THE LABOR MARKET EFFECTS OF CROWDWORK IN INDIA

By
Nóra Gőbel

Submitted to
Central European University
Department of Economics and Business

In partial fulfilment of the requirements for the degree of Master of Arts in Economic Policy in Global Markets

Supervisor: Yusaf Akbar
External supervisor: Uma Rani Amara

Budapest, Hungary

2019

Abstract

Crowdwork refers to cases where the task is outsourced to the crowd, an undefined group of people via online labor platforms instead of a specific individual. The possibility of flexible work in terms of location and time is appealing to the workers, however, previous studies show a more negative impact of this work including low wages, skill mismatch, limited flexibility, lack of transparency, asymmetric access to data as well as the lack of regulation and social protection. Negative effects of crowdwork on wages have been measured in the USA and in the EU, but very few studies have looked at the labor market effects in the context of developing economies, and especially at the country level.

Based on the ILO Crowdwork survey and the NSSO Periodic Labor Force Survey in India, this thesis compares the hourly wages of workers in digital labor platforms to those with similar characteristics in the traditional labor force and investigates whether working on microtask platforms has any effect on the workers' earnings and working conditions. The results of the OLS and 2SLS models suggest that the microtask platforms lead to lower wages and more precarious working conditions in the country in 2017 when domestic work is used as an instrumental variable. This wage gap is even more prevalent among older workers as well as among crowdworkers who hold an additional job.

Table of Contents

List of Tables

List of Figures

1. Introduction

The implications of digitalization on the labor market and the economy has sparked the interest of researchers as it is fundamentally changing the way work is organized. Businesses can access broad set of skills at low cost with the help of the online labor platforms while platform providers benefit as intermediaries, shifting the risk to other parties, often to the workers themselves. Crowdwork, a new form of work on digital labor platforms should be understood as part of the transformations that are happening today on the labor market including the shift towards more atomized hiring and management processes as well as more precarious and informal labor.

Scholars refer to crowdwork as "*human computation*" (Schmidt, 2017), "*invisible work*" (Irani and Silberman, 2013), "*human-cloud*" (De Stefano and Aloisi, 2018), "*humans-as-a-service*" (Irani and Silberman, 2013), "*human intelligence work*" (Kässi and Lehdonvirta, 2018a). Most of the expressions capture the human factor of crowdwork that is mediated through technology. There is a consensus among authors that crowdwork refers to cases where the task is outsourced to the crowd, an undefined group of people via online labor platforms instead of a specific individual (Berg et al., 2018; De Stefano, 2015; Schmidt, 2017). The crowd, that consists of physically dispersed labor, undertakes tasks which are of a very short durations ranging from a couple of minutes to few hours allowing firms to "*hire*" workers only for such a short period of time and reward them with low remunerations (Berg et al., 2018).

On the positive side, this form of work creates flexible work arrangements for individuals especially for those otherwise excluded from the labor market, so that they can earn additional income or access the labor market. On the negative side, though, the lack of regulations leads to more precarious, informal work conditions and creates new challenges for policy makers.

Understanding and capturing the challenges and impacts of crowdwork is highly relevant since it can help policymakers address the difficulties that are already present or may arise in the future when these forms of work become more rampant.

There is an ambivalence of innovation and policy making. Given the high number of positive impact that innovation has brought in different areas, policies should foster an environment where innovation is incentivized. Existing legislative frameworks, in some areas, are also difficult and takes a long time to change. In such a context, digital labor platforms can grow and avoid regulation. However, policies should also ensure that digitalization does not undermine basic rights, fairness, social protection or equal opportunities. It is argued that technological change has brought more "*creation*" than "*destruction*" in the labor markets so far, based on the suggestion that it has increased productivity and created more new employment opportunities as it has abolished. However, the net impact of digitalization will depend on the level of development and digital readiness of countries as well as the policies adopted and implemented at different levels (Aubert-Tarby et al., 2018; UNCTAD, 2019).

It is hard to estimate the size of the platform economy since there is no official data published by the platforms themselves. Based on the figures estimated by previous studies, it is not an overstatement that this form of work has been gaining importance and exhibited such a growing tendency that policy makers should not overlook (Behrendt et al., 2019). As of 2016, around 300 online labor platforms operated around the world (Evans and Gawer, 2016). The number of platform workers in 2014 was estimated to be approximately 45 million only in the developed world (Florisson and Mandl, 2018), while based on the Oxford Internet Institute's estimation, there are about 46 million registered workers on 142 freelance platforms globally (2019)

The impacts of this form of work is not yet fully understood; thus, there is a growing interest among those researching this particular topic to explore different aspects of it. Several authors have contributed to the classification and typology of labor platforms (Durward et al., 2016;

Florisson and Mandl, 2018; Howcroft and Bergvall-Kåreborn, 2019; Nischak et al., 2017; Schenk and Guittard, 2011; Schmidt, 2017; Seppälä et al., 2015), the economics of the new platform economy (Adebambo and Bliss, 2015; Chen, 2017; Evans et al., 2018; Horton and Chilton, 2010; Jacques and Kristensson, 2019; Mason and Watts, 2009; Shafiei Gol et al., 2019; Spulber, 2019), legal aspects and the need of regulation (Adams et al., 2015; De Stefano, 2016, 2015; Felstiner, 2011; Körfer and Röthig, 2017; Todolí-Signes, 2017) as well as social protection (Behrendt et al., 2019; De Stefano, 2015; Forde et al., n.d.). There is a growing number of descriptive analysis of crowdwork on a global scale including workers' characteristics, motivation and working conditions (Berg et al., 2018; Brewer et al., 2016; Kaufmann et al., 2011; Machine and Ophoff, 2014; Naderi, 2018; Seppälä et al., 2015; Vakharia and Lease, 2013; Zyskowski et al., 2015).

However, there are other areas where the literature is very scant such as the impact on workers' skills development and career prospects (Rani and Furrer, 2019). Furthermore, there is only a limited number of studies investigating the impacts of work through digital labor platforms in developing countries (Graham et al., 2017; Rani and Singh, 2019; UNCTAD, 2019). Finally, the number of studies that attempt to measure the labor market impacts on the traditional sectors as well as the scale of the increase of labor market platforms is even more limited (Cantarella and Strozzi, 2018; Jacques and Kristensson, 2019) due to limited data availability and the mismatch between conventional economic indicators and work via online platforms.

This thesis seeks to fill this gap and explore the unique characteristics of crowdworkers in India as well as the effects of undertaking crowdwork on wages and working conditions. The research questions of this thesis are the following: 1) What is the nature of the relationship between the hourly earnings of workers undertaking similar work on microtask platforms and the traditional labor market in India? 2) Does crowdwork as a new form of work affect hourly wages and working conditions in the Indian labor market?

In order to answer these questions, the paper offers detailed literature review which is necessary to understand the relevance of the topic and the operation of digital labor platform processes that serve as potential mechanisms in the hypothetical causal relationship between performing crowdwork and hourly earnings. Based on the literature review and the conceptual framework, the hypothesis is that workers have more precarious working conditions as well as lower wages in the microtask platforms than workers with similar characteristics performing similar work in the traditional labor market in India. The dataset for the analysis comes from the ILO Crowdwork survey and the NSSO Periodic Labour Force Survey in India.

Since workers in the traditional labor force form a heterogeneous group, the analysis is restricted to the information services sector on the assumption that these workers are more similar to crowdworkers in terms of their skills and characteristics. The tasks of crowdworkers based on ILO's Digital labor platforms report (Berg et al., 2018) were matched with most similar traditional tasks in the information technology sector and identified using the National Industrial Classification (Ministry of Statistics and Programme Implementation 2008).

The descriptive analysis shed light on the characteristics of workers in the information technology traditional sectors and crowdworkers in India and capture similarities and differences between groups. The OLS model measures the relationship between hourly wages and being engaged in crowdwork activities which is further investigated in the 2SLS model using domestic work as instrumental variable.

The paper is structured as follows. The first part of the literature review provides a detailed overview of the operation of digital labor platforms and its relevance to the labor markets. This includes matching demand and supply, algorithmic management, working conditions and information asymmetry. The second part of the literature review summarizes some of the relevant experiences from crowdwork focusing on Indian workers. The following chapter places the literature on digital labor platforms to the context of labor market economics in order to develop the main hypotheses. The fourth chapter introduces the dataset and the methodology used for the analysis, and the results of the descriptive analysis, OLS and 2SLS models are presented in chapter 5. Finally, based on the findings, policy recommendations are given in the context of the country.

2. Literature review

2.1 Breaking down crowdwork

2.1.1 The work: transformation of work and the emergence of platform economy and its new challenges

The paradigm of “*work*“ being a location-based activity is challenged by new forms of work arrangement that have been emerging as a result of the penetration of different types of digital technologies including atomization, artificial intelligence (AI), the internet of things (IoT), cloud computing as well as new business models such as digital platforms (UNCTAD, 2019). The extent to which digital labor platforms represent a new form of work is still debated. It can be argued that the work itself is the same, thus, it is too radical to talk about a new form of work. This is only true in a sense that the tasks themselves such as transcription or data cleaning are similar to traditional form of work. What is fundamentally new is that platforms allow business processes to be outsourced without the mediation of formal BPO organizations (Graham et al., 2017).

Digital labor platforms’ contribution to the economy is often claimed to be remarkable as they drive up productivity by highly efficient matching and asset utilization (Evans and Gawer, 2016). It is believed that platforms are efficient in matching of demand and supply not only in the labor market but in other areas such as E-commerce marketplaces or professional networks. In addition to this, platforms have been successfully attracting investment from venture funds and they served as a source of innovation in the US as indicated by the high number of patents (Evans and Gawer, 2016).

Even though the potentials seem bright and boundless, if the current operation of digital platforms reflects the direction where work is evolving to, there is a reason to worry. Previous studies have demonstrated some of the major challenges that have emerged with digital labor platforms such as skill mismatch, low wages, absence of work-life balance, lack of regulation, lack of transparency, monopolization and asymmetric access for data (Berg et al., 2018; Collier et al., 2017; De Stefano, 2015; Rani and Singh, 2019; UNCTAD, 2019).

The new jobs created are more likely to be remote work, temporal, flexible working arrangements, and casual contracts (Aubert-Tarby et al., 2018). Workers on digital labor platforms are not entitled to social security benefit given the lack of regulation on hourly wages and the fact that workers are typically self-employed (Behrendt et al., 2019; Jäger et al., 2019). Another crucial issue which has been arising in the platform economy is its monopolist nature and lack of competition.

Network effects feed the monopolization of the markets since online workers would rather join labor platforms that already have an extended customer circle and well-established system. Furthermore, due to path dependency, the costs to users of switching to an alternative platform is high as they need to build up their profile, reputation and rating (Farrell and Klemperer, 2006). Another contributing factor is the asymmetric access for data. Given their intermediator role, the platforms are able to access and analyze larger amounts of data through interactions. compared to other actors in the market which gives them major competitive advantage over non-platform companies (UNCTAD, 2019).

2.1.2 The crowd

The term "crowdsourcing" which denotes the act of outsourcing work to "the crowd", has been created by combining the words "crowd" and "outsourcing" (Berg et al., 2018). The sociological concept of the crowd has evolved over time (Borch, 2012; Templeton et al., 2015; Wexler, 2011). Wexler (2011) identifies three phrases, through which the crowd develops from an "*irrational herd-like crowd*" to a "*rational social problem generator*" and finally, to a "*problem solver*. In the last and current phrase, the crowd is considered to be completely different than in the previous phrases. For the first time, the crowd can generate useful ideas often with the aid of Web 2.0 technology from which businesses can benefit as well. Howe

(2008) highlights that even though crowdsourcing is intertwined with digitalization, technology is rather the necessary infrastructure allowing these interactions rather than the engine of it. Crowdwork platforms are the digital services that facilitate crowdsourcing. This process of crowdsourcing differs from outsourcing in a way that the latter usually refers to a long-term relationship while the prior results in a short-term project-based contract between unknown individuals (Jinnai, n.d.). Apart from paid tasks, crowdsourcing also refers to unpaid tasks that individuals choose to undertake not necessary for the monetary benefits (Howe, 2008). This behavior seems irrational for conventional economics but looking at it through the lens of sociology may provide valuable insights. Collaboration is a key element in this perspective since the possible intrinsic benefits may arise from the contribution of an individual for a bigger cause or goal or public good.

Even though some scholars agree with the possible intrinsic nature of crowdsourcing, the term 'sharing economy' has been repeatedly criticized. The critics argue that it is a misleading term since it implies positive non-reciprocal social behavior and the increased cooperation between individuals, while in fact the platforms and their services are not involved in sharing in a traditional sense, and involve fee-paying access to goods or assets that individuals often used for economic benefit (Eckhardt and Bardhi 2015; Hamari et al. 2015). Hu (2019) defines sharing economy as a market model that enables and facilitates the sharing of access to goods and services. It is also often described as a "*transformative and disruptive economic model*" since the access to physical goods, assets or services are carried out through rental, sharing or exchange of resources through digital intermediates without any permanent transfer of ownership ((Taeihagh, 2017). Labor platforms crowdsources demands, enabling customers to share the benefit of discounted goods and services (Hu, 2019).

2.2 Novelty in digital labor platforms

Digital labor platforms impact firms in different ways including organization of work, hiring processes, interaction with each other and value creation. Crowdwork, a new form of work is based on crowdsourcing is used as an organization mechanism to reach the physically dispersed crowd and financially compensate them without subcontracting through established firms which leads to increased efficiency at reduced transaction costs (Kittur et al., 2013).

On one side, these platforms provide technical infrastructure for requesters to advertise tasks for the crowd, retrieve and evaluate the results of completed tasks, and pay individual workers. On the other side, they offer services and infrastructure to individuals including a centralized location to identify tasks from different requesters to perform and submit completed tasks and the financial infrastructure to receive payments (Berg et al., 2018).

These digital platforms attract workers with flexible working conditions since individuals can choose the location, time and nature of work. However, as discussed above, they also create new challenges for labor markets and undermine regular employment relations due to the lack of regulation and the rise in informal and non-standard form of work.

The following sections explain in detail what are the new mechanisms that can possibly drive a wage gap between digital and traditional workers. To understand the operation of digital labor platforms, work allocation, value creation, algorithmic management and information asymmetry are discussed in detail. This is essential for the analysis, since apart from the unobservable characteristics, unexplained differences between the wages of crowdwork and comparable conventional occupations occur because of the operation of crowdwork platforms.

2.2.1 Defining digital labor platforms in the digital economy

Crowdwork is also commonly defined as a sub-segment of the platform economy which is commonly mentioned in the literature as gig-economy (De Stefano, 2016; Florisson and Mandl, 2018; Howcroft and Bergvall-Kåreborn, 2019; Schmidt, 2017). Since this is a relatively new, growing field, the terminology often causes confusion. It is important to clarify the meaning of the commonly applied terms since they are not the same across the literature. Due to the inconsistency of definitions, the comparison of previous findings can be difficult as they often refer to different definitions.

