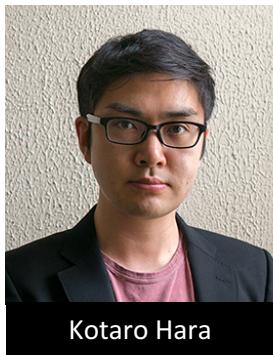


# A Data-Driven Analysis of Workers' Earnings on Amazon Mechanical Turk

Kotaro Hara, Abigail Adams, Kristy Milland, Saiph Savage

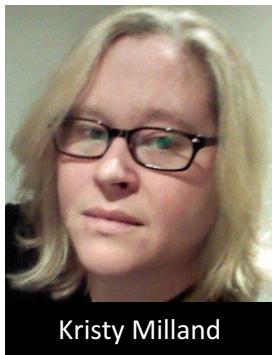
Chris Callison-Burch, Jeffrey P. Bigham



Kotaro Hara



Abigail Adams



Kristy Milland



Saiph Savage



Chris Callison-Burch



Jeffrey P. Bigham

# A Data-Driven Analysis of Workers' Earnings on Amazon Mechanical Turk

Kotaro Hara, Abigail Adams, Kristy Milland, Saiph Savage  
Chris Callison-Burch, Jeffrey P. Bigham



Carnegie  
Mellon  
University



McMaster  
University

W West Virginia  
University

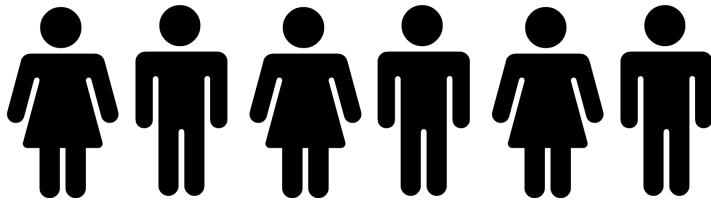
UPenn  
UNIVERSITY OF PENNSYLVANIA



CHI 2018  
Engage with CHI

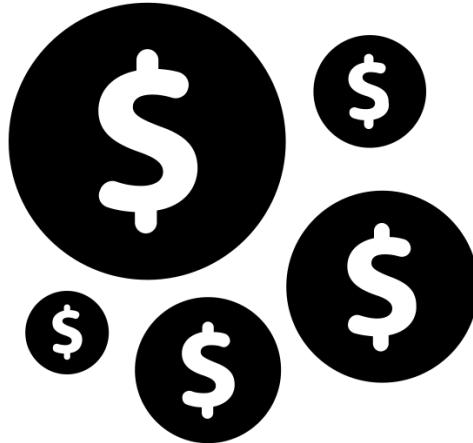
# 600k

online workers and counting



Online outsourcing  
industry generated

\$2 b





Are workers treated fairly? How does this new work style affect their lives?



Incom  
econo  
from  
  
Janine

## Being A

**David Martin, Benjamin V. Hanrahan, Jacki C**

Xerox Research Centre Europe

6 chemin de Maupertuis, Grenoble France

{david.martin, ben.hanrahan, jacki.oneill}@xrce.x

### ABSTRACT

We conducted an ethnomet hodological analysis of publicly available content on Turker Nation, a general forum for Amazon Mechanical Turk (AMT) users. Using forum data we provide novel depth and detail on how the Turker Nation members operate as economic actors, working on which Requesters and jobs are worthwhile to them. We show some of the key ways Turker Nation functions as community and also look further into Turker-Requester relationships from the Turker perspective – considering practical, emotional and moral aspects. Finally, following Star and Strauss [25] we analyse Turking as a form invisible work. We do this to illustrate practical and ethical issues relating to working with Turkers and AMT, and promote design directions to support Turkers and their relationships with Requesters.

### Author Keywords

Ethnomet hodology; content analysis; crowdsourcing; microtasking; Amazon Mechanical Turk; Turker Nation

### ACM Classification Keywords

H.5.3 Group and Organizational Interfaces – Computer Supported Cooperative Work

### General Terms

Human Factors

### INTRODUCTION

The concept of crowdsourcing was originally described by Jeff Howe of Wired Magazine as “the act of a company or institution taking a function once performed by its own employees and outsourcing it to an undefined (and generally large) pool of people in the form of an open call.” The ‘undefined network of people’ is the key topic of this article. We present the findings of an ethnometric analysis of posts and threads on a crowdsourcing platform called Turker Nation<sup>1</sup>. We have sought to understand the reasons, concerns, and relationships with requesters as they are shown in their posts on the forum – present them as faithfully as possible, in their own words:

For years, such labor was mediated by annotation and sub-

Permission to make digital or

classroom copies is granted by the copyright holder for non-commercial educational personal use only. To copy otherwise, to republish, to post on servers, or to redistribute to lists, requires prior permission from the author(s). Publication rights licensed to ACM.

## Examining Cro His

Ali A

{ali.al

### ABSTRACT

The internet is empowering the rise and other forms of on-demand learning body of scholarship has attention shifted to the technical outcomes of this shift, e.g., 1) What are the implications of this shift? 2) How far can work be distributed?, 3) What will work like for workers? In this paper, we argue that workers on piecework – a similar 20th century – to understand how with modern on-demand work, identify whether on-demand work might differentiate itself. This grounding that can help address questions in crowd work, and that learn from history rather than from the past.

### ACM Classification Keywords

H.5.3. Information Interface Design; Group and Organization Interfaces

### Author Keywords

Crowd work; gig work; on-

### INTRODUCTION

The past decade has seen a shift towards mediated labor. A framming components enables coordination of workers at scale [68], workers engage in work with little to no awareness and often with fleeting

permissions

Designers are more than those who seek to move from current states to preferred ones. Designers move from relatively high rung in hierarchy to low rung in hierarchy.

## Stories We Tell About Labor: Turkopticon and the Trouble with “Design”

M. Six Silberman

IG Metall

60329 Frankfurt, Germany

michael.silberman@igmetall.de

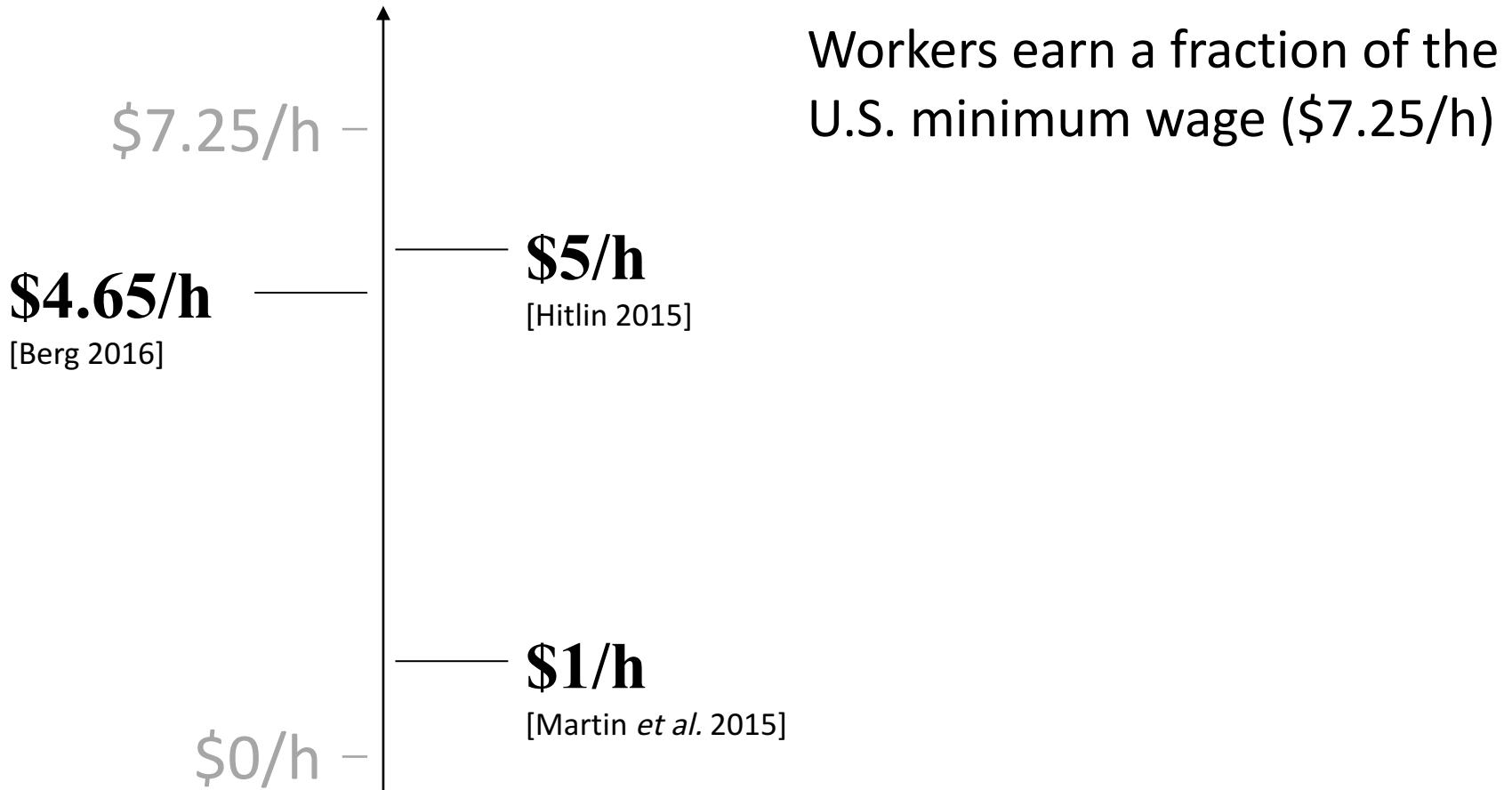
projects. The World Bank, for example, cites design as an engine of “new value chains” in the face of global competition that drives existing commodity profit margins to zero [16]. Design is core to economic growth policies in China [49], and India [68]. American Economic Policy looks to hacking, 3-D printing, and STEAM (Science, Technology, Engineering, Arts, and Math) education to transform workers into citizens who can both generate new sources of financial value and improve material conditions for living.

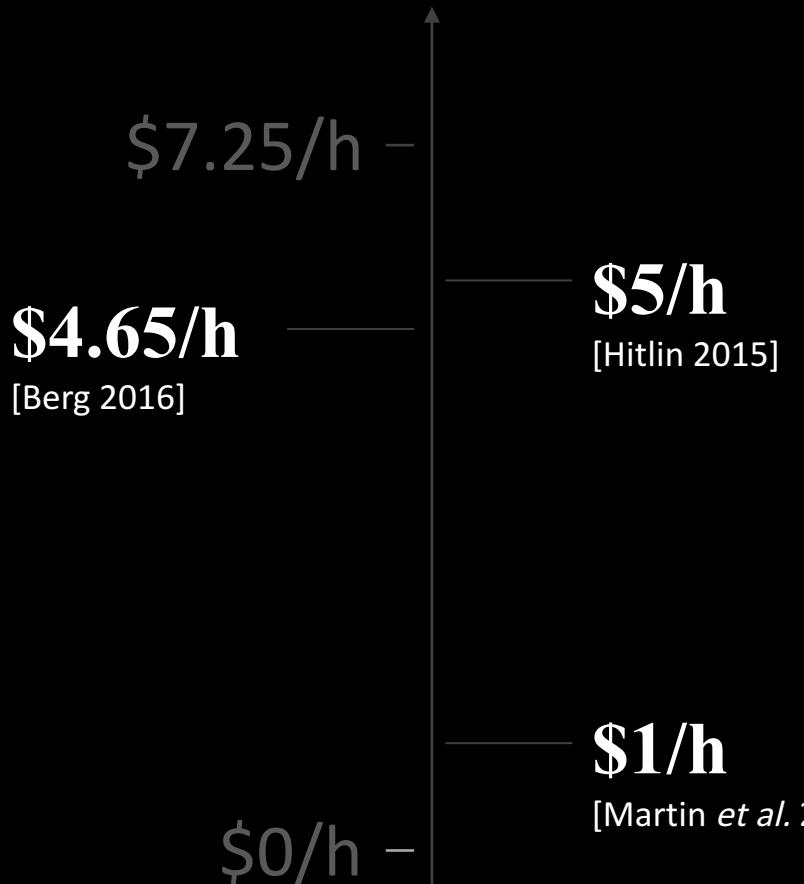
Within such a milieu, designers and HCI practitioners have a privileged place as a research community that self-consciously attempts to generate both the futures of pervasive technologies and methods for generating those futures. We are not simply Herbert Simon’s designers in pursuit of preferred states [77:111], but privileged economic actors. These stories of economic and social progress sustain us institutionally, but they also become complicities and liabilities for those who wish to redistribute power through design practice. We encountered these problems as designers of Turkopticon, an activist intervention into Amazon Mechanical Turk (see [45]). In this paper, we explain how cultural and economic understandings of design shaped how broader publics interpreted our intervention, with problematic consequences for the workers the project sought to support. We describe the conflict between “design” as a cultural position to speak from and the projects’ labor politics. We then describe how we expanded our tactics beyond design itself to sustain the projects’ goals to improve digital microwork.

