.087	0.284	0.236	0.112
Federal/member state controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Education controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Health controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Occupation controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Experience controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Standard errors clustered by federal and member state in parentheses. Comparison sample restricted to employed and self-employed individuals in working age, excluding freelancers working from home. [†]Adjusted for time spent in unpaid activities.
* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

Table 4. Determinants of participation in online crowdwork in the US and EU (OLS)

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
		US		US [‡]		EU		EU [‡]	US+EU	US+EU [‡]
Working in crowdwork										
Female	0.011 (0.018)	0.022 (0.027)	0.028 (0.026)	0.028 (0.026)	-0.008* (0.004)	-0.015*** (0.005)	-0.010** (0.004)	-0.010** (0.004)	-0.010 (0.006)	-0.010 (0.006)
Age	-0.009*** (0.001)	-0.010*** (0.001)	-0.007*** (0.001)	-0.007*** (0.001)	-0.002*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.002*** (0.000)	-0.002*** (0.000)
No. of people in household		-0.027** (0.011)	-0.016** (0.007)	-0.016** (0.007)		0.004** (0.002)	0.002* (0.001)	0.003** (0.001)	-0.002 (0.002)	-0.001 (0.002)
Married or living with a partner		0.015 (0.037)	0.044 (0.032)	0.044 (0.032)		-0.031*** (0.008)	-0.013*** (0.004)	-0.016*** (0.004)	-0.011* (0.006)	-0.012** (0.006)
Married/living with partner × female		0.086** (0.040)	0.038 (0.048)	0.038 (0.048)		0.032*** (0.012)	0.019** (0.008)	0.023*** (0.007)	0.027*** (0.010)	0.031*** (0.010)
Main earner in household		0.200*** (0.034)	0.174*** (0.026)	0.174*** (0.026)		0.032** (0.012)	0.021*** (0.007)	0.026*** (0.007)	0.040*** (0.009)	0.044*** (0.009)
Monthly hh. income covers expenses		0.120*** (0.022)	0.097*** (0.022)	0.097*** (0.022)		0.023*** (0.006)	0.013*** (0.003)	0.020*** (0.004)	0.024*** (0.005)	0.030*** (0.005)
Observations	3217	3085	3085	3085	27,741	27,565	27,565	27,925	30,650	31,010
Adjusted R-squared	0.203	0.249	0.433	0.433	0.046	0.056	0.363	0.283	0.449	0.424
Federal/member state controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Education controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Health controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Occupation controls	No	No	Yes	Yes	No	No	Yes	Yes	Yes	Yes
Experience controls	No	No	Yes	Yes	No	No	Yes	Yes	Yes	Yes

Notes: Standard errors clustered by federal and member state in parentheses. Comparison sample restricted to employed and self-employed individuals in working age, excluding freelancers working from home. [‡]Comparison sample extended to include all individuals in the labor force, excluding freelancers working from home.

* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

Table 5. 2SLS estimates of the effect of online platform work on earnings in US and EU

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	(3)			(4)	(6)		(7)	(8)
	Caregiving (15 h)			Caregiving (40 h)	Caregiving (40 h)		Caregiving (40 h)	Caregiving (40 h)
	Hourly earnings							
	OLS Full sample	OLS Female only	2SLS Full sample	2SLS Female only	2SLS Full sample	2SLS Female only	2SLS Full sample	2SLS Female only
Working in crowdwork	−1.006*** (0.054)	−1.068*** (0.068)	2.214 (1.703)	2.393 (1.877)	−0.846** (0.337)	−0.732** (0.300)	−0.985*** (0.338)	−0.841*** (0.305)
Female	−0.165*** (0.062)		−0.270*** (0.069)		−0.186*** (0.049)		−0.192*** (0.049)	
Crowdwork × female	−0.040 (0.053)							
EU × female	0.126** (0.057)	−0.060 (0.040)	0.249*** (0.064)		0.146*** (0.049)		0.147*** (0.048)	
Age	0.011*** (0.004)	0.011** (0.005)	0.014*** (0.004)	0.009 (0.006)	0.012*** (0.004)	0.011** (0.005)	0.011*** (0.003)	0.011** (0.006)
Age squared	−0.000** (0.000)	−0.000 (0.000)	−0.000** (0.000)	−0.000 (0.000)	−0.000** (0.000)	−0.000 (0.000)	−0.000** (0.000)	−0.000 (0.000)
No. of people in household	0.000 (0.006)	−0.002 (0.008)	−0.001 (0.005)	−0.004 (0.008)	0.000 (0.006)	−0.003 (0.007)	0.002 (0.005)	0.000 (0.007)
Married or living with a partner	0.093*** (0.016)	0.070*** (0.019)	0.113*** (0.019)	0.045** (0.022)	0.093*** (0.016)	0.068*** (0.019)	0.092*** (0.017)	0.070*** (0.018)
Married/living with partner × female	−0.021 (0.025)		−0.067** (0.029)		−0.022 (0.025)		−0.017 (0.025)	
Main earner in household	0.151*** (0.016)	0.130*** (0.017)	0.098*** (0.027)	0.064* (0.036)	0.148*** (0.019)	0.123*** (0.020)	0.150*** (0.018)	0.127*** (0.020)
Main source of income	0.197*** (0.039)	0.174*** (0.056)	2.793** (1.374)	2.817* (1.461)	0.339 (0.292)	0.431* (0.250)	0.314 (0.291)	0.436* (0.254)
Monthly hh. income covers expenses	0.179*** (0.014)	0.160*** (0.025)	0.160*** (0.020)	0.125*** (0.035)	0.178*** (0.014)	0.157*** (0.025)	0.180*** (0.014)	0.161*** (0.024)
Observations	30,647	15,827	30,647	15,827	30,647	15,827	30,647	15,827
Adjusted R-squared	0.418	0.408	−0.024	−0.125	0.304	0.280	0.339	0.315
Federal/member state controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Education controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Health controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Occupation controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Experience controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
F-test			2.427	3.353	10.09	17.77	10.09	17.77
First stage R ²			0.801	0.783	0.804	0.789	0.804	0.789

Notes: Standard errors clustered by federal and member state in parentheses. Control sample restricted to employed and self-employed individuals in working age, excluding freelancers working from home. [†] Adjusted for time spent in unpaid activities.

* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

socio-economic characteristics can really account for self-selection into crowdwork. To better understand who crowdworkers really are, and how they differ from traditional workers, in Table 4 participation in crowdwork is regressed on the same set of regressors as in Table 2.¹⁰

Overall, crowdwork appears to be negatively associated with age, and the linear probability of entering in online labor markets is reduced by 0.2% as a worker's age increases by 1 year.¹¹ Also, family income always appears to be correlated with participation in online crowdwork, supporting the hypothesis that individuals enter these markets to seek auxiliary earnings. The fact that crowdworkers are also more likely to be the main earner in their household, holding marital status as fixed, reinforces this finding.

In the Europe, married individuals, or individuals living with a partner have, in general, a lower (−1.3%) probability of working in crowd employment platforms. However, this correlation is reversed if the individual is female. In United States, this correlation with gender and marital status is only present until occupational controls are introduced.

As denoted from the increase in the model's goodness of fit between columns (2) to (3), and (6) to (7), the introduction of occupational and experience controls seem to explain a large portion of the unexplained variation in participation in crowd employment. However, Table 4 provides little insight on the skill determinants of access to crowdwork. For ease of exposition, occupations are grouped again by major ISCO-08 groups in Figure 4, in which the marginal effect of occupational history and educational attainment is shown, holding other factors as fixed. To allow for a comprehensive discussion of how experience in traditional labor markets affects access to crowdwork, the comparison group again includes the entire labor force.

Workers with no experience in their last traditional occupation—that is, the unemployed—are 33% more likely to enter crowd employment than individuals with less than one year of experience in their last traditional occupation. The fact that participation in crowdwork peaks among the unemployed, and is lowest among workers with less than 3 years' experience in their last occupation, reinforce our earlier findings suggesting that participation in crowdwork might result from slack in labor markets.

Crowdwork then appears to be as an alternative to traditional work for many individuals estranged from the labor market, offering them a source of auxiliary income. This provides us with the opportunity to investigate differential effects in participation conditional on household income. Little to no effect can be found across individuals already in the workforce. For the unemployed, however, participation appear to be higher among those households in which income is enough to cover basic expenses. However, whether this correlation should be explained by the fact that earnings in crowdwork allow the household to meet its needs is unclear.