The platform economy is typically divided into two sub-categories, the web-based platform often referred to as crowdwork and app-work where local labor is allocated though mobile application. Some authors refer to the categories of platform work as cloud work and gig work (Schmidt 2017), some call the same work crowd work and on-demand work via apps (De Stefano 2015) or online gig platforms and local gig platforms (Kässi and Lehdonvirta, 2018a) while others consider app work as a subcategory of crowdwork (Howcroft and Bergvall-Kåreborn, 2019).

Following ILO's classification, this thesis refers to the two sub-categories of platform economy as 1) crowdwork done on digital labor platforms and 2) locally based platforms (Berg et al., 2018). The two categories of digital platform work differ in the nature of work and the location. App work (like Uber or Airbnb) has to be completed at a specific location and time, by a specific person. The digital sphere serves as a place for work allocation while the work itself is executed in the physical world, thus, it becomes more visible than crowdwork. This kind of work involves traditional activities such as driving, housing or delivery services that are assigned through the mobile applications (Schmidt, 2017). Schmidt (2017) and De Stefano (2015) emphasize the global nature of crowdwork.

Crowdwork can be further divided into different levels depending on the complexity and nature of tasks and relations between the parties that can vary largely (Howcroft and Bergvall-Kåreborn, 2019). Microtasks are defined as task divided into tiny units, each rewarded by equally little amount (Berg et al., 2018). As an alternative, crowdwork can involve bigger, more complex tasks or can be assigned to several individuals at the same time while only one of them is used and paid for which is called contest-based crowdwork (Schmidt, 2017). Berg et al. (2018) point out the common characteristic of microtasks that they are simple tasks rewarded with very low remuneration. These tasks emerged as a result of the need of human intelligence in cases when AI cannot solve particular problems and humans are needed to fill the gap (Berg et al., 2018). Microtask platforms are discussed in the following sub-chapters in more detail as they form the main focus of this paper.

2.2.2 Microtasks on digital labor platforms

Based on five microtask platforms, Amazon Mechanical Turk (AMT), Clickworker, CrowdFlower, Microworkers and Prolific, Berg et al. identified the most common tasks that are completed by crowdworkers including 1) data collection, 2) categorization, 3) verification and validation, 4) content moderation, 5) market research and reviews, 6) artificial intelligence and machine learning, 7) transcription, 8) content creation and editing as well as 9) surveys and experiments (Berg et al., 2018). Even though, some tasks give the illusion that high skilled labor-force is required for it such as machine learning, in reality these tasks require no particular skills. The workload that crowdworkers experience is determined by the task-related attributes such as difficulty, complexity and bad design (Naderi, 2018). The platforms are often specialized to specific tasks such as data cleaning, tagging, sentiment analysis, content creation, research surveys, image tagging, marketing campaign reactions, video quality rating transcription, text production product classification, AI training data, data management and business feedback (Berg et al., 2018). Rani and Furrer (2019) specify two broader categories, 1) AI and machine learning that includes data collection, categorization content moderation, audio and image recoding, verification as well as transcription and 2) promoting products and services including content access, market research/reviews, surveys and experiments, content creation and editing.

2.2.3 Actors, work allocation and value creation on digital labor platforms

The core novelty in platform economy is the work allocation among parties. The digital marketplaces involve at least three parties; thus, at least three stakeholders, namely the clients, workers and the platform providers (Leimeister et al., 2016). The client can be a company, an institution, or an individual that outsources or requests tasks that are undertaken by workers who are mostly individuals. This exchange between supply and demand takes place on digital platforms that provide the infrastructure. According to the definition of Choudary et al.(2016), the platform itself is a business "*based on enabling value-creating interactions between external producers and consumers*" (Choudary et al., 2016) that provides an open, participative infrastructure for these interactions and sets governance conditions for them.

This is a two-sided market where platforms play a more inevitable role than merely intermediator between the parties. The role of platforms is matching supply and demand using algorithms as well as the facilitation of goods, services, or social currency that enables value creation for all participants (Choudary et al., 2016; Florisson and Mandl, 2018; Schmidt, 2017). This matching process requires novel approaches to identify opportunities and address challenges. In contrast to the "*pipeline*" or "*linear value chain*" used for traditional businesses where producers are at one end and consumers at the other, digital platforms allow value to be created, changed, exchanged, and consumed in multiple ways and places. Instead of relying on supply-side economies of scale, digital platforms rely on demand-side economies of scale

(Choudary et al., 2016). The value created by digital platforms is largely dependent on their achieved network effect since they become more attractive for potential users as the number of their members increase due to the increased opportunities (Florisson and Mandl, 2018).

The digital labor platforms reduce transaction costs for firms, and shift costs and risks to other actors in the marketplace, often to workers or to business depending upon the situation.

2.2.3 Algorithm management

In the platform economy, the conventional management processes are replaced by algorithm-based processes. The algorithms are often called invisible since it is not always clear for the users how the algorithms create matches, affect ratings, wages and working conditions (Mäntymäki et al., 2019).When algorithms are not transparent, workers might not fully understand the reasons behind the fluctuation of their earnings, their ratings, the reason their task is rejected or the type of tasks that are allocated to them or they can undertake.

Mantymaki, Baiyere and Islam (2019) conclude that the work relations on digital platforms show unique characteristics compared to traditional workplaces. By applying the nonstandard work theory, they identify digital temporality instead of solely temporal attachment, algorithmic administratively instead of administrative and, finally, virtual proximity instead of physical one. The new form of atomized management assign each individual task to multiple workers and use a *"quorum system to compare and evaluate automatically which responses are 'correct' in case of any disagreements*" (Berg et al., 2018). As a consequence, crowdworkers' performance is evaluated in a form of rating (Gupta et al., 2014). Workers have to improve their rating and build up their reputation to get more tasks on platfroms. In order to do so, they often have to undertake unpaid tasks so that they would increase their ratings and thus, their chances for better tasks.

A serious concern with regards to digital platforms is that pricing and thus the income of the workers does not simply depend on the hours worked, the platforms' *"invisible algorithms*" affect the rates and working time as well (Mäntymäki et al., 2019). However, without ratings, the platforms would not have sufficient information to distinguish between "good" and "bad" workers, and would assume each worker is a bad one which would have detrimental effect on the workers. Reputation systems, thus, are intended to decrease the level of uncertainty the employer experience (Kokkodis et al., 2015), but at the same time, it increases the uncertainty of the workers as the rating system is often not transparent.

Algorithms and the design of the platform impacts the marketplace enormously, including the earnings and working hours through ratings, the ways tasks are created and managed, the available types of tasks as well as workers' access to tasks (Gupta et al., 2014). Furthermore, algorithms also limit workers' flexibility when they penalize task rejection (Berg et al., 2018).

2.2.3 Information asymmetry and power imbalance

Information asymmetry and power imbalance are serious concerns that have been flagged by several previous studies (Baiyere et al., 2019; Berg et al., 2018; Irani and Silberman, 2013; Todolí-Signes, 2017; UNCTAD, 2019). Digital platforms control massive amounts of information about both parties that have no access for such information, maybe not even about themselves (UNCTAD, 2019). Different mechanisms feed the asymmetry including network effects, the three-sided software architecture and the venture capital in case of gig work. As operators lay the infrastructure including the algorithm management, it gives them a stronger controlling role compared to the other two actors in the interactions through online labor platforms (Florisson and Mandl, 2018). Platform owners can even influence the success of producers that use their platform by creating consumer demand based on their extended data and analysis of deep behavior patterns.

Platforms minimize transactions costs between the parties that can benefit from getting together. The businesses, therefore, can reach a diverse crowd with minimum costs, while the risks and costs are often borne by the workers.

In some cases, there is direct relation between the client and the worker while in other cases the only link between them is the platform (De Stefano, 2015). While platform providers are strong, often monopolist or oligarchist organizations, the other two parties are rather disorganized, dispersed physically with low bargaining power. (Schmidt, 2017). By breaking the work into small micro tasks, the microtask platforms ensure that taking collective actions would be difficult, almost impossible for the workers. (Smith and Leberstein, 2015)

The fact that digital labor platforms have not really been accepted as regular work providers allowed them to grow without proper regulations in the shadow, to operate in a way to avoid the consideration of employment relations and, thus, to ”*hire*” someone just strictly for such a short time that is required to complete a particular task.

The employment status of the workers in the digital labor platforms has also led to extensive debates, as platform providers specify different "*classifications*" for them such as "*self-employed*", "*freelancer*", "*participant*". For instance Prolific' categorization suggest that workers are not doing actual work since they refer to the works as participants that receive rewards, not even salary (Berg et al., 2018). However, the solution is not as straightforward as some authors suggest. Some researchers argue that workers should be categorized as employees, while others would stick to self-employed status. The findings of Baiyere et al. (2019) suggest that workers on digital platforms appreciate the temporal dimension of work relations on digital labor platforms and consider self-employed status as an advantage that allows them to fit platform work to their personal lives (Baiyere et al., 2019). However, if workers would be truly self-employed, they should be able to choose the task and have more control over their working condition which is not the case on microtask platforms (Berg et al., 2018).

2.3. Crowdwork experience in the global south

This section provides overview on the current trends of crowdworking that are relevant in the topic based on previous empirical studies.

Optimist voices echo that online platforms can abolish barriers and allow workforce living in developing countries to access labor markets in developed economies and thus, identified digital platform work as potential drivers of development and economic growth (UNCTAD, 2019.

However, there are visible inequalities in digitalization among the developed and developing countries. While in the developed world, most people have internet access, developing countries are lagging behind where only 20 percent of people use internet. In areas such as capabilities for harnessing digital data and frontier technologies, the gap is considerably wider (UNCTAD, 2019).

Even though, crowdwork has global nature, differences have been observed across countries. One would presume the invisible feature of platforms prevents discrimination; however, studies show location-based discrimination is common even in the world of borderless online work (Galperin and Greppi, 2017). Limiting the task for particular countries is common practice that tends to exclude workers from the global south (Berg et al., 2018). Job seekers in developing countries experience lower remunerations and were less likely to find work compared to their counterparts in developed countries (Galperin and Greppi, 2017).

In the US, microtask is mostly considered as low status work, due to the low pay and low skill requirements. Interestingly, previous studies documented different reactions from India with regards to the exact same work. Gupta et al. found that "*Turkers*", crowdworkers on the Amazon Mechanical Turk (AMT) digital labor platform in India are proud that they work for "*American clients*" and are paid "*American money*" as it signals prestige. As a consequence, they talk about their work among friends and family more openly and is not differentiated from other types of work.

Regardless the sense of "*Americanization*" of AMT, Indian crowdworkers still seem to be resistant to take up crowdwork as a primary job since the wage level is problematic and crowdwork is less desirable than a regular job (Barnes et al., 2015). Behrendt et al. (2019) point out that in most cases when crowdworkers have access to social security coverage, it is often provided by additional job or by their spouses or other family members instead of the platforms (Behrendt et al., 2019).

It is important to note that the labor markets in developing countries differ in several characteristics from developed countries. A highly relevant difference is the high proportion of workers in informal sector and the prevalent youth unemployment in developing countries. The lack of employment opportunities is prevalent in the formal sector together with the lack of proper enforcement as well as institutional mechanisms (Rani and Singh, 2019). This creates a different environment for digital labor platforms to operate compared to developed economies, since employers might take advantage of the precarious situation of unemployed or underemployed workers.

Even though the promise of flexible work sounds attractive for people especially if they experience some form of restriction preventing them to participate in conventional labor force, crowdworkers often experience limitation in flexibility in practice. Due to its instability and casual nature, this form of work shares several similarities of non-standard employment such as part-time or temporary work (Berg, 2016). Most tasks on AMT are shared in batches at late night hours in India, presumably morning or work hours in the US. Workers in India tend to stay up late working on these tasks (Gupta et al., 2014a). This time difference makes Indian Turkers work outside of working hours, mostly during late hours. This unusual working schedule can be beneficial for those who are working, studying or have care responsibilities during the day, but can make difficult to harmonize crowdwork with traditional jobs.

3. Theoretical Framework and Hypothesis Development

This section provides a theoretical framework for the labor market analysis of crowdwork. The novelties introduced in the literature review, are incorporated in economic theories allowing its analysis on the labor market. The findings from the literature review are incorporated in the existing models which helps to develop the hypothesis of the paper.

3.1 Actors in the digital market

The traditional separation of supply and demand does not apply any longer to the platform economy. The new economic model works in a circular manner as a feedback loop in which data and interactions form the main resource and source of value. Spulber (2019) identifies the platform model as the actor's participation decisions and the problem of coordination in which platforms address the coordination problem and provide incentives for participation. He also captures one of the main contribution of platform model to the markets in economics is the introduction of the intermediary platforms that makes price and output endogenous by managing exchange, adjusting the prices and sales, providing liquidity and immediacy, coordinating buyers and sellers, guaranteeing quality, and monitoring performance (Spulber, 2019).

3.2 Deskilling by matching

The algorithmic matching of demand and labor through digital labor platform can be considered efficient in a way that transaction costs are reduced, and previously too costly or impossible matches become feasible. However, the allocative efficiency might suffer since it incentivizes skilled workers to accept tasks that are below their education or skill levels.

Since the tasks are divided into tiny fragments, the workers hardly require advanced skills to complete them. By breaking down jobs into tasks, one of the main implications of crowdsourcing for the labor market is the process of the possible "*deskilling*" workers. As a result of debunking tasks and deskilling labor, the concept of human capital may become meaningless as it does not only depend on the formal education, the labor market experiences and skill but on the nature of tasks performed.

Furthermore, breaking down jobs to separate tasks might affect workers' wages. Persons working over 40 hours per week typically receive an overtime premium, while the wage rate in part-time jobs is often lower than the wage rate in full-time jobs (Borjas, 2013). What is more, being an intermediary in the transaction, platforms are able to shift costs and risks to other actors in the marketplace.

Based on the existing literature, crowdwork can be understood as precarious work which is directly related to the wellbeing of workers since it is uncertain, unpredictable, unstable in time and risky. The precarious jobs are characterized by high levels of instability and insecurity, poor quality jobs with a very variable income over time, low income absence of skills production, collective representation, social status and poor opportunities for developing the desired career (Standing, 2011).

3.3 Skills and technology

Blinder suggests (2006) that in the next decades, the critical labor-market distinction may shift from the current education level requirement-based distinction to a new divide between jobs. The new divide will be made based on how easily the work is deliverable through wireless connections with little or no reduction in quality which does not correspond well to the conventional distinction (Blinder, 2006). Canonical production functions draw an implicit equivalence between workers' skills and their job tasks, as noted above. Autor (2013) emphasized that the distinction of skills and tasks is relevant when the assignment of skills to tasks is subject to change (Autor, 2013). This is also the case in the platform economy.