This paper contributes to HCI scholarship in design, systems development, and innovation.

workers with large clusters who want to do those workers as independent contractors; this means they are not entitled to minimum wage or other employment benefits. Turkopticon came out of engagements with workers – to articulate their needs and to find ways to meet them.

Martin et al. 2014; Berg 2016; Irani and Silberman 2016; Alkhateeb et al. 2017





These figures are subjective data based on workers' opinions on an online forum and survey responses

The lack of **reward** and task duration data has prevented us from objectively analysing workers' hourly wage



Crowd  
Workers

**3.8m**  
task records

# Research Questions



How much are workers earning on Amazon Mechanical Turk?

What contributes to the low wage?

# Research Questions



**How much are workers earning on Amazon Mechanical Turk?**

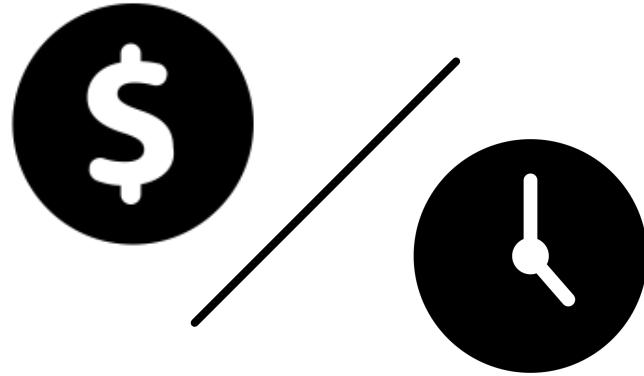
What contributes to the low wage?

# Data

- N=2,676 workers
- Task description
  - title, keywords, description, task IDs, requester IDs, **reward (\$)**
- Task status
  - submitted *vs.* returned
  - **Timestamps (task start, task end, task return)**

# Data

- N=2,676 workers
  - Task description
    - title, keywords, description, task IDs, requester IDs, **reward (\$)**
  - Task status
    - submitted vs. returned
    - **Timestamps (task start, task end, task return)**
- 
- These pieces of information enable us to calculate hourly wage



It is surprisingly hard to get accurate estimation of hourly wage

# Hourly Wage Estimation (Naïve)



Task Reward (\$) / Task Interval = Per-task Hourly Wage

# Hourly Wage Estimation (Naïve)



Worker

$$\frac{\sum \text{ Task Reward } (\$)}{\sum \text{ Task Interval}} = \text{Per-worker Hourly Wage}$$

**Naïve method** of calculating hourly wage

# Hourly Wage Estimation (Naïve)



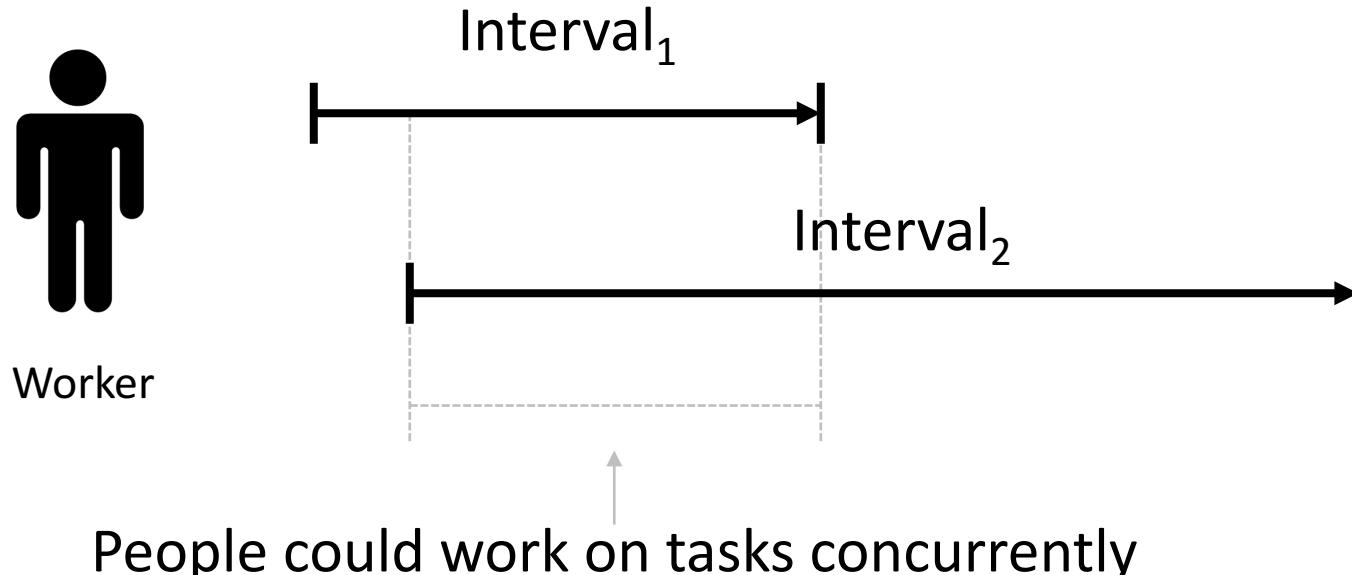
The naïve method may under- or over-estimate the hourly wage

$\sum$  Task Intervals

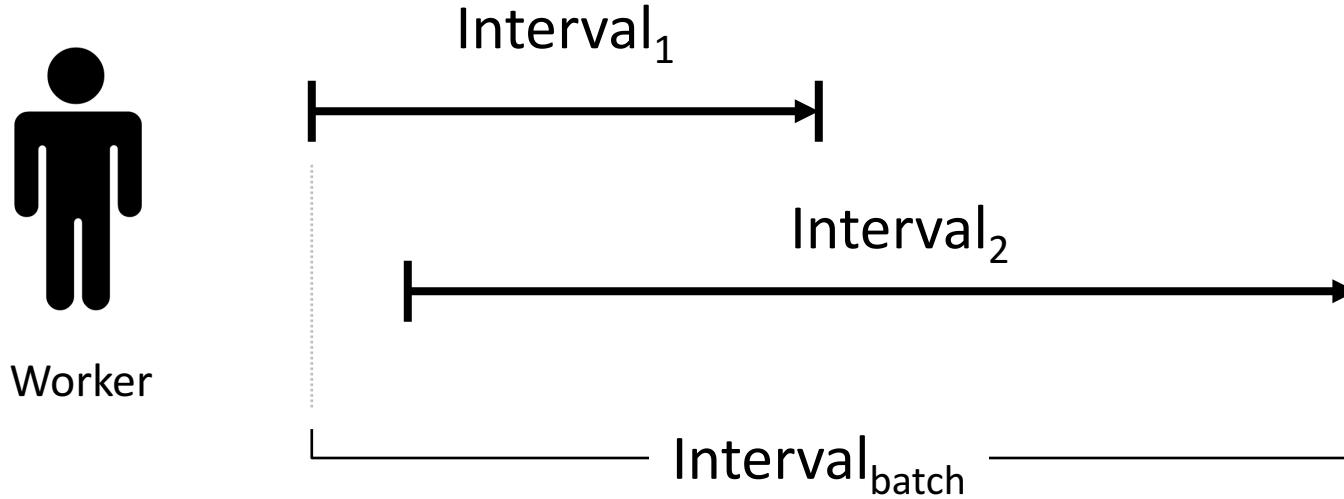
A large, tilted white rectangular box contains this text. The top-left corner features a black stick figure icon. The text "The naïve method may" is in black, while "under- or over-estimate the hourly wage" is in red. Below this, the mathematical expression " $\sum$  Task Intervals" is also in black. The entire box is angled diagonally from the top-left towards the bottom-right.

**Naïve method** of calculating hourly wage

# Wage Under-estimation

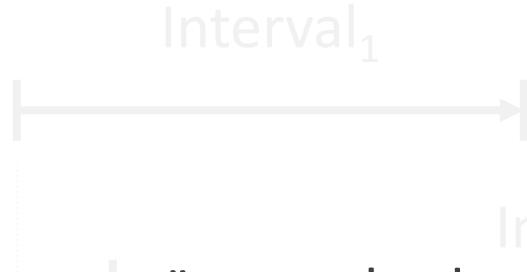


# Wage Under-estimation



$$\text{Interval}_{\text{batch}} < \text{Interval}_1 + \text{Interval}_2$$

# Wage Under-estimation

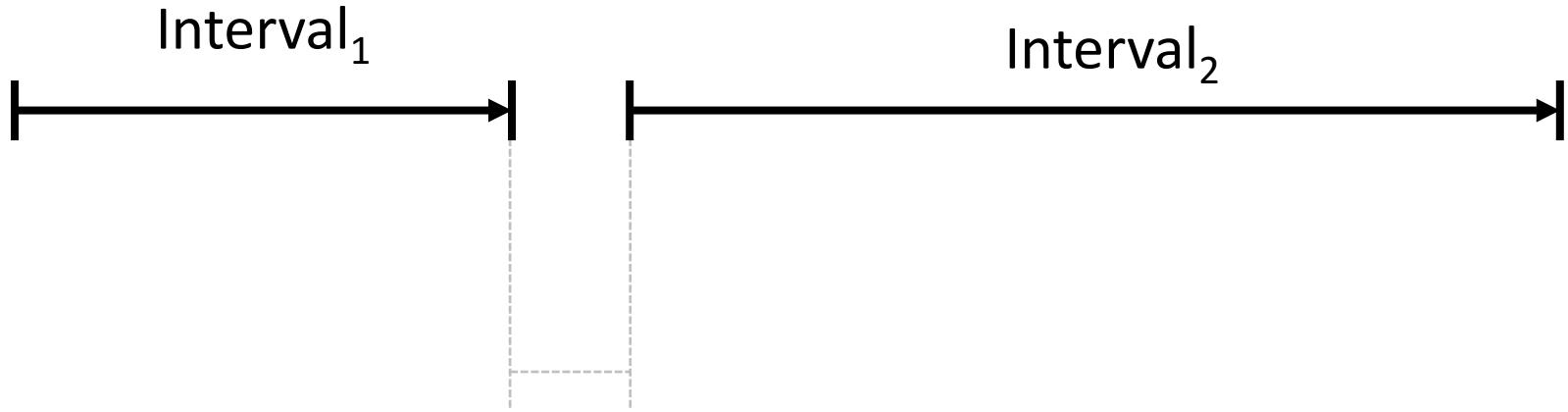


This may cause naïve method to **over-estimate work durations due to interval overlaps** and under-estimate the hourly wage