Looking at other determinants of observable skill, further investigation reveals that participation in crowdwork peaks with individuals holding a bachelor's degree, who are 6.6% more likely to enter these markets than workers with no educational attainment. In the United States, participation similarly peaks among workers with technical and associate degrees. Investigating which occupational groups seems to be most affected, participation appears to be lower, compared to elementary occupations, among workers with history as professionals, in clerical, service, and sales occupations or as manual workers, while prior correlations with professional categories now appear spurious. All in all, these findings suggest that job polarization also plays a part in the growth of crowdwork, as employment levels for individuals with a medium skill level has also seen significant reductions in recent years within both the United States and the Europe (see Autor and Dorn, 2013; Goos *et al.*, 2014).

5. Controlling for self-selection into crowdwork

5.1 Empirical specification

Previous results from our OLS specifications provide a first evidence of the relationship between crowd employment and earnings against more traditional forms of work. We found evidence of self-selection, but no notable effect of skill on earnings in crowdwork. However, it should be considered that unobserved endogenous factors could still affect earnings.

10 Crowdwork-specific regressors—"Main occupation", and "Kept last non-crowdwork occupation"—are omitted.

11 The squared term for age has been omitted from the specification due to its lack of significance.

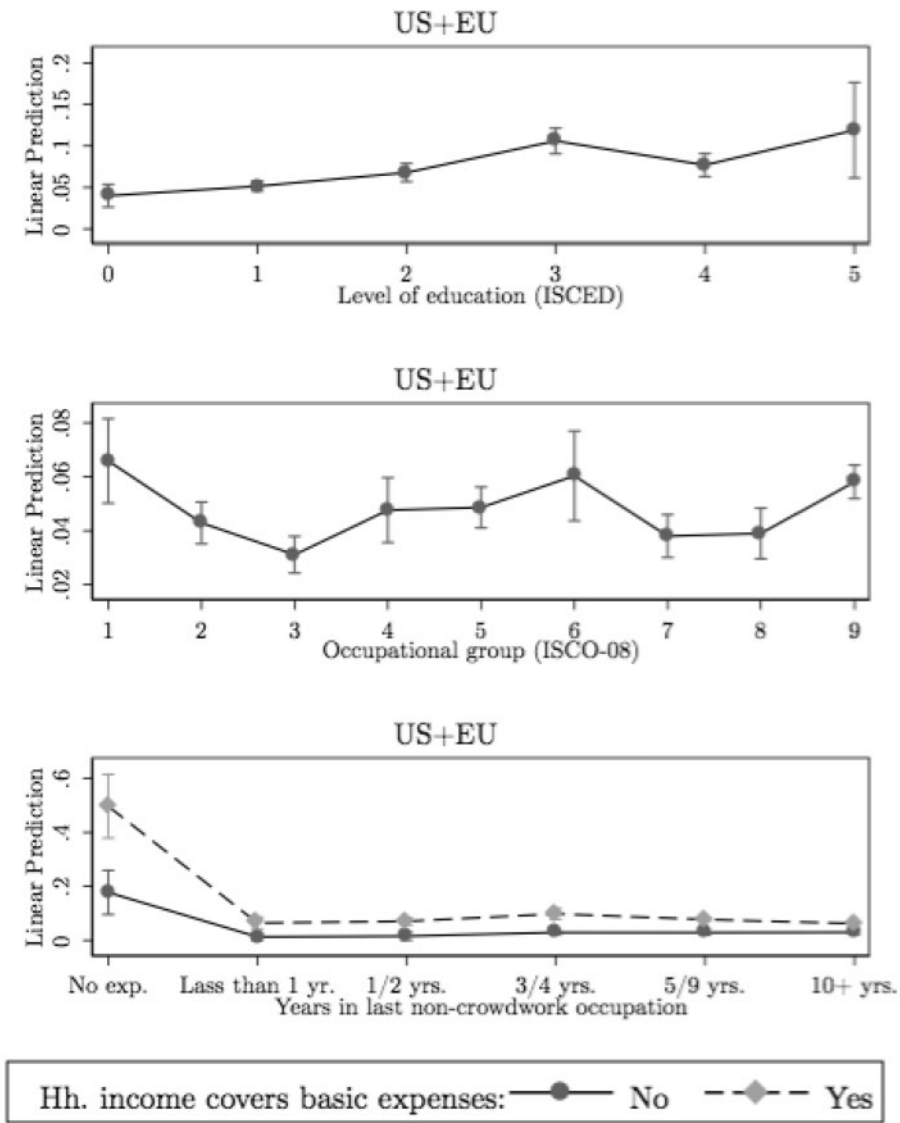


Figure 4. Determinants of participation in crowdwork (US + EU).

Notes: Predictive margins of education, occupation and years of experience with 95% CIs. Levels of education: (0) no high school diploma, (1) high school diploma, (2) technical/associate degree, (3) bachelor's degree, (4) master's degree, (5) higher. Occupational groups: (Baseline) elementary occupations, (1) managers, (2) professionals, (3) technicians and associate professionals, (4) clerical support workers, (5) service and sales workers, (6) skilled agricultural, forestry and fishery workers, (7) craft and related trades workers, (8) plant and machine operators and assemblers.

In this section, we control for this self-selection problem by adopting an instrumental variables approach. Our two-stage least squares (2SLS) specification is the following, where i refers to each individual:

$$Y_{is} = \text{Crowdwork}_{is}\beta + \text{Female}_{is}\varphi_2 + X'_{is}\gamma_2 + \psi_{2s} + \epsilon_{2is} \tag{3}$$

$$\text{Crowdwork}_{is} = Z_{is}\phi + \text{Female}_{is}\varphi_1 + X'_{is}\gamma_1 + \psi_{1s} + \epsilon_{1is} \tag{4}$$

In the second-stage regression (3), Y is the set of earning outcomes, *Crowdwork* is our endogenous variable, *Female* is a dummy which indicates whether the respondent is female and X is the vector of controls

from Section 4.1.¹² The first stage regression (4), estimates the linear probability of working in micro-task platforms. In case *Crowdwork* is assigned exogenously conditionally on Z , then the coefficient β obtained from these predicted values will not suffer from selection bias.

Several exogenous variables, such as age or health condition, are significantly correlated with crowdwork. However, their adoption as instrumental variables would potentially lead to a violation of the exclusion restriction, biasing our estimates downwardly: younger workers typically earn less than older individuals, while workers in poor health may take longer times to complete their work activities, leading to a reduction in hourly earnings.

We then considered a third instrument: time spent in caregiving. This variable is highly correlated with crowdwork. The underlying reasoning is that people may be more involved in crowdwork if they are compelled to stay at home to look after children or elderly relatives: this type of work can be a reasonable source of extra income to these individuals, given their circumstances. This setting allows us to compare outcomes across “compliers”, namely all individuals in caregiving who participate in crowdwork arrangements and all individuals not in caregiving who stay in traditional forms of work.

This identification strategy is not dissimilar from previous studies on part-time employment which instrumented hours of work through household size and fertility (such as Hotchkiss, 1991; Ermisch and Wright, 1993). More generally, Jeon and Ostrovsky (2018) also suggest that the relationship between childcare and women self-employment is likely causal.

Both the ILO and the AWCS-EWCS datasets capture time spent in caregiving, although in different ways. While caregiving appears as a dummy in the ILO dataset, where the respondent is asked whether this activity constituted a full-time commitment immediately before entering crowdwork, it is treated as a continuous variable in the AWCS and EWCS, where the respondent is asked how many hours per week/per day is currently engaged in these activities.

We harmonized the two variables by identifying both a 40 and 15 h-per-week effort as a full-time caring activity, following the findings from the Gallup-Healthways Well-Being Survey. Indeed, according to this study, caregivers working at least 15 h per week have declared that this activity significantly affected their work life.¹³

The inter-temporal difference between the caregiving variable across crowdwork and traditional samples allows us to preserve the causal relationship between caregiving and crowdwork: since respondents in crowdwork are not asked whether they are currently caring for a relative, but rather whether they were engaged in caregiving right before entering crowd employment, potential reverse causality issues are prevented from interfering. Indeed, the working time flexibility characterizing crowd employment might enable individuals already working in online platforms to engage in caregiving.

