Labour Economics 81 (2023) 102305

Contents lists available at ScienceDirect

Labour Economics

journal homepage: www.elsevier.com/locate/labeco

Asymmetric labor-supply responses to wage changes: Experimental evidence from an online labor market☆

Philipp Doerrenberg[a], Denvil Duncan[b], Max Löffler[c,*]

[a] University of Mannheim, CESifo and IZA
[b] Indiana University, Bloomington, ZEW Mannheim and IZA
[c] Maastricht University, CESifo, ZEW and IZA

A R T I C L E I N F O

JEL classification:
J20
J22
J31
D03

Keywords:
Labor supply
Labor supply elasticities
Asymmetric effects of wages
Field experiment
Amazon Mechanical Turk

A B S T R A C T

We test whether labor supply responds symmetrically to wage increases and decreases using a randomized real effort online experiment. The results show that wage increases have smaller effects on labor supply than wage decreases of equal magnitude, especially on the extensive margin where the response to a wage decrease is twice that to a wage increase. This finding suggests that labor-supply responses to wage changes are asymmetric. We discuss the potential mechanisms behind our results including standard models of labor supply, reference dependence in consumption and reciprocity.

1. Introduction

The empirical literature on labor supply often makes the implicit assumption that labor supply responses to wage increases are the same as those for equivalent wage decreases (Bargain et al., 2014; Blundell and MaCurdy, 1999; Meghir and Phillips, 2010). In other words, wage increases and wage decreases of equal magnitude are assumed to have the same effect on labor supply decisions (though with opposite signs). This assumption implies that labor-supply elasticities with respect to wages do not depend on the sign of the wage variation. However, various theoretical approaches predict *asymmetric responses* to wage increases and decreases. Although this result has important implications for the empirical literature, there is little direct empirical evidence regarding the symmetry of the effect of wages on labor supply.

In this paper, we investigate the symmetry of labor-supply responses to wage changes. Our precise research question is: do wage increases and decreases of equal magnitude have symmetric effects on the supply of labor by employees to their current employer?[1] Answering this research question requires a set-up that introduces (quasi-)randomly assigned wage increases and decreases at the same time for comparable individuals. Finding such types of experiments in a "natural" setting is difficult, if not impossible, and thus may partly explain the sparse literature on the symmetry of labor-supply responses to nominal wages.

We address this empirical challenge using a real effort experiment implemented in an online labor market, where we randomly assign wage increases and decreases of equal magnitude to workers. Specifically, we set up a real effort task and invite workers to work on this task in an actual online labor market, namely Amazon's Mechanical Turk (henceforth mTurk). Our task requires workers to transcribe pictures with scanned German-language text. We announce a piece rate of $0.15 per transcribed picture and workers complete a batch of six transcriptions for the announced wage. Workers are randomly assigned to one of three groups: (i) the wage increases by 20%, (ii) the wage decreases by 20%, or (iii) the wage remains constant (control group). We present these updated piece rates to workers after having transcribing the first batch of images. Workers can then select to either stop working on our task or

☆ David Vjazner provided excellent research assistance. We are grateful to Ernst Fehr, Daniel Hamermesh, Johannes Hermle, Rafael Lalive, Harald Lang, Thomas Meissner, Florian Morath, Andreas Peichl, Arno Riedl, Justin Ross, Jan Schmitz, Sebastian Siegloch, Peter Werner, Matthias Wibral, Christopher Zeppenfeld, and various seminar/conference participants for valuable comments and suggestions. The Online Appendix to this paper can be found on www.denvilrduncan.com. Replication data is available via doi:10.7910/DVN/YDK1ZI.

* Corresponding author.
E-mail addresses: doerrenberg@uni-mannheim.de (P. Doerrenberg), duncande@indiana.edu (D. Duncan), m.loeffler@maastrichtuniversity.nl (M. Löffler).

[1] Technically we focus on employees' supply of labor to their current employer, i.e., on a firm-specific labor supply rather than a more general characterization of labor supplied to a labor market.

https://doi.org/10.1016/j.labeco.2022.102305
Received 29 August 2022; Received in revised form 22 November 2022; Accepted 25 November 2022
Available online 7 December 2022
0927-5371/

keep transcribing pictures. We identify the symmetry of labor-supply responses by comparing working behavior between the three randomly assigned groups.