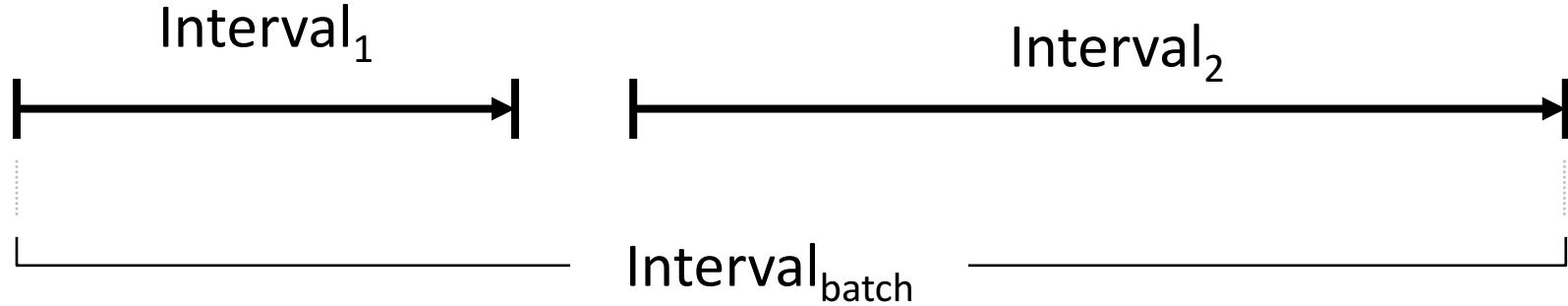
$$\text{Interval}_{\text{batch}} < \text{Interval}_1 + \text{Interval}_2$$

# Wage Over-estimation



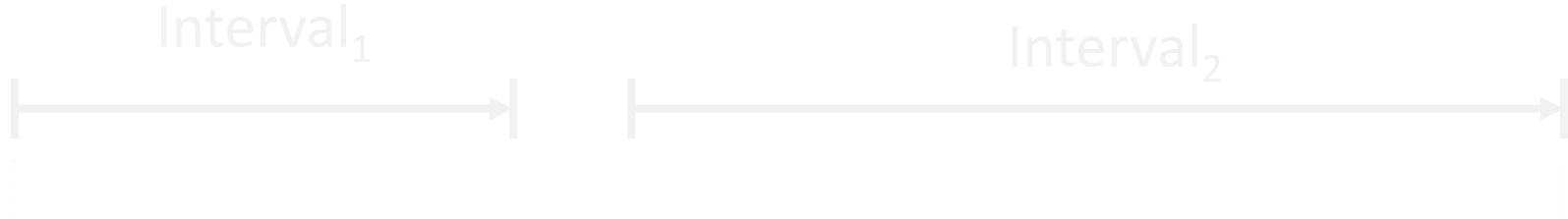
There could be a short gap between two tasks  
(e.g., time to search for a task)

# Wage Over-estimation



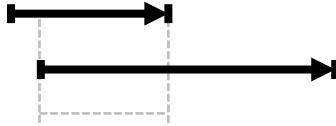
$$\text{Interval}_{\text{batch}} > \text{Interval}_1 + \text{Interval}_2$$

# Wage Over-estimation

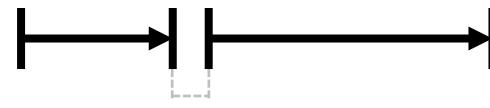


The naïve method may under-estimate a work interval due to  
**time between tasks** and over-estimate the hourly wage

$$\text{Interval}_{\text{batch}} > \text{Interval}_1 + \text{Interval}_2$$



Interval overlap

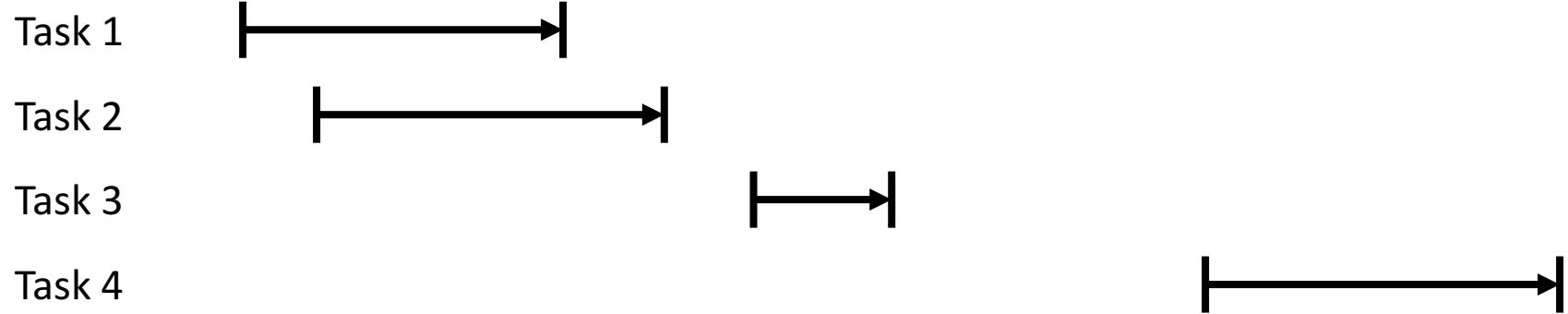


Time between tasks

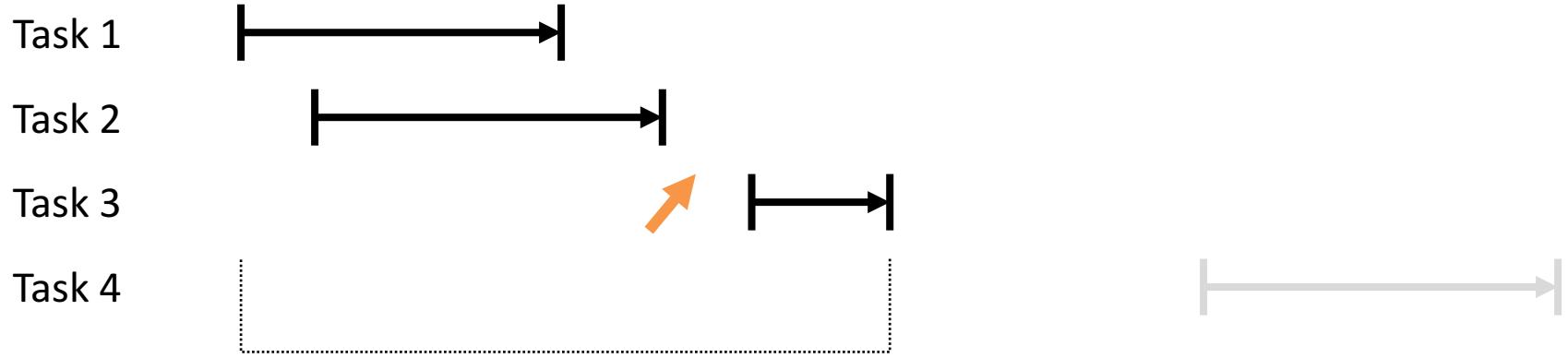
Wage over- and under-estimation may affect the accuracy of hourly wage calculation

To reduce the effects of interval overlaps and time between tasks, we used a **temporal clustering method** to compute hourly wage

# Temporal Clustering

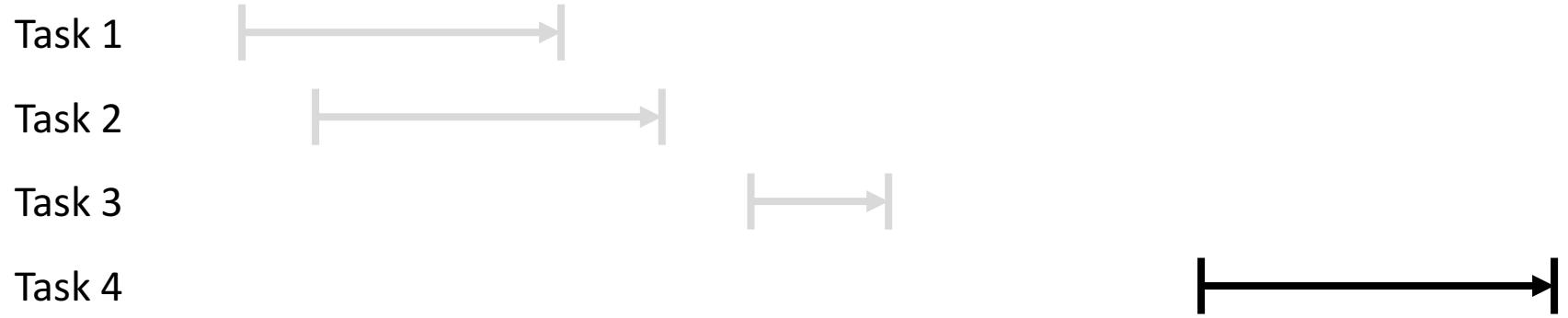


# Temporal Clustering



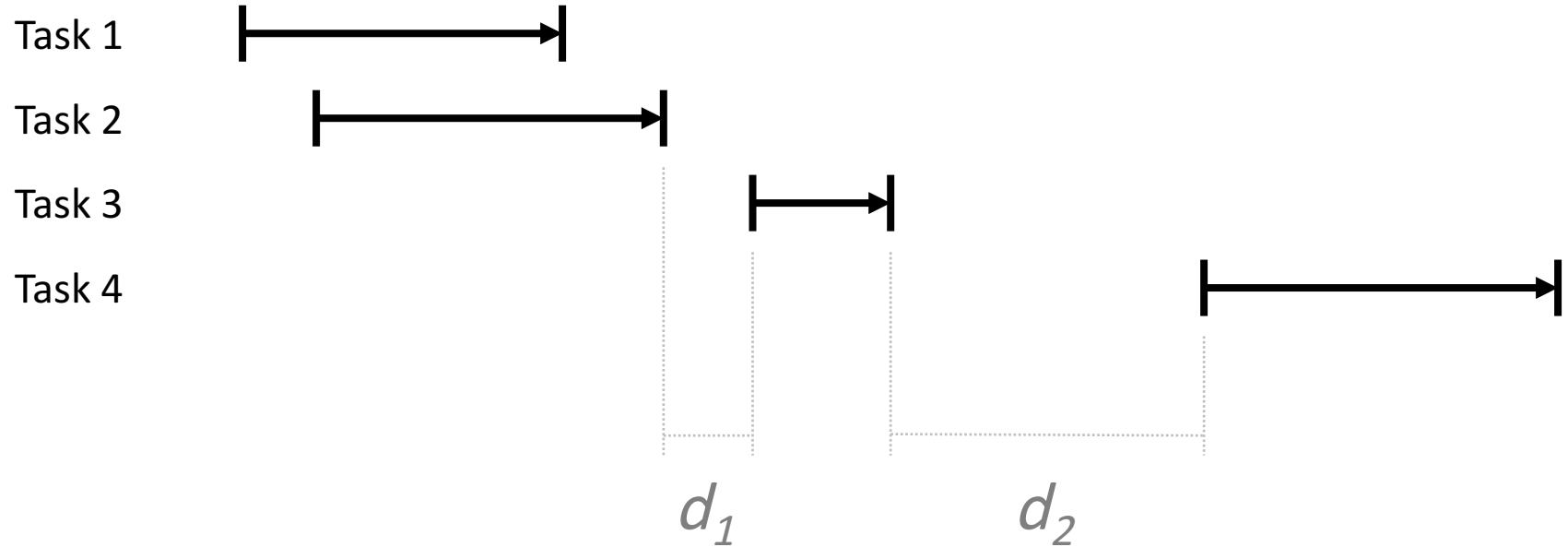
We want to cluster temporally close  
tasks together to ignore this **small gap**