Caregiving appears to be highly correlated with crowdwork in our US sample. This relationship is similar in Europe where caregiving also reveals itself as a significant predictor of platform work, but only at the full time threshold. These differences hint at the possibility of welfare-biased differential effects of caregiving, as caregivers may have access to more labor law safeguards in Europe than in United States, reducing the need for auxiliary earnings from crowdwork. Evidence from Germany (Bick, 2016), indicates that a large fraction of working mothers in part-time would work full-time if they had greater access to subsidized childcare. It is then reasonable to expect labor market institutions to influence participation in crowdwork similarly (see Appendix C).

While the connections between crowdsourcing and caregiving are theoretically plausible and empirically proven, this choice of instrument prevents its use on outcomes other than hourly earnings, which are clearly affected by caregiving. Nonetheless, concerns about the instrument’s endogeneity and the risk of violation of the exclusion restriction persist and are discussed below.

5.2 Gender bias in caregiving

Transition into caregiving might be gender-biased, and the correlation between crowdwork and caregiving is stronger among women (0.246) than among men (0.135). This differential may support prior evidence on men’s caregiving being a complex phenomenon influenced by endogenous socio-economic determinants, uncovering a potential obstacle in our identification strategy. Indeed, while several studies found caregiving to be exogenous to the female

12 The need for the gender dummy appearing outside the X vector will be explained in Section 5.2.

13 For details about the Gallup-Healthways Well-Being Index, see <https://www.gallup.com/175196/gallup-healthways-index-methodology.aspx>.

population (see, as discussed later [Ciani, 2012](#); [Schmitz and Westphal, 2017](#)), the effect on the male population is less unambiguous.

The peculiarities of crowdwork markets come in our favor. Indeed, the anonymity offered on these platforms allows us to exclude the hypothesis that, on the demand side, direct gender discrimination might take place on these markets, being male and female workers virtually indistinguishable. However, on the supply side, while structural pay gaps might persist, further gender differences might be generated by gender-biased self-selection in crowdwork. These complications can be overcome by checking whether crowdwork generates any specific effect on gender-dependent outcomes. Our evidence suggests that this might not be the case (see [Table 5](#)), and that earnings in crowdwork only reaffirm preexisting structural gaps, corroborating results from [Adams \(2020\)](#), without generating further gender disparities.

As this assumption holds, the interaction between gender and the selected instrument can be added to the instrument pool in the first stage of the estimation process. In other words, if the interaction term between gender and crowdwork yields a zero effect on earnings, said interaction can be added to the instrument pool without expecting gender-biased violations of the exclusion restriction.

Alternatively, the 2SLS estimates that can be drawn from the pool of female workers can be also said to hold for the rest of the sample. Given that earnings are estimated by a log wage equation, the non-significance of interaction effect (which we will denote as ζ) allows the non-interacted gender effect to be fully absorbed by the constant term in the split sample estimate, leading, after controlling for all observables, to:

$$\beta \approx \beta_f + \zeta \quad (5)$$

meaning that β_f , the estimated effect of crowdwork on the female population, will approximate the full sample coefficient β minus the interaction term ζ . If ζ is not statistically different from zero, β_f will also closely approximate the baseline effect of platform work on the outcome.

This approach differs from traditional split sample instrumental variable models (see [Angrist and Krueger, 1995](#)). Firstly, both the first and second stage of the IV equation are estimated on the same female subsample. Secondly, the two subsamples do not need to be homogeneous, as gender differences in wage determination processes are already modeled. As discussed, our results suggest self-selection by gender in crowdwork to either be absent, or at least have no impact on earnings. Still, the possibility that randomization through caregiving could only be achieved in the female sample motivates our approach.

Even if caregiving is exogenous to the female population, a final concern might arise in terms of violation of the exclusion restriction: in other words, the amount of time a worker spends in caregiving might be endogenous to the hourly wage he/she could earn in the market. Evidence from the literature allows us to appease these concerns, as [Leigh \(2010\)](#) finds a direct effect on labor market participation, while observing no impact on hourly wages, while [Schmitz and Westphal \(2017\)](#) find female caregiving to be related to a decrease in probability in working full time, with no short-term effect on hourly earnings. Nonetheless, exclusion restriction concerns are further discussed in [Appendix B](#), and an informal test is provided in the next section.

5.3 Results

We now turn to our IV estimates for the effect of platform work on hourly earnings, which are displayed in [Table 5](#), for both the full sample and a female-only sample. OLS estimates for the full sample and the female-only sample are also shown, and include the interaction term *Crowdwork* \times *Female*.

In column (1), controlling for all other observables, the interaction term between gender and crowdwork is not statistically different from zero, while the coefficient on gender retains its prior magnitude and significance, showing a negative linear effect on earnings (-15.2%). Similarly, there is no notable variation in the crowdwork coefficient between the specifications (7) from [Table 2](#) and (1), (2) from [Table 5](#), in which the gender-specific effects were introduced. This finding provides support to our hypothesis that crowdwork platforms do not generate any notable intrinsic gender discriminatory effect other than reaffirming common structural gaps and, as discussed earlier, provide support to our identification strategy.

Looking at the output from the 2SLS regressions, the estimates for the full sample and the female sample show both weak predictive power when instrumenting caregiving with a 15 h weekly threshold (columns 3 and 4): the

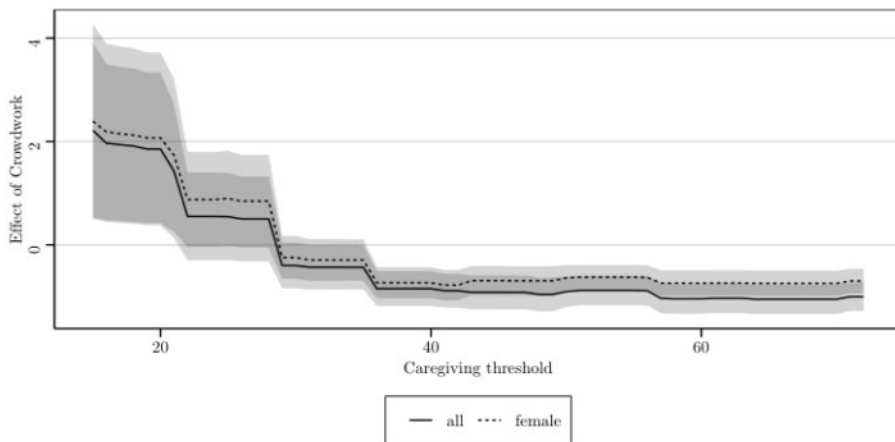


Figure 5. Estimated 2SLS coefficients from varying full-time caregiving thresholds (US + EU).

Notes: Second-stage coefficients for the “Working in crowdwork” instrumented through a caregiving instrument with increasing weekly hours threshold. Comparison sample restricted to employed and self-employed individuals in working age, excluding freelancers working from home.

crowdwork coefficient is never statistically different from zero and the instrument always fails to pass the [Sanderson and Windmeijer \(2016\)](#) multivariate F test of excluded instruments.

Instead, the 40 h threshold generates much more reasonable results for the effect of working in crowdwork (columns 5 and 6), predicting a general and statistically significant reduction (−57.1%; coeff.: −0.846) in hourly earnings. This is a notable change from the −63.4% estimate obtained with the OLS model (column 1), suggesting that self-selection through unobservable characteristics was indeed present.

Restricting our study to the female population to account for endogenous caregiving in the male sub-sample, working in crowdwork reduces earnings by 51.9% (column 6, coeff.: −0.732) over working age women, all else being equal. This is well below the −1.055 (−65.1%) log points that the least squares model would predict over the female sample (column 2). In both cases all instruments pass the F score test for excluded instruments, with the first-stage R^2 also yielding satisfying results. Complete first stage regressions are shown in the [Supplementary Appendix C, Table C.2](#).

As discussed earlier, if caregiving is exogenous only to the female population, the −0.732 coefficient could be considered to be close to an unbiased parameter of the local average treatment effect of crowdwork on earnings within the female population. The lack of a platform-specific gender gap in hourly earnings should allow for generalizing the split-sample estimates as in [equation \(5\)](#), to obtain a baseline reduction in earnings of 53.4%, raising our confidence in the results from the previous full sample specification.