Traditionally, manual, routine labor has been more exposed to technological change since these processes do not require creativity. Today, as a result of digitalization and artificial intelligence, many other jobs are susceptible to replacement and this process is likely to continue in the future (Jäger et al., 2019). Autor et al. (2003) and Goos et al. (2014) point out that digital innovations do not generally lead to an increasing demand for skilled workers relative to unskilled ones, rather to a high substitution susceptibility of routine tasks since the commonly used Skill-Biased Technological Change theory - that argues labor demand is shifting in favor of more educated workers - cannot explain the recent phenomenon of job polarization (Goos et al., 2014). These authors also argue that the current technological change is biased toward the routine-biased technological change (RBTC), and that there is task offshoring (which is also influenced by technological change), and that both of these forces decrease the demand for middling relative to high-skilled and low-skilled occupations.

3.4 Low bargaining power

Palier and Thelen (2012) discuss that political and social-structural reasons has also led to the increase of problematic working conditions and low wages in the service sectors that are also likely to occur in crowdwork. The political argument suggests that for historic and structural reasons the new service jobs are less regulated and organized. While in the old industrialized countries unions have been powerful enough to achieve middle-class wages for medium- and even low-skilled work, the low bargaining power of online workers and the lack of unions suggest that wages and working conditions suffer on digital labor platforms (Palier and Thelen, 2012).

Furthermore, new service jobs are likely to be filled by newcomers or outsiders to the labor markets such as women, young people beginning their careers, migrants, older workers and people with discontinuous careers. These groups may not only lack experience in the labor market but also alternatives for alternate employment, thus, more willing to accept precarious working conditions. Durward argues (2015) that political efforts are needed to increase low-paid and precarious work since neither the market nor technological progress will automatically improve working conditions. As a new kind of service work, crowdwork is attracting newcomers with low bargaining power, it is expected to offer lower wages and working conditions than similar jobs in the conventional labor market.

3.5 Piece-rate wages

Among other factors, an important difference in the traditional jobs and microtasks is the determination of pay rate. The bulk of the traditional regular employment contract are created in the nature of spot labor markets where firms decide in each period how many workers to hire at given wages; workers decide how many hours to work; and the equilibrium wage and employment are determined by the interaction of workers and firms. In these labor markets, the wage equals the worker's value of marginal product.

When measuring the workers' true productivity proves to be difficult or too costly, employers assume that workers desire high salary with as little effort as possible , thus the type of labor market contract matters (Borjas, 2013). The model of crowdwork platforms is analogous to piece rate where instead of time rate, as worker's salary depends strictly on the amount of output produced. Between the 18th century and the end of the 20th century, piece rate was a dominant payment method in the manufacturing and production industries. Ever since, workers productivity became easier to monitor with the help of the technology especially on crowdwork platforms when the entire work is undertaken on the platform. Regardless, digital labor platforms still apply piece-rate system even though they are able to monitor the diverse, geographically dispersed crowds constantly.

Borjas (2013) argues that risk-averse workers are more likely to stay in the traditional time-rate system as they dislike the fluctuations in their weekly or monthly earnings over time. He also suggests that it means piece-rate firms have to compensate workers for the disutility caused by fluctuations in salaries.

If platforms offer compensation for the fluctuation of salaries, the wages of workers with comparable characteristics should be able to earn similar income in the traditional labor market and digital labor platforms. However, given the power imbalance captured in the literature, the digital labor platforms might avoid paying this compensation and provide overall lower wages for crowdworkers.

3.6 Hypothesis

Given the deskilling effect which is further deteriorated by the low bargaining power experienced by crowdworkers, the hypothesis of this paper is that workers are not able to benefit from the premium full-time workers receive and, thus, earn less than their counterparts in full-time traditional employment. Furthermore, as the tasks are simple, short and the employment shows precarious features, workers cannot accumulate skills and experience in the same way as they would do in the traditional labor marker which might prevent them from career development. Lastly, if the current skill-based separation of jobs is shifting towards a new distinction how quickly tasks can be performed and delivered in the labor market, education level may not affect earnings significantly.

4. Data and methodology

This chapter introduces the datasets used for the analysis, describes how the information service activities sector is selected, introduces the regression models used for the analysis and finally, provides key statistical estimations of the key variable used for the regressions.

4.1 Data

The dataset used for this thesis consists of the Indian sample of the Labour Organization (ILO) Crowdwork survey and the Periodic Labour Force Survey (PLFS) by the National Sample Survey Office (NSSO) in India. Both of these surveys were conducted in 2017. Given the similarities in the design of the surveys, the Crowdwork and PLFS datasets have been easily harmonized and were merged to a single dataset.

4.1.1 ILO Crowdwork Survey

The ILO Crowdwork surveys was conducted on a number of crowdwork platforms, namely on AMT, CrowdFlower, Clickworker, Microworkers and Prolific in 2017. The questionnaire was published as a paid task on micro platforms that ensured responders are active crowdworkers. Apart from questions regarding demographics, earnings and hours worked, the questionnaire captures working conditions, motivations, skills and thus offers a very unique and detailed dataset of 2,350 crowdworkers around the world.

4.1.2 NSSO Periodic Labor Force Survey

The Periodic Labor Force Survey (PLFS) by NSSO provides the data on the traditional labor force in 2017. NSSO provides the key employment and unemployment surveys in the country and, thus, the main source for analyzing labor force, activity participation in the country. The PLFS includes labor indicators including Usual Principal Activity Status, National Industry

Code (NIC), hours spent with work, daily earnings as well as demographic information such as age, gender, education level and household size.

Given the fact that PLFS reports individuals' activity status and the details of undertaken economic activities on a daily basis, it is possible to extract detailed information about individuals even if they hold multiple jobs or belong to more than one activity status. The survey also captures the participants NIC and activity status on a daily and weekly basis, participants engaging in more than one job could be included in the sample as long as they fall into the industry under investigation. Based on their activity status, individuals that indicated to be a regular, casual or own-account worker at least on one day in a week were kept in the sample.

4.1.3 Information service activities as the closest sector to microtasks in India

To minimize the unobservable heterogeneity and enable the comparison of crowdwork and the traditional work, only the selected industry, division 63 "*Information service activities*" is included in the final data from the NSSO data. Multiple approaches, sectors and occupations were taken into consideration for this decision. Each task undertaken on microtask platforms is matched to the nearest NIC sub-class based on their descriptions.

In the information service activities sector, "*Data processing*" and "*Providing data entry services*[1]" apply to most cases since these tasks require computer skills, they can be disintegrated and outsourced. The microtask descriptions and the matched closest activities with their NIC are listed in Table 4.1 on the next page. Division 82, "*Office administrative, office support and other business support activities*" is used for robustness check since there are similarities in the nature of tasks undertaken but the characteristics of these workers are assumed to be more distinct from crowdworkers.

[1] Data entry is „the work of entering a specific type of data into computer or any other electronic device using specific software. This work is done by expert data entry operators who have complete knowledge regarding data entry". Source: What is Data Entry? Types of Data Entry Services in India https://www.omdataentryindia.com/blog/what-is-data-entry-types-of-data-entry-in-india.html (accessed 26 November 2019).

Table 4.1: Microtask and NIC descriptions

Microtask	Description	Similar NIC with description
Artificial intelligence and machine learning	• Collection of data and other information to train machine-learning algorithms; • Tasks related to programming and coding or to solve mathematical or logical problem	• 63114 Providing data entry services • 63111 Data processing activities including report writing • 62011 Writing, modifying, testing of computer program to meet the needs of a particular client excluding web-page designing
Categorization	• Classification of entities into groups e.g. bookmarking, tagging, pinning.	• 63114 Providing data entry services
Content access	• Promotion of a specific product; • Search engine optimization by fake traffic creation; • App testing.	• 63121 Operation of web sites that use a search engine to generate and maintain extensive databases of internet addresses and content in an easily searchable format
Content creation and editing	• Create new content; • Proofread, edit or translate existing materials (mostly text); • Might be time consuming.	• 63111 Data processing activities including report writing
Content moderation	• Review contents including text, images and videos • detect if any of the material posted on the website might violate local laws, social norms or the platform's guidelines.	• 63111 Data processing activities including report writing • 63999 Other information service activities n.e.c.
Data collection	• Specific metadata collection; • Information needs to be found, copied and pasted; • Gathering information from specific geographic locations.	• 63114 Providing data entry services • 63121 Operation of web sites that use a search engine to generate and maintain extensive databases of internet addresses and content in an easily searchable format
Market research and reviews	• Review or rating of a product, service or location (imaginary); • from "mystery shopping" to reviewing and testing apps.	• 73200 Market research and public opinion polling
Surveys and experiments	• Surveys from academic researchers; • Broad range of topics; • There may be some overlap with market research.	• 73200 Market research and public opinion polling • 72200 Research and experimental development on social sciences and humanities • 63114 Providing data entry services
Transcription	• Transcription from different types of media, such as audio, text, photos, or videos, into written form.	• 82192 Document preparation, typing, word processing and desktop publishing sevices • 82199 Other specialized office support services activities
Verification and validation	• Verify and "clean" existing data or classifications, or to confirm the validity of some content.	• 63111 Data processing activities including report writing

The NIC description of this sector is in the Appendix (Table A.1).

4.2 Methodology

4.2.1 Ordinary Least Squares model

A number of key labor market indicators were selected to investigate the relationship between the dependent variable, hourly earnings and the independent variable, crowdwork. The Ordinary Least Squares (OLS) model estimation is essential when investigating an effect since correlation is a prerequisite for a potential causation.

The dependent variable in the model is the logarithmic transformation of hourly earnings. The regressor, crowdwork is a binary variable which takes up value 1 if the person in involved in crowdwork and 0 otherwise. Age, gender, education level, marital status, household size and location were used as controls. Education is included as a categorical variable in the model, therefore dummy variables are used taking up either value or 0 or 1. In the regression equation, the base category, thus the highest education category is omitted to avoid collinearity. Table 4.2 displays the list of variables and their short description.

Table 4.2: List of variables used in the OLS model

Variable	Short description	Type
Log hourly earnings	The logarithmic transformation of hourly earnings	integer
Crowdwork	1 if individual undertakes crowdwork, 0 otherwise	binary
Age	Age of individuals	integer
Female	1 if female, 0 if male	binary
Education -low	1 if below high school, 0 otherwise	binary
Education -medium	1 if high school and technical education, 0 otherwise	binary
Education -high	1 if bachelor’s degree, 0 otherwise	binary
Education -highest	1 if master’s and PhD, 0 otherwise	binary
Married	1 if married or lives with partner, 0 otherwise	binary
Household size	Number of individuals in the household	integer
Urban	1 if lives in urban area, 0 if lives in rural area	binary

4.2.2 Two Stage Least Squares model

Endogeneity may occur as a result of omitted variables, measurement errors or reverse causality. The direction of causality between crowdwork and hourly income can be questioned, since the income level can also affect workers' decision to undertake crowdwork. If the explanatory variable, crowdwork is endogenously determined in a regression on hourly earnings, it violates one of the key assumptions for OLS and leads to a biased and inconsistent estimation.

One way to avoid the potential bias in the estimation is to instrument it or in other words replace it with fitted values in the second stage of a 2SLS model (Angrist and Pischke, 2008). To do so, an instrumental variable is introduced to the model which is an exogenous regressor of the problematic, endogenous variable. The first-stage regression provides the predicted value of $\hat{y}_2$ by estimating an OLS against all exogenous variables, including the instruments. In the second stage of the regression model, $\hat{y}_2$ is used in the place of y_2 to estimate y_1 against $\hat{y}_2$ and all of exogenous independent variables but the instrument.

Instruments are valid if two requirements, the instrument exogeneity and the instrument relevance are satisfied. Instrument exogeneity means that valid instruments are uncorrelated with the error term. For this condition, the theoretical argument has to rule out any direct or indirect effect of the instruments on the dependent variable as well as any reverse effect of the dependent variable on the instruments. Furthermore, it has to describe why the instruments influence the endogenous regressors, even after controlling for the effect through the included exogenous regressors. Lastly, to satisfy the instrument relevance condition valid instruments have to be highly correlated with the endogenous regressors even after controlling for the exogenous regressors (Angrist and Pischke, 2008).

The main challenge and difficulty with this model is selecting the right instrument. While, the instrumental relevance can be empirically tested in a first stage regression, the condition of instrument exogeneity requires longer theoretical argument and it is rarely possible to certainly prove that the selected instrument is exogenous. The instrument is weak if its correlation with the included endogenous regressor is weak. In this case, the 2SLS estimator is biased towards the direction of the OLS (Angrist and Pischke, 2008).

The instrumental variable selected for the model is the domestic work dummy variable. The detailed information including main motivation of crowdworkers as well as qualitative answers in the ILO dataset helps identify the instrumental variable. One of the main motivations for crowdworkers to undertake work on digital labor platforms is the promise of flexibility that allows them to work adjust their schedule to take care of domestic duties and care responsibilities as well as work from home. Furthermore, crowdworkers can choose where they would like to work, thus, when they work from home, they are more likely to get involved in domestic duties than their traditional counterparts with less flexible conditions. The variable of undertaking domestic work should not be correlated with hourly earnings since it does not affect work-related activity directly given the fact that domestic work takes place outside of working hours. Studies have argued that unpaid care work is the missing link in the gender gap analysis of labor market outcome, influencing labor market participation rate of women (Ferrant et al., 2014), but this link is under argument and the regressions in the data showed no correlation between the two variables in question.

Similar instrument was used by Cantarella_and Strozzi (2018) for a similar analysis in the USA and EU; however, the authors derived this variable slightly differently than this paper does, referring to those who engaged in care work activities before doing crowdwork which has some limitations.