# Temporal Clustering

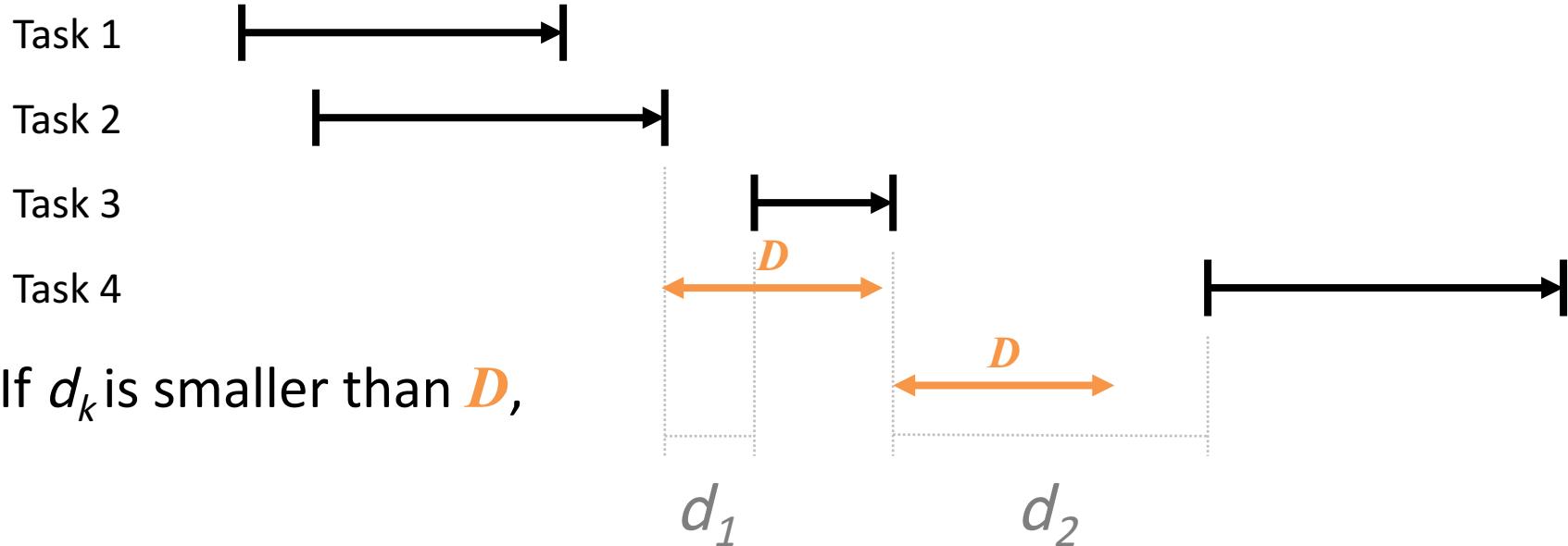


While keeping this isolated task disjoint

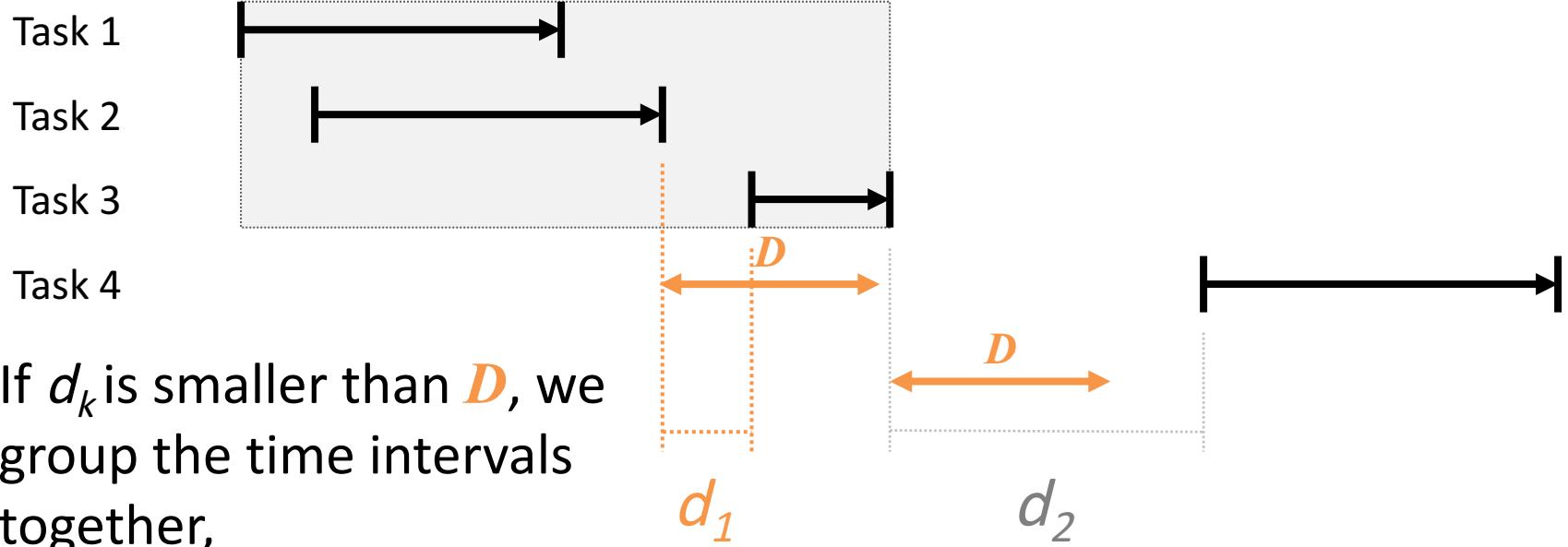
# Temporal Clustering



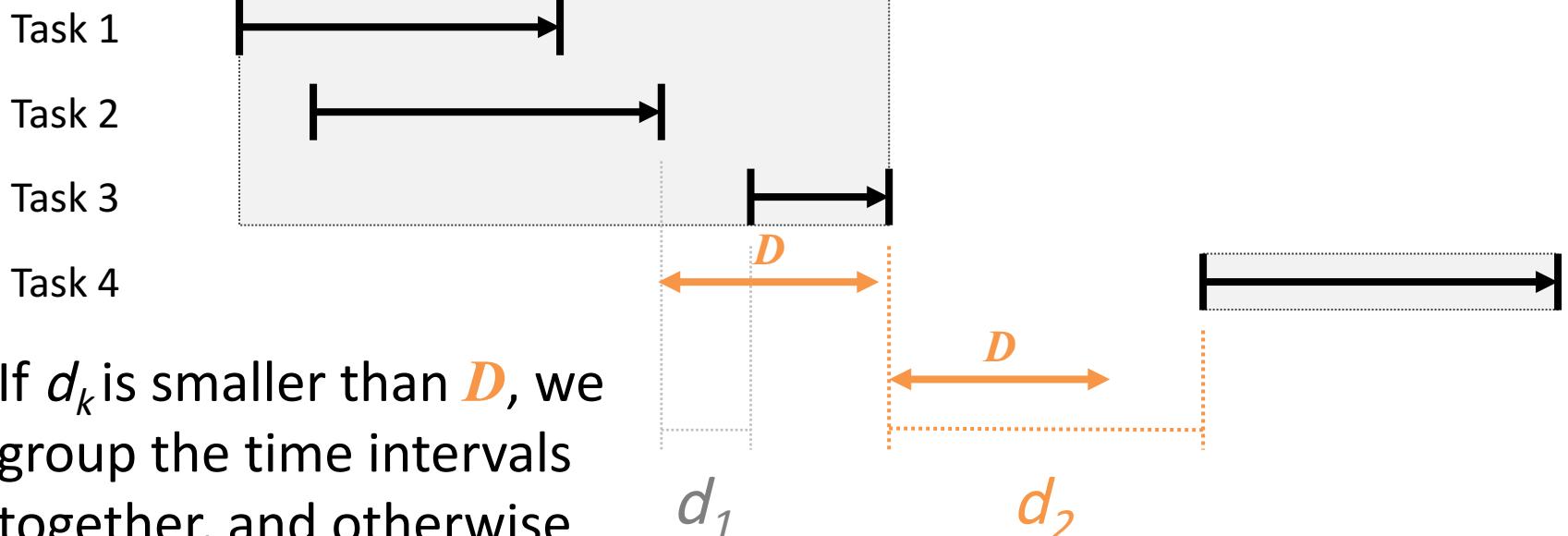
# Temporal Clustering



# Temporal Clustering

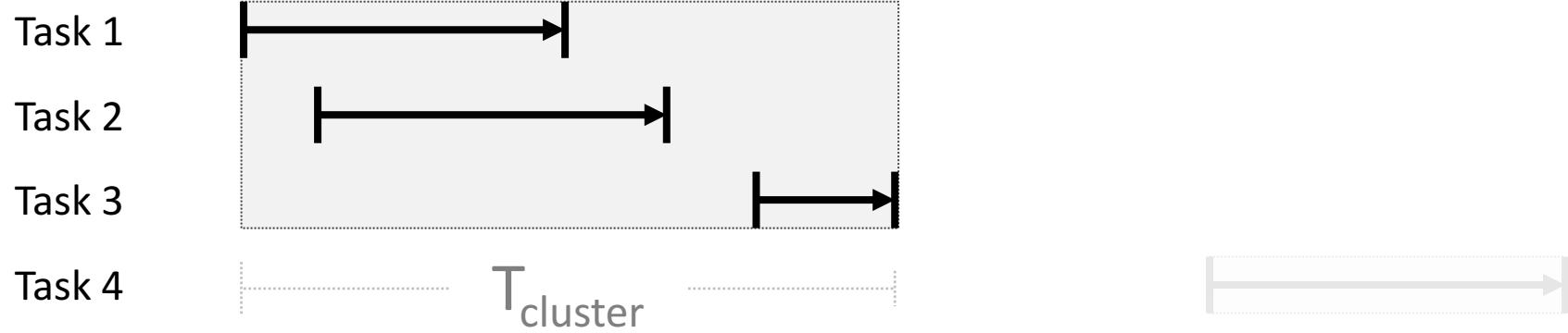


# Temporal Clustering

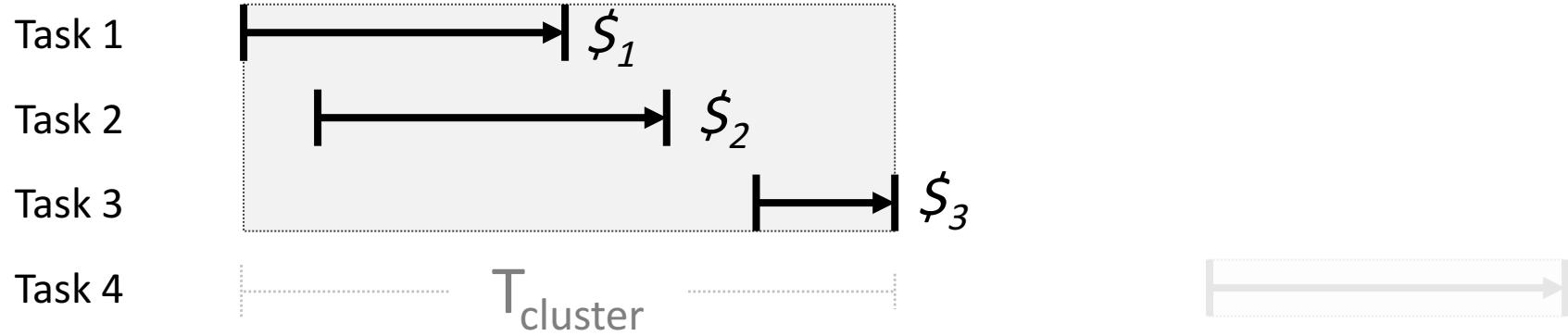


If  $d_k$  is smaller than  $D$ , we group the time intervals together, and otherwise keep them disjoint

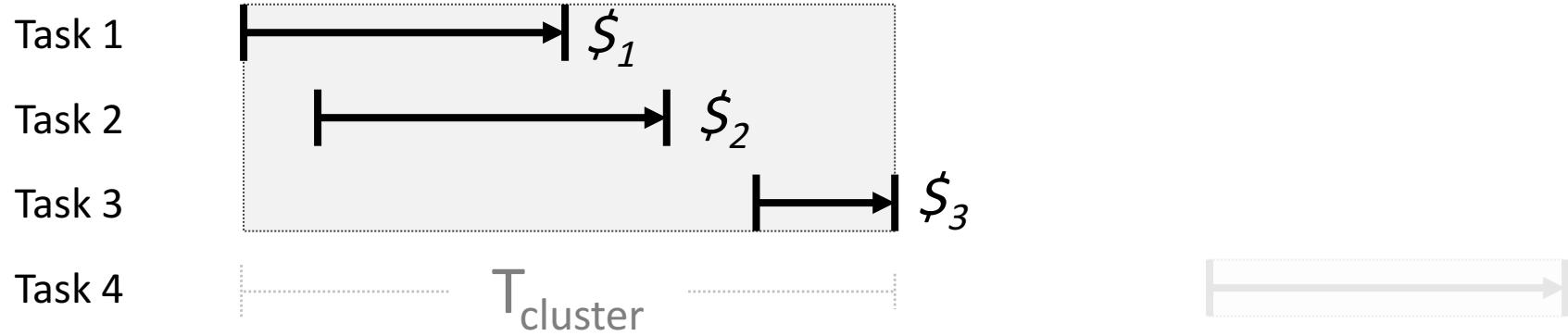
# Temporal Clustering: Cluster-based Hourly Wage