Finally, hourly earnings are again adjusted for time spent in unpaid activities in columns (1) and (2), using full-time (40 h) caregiving as an instrumental variable. This time, the IV estimates point at a 62.6% reduction in hourly nominal earnings. This effect is 56.8% for the female group in column (2) (−58.5% after adjusting for [equation \(5\)](#)). Again, the IV model captures approximately 7.3% of the magnitude of the corresponding OLS coefficient, pointing at the presence of unobserved self-selection bias in the initial least squares model.

As a further check, in order to achieve a better understanding of the variability of the 2SLS estimates as the instrument changes its threshold, and to reduce the conceptual differences between the definitions of full time caregiving between the two groups of platform and traditional workers, [Figure 5](#) plots the selected caregiving threshold against the estimated effect of working in crowdwork, together with their significance level. It is evident from the figure that, with caregiving becoming a significant predictor of crowdwork at its 36 h per week threshold, the estimated coefficients also follow a more reliable pattern with little variation in their sign and statistical significance. Most importantly, full and split sample estimates conform to similar trends, providing evidence that our instrument choice adequately controls for gendered bias in caregiving.

As discussed in Section 5.2 of the [Supplementary Appendix](#), a number of issues might lead to a potential violation of the exclusion restriction. A final informal test for our instrument is provided in [Table 6](#), where hourly earnings are

Table 6. Effect of caregiving on hourly earnings conditional on education in US and EU (OLS)

Variables	(1)	(2)	(3)	(4)	(5)	(6)
	Full sample			Female only		
	C + T	T	C	C + T	T	C
Caregiving (40 h)	−0.140 (0.112)	−0.080 (0.122)	−0.582** (0.229)	−0.150 (0.103)	−0.043 (0.073)	−0.915** (0.374)
Caregiving (40 h) × High school diploma	0.116 (0.114)	0.094 (0.131)	0.689*** (0.235)	0.152 (0.105)	0.048 (0.090)	1.206*** (0.405)
Caregiving (40 h) × Technical/associate degree	0.050 (0.110)	0.053 (0.127)	0.572** (0.243)	0.012 (0.111)	−0.005 (0.076)	0.804** (0.338)
Caregiving (40 h) × Bachelor’s degree	0.039 (0.105)	0.080 (0.125)	0.659*** (0.211)	0.030 (0.108)	−0.002 (0.085)	1.019*** (0.361)
Caregiving (40 h) × Master’s degree	0.080 (0.139)	0.212 (0.139)	0.534* (0.304)	0.137 (0.133)	0.215** (0.100)	0.903** (0.439)
Caregiving (40 h) × Higher	0.179 (0.225)	0.371 (0.275)	0.832*** (0.266)	0.148 (0.317)	0.209 (0.317)	1.276*** (0.402)
Control covariates	Yes	Yes	Yes	Yes	Yes	Yes
Observations	30,647	28,683	1964	15,827	14,914	913

Notes: SE clustered by state in parentheses. “C+T” (crowdwork and traditional work samples), “C” (crowdwork sample), “T” (traditional work sample). Dependent variable: natural logarithm of hourly PPP-adjusted nominal earnings (US dollars). Sample restricted to the traditional work (AWCS + EWCS) groups in (2) and to the crowdwork (ILO) group in (3). Covariate list: age, age squared, number of people in household, main earner, main source of income, education, marital status, health status and state controls.

* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

regressed over the instrument, interacted with education, and the full set of covariates—except for the “Working in crowdwork” variable—across partitions of our sample. The analysis is performed for the 40 h caregiving threshold specification. If heterogeneous ability factors which we cannot already control for affect caregiving decisions and, in turn, hourly earnings, then we should see differential effects of observable skills across traditional workers in caregiving.

We find no evidence of heterogeneous returns to education for caregivers over both the full and female-only samples. We conclude that, if caregiving does not influence the hourly rate of pay between individuals of similar observable ability, it is reasonable to assume that a similar pattern should also emerge from unobservable ability.

The positive returns to education for caregivers in crowdwork are expected, as they motivate our identification strategy. While regressing the outcome on the instrument variable, conditional on the endogenous variable, is usually not a sufficient test for the exclusion restriction, the results from Table 6 can shed some light on specific concerns about potential earnings that we raised in Appendix B: if potential earnings were not accounted for, non-caregivers in traditional occupations should have earned more, on average, than their counterparts in caregiving.¹⁴

These findings increase our confidence in the goodness of the instrument. The informal tests in columns (2) and (5) are equivalent to the zero first stage test presented in Bound and Jaeger (2000), Altonji *et al.* (2005), as the exclusion restriction is tested on a subsample where no crowdworkers are present.¹⁵

If the income bias described in Appendix B is accounted for, the exogenous variation left by the caregiving instrument will then yield the income-indifferent individual propensity to assist a relative needing for care. In other words, once all observables characteristics are controlled for, full-time caregiving should retain its predictive qualities on participation in crowdwork markets, while having no significant effect on hourly earnings.

A number of robustness checks are presented in Supplementary Appendix C, testing the robustness of the instrument to the years of experience on the platform, and the validity of the model specification.

14 Models without education interaction were also tested, with similar results.

15 Recall that little to no crowdworkers could be found in the traditional employment subsamples, as discussed in Section 3.

6. Policy implications

Understanding how much individual ability affects the earning differential between crowd employment and traditional work can have deep ramifications in terms of policy. Should, then, regulators treat crowd employment markets as traditional markets? There is no simple answer to this question.

Our results suggest that regulatory action, if taken, should then target both platform markets and traditional markets. It is clear that participation in crowdwork is linked with slack in traditional labor markets, especially when considered that workers in crowdwork tend to be unable to access traditional jobs, or come from skill and occupational groups which have experienced contractions over recent years. This reflects larger issues correlated with the business cycle or momentous transformational processes such as job polarization, as our results suggest that participation in crowd-employment is more common in economies in which labor protections are less strict, even if little differences can be found between US and EU crowdworkers in terms of hourly earnings.

On the other hand, it is clear that the resulting heterogeneity in the elasticity of labor supply between traditional and crowdwork markets leaves platform and clients with a considerable monopsony power which enables them to exercise a considerable mark-up over labor on the platform. Policy action over these markets should then balance the need to address these imbalances with the idiosyncrasies of traditional labor markets.

More generally, regulation on online labor markets should spur from considerations on how much do these markets actually offer opportunities for social dumping. The demographic and skill composition of crowdwork is too distinct from traditional employment for crowdworkers to be really considered as direct substitutes of traditional workers. Instead, potential complementarities can be appreciated inasmuch as crowdwork offers the opportunity to enter employment to individuals with weaker attachment to the labor market, and auxiliary earnings for all others. Nonetheless, if hourly earnings are significantly lower in platform work even after controlling for all determinants of individual productivity, it is clear that crowdwork platforms are in a unique position of power, and it is worth wondering how much all platforms mediating work can exercise this same power.

From a purely economic perspective, it is difficult to determine whether or not crowdwork should be defined as work (De Stefano, 2016). The stark differences in skills and experience suggest that the threat of job replacement is overstated for micro-task crowdwork, and that the growth of these markets might instead be symptomatic of slack in traditional labor markets. Nonetheless, the environment of online platforms fosters monopsony through information asymmetries which, together with the lack of minimum wages and access to collective bargaining, offers opportunities for wage discrimination, as also suggested by Dube *et al.* (2020). These imbalances should be addressed, possibly within a “harmonized frame of reference” (De Stefano, 2018) for all forms of online platform work.

Further regulations should be addressed on a case to case basis, depending first and foremost on legal arguments concerning the nature of the relationship between the platform and the service provider (Prassl and Risak, 2015), and then considering whether the substitution effect offsets the complementarity effect or, in other words, whether social dumping opportunities prevail over ones of value creation (Kenney and Zysman, 2019).

Legislators should be aware of these caveats also taking into consideration how the recent COVID-19 pandemic could affect online labor markets. Given the negative labor demand shock generated by lockdown policies, labor supply in online labor markets could increase, as many workers displaced from traditional labor markets may choose to transit into online platforms. On the one hand, these platforms can thus help dampen the displacement effects of the lockdowns; on the other hand, the excess supply of online labor could lead to further deterioration of working conditions in terms of precariousness and quality of work, especially if the demand for online platform work remain unchanged. Whether this transition will be temporary or not is yet to be evaluated, but it is clear that if workers are seeking earnings through the gig-economy because of the displacement caused by the pandemic crisis, the issue over the fairness of working conditions on online platforms gains new relevance as non-standard forms of work affirm themselves as the new standard. If skill-biased self-selection disappears because of these developments, the wage rigidity to skill levels, coupled with the monopsonistic structure of these markets, becomes a serious problem that policy-makers should address. Clarifying the employment status of remote and in-person platform workers and addressing the imbalances arising from the monopsonistic nature of these markets becomes of utmost priority, especially if platform work becomes an alternative to traditional work.