4.3 Descriptive statistics

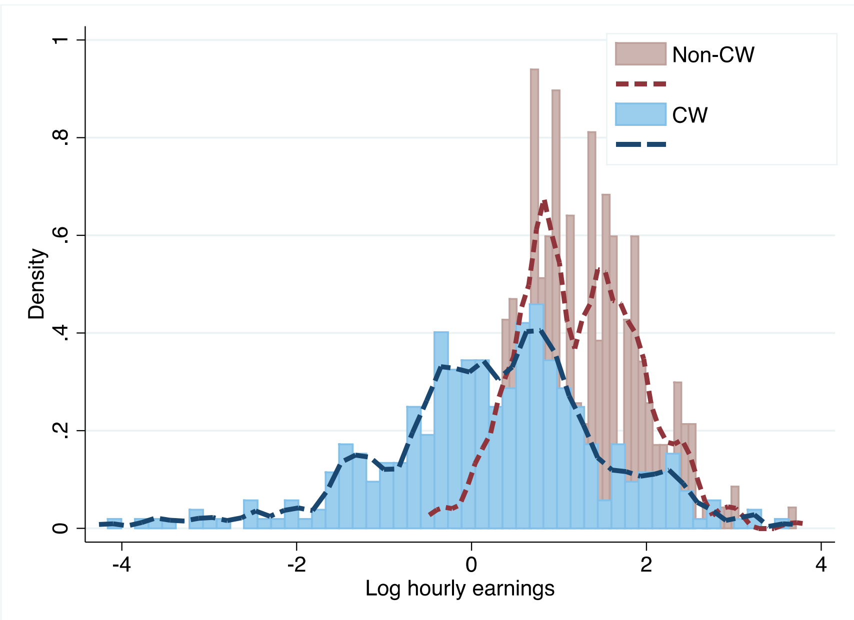
Descriptive statistics provides detailed description of the key variables used in the analysis. The section compares the key statistical estimates of crowdworkers and traditional workers to observe how similar these two groups are and captures the differences that may occur. For the analysis, the two datasets were adjusted and merged to a single dataset of 622 observations from which 336 is coming from the ILO crowdwork dataset and the remaining 286 from the NSSO survey.

4.2.1 Dependent variable: hourly earnings

The average hourly earnings of traditional workers are 135 percent or 2.2 USD higher than the average hourly earnings of crowdworkers in the sample.

In the crowdwork dataset, hourly earnings were calculated from the typical weekly earnings and typical hours worked per week. Earnings in INR were converted to USD at a rate of 69.956 in both datasets. Outliers with higher than 140 and lower than 0.001 USD hourly earnings were dropped from the sample.

Figure 4.1 Distribution of log hourly earnings by crowdwork activity



The average hourly income with crowdwork is 2.62 USD while in the information service activities it is 3.45 USD in India. The median of the crowdworker group's hourly income is 1.36, which means this variable is skewed to the right. As Figure 4.1 shows, the distribution somewhat different, the hourly wages of traditional workers are more concentrated, while the log earnings of crowdworkers vary in a much wider range.

The hourly earnings of crowdworkers vary by platform. The highest rates are provided by Prolific with 4.60 USD on average followed by CrowdFlower with 3.68 USD and AMT with 2.71 USD.

Hours worked

Traditional workers spend significantly more time on their work than crowdworkers in India. Since the hourly earnings of traditional workers were calculated from the daily wages and working hours, it is important to look at this variable carefully.

Crowdworkers spend 32 hours with work in a week from which 25 hours are paid the remaining 7 hours are unpaid. The total weekly hours were calculated by summing up these two numbers. The unpaid tasks are included in this calculation because in many cases, individuals have to undertake low-paid or unpaid tasks in order to increase their ranking and thus get paid tasks. Furthermore, as a part of the risks shifted by the platforms which is discussed in the literature, microtask workers do not receive trainings from the website, so that if they require any new skill, they have to learn it on their own which would count as unpaid hours. Based on their answers, 93 percent of the crowdworkers in the sample would like to work more hours on microtasks but the lack of available work is a serious concern, half of these induvial said this is the main hurdle.

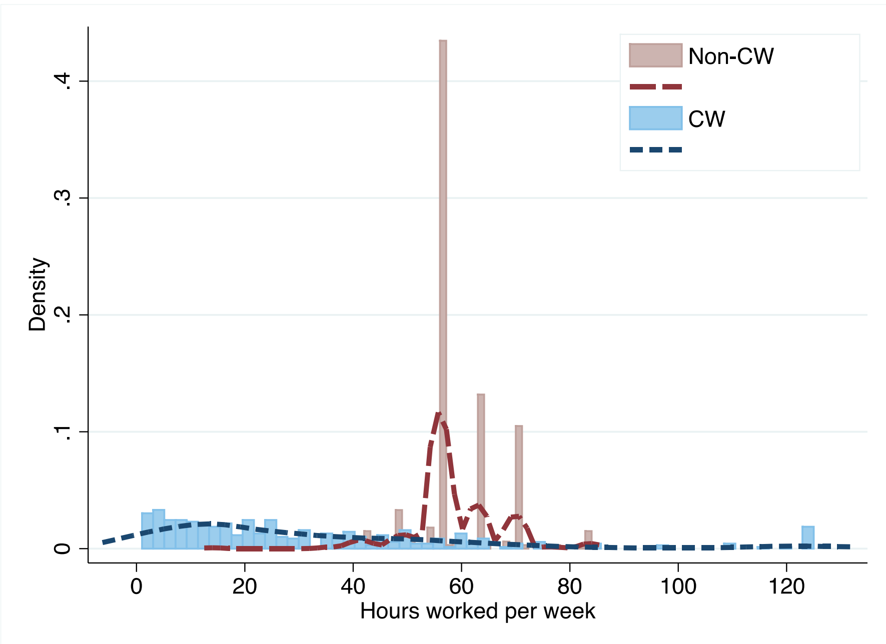
In the NSSO survey, the variable "hours actually worked" is determined on daily basis and includes the total time individuals spend directly on production activity, indirectly related to the production activity, the hours accounted for unavoidable 'in-between time' as well as breaks and short rest or refreshment. The average working hour including breaks and indirect activities is still extremely high, 58 hours per week.

It crucial to note that the survey asks about time spent with unpaid task not including breaks. Therefore, the values reported in the NSSO survey that include breaks as well might overestimate the time spent with work if the workers spend time with rest which is not necessary the case in this context.

Almost all workers in the NSSO survey reported to work seven days a week, while crowd workers tend to work on fewer days, 6 days on average, but 40 percent of them spend time with crowdwork every day a week. As Figure 4.2 shows, individuals were likely to report full hours every day. Looking at each day separately, half of the respondents reported exactly 8 hours in each case.

In the case of crowdwork, the mean value is much lower, but there are considerable number of workers that work over 120 hours. Most of the workers (70 percent) who work over 120 hours a week do not have other paid job. On average, this crowdworker group tends to work more weekly compared to the one that has other paid job, 32.9 and 30.8 hours correspondingly.

Figure 4.2: Distribution of weekly working hours



4.2.2 Explanatory variable: crowdwork

In the final sample 54 percent of the individuals undertake crowdwork. According to the ILO Survey, crowdworkers in India work in all of the observed platforms including Amazon Mechanical Turk, CrowdFlower, Prolific Academic and Microworkers. The majority, 73 percent of Indian crowdworkers in the sample consider Amazon Mechanical Turk as the main platform they work on, one of the largest microtask platforms in the world. The second most popular microtask platform is Microworkers with 16 percent of the workers considering it as the main platform they work on.

A part of the crowdworkers have other source or engage in other paid jobs. Therefore, it can be insightful to capture possible differences between those who only engage in crowdwork and those who hold additional paid work as well. More than half of the respondents in the sample have no other paid job, 57.7 percent of the individuals participating in the survey said the only paid work they do is crowdwork, 10.5 percent had another job but considered crowdwork as main job, while 31.8 percent considered another paid job as main job.

Around 65 percent of the crowdworkers have other sources of income such as spouse's income, government support or other work, the other 35 percent are fully dependent on crowdwork.

Interestingly, the main job is not identical with the main source of income. Half of those who only do crowdwork as paid activity, claimed that crowdwork was not their primary source of income. These numbers reflect that the amount they earn might not be sufficient for covering the basic costs.

66 percent of the workers who hold additional paid job(s) said they would like to do more non-crowdwork. Interestingly, over 30 percent of those who said they would not do more hours in the other job said they would rather like to do more crowdwork which means there is some tendency of moving towards crowdwork from traditional work.

Looking at some basic characteristics, there are some small differences between those who hold additional job compared to those that are not. Those who hold additional job tend to be a bit older, less likely to be a woman, more likely to be married, slightly more likely to have domestic duty, more likely to have higher education level. These crowdworkers tend to work less hours and earn comparatively less, 2.5 USD per hour while those who only do crowdwork earn 2.7 USD per hour on average.

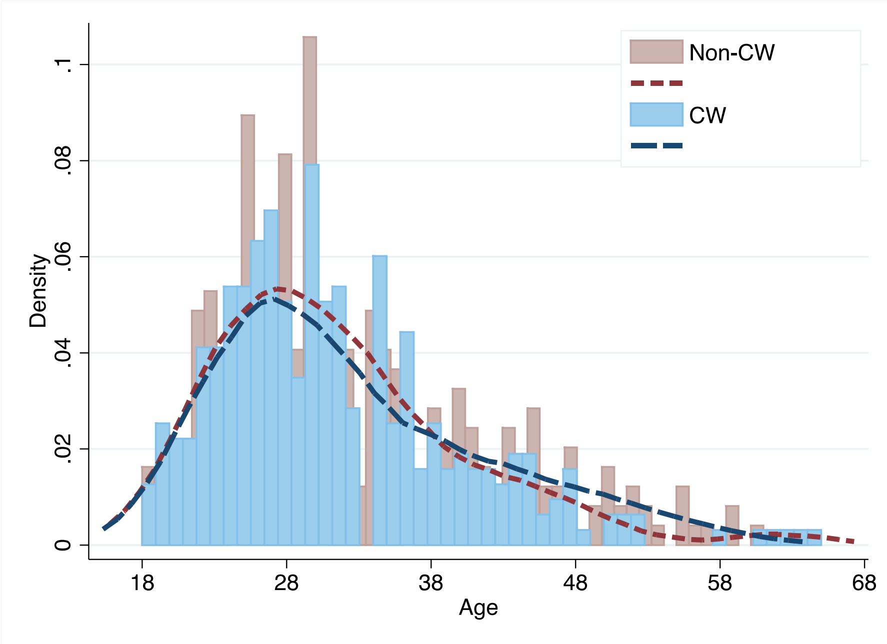
In the group of traditional workers, very few people, only 4.2 percent hold at least one additional job.

4.2.3 Controls: age, education, gender, marital status, household size, urban location

Age

Crowdworkers in India are relatively young, the average age is 31.5 years. It is important to note that not only youngsters engage in microtask platforms, the oldest person that completed the survey is 65 years old. In order to limit the analysis for working-age population, only individuals between 18 and 65 years were kept in the final sample of the PLFS survey. This variable is slightly skewed to the right with the median of 30 years and its distribution is almost identical in the two groups as Figure 4.3 illustrates. The average age of the corresponding traditional sector is very similar to the crowdworker group, 32.6 years. This variable is slightly skewed, the median is 30 years in both groups.

Figure 4.3: Age distribution by crowdwork status



Education

Both crowdworkers and traditional information service workers have comparably high education level, most of them have at least a Bachelor's degree, while a larger proportion of traditional workers have Master's degree.

There is a clear difference among education level between crowdworkers in India and developed economies. 87.8 percent of the crowdworkers in India holds at least a Bachelor's degree which is well above the crowdworkers' global average of 55.7 percent according to the sample, supporting the previous findings that crowdworkers typically have higher education level in developing countries.

In the traditional information service sector, workers are also highly educated, 70 percent of them hold at least a bachelor's degree. Even though both groups hold a degree, half of the tradition workers hold a Master's degree, while one third of the crowdworkers have the same education level.

Gender

The gender imbalance in the sample is prevalent among microtask platforms and traditional market in India.

The majority, 80 percent of the Indian crowdworkers are men in the sample which is not different from the pattern that developing countries exhibit. In the developed world, this gap is not as wide, almost half of the crowdworkers are women.

In the traditional information service sector in India, the gender composition is very similar, 76 percent of the respondents are male. This shows that there are slightly more women in the traditional information service sample compared to the crowdwork sample in India, but the difference is small.

Surprisingly, in this sector, women earn more than men on average. On microtask platforms, men earn 2.6 USD, while women earn 2.7 USD per hour, while in the traditional information sector, their counterparts earn 4.3 USD and 4.80 USD per hour respectively.

In both groups, women are more likely to hold a degree, 91 percent of the women working on microtask platforms and 71 percent of their counterparts in the traditional labor force hold at least Bachelor's degree.

Women working on microtask platforms tend to be on average 5 years older than male crowdworkers, while, as a contrast, women in the traditional information service sector are more than a year younger on average than their male counterparts.

Urban location

The urban dummy variable shows whether the individual lives in urban area or not. The percentage of individuals living in urban areas are high in both datasets, 77.7 percent of the crowdworkers and 81.1 percent of traditional information service worker live in cities. This is much higher than the national average. In 2017, around 30 percent of the population lived in urban areas (World Bank, 2018).

Marital status and household size

Around half of the individuals are married or live with a partner in both groups. In the traditional group this proportion is 55 percent, while among crowdworkers it is 54 percent.

The average household size is also very similar, 4.4 and 4.1 persons respectively. Those that do not undertake crowdwork are more likely to live in larger households, 21.7 percent of them live in a household that consists of more than 5 people while this proportion is only 11.9 percent among crowdworkers.

4.2.4 Instrument: Domestic work

In the ILO crowdwork sample, the selected domestic work instrument takes up value 1 if individuals indicated they currently undertake domestic duties including care for children, elderly, ill people or other related activities and 0 otherwise. In the NSSO data, individuals who undertake domestic responsibilities at least on one day per week were selected based on the daily activity status code.

In India, larger proportion of crowdworkers have domestic responsibilities compared to the rest of the world, regardless of gender. The survey includes a question about the type of domestic responsibilities they undertake including taking care of children, elderly, adult in need and other domestic work. This variable is selected as instrumental variable in the 2SLS model. According to this indicator, 14 percent in workers in India undertake domestic work, and interestingly, the gender gap is narrow, 16.2 percent of female and 13.4 percent of the male crowdworkers have domestic commitments.

This pattern can be also captured in other variables in the crowdwork sample. Looking at their motivations, 6.71 mentioned family, taking care of ill relative or spending time with kids was a part of their decision to undertake crowdwork. What is more, 21 percent of the respondents was engaged in caretaking responsibilities right before starting working on microtasks.

4.2.5 Other important characteristics

Skills based on other job

Almost half, 41.76 percent of the individuals in the sample hold at least one paid job apart from crowdworking. In the Appendix, Table A.3 displays their occupation in their other paid job. In case of more additional paid jobs, the main one was taken into account. A large proportion of the respondents do high skilled work. Almost half, 40.85 percent of those crowdworkers that hold additional paid job belong to the professional category which falls into the highest, 4th skill level according to the International Standard Classification of Occupations (ISCO) (International Labour Organization, 2007). This indicates that these workers probably undertake complex tasks that require problem-solving, decision-making, creativity and high level of interpersonal communication skills in their other work.