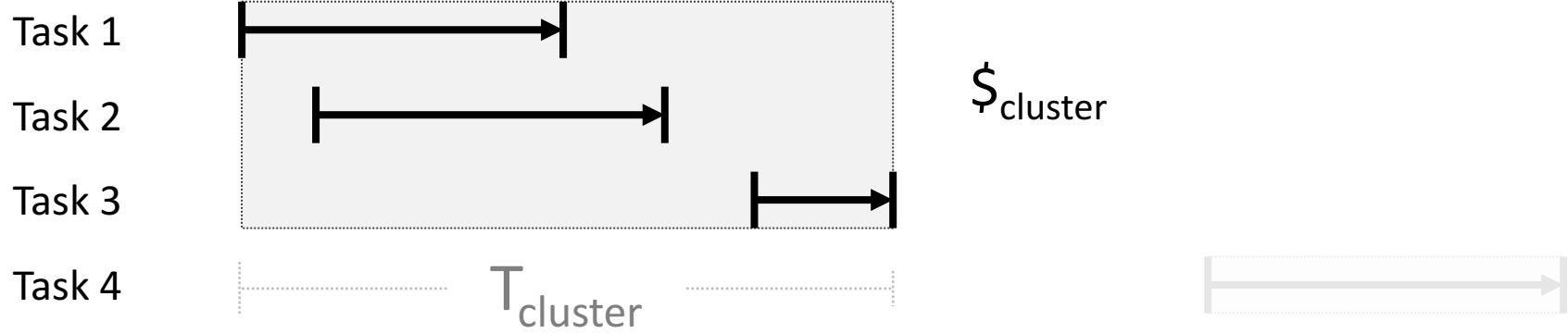
# Temporal Clustering: Cluster-based Hourly Wage



# Temporal Clustering: Cluster-based Hourly Wage

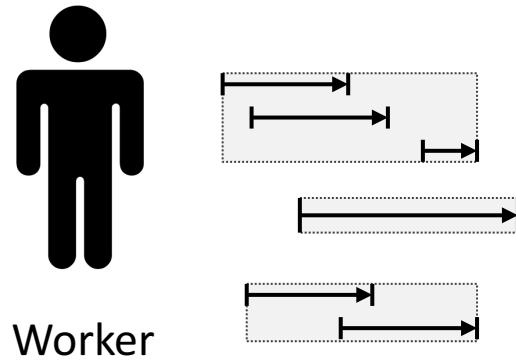


# Temporal Clustering: Cluster-based Hourly Wage



We define **per-cluster hourly wage** as  $\$_{\text{cluster}} / T_{\text{cluster}}$

# Temporal Clustering: Cluster-based Hourly Wage



A black silhouette of a person icon labeled "Worker" is positioned to the left of three horizontal bars. Each bar is enclosed in a dotted rectangle and contains a double-headed arrow indicating its length. The first bar is short, the second is medium, and the third is long.

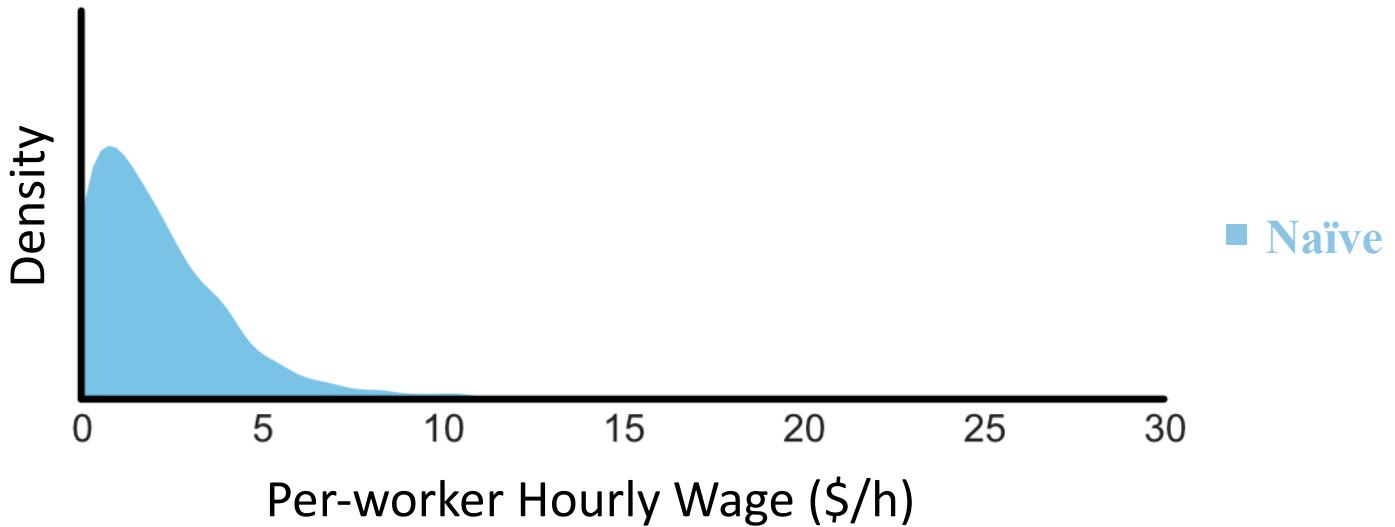
$$\frac{\sum \text{Cluster Reward} (\$)}{\sum \text{Cluster Interval}} = \text{Per-worker Hourly Wage with Clustering}$$

Because different choice of  $D$  yield different sets of clusters, we use  $D=0$  and  $D=1$  minute and see their effects on cluster-based hourly wages

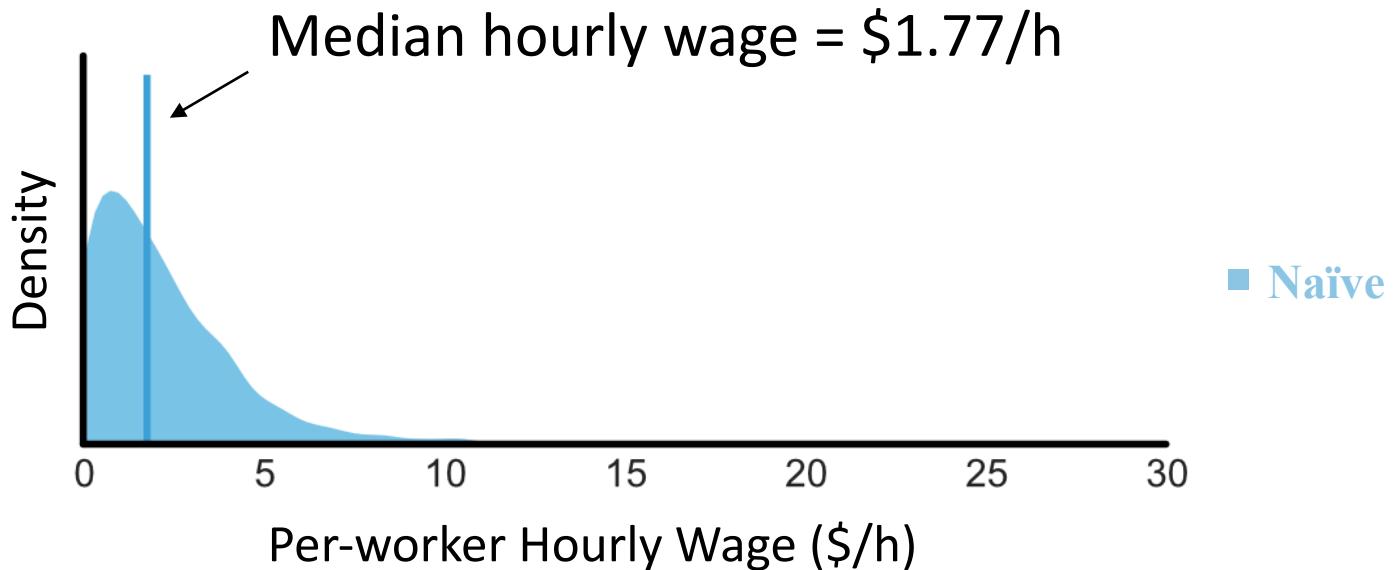
# Worker Hourly Wage: Result (Naïve)



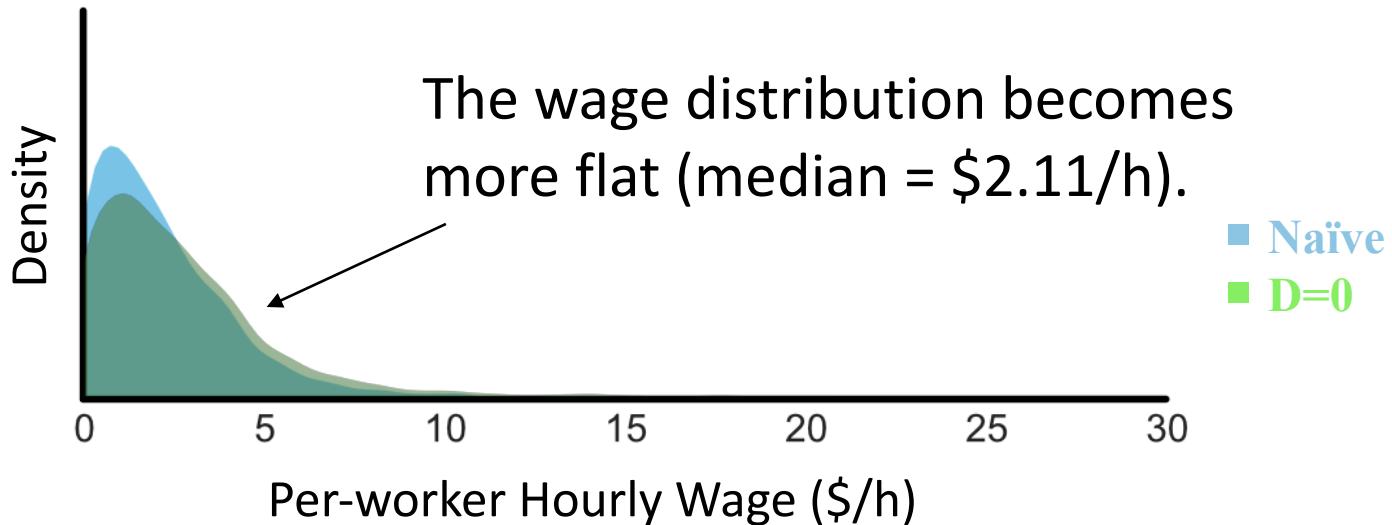
# Worker Hourly Wage: Result (Naïve)