Our results can be summarized as follows. First, we find that wage *increases* have a positive effect on labor supply whereas wage *decreases* reduce labor supply in our task. This finding provides clear support for a positive relationship between labor supply and wages. Second, labor-supply responses to wage increases and decreases are asymmetric; workers react more strongly to wage decreases than wage increases of equal magnitude (in absolute terms). We find clear statistically evidence of asymmetry using non-parametric tests for the equality of the distributions comparing (i) absolute differences between the control and wage-*increase* group to (ii) absolute differences between the control and wage-*decrease* group, respectively. Using conventional Wald-tests, we further find statistically significant evidence on the extensive margin, which we define as the share of workers who quit our task immediately after seeing the treatment notification.[2] Workers in the wage-decrease group have a 18%p *higher* probability to quit the labor task (compared to the control group), while workers in the wage-increase group are 8.5%p *less* likely to quit (relative to the control group). Third, our results further show that neither wage increases nor wage decreases have an effect on the average time spent per transcribed picture or the quality of transcriptions, which is around 97 percent in all groups.

Our results are consistent with a simple standard labor supply model where utility is concave in consumption. We also discuss various alternative mechanisms behind our results, none of which is able to explain all of our empirical findings. In particular, we argue that models of reservation wages, dynamic models with learning about effort costs, and models with reference-dependence in consumption or reciprocity concerns are only partly in line with our results while some key predictions are at odds with the data.

Contribution to the Literature. We make the following contributions to the literature on labor-supply effects of wage changes. First, we contribute to the general labor-supply literature (see Keane, 2011, for a survey). Many of the existing studies use panel-data and pool upward and downward variation in wages to estimate the wage elasticity of labor supply. Because the elasticity estimated by these studies represents an average of responses induced by wage increase and decreases, our results suggest that existing estimates likely overstate the effect of wage increases while understating the effect of wage decreases. Our results thus raise questions about the comparability of labor-supply elasticities across studies that differ in the sign of the wage changes used for identification. It cannot be concluded from the estimated elasticities that workers are more responsive in the one setting relative to another without knowing whether the sign of the wage changes is the same. This is especially important for meta-analysis studies on labor supply (e.g., Evers et al., 2008).[3]

Second, our paper is related to three studies investigating potentially asymmetric responses. Most importantly, Kube et al. (2013) conduct a field experiment with students working in a library for a given period of time. They find that wage cuts decreases work effort whereas wage increases have no effect. While these results are broadly consistent with our findings, our setup differs from theirs in the design of the experiment as well as the labor market institution, which has important implications for the interpretation and application of our findings. We pay workers for each transcribed picture instead of for a predetermined number of hours; this implies that we study a situation where workers have less scope to shirk as a means of punishing their employer. In addition we allow workers to quit the labor task whenever they choose to do so. While they are in a gift-exchange setting, our task offers little room for either positive or negative reciprocity. Another related paper is Falk et al. (2006) who find in a laboratory experiment that reservation wages respond asymmetrically to the introduction and removal of minimum wages. More broadly, we explicitly test for asymmetry in behavioral responses, and thereby relate to a recent paper by Benzarti et al. (2020) who use observational data to document that increases in value added taxes have larger effects on prices than VAT reductions.

Third, we add to the experimental literature on the effect of wages on effort and labor supply. These studies provide credible randomized evidence in the absence of (discrete) work-time constraints, something which is difficult to obtain using observational data. Papers based on laboratory experiments provide robust evidence that labor effort and wages are characterized by a positive relationship (see the survey by Charness and Kuhn, 2011), which is consistent with our findings. However, laboratory experiments are subject to the usual concern that they cannot easily be generalized to real-world situations. Field experiments with higher external validity find mixed effects regarding the relationship between wages and effort. While some field experiments find a positive effect of wages on effort/labor supply (DellaVigna and Pope, 2018; Fehr and Goette, 2007), other studies find either no relationship (Hennig-Schmidt et al., 2010), short-run temporary effects which do not make a difference for final work outcomes (Gneezy and List, 2006), or (positive) effects for only certain types of workers (Cohn et al., 2015). Our results add to the (ongoing) discussion on the wage-effort relation by providing evidence of a positive relationship between wages and labor effort in online labor markets.

The paper is organized as follows. Section 2 describes the real labor task and its implementation in Amazon's Mechanical Turk. We present the empirical analysis and the results in Section 3. We discuss the potential economic mechanisms behind our findings, as well as their implications and generalizability, in Section 4. Section 5 concludes.

2. The experiment

This section outlines our experiment. We begin by describing the labor task and the treatment design in Section 2.1. Section 2.2 provides more details on the implementation in mTurk and Section 2.3 describes the estimation sample used in our analysis.