Motivations

The ILO survey asked questions about crowdworkers motivations to engage in this kind of activity. Preference to work from home was the most common motivation with about half of the crowdworkers mentioning that it influenced their decision and 26 percent of crowdworker claimed it was the most influential factor. Interestingly, many crowdworkers in India seem to enjoy this form of work since this option was selected by 60 percent of them. For about 14 percent, the decided to perform crowdwork because they enjoyed this form of work. What is more, over third of the individuals decided to undertake crowdwork either because they prefer or can only work from home which is in line with the previous literature suggesting that flexibility is a key characteristic of crowdwork. The table with the main motivations is included in the Appendix (Table A.4). Another common reason is complementing pay from other jobs, with 20 percent of the respondents selecting this option. Interestingly, not everyone holds additional jobs from those who selected this option as the most important reason. Even though, there are some differences among the region, the overall pattern seems similar globally.

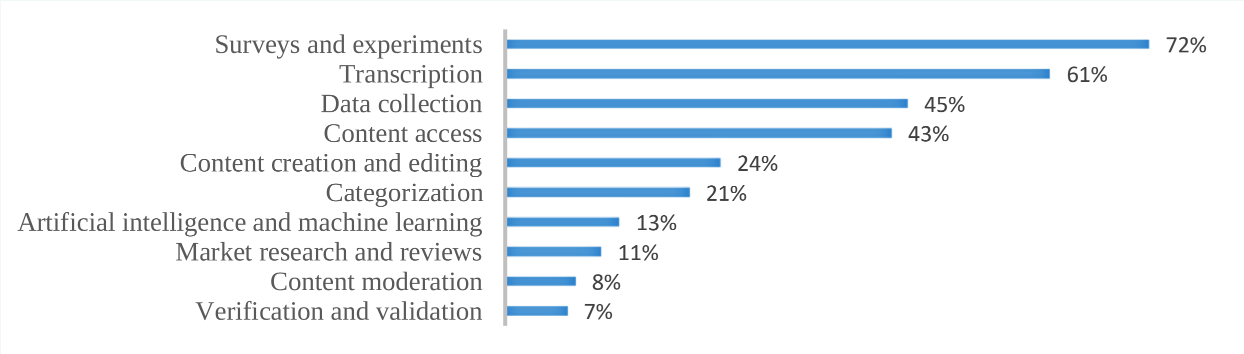
Employment status and tasks

Immediately before engaging in crowdwork, a significant proportion, 61 percent of the survey participants were employees which also supports the assumptions that the crowdwork group is similar to the traditional information service sectors.

In the traditional group, 73 percent of the workers are regular employees, and the remaining 27 percent is mostly self-employed. Since there are no significant differences across these groups in terms of their earnings or hours worked, neither self-employed individuals nor employees were dropped.

The data supports the previous findings that crowdworkers undertake a large variety of tasks, but rarely any with high skill requirement. As Figure 4.4 shows, the most common tasks were survey (72.4 percent), transcription (62.0 percent) and data collection (45.2 percent) among the Indian crowdworkers. Over 70 percent of the workers reported to be engaged in at least two kind of tasks. The task distribution does not differ significantly by gender. In 7 out of the 11 categories, the proportion of women that undertake the given type of task outnumber the proportion of men in the same category. A possible explanation is that women undertake different kinds of task than their male counterparts.

Figure 4.4: Percentage of individuals by task types on microtask platforms



4.2.6 Summary of descriptive statistics

The following table (Table 4.3) summarizes the main variables for the group of people working on microtask platforms, in the traditional information service sector.

Table 4.3 Summary statistics of key indicators

	Crowdwork		**Information service activities**	
	Mean	*SD*	Mean	*SD*
Mean hourly income (USD)	2.61	*3.97*	4.43	*3.94*
Mean log hourly income (USD)	0.23	*1.32*	1.23	*0.69*
Weekly working time (hours)	32.01	*29.62*	58.38	*8.65*
Age	31.46	*8.62*	32.55	*9.34*
Female	0.2024	*0.4023*	0.2412	*0.4286*
Education – No high school	0.0148	*0.0066*	0.0105	*0.0060*
Education – High school	0.0772	*0.2676*	0.2483	*0.4328*
Education- Technical	0.0297	*0.0298*	0.0420	*0.0420*
Education-Bachelor’s	0. 5460	*0.4987*	0.0420	*0.2008*
Education-Master’s	0.3116	*0.4642*	0.4930	*0.5001*
Education- PhD	0.0208	*0.0208*	0.1643	*0.1643*
Married	0.5359	*0.4994*	0.5489	*0.4985*
Household size	4.10	*1.25*	4.42	*1.67*
Urban	0.7754	*0.4179*	0.8119	*0.3920*
Domestic work	0.1399	*0.3474*	0.0279	*0.1652*
Number of observations	336		286	

The traditional workers in the sample are slightly older with more education. Crowdworkers are more likely to hold a bachelor’s degree while their traditional counterparts rather have master’s degree or PhD. In both groups, most workers, 78 percent of crowdworkers and 81 percent of traditional workers live in urban areas. The gender imbalance is prevalent in each group in the analysis including the sample of crowdworkers, around 20 percent. The percentage of women is slightly higher among crowdworkers supporting the argument to some extent that this type of work can include traditionally underrepresented groups to the labor market.

One major difference is the number of working hours. Workers in the traditional information sector work exceptionally high working hours, many of them spend 8 hours with work on 7 days a week. Meanwhile, crowdworkers work considerably less, while they are more likely to work either extremely low or high working hour. Traditional workers are mostly regulated employees which explains the similarly high and regular working hours. The impact of the absence of regulation as well as the lack of sufficient work is reflected in the irregular, fluctuating and often lower working hours than the individuals would prefer.

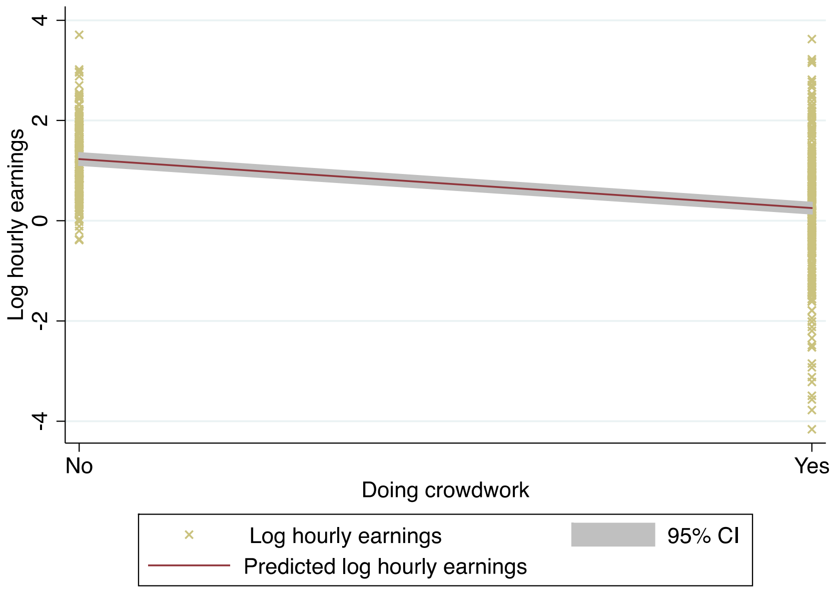
5. Results and discussion on crowdwork and the labor market outcomes in India

This chapter introduces the results of the models used for the analysis. First it discusses the Ordinary Least Squares Estimations, followed by the different specifications of the Two-Stage Least Squares model. Robustness-check as well as the discussion of limitations are added to the chapter which ends with the discussion of the results presented.

5.1 Ordinary Least Squares Estimations

The dependent variable is the individual's log hourly earnings in 2017. The regressor of interest is the crowdwork variable, CW which takes up value of 1 if the individual undertakes any activity on microtasks and 0 otherwise.

Figure 5.1 Predicted relationship between doing crowdwork and hourly earnings



As Figure 5.1 and the coefficients illustrate, the relationship between hourly earnings and doing crowdwork is negative and significant throughout the OLS regressions regardless the number of controls. Table 5.1 shows different regression models with an increasing number of controls, with an initial sample including a total of 622 individuals. In the model without controls, individuals that undertake crowdwork are associated with 98 percentage lower wages on average compared to the workers in the information service sector on the traditional labor market. After controlling for different characteristics, this percentage increases to as high as 112 percentage in model (7), as it is shown in Table 5.1. By adding more controls, the R_2 value also increases from 0.182 to 0.232.

Table 5.1: OLS estimations: Crowdwork on log hourly earnings

VARIABLES	(1) OLS No control	(2) OLS Female	(3) OLS Female, Age	(4) OLS Female, Age, Education	(5) OLS Female, Age, Education, Married	(6) OLS Female, Age, Education, Married, Household	(7) OLS All controls
Crowdwork	-0.98***	-0.98***	-0.98***	-1.10***	-1.11***	-1.13***	-1.12***
	(0.080)	(0.080)	(0.079)	(0.105)	(0.104)	(0.104)	(0.103)
Female		-0.05	-0.07	-0.06	-0.06	-0.07	-0.10
		(0.104)	(0.103)	(0.100)	(0.099)	(0.100)	(0.099)
Age			0.09***	0.08**	0.06*	0.06*	0.06*
			(0.031)	(0.033)	(0.035)	(0.035)	(0.035)
Age-squared			-0.00***	-0.00**	-0.00	-0.00*	-0.00
			(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Education-Low				0.07	0.05	0.07	0.10
				(0.404)	(0.399)	(0.392)	(0.381)
Education-Medium				-0.29***	-0.29***	-0.27***	-0.24**
				(0.100)	(0.099)	(0.099)	(0.099)
Education - High				0.14	0.16	0.16	0.18
				(0.127)	(0.128)	(0.127)	(0.127)
Married					0.15	0.17*	0.16*
					(0.098)	(0.099)	(0.098)
Household						-0.04*	-0.04
						(0.026)	(0.025)
Urban							0.34***
							(0.098)
Constant	1.23***	1.24***	-0.44	-0.13	0.16	0.33	0.07
	(0.041)	(0.047)	(0.523)	(0.550)	(0.588)	(0.602)	(0.599)
Observations	622	622	622	622	622	622	622
R-squared	0.182	0.182	0.198	0.211	0.214	0.217	0.232

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Education reference category: Highest

This model includes control variables such as age, gender, marital status, education level, urban location dummy and suggests that crowdworkers are expected to earn 112 percentage less on average than their traditional counterparts in the information service sector with the same age, gender, education level, marital status, household location, urban location. Both outcomes are significant at 99 percent significance level. These results are in line with previous findings that suggest crowdworkers earn less compared to the workers on conventional labor market (Berg et al., 2018; Cantarella and Strozzi, 2018). Interestingly, some controls are not significant such as gender, marital status, household size and a part of the education categories.

The coefficient of female is negative, but not significant, which implies that there is no significant relationship between gender and hourly earnings in the model under investigation. An explanation for this could be the low number of female workers on the sample (22 percent), but it is also possible that there is no significant wage gape between men and women which could be explained by the high education level of women in the sector or the fact that in microtask platforms, gender is not revealed.

The positive, significant coefficient of urban variable shows that individuals ceteris paribus earn 34 percent more on average than their counterparts living in rural areas. In rural areas, poor either poor internet connection or connectivity could be a reason for lower earnings or the fact that living costs are typically lower in rural areas which might make increase workers' willingness to accept lower paying jobs.

The reference education category used for the OLS regressions is highest education level including individuals with master's degree and above. Only one category, the Medium education group shows significant correlation with the dependent variable, which includes high school and technical education. However, a reason might be the unequal distribution of workers among these categories, since there are more traditional workers who hold master's degree.

The strong correlation between crowdwork and hourly earnings does not necessarily imply causal relationship, since other factors may influence the outcome, many of which might be unobservable.

5.2 Two-Stage Least Squares Estimations

In order to estimate the causal effect of crowdwork on hourly earnings, 2SLS model is used which relies on the domestic work as instrumental variable. In order to capture possible differences across crowdworkers, different models are added to the analysis. In the first case, the crowdworkers are divided to two groups, one with crowdworkers that do not engage in any other paid economic activity other than crowdwork and one with crowdworkers that do not have additional job(s). The information sector workers form the group of traditional workers.

In the second case, the entire sample is divided by age into wo groups, younger workers below 30 years and older ones above.

5.2.1 2SLS Estimations, by crowdworkers holding additional job

The coefficient of domestic work in the first-stage regression is 0.19 in model (1) including all crowdworkers, which implies that workers that undertake care responsibilities are 19 percent more likely to do crowdwork than those individuals with comparable characteristics who do not engage in domestic responsibilities. When crowdworkers are divided into two groups, those who only do crowdwork and those that have other jobs a s well, the results remained very similar, 22 and 23 percent respectively.

To prove that the instrument is valid, the coefficient must also be significantly different from zero, which can be tested with a F-statistics. Domestic work proved to be a good candidate, since the F statistics are high in both cases, above 40, with the p-value of 0.

Table 5.2 First-stage results of the 2SLS model

VARIABLES	(1) 2SLS CW all	(2) 2SLS CW only	(3) 2SLS CW & other work
Domestic work	0.19***	0.22***	0.23***
	(0.049)	(0.071)	(0.071)
Constant	1.14***	1.45***	0.45*
	(0.229)	(0.259)	(0.230)
Observations	622	479	427
R-squared	0.320	0.348	0.398
F test model	47.33	40.36	42.69
P-value of F model	0	0	0

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 5.3: Crowdwork and log hourly earnings OLS and 2SLS results

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	2SLS	OLS	2SLS	OLS	2SLS
VARIABLES	All crowdwork		Only perform crowdwork		Crowdwork and other job	
Crowdwork	-1.12***	-1.65**	-1.06***	-1.25	-1.14***	-1.45**
	(0.103)	(0.778)	(0.127)	(0.782)	(0.149)	(0.720)
Female	-0.10	-0.08	-0.06	-0.04	-0.07	-0.07
	(0.099)	(0.106)	(0.108)	(0.114)	(0.095)	(0.104)
Age	0.06*	0.06*	0.08*	0.07**	0.01	0.02
	(0.035)	(0.032)	(0.040)	(0.036)	(0.036)	(0.039)
Age squared	-0.00*	-0.00*	-0.00*	-0.00**	-0.00	-0.00
	(0.000)	(0.000)	(0.001)	(0.000)	(0.001)	(0.001)
Education-Low	0.10	0.23	-0.34	-0.30	0.57	0.63
	(0.381)	(0.420)	(0.330)	(0.431)	(0.443)	(0.421)
Education-Medium	-0.24**	-0.27**	-0.22**	-0.22**	-0.39***	-0.42***
	(0.099)	(0.124)	(0.097)	(0.110)	(0.097)	(0.135)
Education - High	0.18	0.48	0.17	0.29	0.12	0.32
	(0.127)	(0.454)	(0.155)	(0.523)	(0.172)	(0.486)
Married	0.16	0.20*	0.25**	0.26**	0.09	0.11
	(0.098)	(0.119)	(0.105)	(0.122)	(0.098)	(0.114)
Household size	-0.04	-0.06	-0.03	-0.04	-0.04	-0.04
	(0.027)	(0.035)	(0.028)	(0.032)	(0.026)	(0.031)
Urban	0.34***	0.33***	0.45***	0.44***	0.34***	0.33***
	(0.098)	(0.105)	(0.104)	(0.110)	(0.104)	(0.110)
Constant	0.07	0.34	-0.46	-0.33	0.79	0.72
	(0.600)	(0.706)	(0.666)	(0.811)	(0.637)	(0.631)
Observations	622	622	479	479	427	427
R-squared	0.231	0.194	0.291	0.286	0.270	0.257

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Education reference category: Highest

After including the selected instrumental variable, the crowdwork coefficient becomes even more negative than in the OLS models. In the 2SLS model, the results indicated that crowdwork is associated with 165 percent lower wages than traditional workers in the observed sector on average in 2017. The coefficients from the 2SLS models are presented in Table 5.3.