7. Conclusions

In this article, we have provided an empirical analysis of the effect of crowdwork on working conditions in both the United States and Europe by comparing wage and labor market conditions across platform workers and workers in traditional occupations. The aim was to analyze the effects of individual ability on earnings in the platform economy, controlling to the best of our ability for skill-biased self-selection into crowdwork.

Our findings cast a dark shadow over online platform work, confirming the results of available descriptive evidence on crowdworkers' working conditions. On an hourly basis, crowdworkers earn circa sixty percent less than comparable workers in terms of ability, while spending nearly as much time working on the platform as their counterparts do in traditional occupations. Salary conditions worsen after accounting for unpaid work time. Moreover, crowdwork offers essentially no return to skills, meaning that crowdworkers' earnings are infinitely elastic to the respondent's educational level, occupational history, and idiosyncratic productivity. Finally, the labor force in crowdworking arrangements appears to be highly under-utilized, with crowdworkers being more likely to be left wanting more work than comparable individuals.

While we find evidence for endogenous self-selection within these markets, salary estimates are largely unaffected by skill-biased selection. The analysis suggests that the most observed disparities in labor market conditions between crowdworkers and traditional workers should be attributed to factors other than individual ability. This leads us to believe that this effect could be better explained by other factors, such as excess supply spurring from competition from equally skilled but cheaper labor from other countries within the same platform, the scarcity, and heterogeneity in demand for these kind of activities and/or the monopsonistic nature of platforms, linked with the general lack of labor standards, enabling the imposition of a heavy mark-up over online workers.

Along with the fact that online platform workers do not appear to be looking for other jobs more than traditional workers, and that they tend to have no experience in traditional labor markets, these findings suggest that crowdworkers belong to a new category of mostly idle workers whose human capital is not being fully utilized nor adequately compensated. In any case, the poor working conditions crowdsourcers have to live with are the result of an interplay between the aforementioned factors, and it is up to future research to test each of these hypotheses individually, disentangling the influence of each potential determinant from the others.

To conclude, while our results refer to online platform workers, it is worth pointing out that the external validity of our estimates is limited by the nature of the crowdwork platforms themselves and, while our conclusions may be extended to other routine-task intensive platforms such as Crowdfunder or Clickworker, they may not hold in other contexts in which more diversified tasks, requiring specific skills and creative input from service providers, are offered, such as in the case of online freelance marketplaces like Upwork. In these cases, our analysis should be complemented by specific research on these platforms.

The validity of our considerations also rests within the static nature of our analysis. On the theoretical side, while our labor supply model could be extended to a multi-period setting with relative ease, the magnitude of the platform effect estimated empirically might change over time. Average salary conditions in crowdwork might change because factors of supply and demand change with time. As we have discussed, as more tasks performed in traditional occupations become offshorable, demand for crowdwork might increase, and at the same time, labor supply might vary as a result of traditional labor market shocks all over the world. Nonetheless, unless the nature of the tasks performed on micro-task platforms changes radically over the years, these considerations do not invalidate our findings on the nature of skill-biased self-selection into crowd employment.

Supplementary data

[Supplementary materials](#) are available at *Industrial and Corporate Change* online.

Acknowledgments

The authors would like to thank Janine Berg and the International Labour Organization for providing access to the survey data from Berg (2015) and Berg *et al.* (2018). The authors are also grateful to the participants in the 2nd IZA/OECD Workshop "Labor Productivity and the Digital Economy," the AIEL XXXIII National Conference of Labour Economics, the BOMOPAV Economics

Meeting 2019, and the 2nd IDSC of IZA/CAIS Workshop “Matching Workers and Jobs Online” for their useful comments. The authors would like to thank Alessandro Franconi for his suggestions and advice.

Funding

This work was supported by the University of Modena and Reggio Emilia [UNIMORE FAR GRANT 2016 “Gender, Family and Culture, across Time and Space”]. We wish to confirm the funding source did not have any involvement in study design, in the collection, analysis and interpretation of data, in the writing of the report, and in the decision to submit the article for publication.

References

- Adams, A. (2020), ‘The gender wage gap on an online labour market: the cost of interruptions,’ *CEPR Discussion Paper Series 14294*. Centre for Economic Policy Research.
- Altonji, J. G. and R. M. Blank (1999), ‘Race and gender in the labor market,’ in O. Ashenfelter and D. Card (eds), *Handbook of Labor Economics*, vol. 3, part C. Elsevier: Amsterdam, pp. 3143–3259.
- Altonji, J. G., T. E. Elder and C. R. Taber (2005), ‘An evaluation of instrumental variable strategies for estimating the effects of catholic schooling,’ *The Journal of Human Resources*, XL(4), 791–821.
- Angrist, J. and V. Lavy (2009), ‘The effects of high stakes high school achievement awards: evidence from a randomized trial,’ *American Economic Review*, 99(4), 1384–1414.
- Angrist, J. D. and A. B. Krueger (1991), ‘Does compulsory school attendance affect schooling and earnings?,’ *The Quarterly Journal of Economics*, 106(4), 979–1014.
- Angrist, J. D. and A. B. Krueger (1995), ‘Split-sample instrumental variables estimates of the return to schooling,’ *Journal of Business & Economic Statistics*, 13(2), 225.
- Autor, D. H. and D. Dorn (2013), ‘The growth of low-skill service jobs and the polarization of the US labor market,’ *American Economic Review*, 103(5), 1553–1597.
- Berg, J. (2015), ‘Income security in the on-demand economy: findings and policy lessons from a survey of crowdworkers,’ *Comparative Labor Law and Policy Journal*, 37(543), 543–576.
- Berg, J., M. Furrer, E. Harmon, U. Rani and M. S. Silberman (2018), *Digital Labour Platforms and the Future of Work: Towards Decent Work in the Online World*. International Labour Office: Geneva.
- Berinsky, A. J., G. A. Huber and G. S. Lenz (2012), ‘Evaluating online labor markets for experimental research: Amazon.com’s Mechanical Turk,’ *Political Analysis*, 20(3), 351–368.
- Bick, A. (2016), ‘The quantitative role of child care for female labor force participation and fertility,’ *Journal of the European Economic Association*, 14(3), 639–668.
- Blundell, R. and T. MaCurdy (1999), ‘Labor supply: a review of alternative approaches,’ in O. Ashenfelter and D. Card (eds), *Handbook of Labor Economics*, vol. 3, part C. Elsevier: Amsterdam, pp. 1559–1695.
- Bogliacino, F., C. Codagnone, V. Cirillo and D. Guarascio (2020), ‘Quantity and quality of work in the platform economy,’ in K.F. Zimmermann (ed.), *Handbook of Labor, Human Resources and Population Economics*. Springer International Publishing: Berlin, pp. 1–28.
- Borchert, K., M. Hirth, M. Kummer, U. Laitenberger, O. Slivko and S. Viete (2018), ‘Unemployment and online labor,’ *ZEW Discussion Paper No.*, 18–023.
- Bound, J. and D. A. Jaeger (2000), *Do Compulsory School Attendance Laws Alone Explain the Association between Quarter of Birth and Earnings?*, Vol. 19. Emerald Group Publishing Limited: Bingley, pp. 83–108.
- Card, D. and A. B. Krueger (1992), ‘Does school quality matter? Returns to education and the characteristics of public schools in the united states,’ *Journal of Political Economy*, 100(1), 1–40.
- Chang, Y., S.-B. Kim, K. Kwon and R. Rogerson (2011), ‘Interpreting labor supply regressions in a model of full- and part-time work,’ *The American Economic Review*, 101(3), 476–481.
- Ciani, E. (2012), ‘Informal adult care and caregivers’ employment in Europe,’ *Labour Economics*, 19(2), 155–164.
- CIPD (2017), ‘To gig or not to gig? stories from the modern economy,’ Survey report, Chartered Institute of Personnel and Development.
- Collins, B., A. Garin, E. Jackson, D. Kousta and M. Payne. (2019), ‘Is gig work replacing traditional employment? Evidence from two decades of tax returns,’ *Unpublished paper, IRS SOI Joint Statistical Research Program*.
- Dagsvik, J. K., Z. Jia, T. Kornstad and T. O. Thoresen (2013), ‘Theoretical and practical arguments for modelling labor supply as a choice among latent jobs,’ *Journal of Economic Surveys*, 28(1), 134–151.
- De Stefano, V. (2016), ‘The rise of the just-in-time workforce: on-demand work, crowdwork, and labor protection in the gig-economy,’ *ILO Conditions of Work and Employment Series*, No. 71.
- De Stefano, V. (2018), *European Legal Framework for “Digital Labour Platforms”*. European Commission. Joint Research Centre.