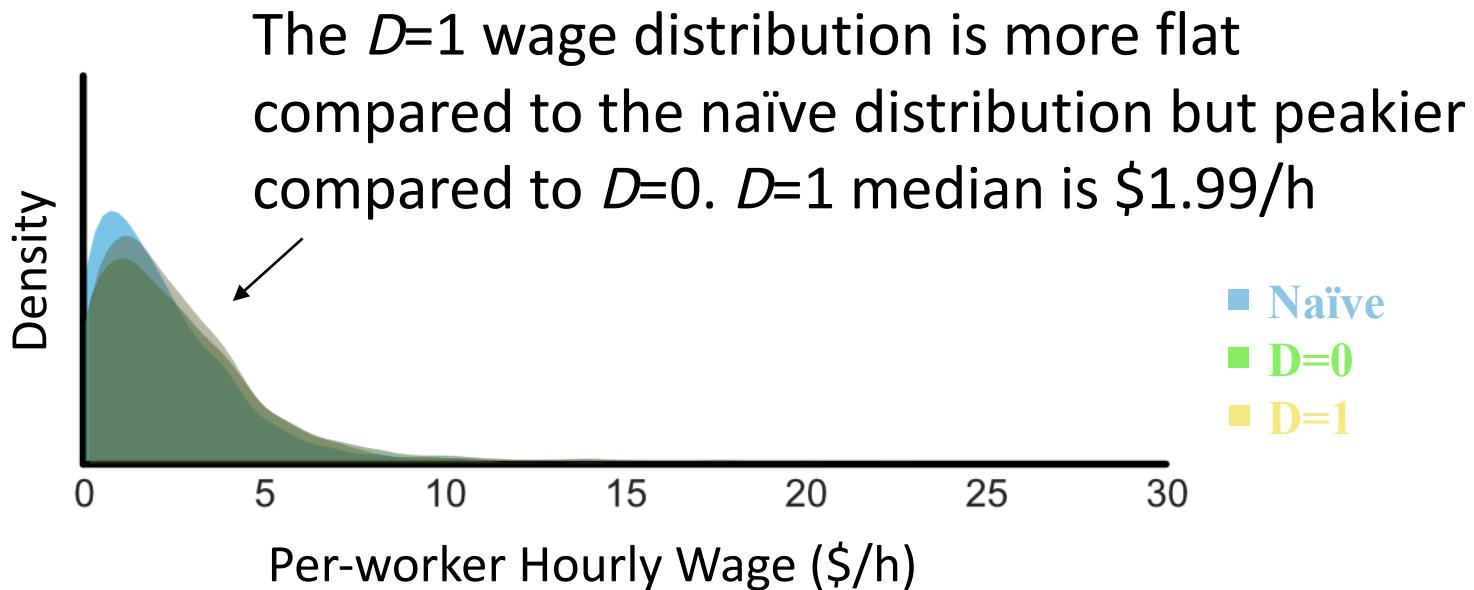
# Worker Hourly Wage: Result (Naïve)



# Worker Hourly Wage: Result (Clustered)

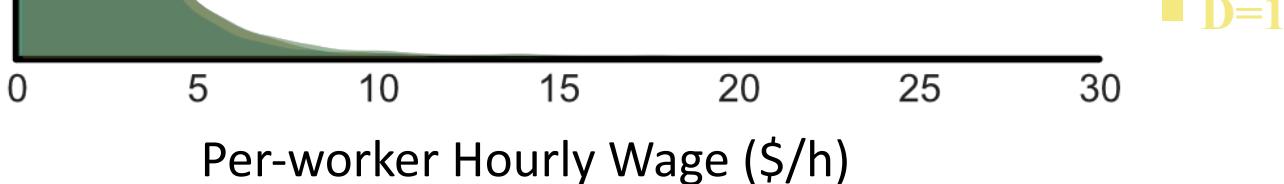


# Worker Hourly Wage: Result (Clustered)



# Worker Hourly Wage: Result

Median worker hourly wage is around **\$2/h**. Naïve estimation method under-estimates the hourly wage by approximately 12% (compared to  $D=1$ ).



# Takeaway 1

The majority of workers on Amazon Mechanical Turk work with **hourly wage below \$2/h**

# Research Questions



**How much are workers earning on Amazon Mechanical Turk?**

What contributes to the low wage?

# Research Questions



How much are workers earning on Amazon Mechanical Turk?

**What contributes to the low wage?**

# What contributes to the low wage?



Unpaid work



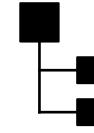
Low reward



Requesters



Qualifications



Task types

# What contributes to the low wage?



**Unpaid work**



Low reward



**Requesters**



Qualifications



Task types

# What contributes to the low wage?



**Unpaid work**



Low reward



Requesters



Qualifications



Task types

## Being A Turker

**David Martin, Benjamin V. Hanrahan, Jacki O'Neill**  
Xerox Research Centre Europe  
6 chemin de Maupertuis, Grenoble France  
{david.martin, ben.hanrahan, jacki.oneill}@xrc.eurox.com

**Neha Gupta**  
University of Nottingham  
University Park NG7 2TD Nottingham  
neha.gupta@xrc.eurox.com

### ABSTRACT

We conducted an ethnmethodological analysis of publicly available content on Turker Nation, a general forum for Amazon Mechanical Turk (AMT) users. Using forum data we provide novel depth and detail on how the Turker Nation members operate as economic actors, working out which Requesters and jobs are worthwhile to them. We show some of the key ways Turker Nation functions as a community and also look further into Turker-Requester relationships from the Turker perspective – considering practical, emotional and moral aspects. Finally, following Star and Strauss [25] we analyse Turking as a form of invisible work. We do this to illustrate practical and ethical issues relating to working with Turkers and AMT, and to promote design directions to support Turkers and their relationships with Requesters.

**Author Keywords**  
Ethnomethodology; content analysis; crowdsourcing; microtasking; Amazon Mechanical Turk; Turker Nation.

**ACM Classification Keywords**  
H.5.3 Group and Organizational Interfaces – Computer, Supported Cooperative Work

**General Terms**  
Human Factors

**INTRODUCTION**  
The concept of crowdsourcing was originally defined by Jeff Howe of Wired Magazine as “*the act of a company or individual taking a problem inside their四合院 and outsourcing it to an undefined (and generally large) network of people in the form of an open call*” [8]. This ‘undefined network of people’ is the key topic of this article. We present the findings of an ethnmethodological analysis of posts and threads on a crowdsourcing forum called Turker Nation<sup>1</sup>. We have sought to understand members of the crowd – their reasoning, practices, concerns, and relationships with requesters and each other – as they are shown in their posts on the forum. We seek to present them as faithfully as possible, in their own words, in

order to provide more definition to this network of people. We believe that this will be beneficial for researchers and businesses working within the crowdsourcing space.

Crowdsourcing encompasses multiple types of activity: invention, project work, creative activities, and microtasking. This latter is our focus here. The most well-known microtask platform is Amazon Mechanical Turk (AMT)<sup>2</sup>, and the Turker Nation forum that we studied is dedicated to users of this platform. The basic philosophy of microtasking and AMT is to delegate tasks that are difficult for computers but easy for the workforce. This has been termed ‘artificial artificial intelligence’ and includes image tagging, duplicate recognition, translation, transcription, object classification, and content generation are common. ‘Requesters’ (the AMT term for people who have work to be completed) post multiple, similar jobs as Human Intelligence Tasks (HITs), which can then be taken up by registered ‘Turkers’. Turkers (termed ‘Providers’ by AMT) are the users completing the HITs, which typically take seconds or minutes paid at a few cents at a time.

For Amazon, the innovative idea was to have an efficient and cost effective way to curate and manage the quality of content on their vast databases (weeding out duplicates, vulgar content, etc.). While Amazon is still a big Requester, AMT has been deployed as a platform and connects a wide variety of Requesters with up to 500,000 Providers. However, Fort et al. [6] have performed an analysis on the available data and suggest that real number of active Turkers is between 15,059 and 42,912; and that 80% of the tasks are carried out by the 20% most active (3,011–8,582) Turkers. While these numbers are useful, the research community still has little deep qualitative knowledge about this workforce. Questions remain unanswered such as: how and what do they look for in jobs; what are their concerns; and how do they relate to requesters?

**LITERATURE REVIEW**  
To date much of the research on AMT takes the employers’ perspective, e.g. [14, 15, 17, 18], and this has in turn been highlighted [6, 16]. Silberman et al. [23] note that this mainstream research looks at how: “[t]o motivate better, cheaper and faster worker performance [...] to get good

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, require prior specific permission and/or a fee. CSCW '13, February 23–27, 2013, San Antonio, Texas, USA.  
Copyright 2013 ACM 978-1-4503-1331-5/13/02 \$15.00

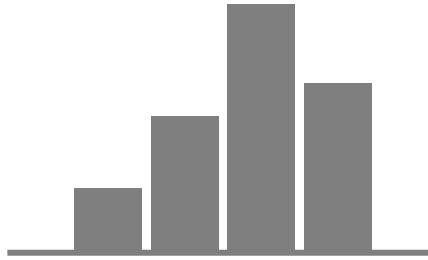
<sup>1</sup> <http://turkernation.com/forum.php>

<sup>2</sup> <http://www.mturk.com>

[...] aspects of turking [(working on Amazon Mechanical Turk)] like simply searching for jobs can take a considerable amount time.

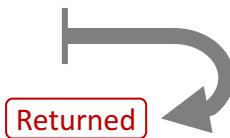
The time spent learning and searching are clear examples of invisible [(unpaid)] work that Turkers must engage in [...].

Martin *et al.*, (2014) *Being a Turker*, CSCW 2014



The issue of unpaid work has been identified in prior work,  
**but its effects are not quantified**

We quantify three types of unpaid work



Time spent on  
**returned tasks**

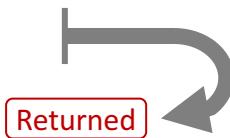


Time spent on  
**rejected tasks**



Time  
**between tasks**

We quantify three types of unpaid work



Time spent on  
**returned tasks**



Time spent on  
rejected tasks

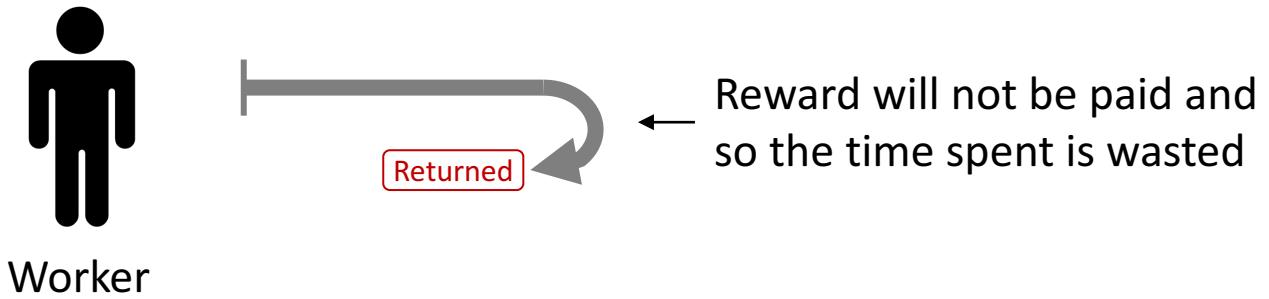


Time  
between tasks

# Task Submit and Return

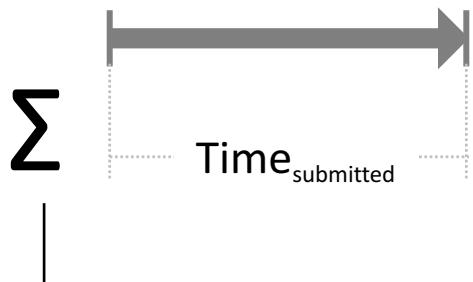


# Task Submit and Return



# Time Spent on Returned Tasks

Submitted

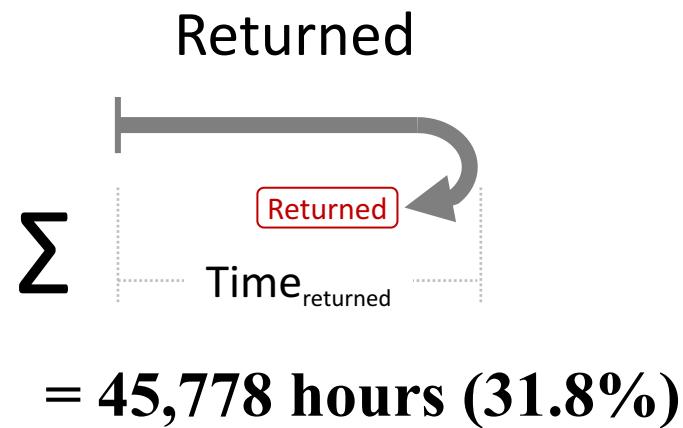
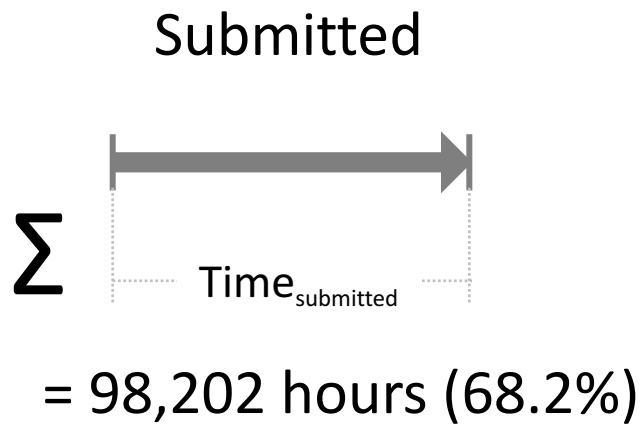