This result is very similar in the third model where the coefficient is-1.45, significant at 95 percent. Interestingly, in the second 2SLS model where those crowdworkers are included that do not have other job, the negative result is not significant anymore. These results imply that the negative relationship does not necessary hold when crowdworkers do not have another job.

In both 2SLS models, female workers are associated with 4 to 10 percent lower wages depending on the model but these coefficients are not significant, thus, there is still no significant gender gap in the hourly earnings. As discussed above, there are different possible explanations for this including the possibility of the absence of significant gender pay gap in the sector as the higher hourly earnings also suggested in the Descriptive statistics section.

Being married is associated with positive earnings on average, but it is not significant when workers hold another job. The household size is negative but not significant in any of the models.

The coefficient of medium level education remains negative and significant, but the high level of education (bachelor's level) is still positive compared to the highest education level (master's and PhD). The correlation of log wage and low education level has different direction in different models when all controls are included. When all crowdworkers are included, workers with the lower education lever are expected to earn 10 percent in the OLS and 23 percent more in the 2SLS model than their counterparts with the highest education level. The pattern is the same with crowdworkers that have another job. The only case when workers with lower education are associated with lower hourly earnings is when crowdworkers do not have additional job. However, these results are not significant, only the medium level education which is associated with 22-42 percent lower hourly earnings on average.

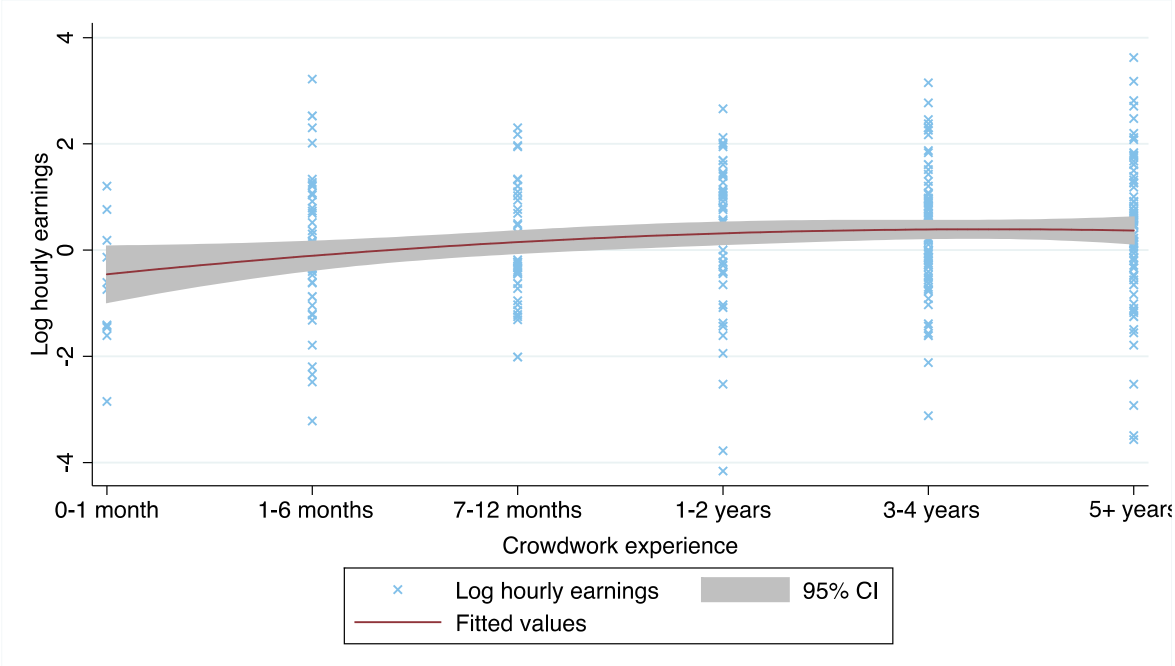
The urban variable which is significant at 99 percent level does not change across the models, individuals living in urban areas are expected to earn 33-45 percent more than their counterparts in rural area but have the same characteristics otherwise.

On the traditional market, older age which is often used as an indicator of more work experience is typically associated with higher wages. Age is positive and significant variable, but it is close to 0 and when all of the controls are included, and its significance level drops to 90 percent. Furthermore, its squared value is 0 which suggests linear relationship. Since this variable has highly relevant implications for skills and career development, the next section further investigates the relationship between crowdwork and hourly earnings by age groups.

5.2.2 2SLS Estimations, by age groups

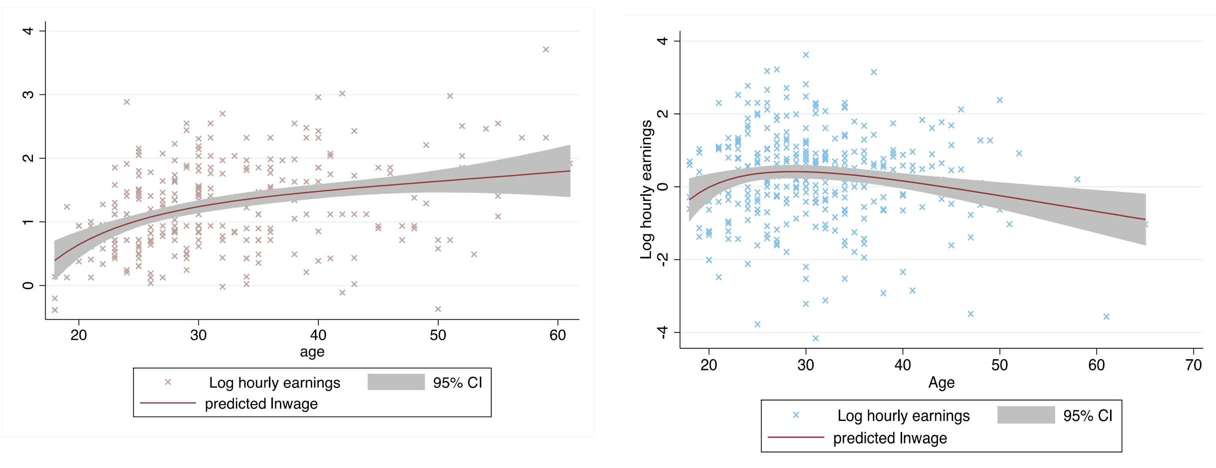
Conventional jobs typically reward experience with higher wages, but this does not necessary apply to crowdworkers. The ILO survey includes data about individuals experience in crowdwork. Figure 5.2 illustrates that the log hourly earnings do not differ significantly for the newcomers and the experiences crowdworkers. After an initial increase in the first months which might cover the rating building period, the log hourly earnings remain flat. In the OLS model, when the previously used controls are included, the coefficients show positive correlation in each group when the reference group is 0-1 month experience, but the positive relationship at 3-4 years of experience does not increase any further, as the coefficient for 5+ years experience group is lower than the coefficient of previous experience group.

Figure 5.2: Log hourly earnings by crowdwork experience



The NSSO survey does not include data about the work experience of workers, therefore, their age is examined as the indicator of experience. In the traditional information service sector, individuals can expect higher wages as they get older, as Figure 5.3 illustrates. In microtask platforms, above 30, older workers are associated with decreasing wages. As discussed in the theoretical framework and illustrated in Figure 5.2 crowdworkers might not benefit from more experience on the microtask platforms after a certain time since the work does not require any skills and there are no perspectives for promotion.

Figure 5.3.: Log earnings and age, on traditional labor market and microtask platforms



To check whether there is any difference by age, the same 2SLS model is applied for two different age groups. The younger group consists of individuals under 30 years, while the workers in the older group is at least 30 years old. This threshold was selected not only because it is the median for both groups, but also because workers above this age typically have more work experience and are more likely to have a family and undertake domestic and care work. Domestic work proves to be a strong instrument in this model as well, with positive at significant coefficient at 95 and 99 percent, and with F-statistics of 22.2 and 36.2 in younger and older group respectively.

Table 5.4 First step results of 2SLS model by age groups

	(1)	(2)	(3)
	2SLS	2SLS	2SLS
VARIABLES	All age groups	Under 30	Above 30
Domestic work	0.19***	0.20**	0.18***
	(0.049)	(0.078)	(0.059)
Constant	1.14***	5.02***	1.08**
	(0.229)	(1.846)	(0.507)
Observations	622	283	339
R-squared	0.320	0.352	0.343
F test model	47.33	22.22	36.18
P-value of F model	0	0	0

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

As Table 5.5 shows, the crowdwork coefficient is significant and negative in both age groups, but the relationship is stronger in the older group. In the younger group, workers on microtask platforms are expected to earn 37 percent than their traditional counterparts with the same characteristics, which the smallest difference observed in this chapter. Furthermore, this result is not significant, therefore, it is possible that crowdwork has no effect on the younger population's hourly earnings in the information services sector. For the older group, the results show 233 percent gap in the hourly earnings between crowdworkers and traditional workers in the traditional information service sector which is significant on 95 percent level.

Table 5.5 Crowdwork and hourly earnings by age groups, OLS and 2SLS estimates

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	2SLS	OLS	2SLS	OLS	2SLS
VARIABLES	Both age groups		Under 30		Above 30	
Crowdwork	-1.12***	-1.65**	-0.62***	-0.37	-1.46***	-2.33**
	(0.103)	(0.778)	(0.145)	(1.196)	(0.142)	(1.074)
Female	-0.10	-0.08	0.07	0.09	-0.14	-0.04
	(0.099)	(0.106)	(0.138)	(0.173)	(0.135)	(0.192)
Age	0.06*	0.06*	0.55*	0.64	0.04	0.04
	(0.035)	(0.032)	(0.301)	(0.514)	(0.079)	(0.071)
Age squared	-0.00*	-0.00*	-0.01	-0.01	-0.00	-0.00
	(0.000)	(0.000)	(0.006)	(0.011)	(0.001)	(0.001)
Education-Low	0.10	0.23	-0.49	-0.57	0.58	0.69
	(0.381)	(0.420)	(0.608)	(0.691)	(0.456)	(0.525)
Education-Medium	-0.24**	-0.27**	-0.19	-0.20	-0.20	-0.38
	(0.099)	(0.124)	(0.176)	(0.192)	(0.128)	(0.275)
Education-High	0.18	0.48	-0.03	-0.19	0.29	0.76
	(0.127)	(0.454)	(0.181)	(0.751)	(0.184)	(0.598)
Married	0.16	0.20*	0.07	0.06	0.29**	0.34**
	(0.098)	(0.119)	(0.141)	(0.150)	(0.127)	(0.168)
Household size	-0.04	-0.06	-0.04	-0.03	-0.05	-0.07
	(0.027)	(0.035)	(0.037)	(0.048)	(0.042)	(0.053)
Urban	0.34***	0.33***	0.31**	0.32**	0.40***	0.37**
	(0.098)	(0.105)	(0.144)	(0.148)	(0.138)	(0.155)
Constant	0.07	0.34	-6.42*	-7.58	0.51	0.98
	(0.600)	(0.706)	(3.622)	(6.582)	(1.570)	(1.559)
Observations	622	622	283	283	339	339
R-squared	0.231	0.194	0.187	0.177	0.308	0.225

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Education reference category: Highest

5.3 Robustness check

For robustness check, Division 82 Office administrative, office support and other business support activities is used in the first specification since this is another comparable sector to crowdworkers based the matching completed in chapter 4. The following table introduces the result of the same regression model used for the analysis on this sector. Similarly, to the previous results, the coefficients are negative and significant.

Table 5.6: Main statistics of Office administrative, office support and other business support sector

	Mean	***SD***
Mean hourly income (USD)	4.95	*3.44*
Mean log hourly income (USD)	1.36	*0.74*
Weekly working time (hours)	54.24	*8.73*
Age	39.02	*10.99*
Female	0.2103	*0.4076*
Education – No high school	0.0468	*0.2113*
Education – High school	0.5007	*0.5001*
Education- Technical	0.0163	*0.1268*
Education-Bachelor's	0.0728	*0.0859*
Education-Master's	0.3558	*0.4789*
Education- PhD	0.074	*0.2599*
Married	0.7050	*0.4562*
Urban	0.7444	*0.4363*
Household size	4.54	*1.86*
Domestic work	0.0631	*0.2433*
Number of observations	1,346	

As Table 5.7 shows the outcomes look very similar in the traditional office administrative, office support and other business support sector than in the traditional information sector. In the OLS models, all crowdwork coefficients are negative and significant at 99 percent level. In the 2SLS model, crowdwork coefficients remain negative, and similarly to the pervious results, only the first (including all crowdworkers) and the third (including only those crowdworkers that have other job too) models are significant at 90 percent

The wages of those crowdworkers that have an additional job, suffer a huge, 289 percent wage loss compared to individuals with the same characteristics in the traditional administrative sector. When taking all crowdworkers into consideration, this proportion is more modest, 249 percent. When crowdworkers do not have additional job, they are expected to earn 223 percent less than the traditional office administrative workers, but this result is not significant.