- Degryse, C. (2016), 'Digitalisation of the economy and its impact on labour markets,' *ETUI Research Paper—Working Paper* 2016.02.
- Dehejia, R. H. and S. Wahba (1999), 'Causal effects in nonexperimental studies: reevaluating the evaluation of training programs,' *Journal of the American Statistical Association*, 94(448), 1053–1062.
- Difallah, D., E. Filatova and P. Ipeirotis (2018), 'Demographics and dynamics of Mechanical Turk workers,' *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining*. ACM Press.
- Donald, S. G. and K. Lang (2007), 'Inference with difference-in-differences and other panel data,' *The Review of Economics and Statistics*, 89(2), 221–233.
- Donovan, S. A., D. H. Bradley and J. O. Shimabukuru (2016), 'What does the gig economy mean for workers?,' *CRS Report R44365*. Washington, DC: Congressional Research Service. <https://crsreports.congress.gov/Home/About>
- Drahokoupil, J. and A. Piasna (2017), 'Work in the platform economy: beyond lower transaction costs,' *Intereconomics: Review of European Economic Policy*, 52(6), 335–340.
- Dube, A., J. Jacobs, S. Naidu and S. Suri (2020), 'Monopsony in online labor markets,' *American Economic Review: Insights*, 2(1), 33–46.
- Ermisch, J. F. and R. E. Wright (1993), 'Wage offers and full-time and part-time employment by British women,' *The Journal of Human Resources*, 28(1), 111–133.
- Eurofound (2015), *New Forms of Employment*. Luxembourg: Publication Office of the European Union. <https://www.eurofound.europa.eu/about-eurofound/who-we-are/legal-information>
- Eurofound (2017), 'European working conditions survey, 2015,' *UK Data Service*, doi:10.5255/UKDA-SN-8098–4.
- Farrell, D. and F. E. Greig (2017), *The Online Platform Economy: Has Growth Peaked?* Jp Morgan Chase & Co. Institute: New York.
- Goos, M., A. Manning and A. Salomons (2014), 'Explaining job polarization: routine-biased technological change and offshoring,' *American Economic Review*, 104(8), 2509–2526.
- Hara, K., A. Adams, K. Milland, S. Savage, C. Callison-Burch and J. P. Bigham (2018), 'A data-driven analysis of workers' earnings on Amazon Mechanical Turk,' *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, CHI '18*. New York, NY, USA: ACM, pp. 1–449:14.
- Harris, S. D. and A. B. Krueger. (2015), 'A proposal for modernizing labor laws for twenty-first-century work: The "independent worker",' *The Hamilton Project Discussion Papers 2015 2*.
- Hersch, J. and L. S. Stratton (2000), 'Household specialization and the male marriage wage premium,' *ILR Review*, 54(1), 78–94.
- Hirsch, B. T. and D. A. Macpherson (2003), 'Union membership and coverage database from the current population survey: note,' *Industrial and Labor Relations Review*, 56(2), 349–354.
- Horton, J. J. and L. B. Chilton (2010), 'The labor economics of paid crowdsourcing,' *Proceedings of the 11th ACM Conference on Electronic Commerce, EC '10*. New York, NY, USA: Association for Computing Machinery, pp. 209–218.
- Horton, J. J., D. G. Rand and R. J. Zeckhauser (2011), 'The online laboratory: conducting experiments in a real labor market,' *Experimental Economics*, 14(3), 399–425.
- Hotchkiss, J. L. (1991), 'The definition of part-time employment: a switching regression model with unknown sample selection,' *International Economic Review*, 32(4), 899–917.
- Huws, U., N. H. Spencer, D. S. Syrdal and K. Holts (2017), *Work in the Gig Economy - Employment in the Era of Online Platforms. FEPS Studies*. Brussels: Foundation for European Progressive Studies.
- International Labor Office (2013), *International Standard Classification of Occupations 2008 (ISCO-08), Structure, Group Definitions and Correspondence Tables*. International Labor Office: Geneva.
- Jackson, E. (2019), 'Availability of the gig economy and long run labor supply effects for the unemployed,' Working Paper.
- Jeon, S.-H. and Y. Ostrovsky (2018), 'Balancing family and work: transition to self-employment among new mothers,' *Oxford Economic Papers*, 71(1), 47–72.
- Katz, L. F. and A. B. Krueger (2018), 'The rise and nature of alternative work arrangements in the United States, 1995–2015,' *ILR Review*, 72(2), 382–416.
- Kenney, M., P. Rouvinen, T. Seppälä and J. Zysman (2019), 'Platforms and industrial change,' *Industry and Innovation*, 26(8), 871–879.
- Kenney, M. and J. Zysman (2016), 'The rise of the platform economy,' *Issues in Science and Technology*, 32(3), 61–69.
- Kenney, M. and J. Zysman (2019), 'Work and value creation in the platform economy,' in S.P. Vallas and A. Kovalainen (eds), *Work and Labor in the Digital Age (Research in the Sociology of Work, Vol. 33)*. Emerald Publishing Limited: Bingley, pp. 13–41.
- Leigh, A. (2010), 'Informal care and labor market participation,' *Labour Economics*, 17(1), 140–149.
- Lepanjuuri, K., R. Wishart and P. Cornick. (2018), 'The characteristics of those in the gig economy,' Final report, UK Department for Business, Energy & Industrial Strategy.
- Maestas, N., K. J. Mullen, D. Powell, T. von Wachter and J. B. Wenger (2017), *The American Working Conditions Survey Data: Codebook and Data Description*. Santa Monica, CA: RAND Corporation.
- OECD (2016), 'New forms of work in the digital economy,' *OECD Digital Economy Papers 260*, Paris: OECD Publishing.

- OECD (2018), 'Online platforms: a practical approach to their economic and social impacts,' DSTI/CDEP(2018)5. Paris: OECD Publishing.
- Paolacci, G., J. Chandler and P. G. Ipeirotis (2010), 'Running experiments on Amazon Mechanical Turk,' *Judgment and Decision Making*, 5, 411–419.
- Peetz, D. (2019), *The Realities and Futures of Work*. Australian National University Press: Canberra.
- Pesole, A., M. Urzì Brancati, E. Fernández-Macías, F. Biagi and I. González Vázquez (2018), 'Platform workers in Europe. Evidence from the COLLEEM survey,' *Eur - Scientific and Technical Research Reports*, Joint Research Centre, European Commission.
- Prassl, J. and M. Risak (2015), 'Uber, Taskrabbit, and co.: platforms as employers-rethinking the legal analysis of crowdwork,' *Comparative Labor Law and Policy Journal*, 37(619), 619–652.
- Rani, U. and M. Furrer (2019), 'On-demand digital economy: can experience ensure work and income security for microtask workers?,' *Jahrbücher für Nationalökonomie und Statistik*, 239(3), 565–597.
- Sanderson, E. and F. Windmeijer (2016), 'A weak instrument test in linear IV models with multiple endogenous variables,' *Journal of Econometrics*, 190(2), 212–221.
- Schmitz, H. and M. Westphal (2017), 'Informal care and long-term labor market outcomes,' *Journal of Health Economics*, 56, 1–18.
- Wooldridge, J. (2007), 'Inverse probability weighted estimation for general missing data problems,' *Journal of Econometrics*, 141(2), 1281–1301.

Appendix

A. Labor supply in crowd employment

The estimation the wage premium for traditional employment against crowdwork requires a prior reflection over the nature of labor supply in crowd employment. So far, the theory literature has focused on the ex-post features of crowdwork (Horton and Chilton, 2010; Dube et al., 2020), and little attention has been given to workers' decisions to enter these markets, with the exception of Borchert et al. (2018) who, however, only focus on exogenous determinants of participation.