2.1. Design

Labor Task. We selected an online labor task that requires subjects to transcribe German text shown in a series of images.[4] Each image has approximately five lines and 43 words (344 characters). Fig. 1 shows an example. Subjects are randomly shown one of 128 images at a time and are instructed to hit "save picture" when they are done transcribing. A new image is shown after the subject hits "save picture".

Treatment and Groups. We use a between-subjects design in order to identify the effect of wage changes on labor supply. Subjects are randomly assigned to one of three groups: one control group and two treatment groups. Subjects in all three groups work on the labor task described above and are paid a piece rate for each image that they transcribe. The piece rate is set at $0.15 for each of the first six transcribed

[2] We use the term "extensive margin" for the decision to quit our task immediately after the treatment notification. It is possible that subjects who decline our bonus task work on another task on mTurk. In this regard, our usage of the term "extensive margin" should not be understood as describing the decision to work at all or not.

[3] It is sometimes argued that nominal wage cuts are rare and therefore not relevant. While we acknowledge that nominal wage cuts occur less often than increases (see the literature on nominal wage rigidities, e.g., Kaur, 2019), it has been shown that wage cuts do happen; for example during recessions and bankruptcies, and for the self-employed and salary earners (Kahn, 1997). Our study is also relevant for decreases in real wages, which occur more frequently than nominal wage cuts. Our results suggest that inflation induced real-wage decreases may have larger labor supply effects than previously thought.

[4] Horton et al. (2011) use a similar task and motivate it with the following advantages: transcribing text (i) is tedious, (ii) requires effort and attention, and (iii) has a clearly defined quality measure.

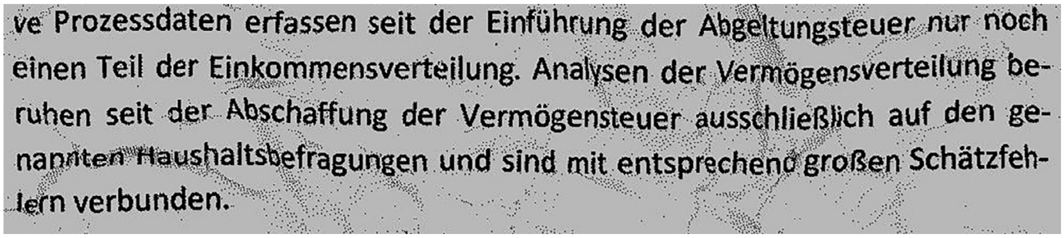

ve Prozessdaten erfassen seit der Einführung der Abgeltungsteuer nur noch einen Teil der Einkommensverteilung. Analysen der Vermögensverteilung beruhen seit der Abschaffung der Vermögensteuer ausschließlich auf den genannten Haushaltsbefragungen und sind mit entsprechend großen Schätzfehlern verbunden.

Fig. 1. Image of Text to be Transcribed. *Notes:* This figure depicts a screenshot of an image of German text that was to be transcribed by the subjects. Subjects were randomly shown one of 128 images. All images are comparable to the image depicted in the figure.