Table 5.7: Crowdwork and log hourly earnings OLS and 2SLS - Office administrative, office support sector

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS- CW all	2SLSCW all	OLS-only CW	2SLS-only CW	OLS- CW and other	2SLS-CW and other
VARIABLES	Crowdwork all		Only perform crowdwork		Crowdwork and other job	
CW	-1.19***	-2.49*	-1.13***	-2.23	-1.22***	-2.89*
	(0.105)	(1.442)	(0.130)	(2.285)	(0.159)	(1.751)
Female	-0.17***	-0.13*	-0.16***	-0.12	-0.16***	-0.15***
	(0.054)	(0.073)	(0.054)	(0.094)	(0.051)	(0.055)
Age	0.03**	0.01	0.05***	0.03	0.03**	0.04**
	(0.015)	(0.027)	(0.015)	(0.045)	(0.014)	(0.016)
Age squared	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Education - Low	-0.69***	-0.77***	-0.78***	-0.81***	-0.71***	-0.77***
	(0.108)	(0.144)	(0.097)	(0.123)	(0.111)	(0.123)
Education - Medium	-0.37***	-0.50***	-0.38***	-0.43***	-0.41***	-0.52***
	(0.039)	(0.149)	(0.038)	(0.112)	(0.037)	(0.117)
Education - High	0.13	1.13	0.12	0.99	0.10	1.44
	(0.129)	(1.102)	(0.160)	(1.812)	(0.180)	(1.405)
Married	0.10*	0.16*	0.12**	0.15*	0.08	0.11*
	(0.055)	(0.083)	(0.055)	(0.084)	(0.053)	(0.067)
Household size	-0.01	-0.02	-0.00	-0.01	-0.00	-0.01
	(0.012)	(0.019)	(0.012)	(0.016)	(0.012)	(0.015)
Urban	0.12***	0.14**	0.12***	0.13***	0.08*	0.11*
	(0.045)	(0.053)	(0.044)	(0.048)	(0.043)	(0.056)
Constant	0.45	1.14	0.15	0.70	0.46*	0.58*
	(0.288)	(0.822)	(0.278)	(1.180)	(0.279)	(0.323)
Observations	1,682	1,682	1,539	1,539	1,487	1,487
R-squared	0.309	0.163	0.315	0.229	0.290	0.124

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1
Education reference category: Highest

.

Interestingly, in this sector, there is a significant gender pay gap of 13-17 percent and the education levels show. Furthermore, higher education level is associated in higher wages which is not this clear in the information service sector. While living in urban area does not have such as strong relationship as earlier, the coefficients are somewhat smaller but still significant. Overall, the model shows very similar results in the office administrative, office support and other business than in the information service sector.

Table 5.8: First-stage results of the 2SLS model - Office administrative, office support sector

VARIABLES	(1) 2SLS India all CW	(2) 2SLS India only CW	(3) 2SLS India has other job
Domestic work	0.06**	0.04*	0.05*
	(0.029)	(0.025)	(0.027)
Constant	1.30***	1.30***	0.87***
	(0.109)	(0.106)	(0.082)
Observations	1,682	1,539	1,487
R-squared	0.491	0.472	0.487
F test model	327.7	130	90.37
P-value of F model	0	0	0

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

The robustness check shows the model hold in a different sector, even though, this sector is larger, the tasks are more general, and the sample include more individuals. The significance level has dropped in the first step of the model but in general, similar results are shown as in the main analysis.

5.4 Limitations

A limitation of this study is that the dataset covers only the year of 2017. The results, thus, should only be interpreted in the given period and in the observed sector. To capture a more general trend in India, more years should be analyzed.

Furthermore, the analysis relies on the assumption that crowdworkers are similar in their unobservable characteristics including abilities, risk aversity, ambition than their counterparts in the traditional information service sector. Additionally, we can never completely exclude the possibility that the instrumental variable, domestic work is correlated with hourly wages in an indirect way.

These results are calculated using five platforms in the crowdwork survey, namely on AMT, CrowdFlower, Clickworker, Microworkers and Prolific. While the results may hold to other microtask platforms, they should not be interpreted for other kinds of digital work such as app-based workers or online freelancers, since the tasks, algorithms, location requirement may differ in these cases.

5.5 Discussion

The results from the 2SLS estimation indicate that in most cases, OLS estimation underestimates the negative relationship between crowdwork hourly wages. Since the instrument used is valid, satisfying the instrument exogeneity and the instrument relevance requirements, and the coefficients are significant, causality is likely to occur in the model under the assumption microtask workers and traditional information services workers in the sample do not differ in their relevant unobservable characteristics. The mechanism through which the potential causal relationship works is that crowdworkers lose full-time premium, bargaining power when they undertake work on digital labor platforms which workers cannot completely control or even understand due the algorithmic management.

These results support the findings of previous studies suggesting that digital labor platforms lead to more precarious working conditions. With the assumption that individuals do not differ significantly in their unobservable characteristics, microtask platforms seem to have negative effect on the Indian information services sector wages in 2017.

The Indian workers on microtask platforms earn 165 percent less on average compared to workers with the same characteristics in the traditional information service sector in the given year.

Workers in the microtask platform work considerably fewer hours than workers in the traditional sector under investigation. However, it might not happen because of unobservable differences. Based on the detailed crowdwork dataset which provides valuable insights on crowdworkers motivation, we can see that nearly all the crowdworkers would like to spend more hours with crowdwork, but a common constraint is the lack of available tasks.

The differentiation in the 2SLS model between crowdworkers who only do crowdwork as paid work and those who hold another job is relevant, since the ILO survey offers information about the other job and thus about the skill level of the latter group. Since the majority of this group holds managerial or professional positions based on the International Standard Classification of Occupations (Interational Labor Organization, 2012), and 65 percent is working as an employee, we can make the assumption that they are not very different from the workers in the information sector in term of their skills and risk preferences.

This first model specification suggests that in 2017, working on microtask platform has negative impact on hourly earnings, those who engage in this kind of work get 165 percent lower hourly earnings than those who did similar work in the information services sector with the same background and characteristics. This gap is wider for those who hold additional job(s) and lower for those whose only paid job is this form of work.

Results of the second specification show that older workers experience higher wage difference than the younger ones. Looking at the total sample, the wage gap between the digital labor platform and traditional employment ceteris paribus is considerably smaller than in the younger group, 233 and 37 percent correspondingly. It is possible that older workers do no acquire the same level of digital skills which is needed for more successful crowdworking. The results also show that even though in crowdwork, experience is associated with higher earnings, the increase stops at the experience group of 3-4 years and the positive relationship weakens in afterwards.

While digitalization has had numerous positive impacts and supported the formalization of work in developing countries, the results suggests that crowdwork does not bring the same positive effects to the traditional information service sector in India. Instead of facilitating formality, the reverse is happening in the microtask platforms. Crowdwork should form a bridge from periods of unemployment and underemployment to formal employment but it rather forms a bridge to precarious, informal employment.

6. Policy recommendations

The results of the analysis suggest that microtask labor platforms have negative effect on hourly wages in the information service sector in India in 2017.

The results also show that the wage gap between workers on microtask platforms and traditional labor market is wider among older workers compared to younger works. These results demonstrate how workers with limited digital skills find themselves at a disadvantage as opposed those who are better equipped for the digital economy. The findings, thus, call for the need for awareness about the significance of skills that are relevant for the information technology and digital economy. Education policies should update their curriculum to make sure individuals are prepared for the shifts towards digitalization.

However, the findings also raise the question whether crowdwork is the best way to utilize high skilled labor force. As it is discussed throughout the paper, this form of employment has negative impacts not only on the wages but on the human capital and possible on the career perspectives as well under the restriction and the assumptions that are made. Therefore, when educating labor is costly, it should be considered to offer more desirable opportunities for highly educated work force that look for more flexibility.

The results show that workers in the information service sector spend way more time with work then the national legislation level of 48 hours a week. Skilled labor in the information technology sector might need more flexibility possible with fewer working hours. Even though workers earn lower wages as crowdworkers, flexibility is a prevalent reason why they decided to undertake this form of work. If workers do not have sufficient opportunities in the formal sector or that does not fit their expectations, they might look for opportunities in the informal sector. India should utilize educated workforce more efficiently by creating development strategy to absorb educated workforce in the information technology sector.

Transparent and global data source such as working time, earnings of the workers would allow consistent monitoring and analysis that could result in policies allowing more suitable labor protection for this kind of work, but the global nature of digital labor platforms makes regulation problematic. Even though platforms have negative impacts on specific sectors in India, these platforms are based in different locations around the world, often in the USA. Therefore, law or regulation regarding digital labor platforms in India might not be able to deal with the core of the issue. This challenge leads to more general question about regulating different kinds of digital platforms.

It is vital to address the possible threats of digital labor platforms before they become urgent issue, and thus, offer anticipatory solution instead of reactive ones. Nonetheless, if there is no action taken, the negative effects are expected to become more pressing in short time. The trends are becoming clearer as the number of studies in the topic grows, offering a better understanding of the areas that need intervention.

7. Conclusion

This thesis identifies the key factors which potentially affect crowdworkers' wages and working conditions as information asymmetry, power imbalance as well as the lack of transparency in algorithm management and the lack of regulation, based on the existing literature.

To compare crowdwork with similar jobs in the traditional labor market, crowdwork tasks are matched with the most similar traditional tasks based on ILO's Digital labor platforms and the future of work report and the National Industrial Classification in India.

The descriptive analysis shows that microtask workers are very similar to the traditional information services workers. A major difference between these two groups are the working hours which is significantly higher in the latter.

The OLS estimations show negative and significant relationship between hourly earnings and crowdwork which becomes even more negative after adding controls such as age, gender, education level, household size. To examine possible causality, domestic work is used in the 2SLS model as instrumental variable which is a strong and relevant instrument. The results of these models suggest that microtask platforms drive hourly wages in 2017 in the Indian information services sector, if the groups under investigations do not differ in terms of their unobservable characteristics. This implies that if traditional information service workers would engage in crowdwork, they would expect to earn 165 percent less on average as they earn in the information service sector.

This wage gap is even more prevalent among older workers, with ceteris paribus 233 percent of wage difference in contrast to the 37 percent difference among younger workers. These findings support the previous suggestions that microtask platforms lead to more precarious conditions for the workers. However, the results are not significant for workers that only undertake crowdwork as paid employment.

As this is a relatively new field, there are different aspects to investigate in the future, such as the implications of poor social security coverage, country-based discrimination on digital labor platforms. Additionally, similar analysis that does not make restriction based on sector, but on the employment status or occupation code instead might provide additional insights for understanding the labor market effects of crowdwork in India.

The results of the thesis results help understand the trends and the impacts that the digitalization of labor markets may have on different sectors. It also has impactions for developing countries where the education level of crowdworkers is typically higher than in developed countries. Overall, this thesis contributes to the literature of understanding and quantifying the possible labor market effects of crowdwork in the context of a developing country.

Bibliography

Adams, A., Freedland, M.R., Prassl, J., 2015. The'Zero-Hours Contract': Regulating Casual Work, or Legitimating Precarity?

Adebambo, B.N., Bliss, B.A., 2015. The Value of Crowdsourcing : Evidence from Earnings Forecasts.

Aguilar, M., Custovic, A., Patel, A., 2018. Crowdsourcing Economic Forecasts. SSRN Electron. J. https://doi.org/10.2139/ssrn.3160795

Al-Ani, A., Stumpp, S., 2016. Rebalancing interests and power structures on crowdworking platforms 5, 19.

Angrist, J.D., Pischke, J.-S., 2008. Mostly Harmless Econometrics: An Empiricist's Companion 290.

Aubert-Tarby, C., Escobar, O.R., Rayna, T., 2018. The impact of technological change on employment: The case of press digitisation. Technol. Forecast. Soc. Change 128, 36–45. https://doi.org/10.1016/j.techfore.2017.10.015

Autor, D.H., 2013. The "task approach" to labor markets: an overview. J. Labour Mark. Res. 46, 185–199. https://doi.org/10.1007/s12651-013-0128-z

Autor, D.H., Levy, F., Murnane, R.J., 2003. The Skill Content of Recent Technological Change: An Empirical Exploration. Q. J. Econ. 55.

Baiyere, A., Islam, A.K.M.N., Mäntymäki, M., 2019. Duality of Work in Sharing Economy Platforms 10.

Behrendt, C., Nguyen, Q.A., Rani, U., 2019. Social Protection Systems and the Future of Work: Ensuring Social Security for Digital Platform Workers. Int. Soc. Secur. Rev. Vol. 72, 17–41.

Berg, J., 2016. Income security in the on-demand economy: findings and policy lessons from a survey of crowdworkers.

Berg, J., Furrer, M., Harmon, E., Rani, U., Silberman, M.S., 2018. Digital labour platforms and the future of work. Decent Work Online World Rep. ILO.

Blinder, A.S., 2006. Offshoring: The Next Industrial Revolution? Foreign Aff. 85, 113. https://doi.org/10.2307/20031915

Boeri, T., Ours, J.C. van, 2008. The economics of imperfect labor markets. Princeton University Press, Princeton.

Borch, C., 2012. The politics of crowds: an alternative history of sociology. Cambridge University Press, Cambridge ; New York.

Borjas, G.J., 2013. Labor economics, 6th ed. ed. McGraw-Hill, New York.

Brewer, R., Morris, M.R., Piper, A.M., 2016. "Why would anybody do this?": Understanding Older Adults' Motivations and Challenges in Crowd Work, in: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16. Presented at the the 2016 CHI Conference, ACM Press, Santa Clara, California, USA, pp. 2246–2257. https://doi.org/10.1145/2858036.2858198

Bührer, C., Hagist, C., 2017. The Effect of Digitalization on the Labor Market, in: Ellermann, H., Kreutter, P., Messner, W. (Eds.), The Palgrave Handbook of Managing Continuous Business Transformation. Palgrave Macmillan UK, London, pp. 115–137. https://doi.org/10.1057/978-1-137-60228-2_5

Cahuc, P., Carcillo, S., Zylberberg, A., 2014. Labor economics, Second Edition. ed. MIT Press, Cambridge, MA.

Cantarella, M., Strozzi, C., 2018. Labour market effects of crowdwork in the US and EU: an empirical investigation 39.