To understand the role of endogenous determinants better, we develop a theoretical framework based on the standard neoclassical model of labor and leisure choice (for an overview, see Blundell and Macurdy, 1999), which can motivate our empirical strategy and provide an intuition of our approach.¹⁶

We begin by assuming that each worker has a utility function $U(C, L, S)$, deriving utility from consumption C , leisure L and individual preferences S .

Utility is assumed to be maximized subject to the budget constraint:

$$C_{it} = w_{it}T - w_{it}L_{it} + I_{it} \quad (6)$$

where w_{it} is the hourly wage at time t for individual i , C_{it} is consumption, L_{it} is the number of hours spent on leisure or other non-paid activities, T is the total amount of hours to be allocated, and I_{it} is unearned income. We assume w_{it} to be determined by a fixed component, i.e. the cost of labor per unit of work, and an idiosyncratic productivity shock. Empirically, earnings are usually modeled through the Mincerian earnings function, which models their logarithm as a function of the vector of schooling, experience and ability X'_{it} , and institutional factors ψ_s :

$$\log(w_{it}) = P(Z_{it}, R(X_{it}, t))' \beta + X'_{it} \gamma + \psi_s \quad (7)$$

- 16 Our theoretical framework is similar in spirit to the one for part-time work proposed by Chang et al. (2011). Due to the cross-sectional nature of our data, we focus our theoretical overview on a static, within-period equilibrium model for labor supply. Nonetheless, our intuitions can easily be transposed to a life cycle setting with Frisch labor supply equations. Furthermore, while we focus on individual choice, holding household income from other members as fixed unearned income, the model can easily be adapted for family labor supply model. Finally, the nature of crowdwork requires a continuous job supply model. As crowd workers can freely adjust their labor supply, discrete choice models between bundles of hours do not apply, and standard theory can offer a more intuitive theoretical outlook of crowd employment. While our model focuses on micro-task crowd employment, our analysis can be extended to other platforms as well, either by offering distinct budget constraints for each platform, or by tuning the wage equation to group all platforms together.

We model salary differences between forms of work through the β coefficients vector. The choice function $P(Z_{it}, R(X_{it}, t))$ states final choice for a specific form of work, and outputs a vector of dichotomous variables for traditional work, crowdwork, and traditional work with auxiliary earnings from crowdwork.¹⁷

This function depends on two factors, the first being preference for certain job characteristics, such as the need to work in atypical or intermittent hours, or to work from home, which we model as the stochastic shock Z_i .

The second one is the restriction function $R(X_{it}, t)$ (as in Dagsvik *et al.*, 2013) which determines the set of accessible forms of work at time t depending on the worker's investment in human capital and ability. Crowdwork features no barriers to entry, whereas the option to access traditional work may not always be available.¹⁸ We will define as job-restricted the workers who cannot access traditional forms of employment. $R(X_{it}, t)$ is unobserved to the econometrician so that, while the output of the restriction function is revealed for traditional workers, it is not possible to determine whether access to crowd employment resulted from a deliberate choice or from the lack of it.

Individuals in traditional employment, be it part-time or full-time, are rarely able to allocate all their available hours to work, and usually need to allocate a minimum number of hours to work. We then assume that time allocated to leisure L_i within traditional employment is bounded by the support $L_{max} \geq L_i > L_{min}$. Instead, crowdworkers can allocate as much time as they want to work.¹⁹

Assuming that a worker will take the highest paying occupation for the utility-maximizing number of hours allocated to leisure, for $L_i > L_{max}$ the worker will choose crowdwork (or not to work) as no traditional jobs are available below this number of work hours, for $L_{max} \geq L_i > L_{min}$ the individual will work in traditional occupations and not crowdwork, while for $L_{min} \geq L_i$ the worker will work in a traditional occupation and seek additional earnings through crowdwork.²⁰

For individuals with access to traditional employment, three budget constraints (for clarity, the time subscript is omitted) are available:

$$\begin{aligned} C_{CWi} &= w_{CWi}(T - L_i) + I_i \text{ for } L_i > L_{max} \\ C_{TRADi} &= w_{TRADi}(T - L_i) + I_i \text{ for } L_{max} \geq L_i > L_{min} \\ C_{AUXi} &= w_{CWi}(T - L_i) + w_{TRADi}(T - L_{min}) + I_i \text{ for } L_{min} \geq L_i \end{aligned} \quad (8)$$

so that C_{CWi} , C_{TRADi} , and C_{AUXi} , define budget constraints for employment in crowdwork, traditional work, and traditional work with auxiliary earnings from crowdwork.

For job-restricted individuals, instead, only the first of the three budget constraints from (8) is available for any value of L_i . The agent maximizes his/her utility by first allocating leisure subject to each of the available budget constraints.

For example, for crowdwork, the optimization problem is solved by the Lagrangian:

$$\ell_{CWi}(L_i) = U(C, L, S) + \lambda(w_{CWi}(T - L_i) + I_i - C_{CWi}) \quad (9)$$

Other value functions are obtained in a similar way. Solving for each occupational choice, then:

- 17 For ease of tractability, this will be our preferred specification, but differences in earnings between traditional work and crowdwork can also be modeled in other ways: for example, assuming linear additivity, estimation procedures based on the Blinder-Oaxaca decomposition can be considered.
- 18 Restrictions to traditional markets do not necessarily correlate with low-ability but are also linked to time-varying factors reflecting saturation within skill groups in the labor market at time t . These factors reflect business cycle fluctuations in markets and could also be related to ongoing processes of job polarization, or to changes in employment legislation.
- 19 We include time spent searching for tasks in the hours of work dedicated to crowdwork. If there are few tasks on the demand side, a worker might still spend their full leisure budget looking for tasks. Differences in task availability conditional on hour of the day, of day of the week, could be modeled through non-linear budget constraints. We avoid this for clarity purposes.
- 20 For simplicity, we hold L_{min} and L_{max} as fixed, but they might be allowed to vary depending on $R(X_i)$ under different conditions.

$$U(C, L, S, P) = \max_{P_i} \{U_{CW_i}, U_{TRAD_i}, U_{AUX_i}\} \quad (10)$$

Workers will choose the value of $P(i, t)$, i.e. the occupational choice, which maximizes their utility $U(C, L, S, P)$ at time t .

Access into crowdwork markets is explained in a variety of ways, holding Z_i as fixed. For w_i low enough, individuals with flat indifference curves might enter crowdwork because they seek auxiliary earnings, or be indifferent to whether to enter the traditional workforce or work through online labor platforms. Instead, individuals with steep indifference curves might enter crowdwork (or not work) in order to dedicate more hours to leisure or other non-paid activities, especially when unearned income is high enough. Job-restricted individuals, instead, will stick to crowdwork no matter their preference for work or leisure.

Together, these issues might explain heterogeneity in the elasticity of labor supply between crowdwork and traditional employment, which is clearly generated within the choice function. However, when final choice is determined by job restrictions, the source of variation in access to crowd employment is endogenous as X_i appears twice in the equation. When participation is affected by Z_i , this source is instead exogenous.

Any comparison in earnings between crowdwork and traditional employment is then bound to generate biased results if these endogenous elements are not considered. If the variation in $R(X_i, t)$ cannot be accounted for by the available set of controls the econometrician must then leverage the variation in Z_i as a source of randomization.

B. Endogenous caregiving and violations of the exclusion restriction

The chosen instrumental variable could also pose as a threat to our identification strategy in terms of violations of the exclusion restriction.

A first potential concern comes from the fact that caregiving could influence skills, and thus the returns to them. In the short term, however, caregiving should affect the opportunity to work more, not the ability of an individual. Nonetheless, it is reasonable to believe that the amount of time a worker spends in caregiving is endogenous to the wage he/she could earn in the market. While this condition clearly affects working hours and—by extension—total earnings, the effects on hourly earnings are less obvious. If the wage is high enough, individuals could purchase care for either a child or a relative, or the opportunity cost of time might be higher. In such a case, transition to caregiving will be biased towards lower salaries.