For all tasks from all workers

Returned



# Time Spent on Returned Tasks: Result



## We quantify three types of unpaid work



Time spent on  
returned tasks

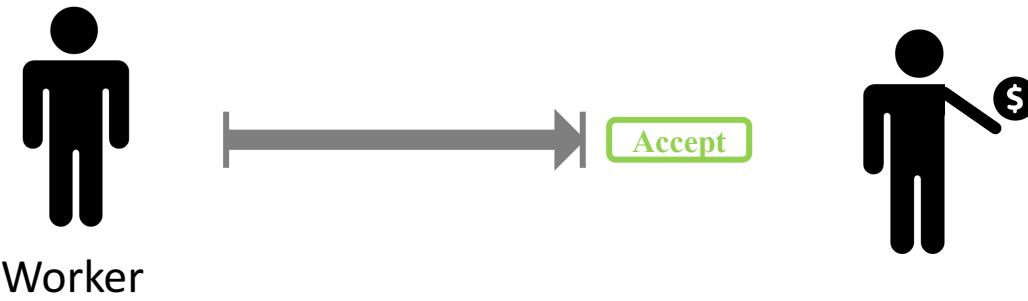


Time spent on  
**rejected tasks**

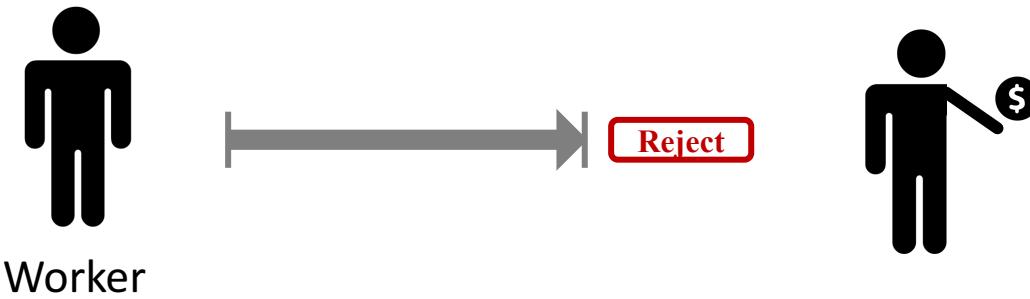


Time  
between tasks

# Task Accept and Reject

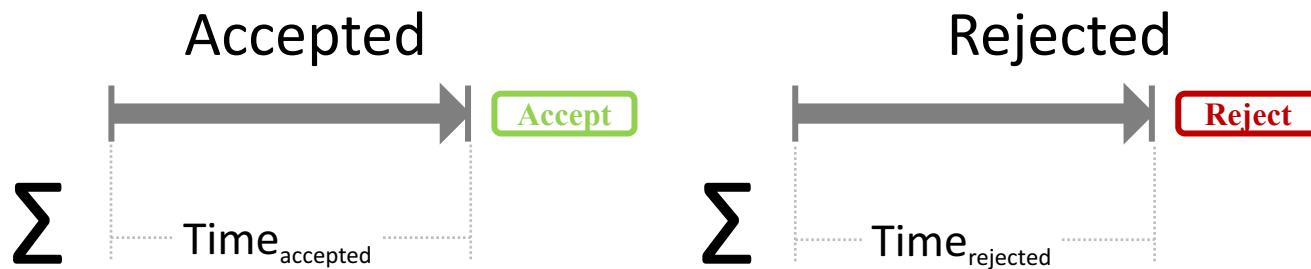


# Task Accept and Reject



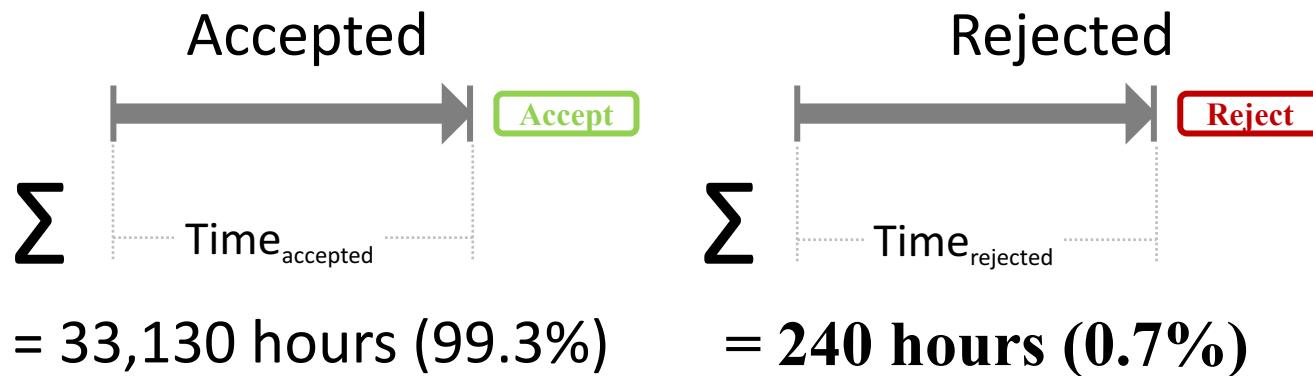
# Time Spent on Rejected Tasks

We had data on task accept *vs.* reject status for 29.6% of the submitted tasks



# Time Spent on Rejected Tasks: Result

We had data on task accept *vs.* reject status for 29.6% of the submitted tasks



We quantify three types of unpaid work



Time spent on  
returned tasks

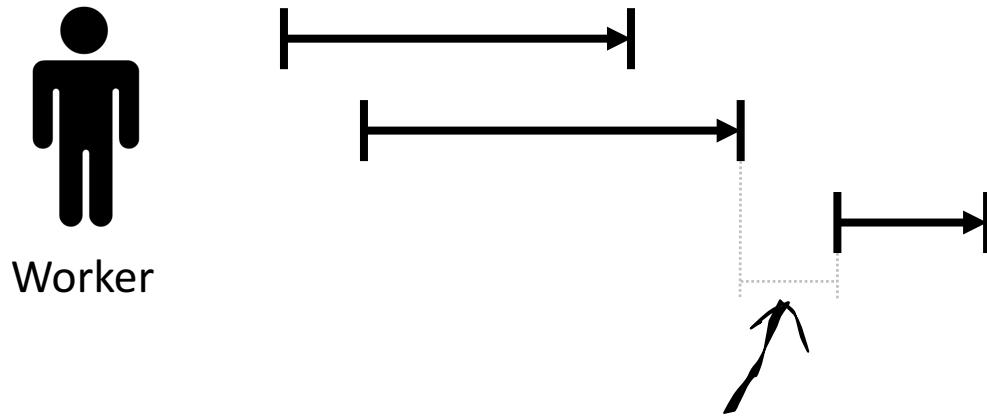


Time spent on  
rejected tasks



Time  
**between tasks**

# Time Between Tasks

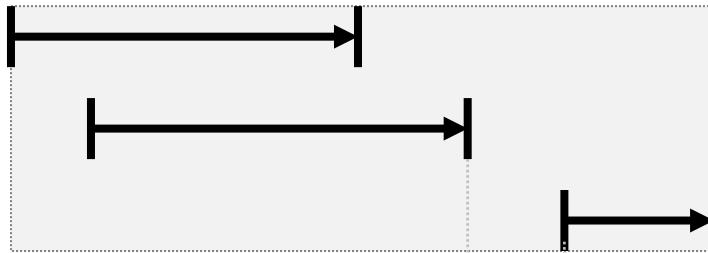


We want to know the effect of  
this small gap between tasks  
(e.g., task search time)

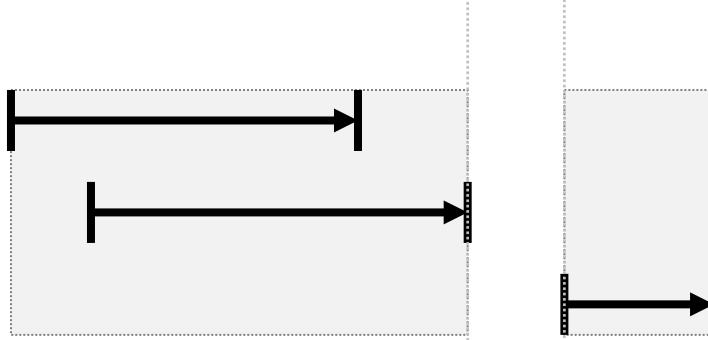
# Time Between Tasks



Worker



Clustering ( $D=1\text{min}$ )



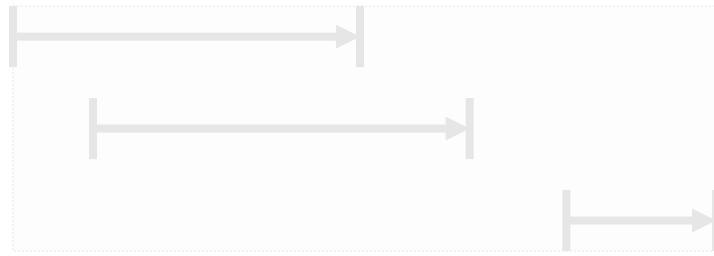
Clustering ( $D=0\text{min}$ )

$$\sum \Delta$$

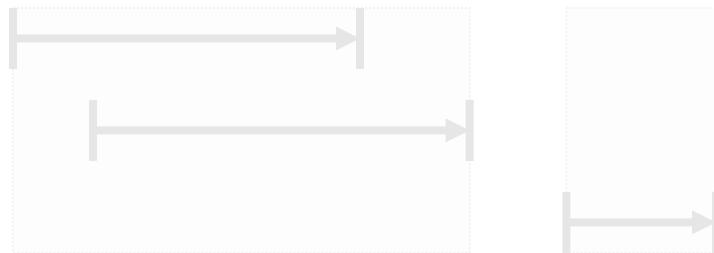
# Time Between Tasks: Result



Worker



Clustering ( $D=1\text{min}$ )

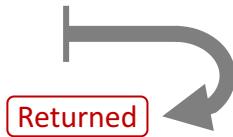


Clustering ( $D=0\text{min}$ )

$$\sum \Delta = 4,603 \text{ hours}$$

# Result

**45,778** hours  
**31.8%** of work



**240** hours  
**0.7%** of work



**4,603** hours  
**4.7%** of work



## Takeaway 2

Returning tasks has the biggest impact on the hourly wage. In fact, **31.8% of work time is wasted due to this unpaid work.**

# What contributes to the low wage?



**Unpaid work**



Low reward



Requesters



Qualifications



Task types

# What contributes to the low wage?



Unpaid work



Low reward



**Requesters**



Qualifications



Task types

< \$2/h

Workers are underpaid.

< \$2/h

Workers are underpaid. Is this because all requesters treat workers unfairly,

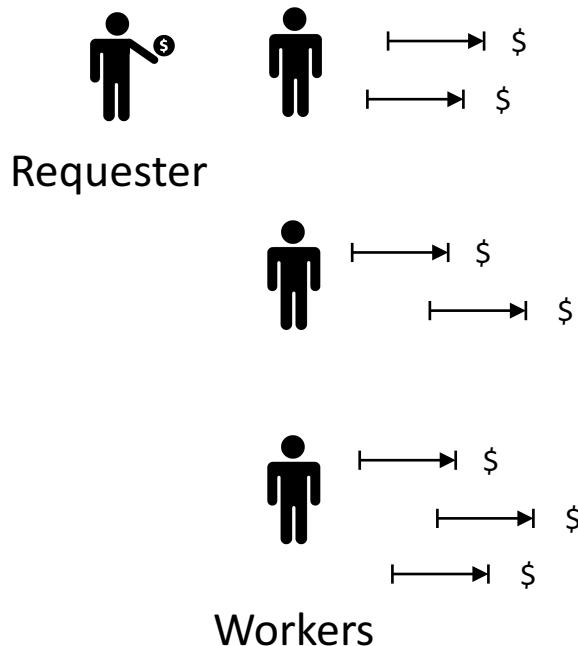
< \$2/h

Workers are underpaid. Is this because all requesters treat workers unfairly, or are there a small number of requesters who post many very low paid tasks?