Chen, D.L., 2017. The Economics of Crowdsourcing: A Theory of Disaggregated Labor Markets. SSRN Electron. J. https://doi.org/10.2139/ssrn.2928786

Choudary, S.P., Van Alstyne, M., Parker, G.G., 2016. Platform Revolution: How Networked Markets Are Transforming the Economy and How to Make Them Work for You 211.
Collier, R.B., Dubal, V.B., Carter, C., 2017. Labor Platforms and Gig Work: The Failure to Regulate. SSRN Electron. J. https://doi.org/10.2139/ssrn.3039742
De Stefano, V., 2016. Introduction: Crowdsourcing, the Gig-Economy and the Law.
De Stefano, V., 2015. The Rise of the "Just-in-Time Workforce": On-Demand Work, Crowd Work and Labour Protection in the "Gig-Economy." SSRN Electron. J. https://doi.org/10.2139/ssrn.2682602
De Stefano, V., Aloisi, A., 2018. European legal frameworkfor"digital labour platforms." Eur. Comm. https://doi.org/10.2760/78590
Deranty, J.-P., 2008. Work and the Precarisation of Existence. Eur. J. Soc. Theory 11, 443–463. https://doi.org/10.1177/1368431008097011
Durward, D., Leimeister, J.M., Blohm, I., 2016. Principal forms of crowdsourcing and crowd work. FEPS - Foundation for European Progressive Studies, Brussels, Belgium.
Evans, J., O'Brien, J., Ch Ng, B., 2018. Towards a geography of informal transport: Mobility, infrastructure and urban sustainability from the back of a motorbike. Trans. Inst. Br. Geogr. 43, 674–688. https://doi.org/10.1111/tran.12239
Evans, P.C., Gawer, A., 2016. The Rise of the Platform Enterprise 30.
Farrell, J., Klemperer, P., 2006. Coordination and Lock-In: Competition with Switching Costs and Network Effects. SSRN Electron. J. https://doi.org/10.2139/ssrn.917785
Felstiner, A., 2011. Working the crowd: employment and labor law in the crowdsourcing industry. Berkeley J. Employ. Labor Law 143–203.
Ferrant, G., Pesando, L.M., Nowacka, K., 2014. Unpaid Care Work: The missing link in the analysis of gender gaps in labour outcomes 12.
Florisson, R., Mandl, I., 2018. Platform work: Types and implications for work and employment – Literature review 131.
Forde, C., Stuart, M., Joyce, S., Oliver, L., Valizade, D., Alberti, G., Hardy, K., Trappmann, V., Umney, C., Carson, C., n.d. The Social Protection of Workers in the Platform Economy 128.
Galperin, H., Greppi, C., 2017. Geographical Discrimination in Digital Labor Platforms. SSRN Electron. J. https://doi.org/10.2139/ssrn.2922874
Goos, M., Manning, A., Salomons, A., 2014. Explaining Job Polarization: Routine-Biased Technological Change and Offshoring. Am. Econ. Rev. 104, 2509–2526. https://doi.org/10.1257/aer.104.8.2509
Graham, M., Hjorth, I., Lehdonvirta, V., 2017. Digital labour and development: impacts of global digital labour platforms and the gig economy on worker livelihoods. Transf. Eur. Rev. Labour Res. 23, 135–162. https://doi.org/10.1177/1024258916687250
Gupta, N., Martin, D., Hanrahan, B.V., O'Neill, J., 2014. Turk-Life in India, in: Proceedings of the 18th International Conference on Supporting Group Work - GROUP '14. Presented at the the 18th International Conference, ACM Press, Sanibel Island, Florida, USA, pp. 1–11. https://doi.org/10.1145/2660398.2660403
Horton, J.J., Chilton, L.B., 2010. The labor economics of paid crowdsourcing, in: Proceedings of the 11th ACM Conference on Electronic Commerce. ACM, pp. 209–218.
Howcroft, D., Bergvall-Kåreborn, B., 2019. A Typology of Crowdwork Platforms. Work Employ. Soc. 33, 21–38. https://doi.org/10.1177/0950017018760136
Howe, J., 2008. Crowdsourcing: Why the Power of the Crowd is Driving the Future of Business. New York, NY, USA.
Hu, M., 2019. Sharing Economy - Making Supply Meet Demand. Springer.
Interational Labor Organization, 2012. International Standard Classification of Occupations.

International Labour Organization, 2007. Resolution Concerning Updating the International Standard Classification of Occupations.
Irani, L.C., Silberman, M.S., 2013. Turkopticon: interrupting worker invisibility in amazon mechanical turk, in: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI ’13. Presented at the the SIGCHI Conference, ACM Press, Paris, France, p. 611. https://doi.org/10.1145/2470654.2470742
Jacques, J.T., Kristensson, P.O., 2019. Crowdworker Economics in the Gig Economy, in: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems - CHI ’19. Presented at the the 2019 CHI Conference, ACM Press, Glasgow, Scotland Uk, pp. 1–10. https://doi.org/10.1145/3290605.3300621
Jäger, G., Zilian, L.S., Hofer, C., Füllsack, M., 2019. Crowdworking: working with or against the crowd? J. Econ. Interact. Coord. 14, 761–788. https://doi.org/10.1007/s11403-019-00266-1
Jinnai, Y., n.d. Understanding Job Matching in Online Labor Markets: Evidence from the World’s Largest Crowdsourcing Service 19.
Kalleberg, A.L., Vallas, S.P. (Eds.), 2017. Probing Precarious Work: Theory, Research, and Politics, in: Research in the Sociology of Work. Emerald Publishing Limited, pp. 1–30. https://doi.org/10.1108/S0277-283320170000031017
Kässi, O., Lehdonvirta, V., 2018a. Online labour index: Measuring the online gig economy for policy and research. Technol. Forecast. Soc. Change 137, 241–248. https://doi.org/10.1016/j.techfore.2018.07.056
Kässi, O., Lehdonvirta, V., 2018b. Online labour index: Measuring the online gig economy for policy and research. Technol. Forecast. Soc. Change 137, 241–248. https://doi.org/10.1016/j.techfore.2018.07.056
Kaufmann, N., Schulze, T., Veit, D., 2011. More than fun and money. Worker Motivation in Crowdsourcing – A Study on Mechanical Turk 12.
Kittur, A., Nickerson, J.V., Bernstein, M., Gerber, E., Shaw, A., Zimmerman, J., Lease, M., Horton, J., 2013. The future of crowd work, in: Proceedings of the 2013 Conference on Computer Supported Cooperative Work - CSCW ’13. Presented at the the 2013 conference, ACM Press, San Antonio, Texas, USA, p. 1301. https://doi.org/10.1145/2441776.2441923
Kokkodis, M., Papadimitriou, P., Ipeirotis, P.G., 2015. Hiring Behavior Models for Online Labor Markets, in: Proceedings of the Eighth ACM International Conference on Web Search and Data Mining - WSDM ’15. Presented at the the Eighth ACM International Conference, ACM Press, Shanghai, China, pp. 223–232. https://doi.org/10.1145/2684822.2685299
Körfer, A., Röthig, O., 2017. Decent crowdwork ? the fight for labour law in the digital age. Transf. Eur. Rev. Labour Res. 23, 233–236. https://doi.org/10.1177/1024258917701364
Leimeister, J.M., Durward, D., Zogaj, S., 2016. Crowd worker in Deutschland: eine empirische Studie zum Arbeitsumfeld auf externen Crowdsourcing-Plattformen.
Machine, D., Ophoff, J., 2014. Understanding What Motivates Participation on Crowdsourcing Platforms, in: E-Skills for Knowledge Production and Innovation Conference, Cape Town, South Africa. pp. 191–200.
Mäntymäki, M., Baiyere, A., Islam, A.K.M.N., 2019. Digital platforms and the changing nature of physical work: Insights from ride-hailing. Int. J. Inf. Manag. 49, 452–460. https://doi.org/10.1016/j.ijinfomgt.2019.08.007
Mason, W., Watts, D.J., 2009. Financial incentives and the “performance of crowds. Knowl. Discov. Data Min.- Hum. Comput. KDD-HCOMP.
Mincer, J., n.d. Human Capital and the Labor Market 8.

Naderi, B., 2018. Motivation of Workers on Microtask Crowdsourcing Platforms, T-Labs Series in Telecommunication Services. Springer International Publishing, Cham. https://doi.org/10.1007/978-3-319-72700-4
Nischak, F., Hanelt, A., Kolbe, L.M., 2017. Unraveling the Interaction of Information Systems and Ecosystems-A Comprehensive Classification of Literature.
Palier, B., Thelen, K., 2012. Dualization and Institutional Complementarities: Industrial Relations, Labor Market and Welfare State Changes in France and Germany. pp. 201–224.
Pénin, J., Helmchen, T.B., 2011. Crowdsourcing of inventive activities: definition and limits. Int. J. Innov. Sustain. Dev. 5, 246. https://doi.org/10.1504/IJISD.2011.043068
Rani, U., Furrer, M., 2019. On-Demand Digital Economy: Can Experience Ensure Work and Income Security for Microtask Workers? Jahrb. Für Natl. Stat. 239. https://doi.org/10.1515/jbnst-2018-0019
Rani, U., Singh, P.J., 2019. Digital Platforms, Data, and Development: Implications for Workers in Developing Economies. Comp. Labor Law Policy J. 101–123.
Schenk, E., Guittard, C., 2011. Towards a characterization of crowdsourcing practices. J. Innov. Econ. 7, 93. https://doi.org/10.3917/jie.007.0093
Schmidt, F.A., 2017. Digital Labour Markets in the Platform Economy 32.
Schriner, A., Oerther, D., 2014. No Really, (Crowd) Work is the Silver Bullet. Procedia Eng. 78, 224–228. https://doi.org/10.1016/j.proeng.2014.07.060
Scribano, A., Lisdero, P., 2019. Digital Labour, Society and the Politics of Sensibilities.
Seppälä, T., Halén, M., Juhanko, J., Korhonen, H., Mattila, J., Parviainen, P., Talvitie, J., Ailisto, H., Hyytinen, K.-M., Kääriäinen, J., 2015. The Platform–History, Characteristics, and the Definition. The Research Institute of the Finnish Economy.
Shafiei Gol, E., Stein, M.-K., Avital, M., 2019. Crowdwork platform governance toward organizational value creation. J. Strateg. Inf. Syst. 28, 175–195. https://doi.org/10.1016/j.jsis.2019.01.001
Smith, R., Leberstein, S., 2015. Rights on Demand.
Spulber, D.F., 2019. The economics of markets and platforms. J. Econ. Manag. Strategy 28, 159–172. https://doi.org/10.1111/jems.12290
Sudoh, O. (Ed.), 2005. Digital economy and social design. Springer, New York, NY.
Taeihagh, A., 2017. Crowdsourcing, Sharing Economies and Development. J. Dev. Soc. 33, 191–222. https://doi.org/10.1177/0169796X17710072
Templeton, A., Drury, J., Philippides, A., 2015. From Mindless Masses to Small Groups: Conceptualizing Collective Behavior in Crowd Modeling. Rev. Gen. Psychol. 19, 215–229. https://doi.org/10.1037/gpr0000032
Todolí-Signes, A., 2017. The ‘gig economy’: employee, self-employed or the need for a special employment regulation? Transf. Eur. Rev. Labour Res. 23, 193–205. https://doi.org/10.1177/1024258917701381
UNCTAD, 2019. Digital Economy Report 2019: value creation and capture - implications for developing countries. United Nations, S.l.
Vakharia, D., Lease, M., 2013. Beyond AMT: An Analysis of Crowd Work Platforms. ArXiv13101672 Cs.
Wexler, M.N., 2011. Reconfiguring the sociology of the crowd: Exploring crowdsourcing. Int. J. Sociol. Soc. Policy 31, undefined-undefined. https://doi.org/10.1108/01443331111104779
Zyskowski, K., Morris, M.R., Bigham, J.P., Gray, M.L., Kane, S.K., 2015. Accessible Crowdwork?: Understanding the Value in and Challenge of Microtask Employment for People with Disabilities, in: Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing - CSCW ’15. Presented

at the the 18th ACM Conference, ACM Press, Vancouver, BC, Canada, pp. 1682–1693. https://doi.org/10.1145/2675133.2675158

Appendix

Table A.1: Office administrative, office support and other business support activities NIC with descriptions

NIC	Task description
Office administrative and support activities (821)	8211 Combined office administrative service activities This class includes the provision of a combination of day to day office administrative services, such as reception, financial planning, billing and record keeping, personnel and mail services etc. for others on a contract or fee basis 82110 Combined office administrative service activities 8219 Photocopying, document preparation and other specialized office support activities This class excludes: - printing of documents (offset printing, quick printing etc.), see 1811 - specialized stenotype services such as court reporting, see 8299 - public stenography services, see 8299 82191 Photocopying, duplicating and bluprinting services 82192 Document preparation, typing, word processing and desktop publishing sevices 82199 Other specialised office support services activities
Other specialized office support services activities (822)	Other specialized office support services activities 8220 Activities of call centres
Organization of conventions and trade shows (823)	Organization of conventions and trade shows This class includes the organization, promotion and/or management of events, such as business and trade shows, conventions, conferences and meetings, whether or not including the management and provision of the staff to operate the facilities in which these events take place. 82300 Organization of conventions and trade shows
Business support service activities (829)	This class includes the collection of payments for claims and remittance of payments collected to the clients, such as bill or debt collection services. This class also includes the activities of compiling information, such as credit and employment histories on individuals and credit histories on businesses, and providing the information to financial institutions, retailers, and others who have a need to evaluate the creditworthiness of these persons and businesses.

	82910 Activities of collection agencies and credit bureaus 82990 Other business support service activities

Table A.2: Percentage of female, male and total workers by tasks undertaken on micro platforms in India

Crowdwork task	Female	Male	Total
Artificial intelligence and machine learning (%)	17.39	11.68	12.83
Categorization (%)	24.64	19.71	20.7
Content access (%)	40.58	44.16	43.4
Content creation and editing (%)	31.88	22.26	24.2
Content moderation (%)	7.25	8.03	7.9
Data collection (%)	**47.83**	**44.53**	**45.2**
Market research and reviews (%)	10.14	10.95	10.8
Surveys and experiments (%)	**79.71**	**70.44**	**72.30**
Transcription (%)	**65.22**	**60.22**	**61.2**
Verification and validation (%)	7.25	6.93	7.0

Table A.3: Crowdworkers' occupation according to the International Standard Classification of Occupations

Occupation	Percentage of crowdworkers	Skill level
Manager	18.31%	3+4
Professional	40.85%	4
Technician and Associate Professional	7.75%	3
Clerical Support	7.04%	2
Services and Sales	6.34%	2
Skilled Agricultural, Forestry, and Fis	0.70%	2
Craft and Related Trades	4.93%	2
Plant and Machine Operators, Assembler	0.70%	2
Elementary Occupation	2.11%	1
Armed Forces	2.11%	1+2+4
Others	9.15%	
Number of observations	142	

Table A.4:Main motivations for doing crowdwork

Most important reason for doing crowdwork	**Percentage of crowdworkers**
Could not find other employment	6.41
I can only work from home	11.08
I prefer to work from home	25.95
Pay is better than other available jobs	8.45
To complement pay from other jobs	24.49
To earn money while going to school	6.71
As a form of leisure	3.50
I enjoy it	13.41
Other	0
Number of obervations	

Table A.5: OLS estimation of experience level and hourly wages among crowdworkers

VARIABLES	(1) OLS Crowdworkers
B1==1-6 months	0.68
	(0.439)
B1==7-12 months	0.69*
	(0.410)
B1==1-2 years	0.86**
	(0.427)
B1==3-4 years	1.16***
	(0.410)
B1==5+ years	1.13***
	(0.423)
Female	-0.14
	(0.188)
Age	0.05
	(0.052)
Age squared	-0.00
	(0.001)
Education-Low	0.34
	(0.435)
Education-Medium	-0.29
	(0.241)
Education-Highest	-0.22
	(0.158)
Married	0.08
	(0.173)
Household size	-0.06
	(0.061)
Urban	0.26
	(0.168)
Constant	-1.11
	(0.970)
Observations	336
R-squared	0.085

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1