To account for this endogenous variation properly, we would need to have access to a measure of unobserved potential earnings, that is, the hourly salary a worker would earn before being engaged in caregiving. Should potential earnings be available, it holds that:

$$E[Y_{1,0}|Y_{0,0}] - E[Y_{1,c}|Y_{0,c}] = \xi \quad (11)$$

This means that the expected value of earnings of workers in caregiving $E[Y_{1,c}]$ would equal the expected value of earnings of individuals not in caregiving $E[Y_{1,0}]$, minus the caregiving bias ξ , once we control for unobserved potential earnings Y_0 .

However, we are only able to observe effective earnings. While we cannot observe potential earnings, we can nonetheless establish a proxy for them through our set of controls X . In such a case, Y_0 can then be replaced by our set of controls X . This should also allow us to filter out other endogenous aspects of caregiving linked, with individual and household characteristics, and, most importantly, occupation. Again, the fact that we can control for the last non-crowdwork occupation—and the time spent in it—held by each respondent provides an excellent proxy for potential earnings.

As shown in Section 5.3, our findings suggest that hourly earnings in traditional work are unaffected by caregiving after interacting it with other determinants of observable skill (see Table 6). This suggests that not only this bias is absent, but as long as there is no reason to suspect ξ to be positive, that the chosen set of controls correctly proxies for potential earnings as in equation (11).

In other words, while access to caregiving might be related to the inability to purchase formal care, our set of controls should correctly predict potential earnings and appease concerns concerning the violation of the exclusion restriction, given that the effect of caregiving on the hourly rate of pay appears indifferent to the level of observable skill, suggesting that the endogenous components of caregiving should be fully absorbed by our set of covariates.

C. Labor market institutions and participation in crowd employment

So far, the inclusion of federal and member states fixed effects in our estimates controls for the effects of cross-country and institutional heterogeneity on participation in crowd employment and on salaries.

Following from Donald and Lang (2007) and Angrist and Lavy (2009), it can be shown how these group controls absorb the effect of state-level variation in institutional factors. By assumption, the presence of institutional variation across federal and member states implies the residual structure of the model to follow the group structure $e_{is} = \epsilon_{is} + \nu_s$. Suppose that we wanted to capture the effect of a vector of group-level predictors K' on hourly earnings. Updating our empirical wage equation, we obtain:

$$Y_{is} = \alpha + Crowdwork_{is}\beta + X'_{is}\gamma + K'_s\delta + e_{is} \quad (12)$$

As $\psi_s = \alpha + K'_s\delta + \nu_s$, this regression model is functionally equivalent to the one presented in equation (1). With the addition of state-level standard error clustering, the grouped structure of the error component is also taken into account. This property of the fixed effect estimator can be shown to hold for the instrumental variable estimates from equations (3) and (4). For similar reasons, institutional factors were also held fixed in the discussion from Appendix A.

While this specification is useful for the estimation of the 'Working in crowdwork' coefficient, the researcher might be interested in estimating the effects of state-specific institutional effects captured by the vector of coefficients δ . Indeed, these parameters can reveal important information concerning the effects of institutions on participation in online labor markets, as the ψ_s will also capture other confounding factors unrelated to labor market institutions.

To assess the role of these institutional factors, the two-step estimation method developed in Donald and Lang (2007) can be used by estimating the individual level equation, from which the estimated fixed effects $\hat{\psi}_s$ are extracted and then regressed over the group-level predictors, as follows:

$$\hat{\psi}_s = \alpha + K'_s\delta + (\nu_s + (\hat{\psi}_s - \psi_s)) \quad (13)$$

This process can be followed for both first and second stage equations. We begin by fitting ψ_{1s} , finding which institutional factors affect participation in crowdwork. We include, in the institutional predictors vector K'_s , the following state-level variables: hourly minimum wage, defined as PPP-adjusted real hourly earnings; paid parental leave, in weeks; union density, defined as the proportion of union members over the total number of employees; collective bargaining coverage, defined as the proportion of employees with the collective right to bargain; unemployment replacement rate over previous income; and strictness of employment protection.²¹

Results are shown in Table C1 (columns 1 to 4) and suggest that strictness of job protection is the greatest institutional predictor for participation in crowd employment after controlling for all individual determinants as in Section 4.2.2. Standardized beta coefficients are shown in column (2) so that the magnitude of the coefficient is directly comparable. Paid parental leave also appears significantly correlated with selection in crowdwork, with the results suggesting that each additional paid week of leave reduces the probability of entering crowd employment by 0.1%. Other institutional predictors are not statistically significant.

These results are largely unaffected when participation in crowd employment is studied across all workforce (columns 3 and 4), and, in general, support the interpretation that crowd employment can be seen as an alternative to employment in context where job security is low and parental leave is limited. This finding does not contrast our identification strategy, but rather enhances it.

It is important to remark that these estimates are far from causal, and state-specific endogenous factors might, in turn, influence parental leave and the level of job protection. These factors cannot be addressed in the model estimated with equation (13) but are nonetheless controlled for in the individual-level regressions though the parameter ψ_s . Recall that all these institutional factors which can affect the decision to enter crowdwork are arguably outside of the control of the individual, so we believe the nature of these heterogeneities to be exogenous to the worker, meaning that most endogeneity concerns can be left aside for what concerns the individual-level analysis. Also, the usefulness of this analysis is limited to the study of institutional determinants for participation in crowd employment.

21 Sources: Minimum wages—author's calculations based on data from US Department of labor and Eurostat; trade union density and collective bargaining coverage—Bureau of labor statistics, ILO, and Hirsch and Macpherson (2003); Paid parental leave—OECD, SSA; Replacement rate—US Department of labor and OECD; Job protection—OECD.

Table C1 Institutional determinants of participation and salaries in crowdwork

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Participation in CW				Salary differentials			
	Employed		In workforce		Hourly earnings		Hourly earnings [†]	
	OLS	Beta	OLS	Beta	OLS	Beta	OLS	Beta
Union density	−0.001 (0.001)	−0.067	−0.001 (0.001)	−0.067	0.005 (0.029)	0.024	−0.005 (0.025)	−0.024
Collective agreement coverage	−0.000 (0.001)	−0.008	−0.000 (0.001)	−0.010	−0.000 (0.011)	−0.003	0.000 (0.010)	0.004
Hourly minimum wage (USD)	−0.007 (0.010)	−0.066	−0.007 (0.010)	−0.061	0.186 (0.312)	0.138	0.173 (0.282)	0.139
Hourly minimum wage set with collective agreement	0.012 (0.045)	0.015	0.014 (0.048)	0.018	1.413 (1.121)	0.124	1.599 (1.020)	0.152
Paid parental leave (weeks)	−0.001*** (0.000)	−0.152	−0.001*** (0.000)	−0.153	−0.001 (0.011)	−0.017	−0.002 (0.010)	−0.036
Strictness of job protection	−0.106*** (0.018)	−0.547	−0.112*** (0.019)	−0.559	−0.912** (0.420)	−0.394	−0.831** (0.388)	−0.389
Net replacement rate in unemployment	−0.000 (0.001)	−0.005	−0.000 (0.001)	−0.007	0.015 (0.037)	0.072	0.018 (0.034)	0.094
Observations	78		78		64		64	
Adjusted R-squared	0.467		0.491		0.467		0.491	
Missing values controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Robust standard errors in parentheses. [†]Adjusted for unpaid tasks.
P* < 0.01; *P* < 0.001.

Repeating this exercise for ψ_{2s} is trivial, but said results would not be indicative of how institutional factors affect crowdworkers’ conditions, as ψ_{2s} averages the group-effect overall earnings, traditional workers included.

Instead, we test whether crowdworkers experience significantly lower earnings than traditional workers conditional on the quality of labor market institutions by estimating [equation \(12\)](#) separately for crowdworkers and traditional workers, then extracting the state fixed effects ψ_{2t} and ψ_{2c} , and estimating [equation \(13\)](#) replacing its left side with the difference between the two estimated group effects.

The results are displayed in [Table C1](#) (columns 5 to 8). Strictness of job protection appears to be the only statistically significant predictor for the earnings differential, meaning that the difference in earnings is greatly reduced with stricter labor markets. This might consistent with the interpretation that less protected job markets generate a stronger skill-biased self-selection into crowdwork, but these results might more simply correlate with the possibility that crowd employment markets grow faster in markets with lower protection. Overall, however, the lack of significance of the other coefficients confirms the suspicion that traditional labor market regulations are generally unable to regulate online labor markets, especially when considering how these markets operate on a global scale.