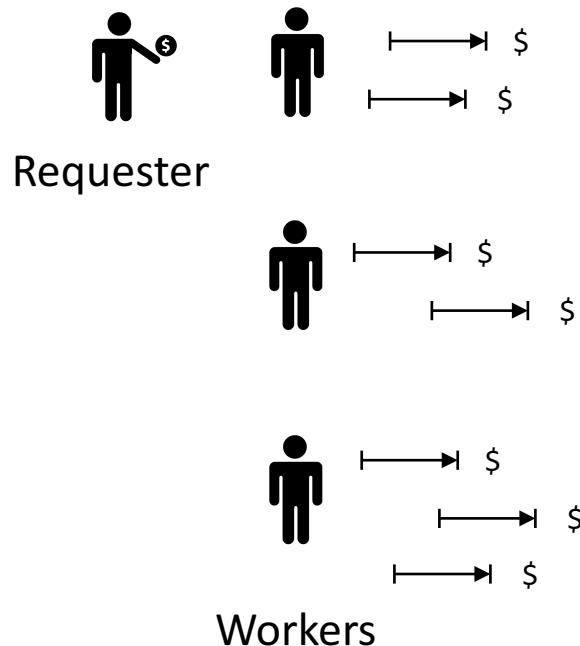


We investigated the distribution of **per-requester hourly payment**

# Per-requester Hourly Payment



# Per-requester Hourly Payment



# Per-requester Hourly Payment



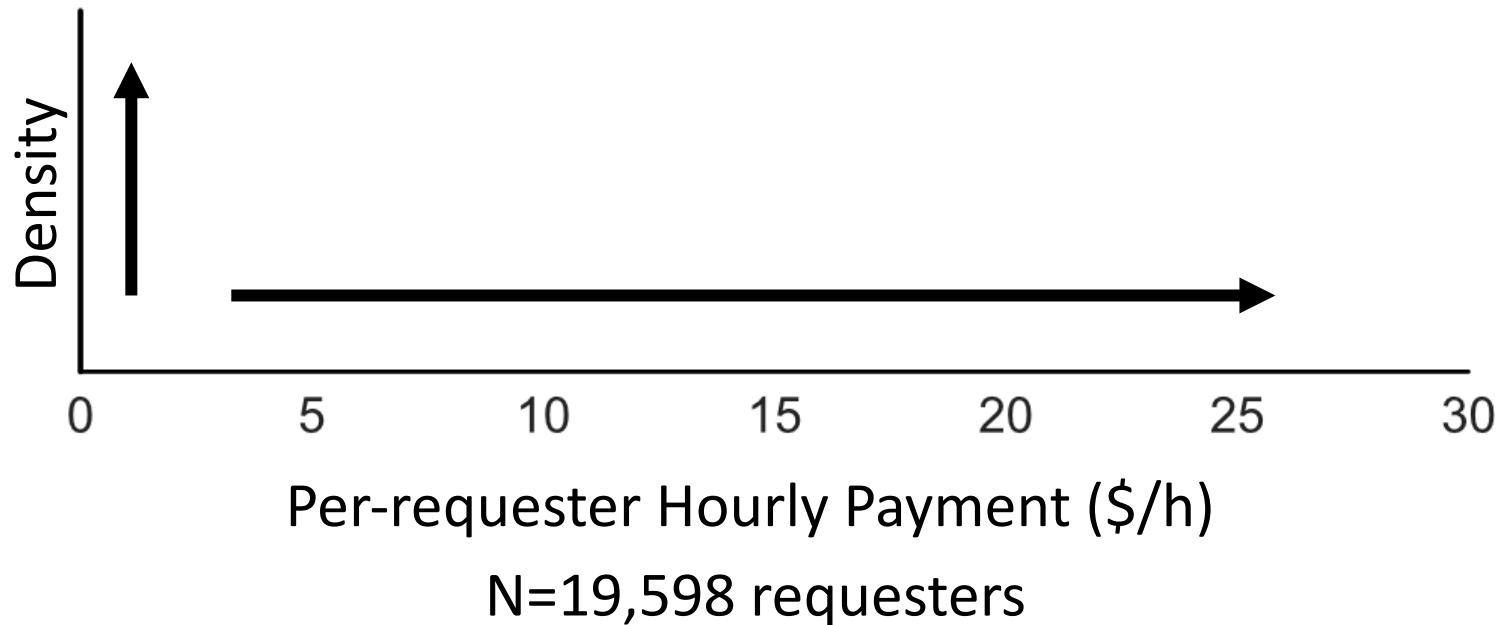
Requester



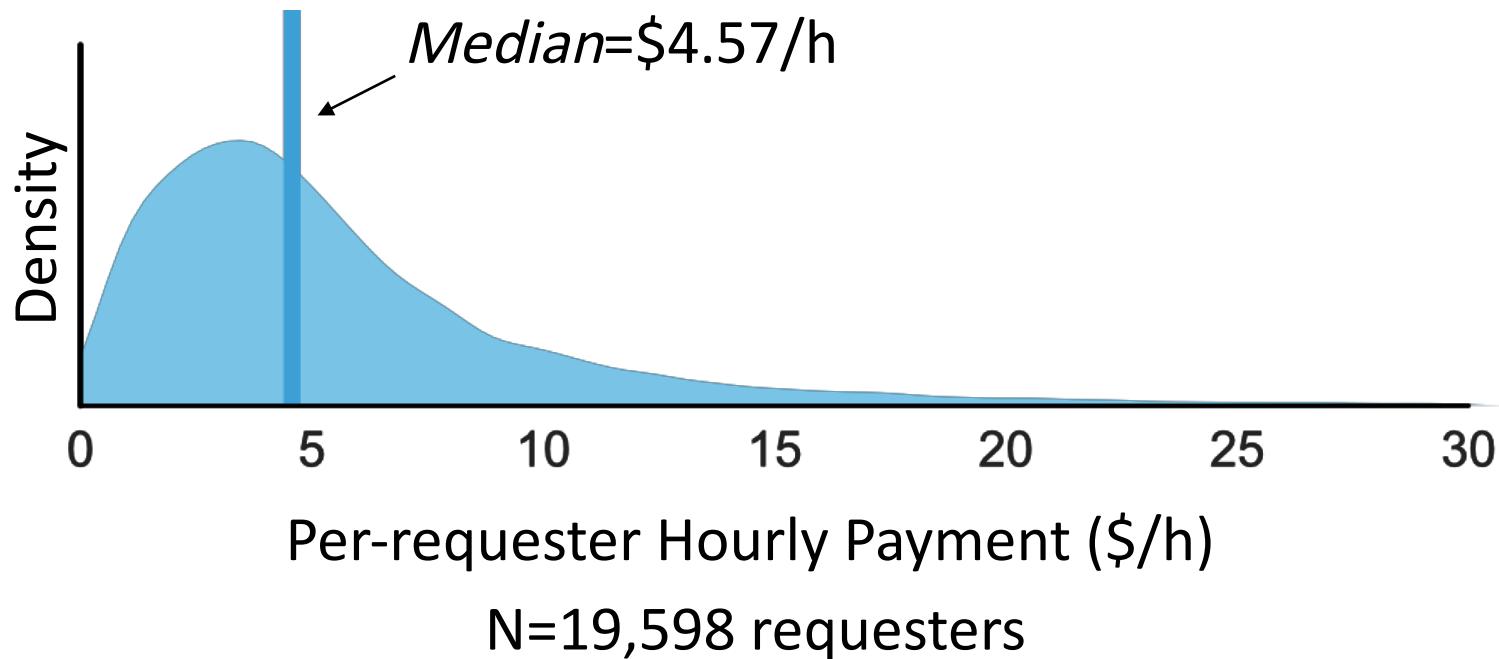
Workers

$$\frac{\sum \text{ Task Payment } (\$)}{\sum \text{ Task Interval}} = \text{ Per-requester Hourly Payment}$$

# Per-requester Hourly Payment: Result



# Per-requester Hourly Payment: Result



## Takeaway 3

About half of the requesters pay below \$5/h, which is below the U.S. federal minimum wage.

# Takeaways

< \$2/h

Crowd workers are underpaid and they often earn below \$2/h



Unpaid work, particularly returning tasks has a large impact on the hourly wage



Majority of the requesters reward workers below \$5/h

# Discussion and Future Work

- Could we create **tools for workers** to make it easier to search for tasks that give them good wage, avoid tasks that are not completable, and find requesters fair wage?
- Could we create **technologies for requesters** to help them pay fairly?

# Discussion and Future Work



Sukrit Venkatagiri  
@thesukrit

Following

Today @VTGSA Research Symposium, I spoke with @VTPamplin Business PhD student about how HIT reward + overhead on @amazonmturk translates to \$/hr, citing @kotarohara\_en's work ([arxiv.org/abs/1712.05796](https://arxiv.org/abs/1712.05796)). She promised she & her colleagues would pay Turkers a fair wage from now on.

9:51 PM - 28 Mar 2018

5 Retweets 14 Likes



5

14



Tweet your reply

How can we nudge more people to pledge and help them keep committing to their promises?

# Limitations

- Sampling bias
- Our analyses did not take bonus payment into account
- No information about demographics

# A Data-Driven Analysis of Workers' Earnings on Amazon Mechanical Turk

Kotaro Hara, Abigail Adams, Kristy Milland, Saiph Savage

Chris Callison-Burch, Jeffrey P. Bigham

The Economic and Social Research Council, Grant ES/N017099/1

We thank the workers who provided the work log data that enabled this research

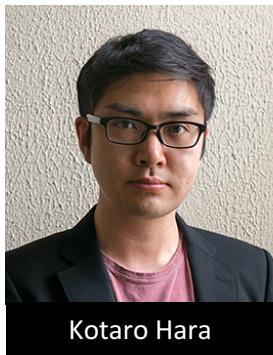
Image credits to Saeful Muslim, Kirby Wu from the Noun Project



# A Data-Driven Analysis of Workers' Earnings on Amazon Mechanical Turk

Kotaro Hara, Abigail Adams, Kristy Milland, Saiph Savage

Chris Callison-Burch, Jeffrey P. Bigham



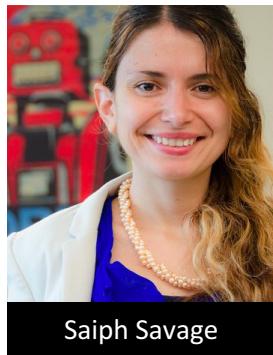
Kotaro Hara



Abigail Adams



Kristy Milland



Saiph Savage



Chris Callison-Burch



Jeffrey P. Bigham

# A Data-Driven Analysis of Workers' Earnings on Amazon Mechanical Turk

Kotaro Hara, Abigail Adams, Kristy Milland, Saiph Savage

Chris Callison-Burch, Jeffrey P. Bigham

# Questions?



# A Data-Driven Analysis of Workers' Earnings on Amazon Mechanical Turk

Kotaro Hara, Abigail Adams, Kristy Milland, Saiph Savage

Chris Callison-Burch, Jeffrey P. Bigham

< \$2/h

Crowd workers are underpaid and they often earn below \$2/h



Unpaid work, particularly returning tasks has a large impact on the hourly wage



Majority of the requesters reward workers below \$5/h

# Takeaways

< \$2/h



Crowd workers are underpaid and they often earn below \$2/h

Unpaid work, particularly returning tasks has a large impact on the hourly wage



Majority of the requesters reward workers below \$5/h



## Takeaway 2

Returning tasks has the biggest impact to the hourly wage. In fact, **31.8% of work time is wasted due to returned tasks.**