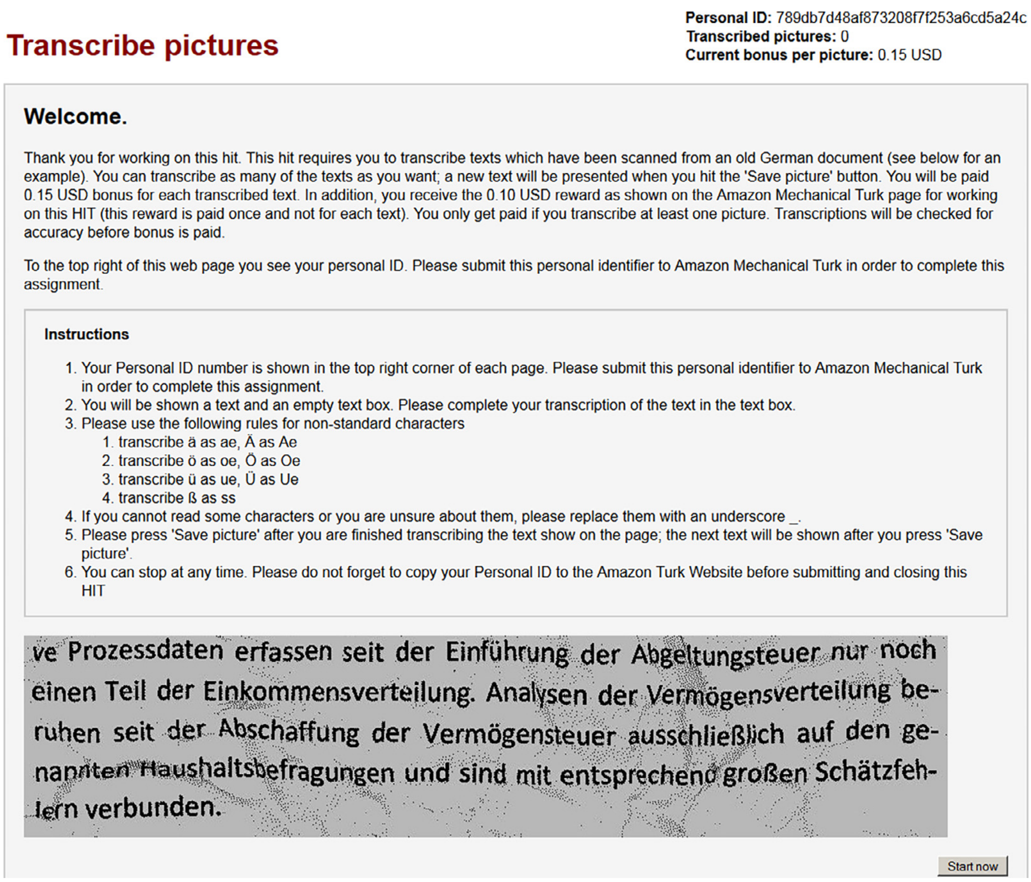


Fig. 2. Treatment Variation. *Notes:* This figure depicts a screenshot of the treatment notification in the "wage decrease" group. The treatment notifications for the "control" and "wage increase" group, respectively, were identical except for the information regarding the piece-rate wage for the subsequent images. The treatment notification popped up after a subject transcribed six images.

images in all three groups.[5] Subjects receive a notification thanking them for transcribing the images after the first six images have been transcribed. They are then told that they can transcribe additional images and that the piece rate for the additional images is either $0.18, $0.15 or $0.12, for the wage-increase, control, and wage-decrease groups, respectively (see Fig. 2 for an exemplary treatment notification). We did not provide workers with a reason for their wage changes in order to keep a neutral framing (Kube et al., 2013).

Wage Expectations. The experiment is designed to establish an exogenous and salient expectation regarding the per-unit wage in the mTurk task. Workers who start working face the announced wage of $0.15 for the first six transcribed pictures (this is also the wage that is announced in the job advertisement). We argue that this design generates the expectation that the per-unit wage will remain constant at $0.15 throughout the entire task. Our experimental design therefore allows us to study how unexpected wage increases and wage decreases affect labor supply. If we had initially told subjects that the wage would either increase or decrease, they could have adjusted their expectations and the labor supply response to varying wages would not have been comparable to real-situations where workers experience unanticipated wage changes. This is consistent with Kube et al. (2013) who argue that deviations from an exogenous expectation capture the key aspects of wage changes (for example, disappointment and the break of trust relation in the case of wage cuts).

We explicitly address potential concerns of deception since the initial job description does not notify subjects of the possibility that the wage may increase or decrease after a certain number of transcribed pictures. To this end, we include the following pieces of information in the treatment notification (see Fig. 2). First, we thank workers for

[5] The piece rate is called *bonus* in the experiment. This is the usual wording if one is to implement per-piece payment within the same task in the mTurk labor market.

completing the transcription task and remind them that, as promised, they will be paid $0.15 for each of the pictures they had already transcribed. Next, we inform them that they have the option to transcribe *additional* images and that the piece rate for these additional transcriptions is different from that for the first batch of transcriptions. Finally, we make it clear that they can stop and exit the task at this point if they wish and instruct them on what to do next to ensure payment.[6] We argue that these design features give the impression to workers that the task consists of two parts and ensures that we did not deceive them regarding the wage in both parts.

2.2. Implementation

Labor Market and Recruitment. The experiment is implemented with workers on Amazon's Mechanical Turk, an online labor market where job offers are posted and workers choose jobs for payment. mTurk is particularly well known for small text transcription and image recognition tasks, which are easily carried out by humans but difficult for computers. We are thus able to identify the effect of wage changes in a "naturally" occurring labor market. While online labor markets are still clearly different to offline settings, the behavior of online workers has been shown to be comparable to those of subjects in laboratory studies (Horton et al., 2011). In contrast to the lab, we avoid experimenter effects because subjects do not know that they participate in an experiment (Buhrmester et al., 2011; Horton et al., 2011; Mason and Suri, 2011; Paolacci et al., 2010).

To implement our experiment, we first create a "human intelligence task" (HIT) that is advertised on mTurk, showing a brief description of the task and details on the compensation. We restrict the subject pool to workers from the US in order to ensure that the labor costs are non-zero.[7] Subjects who start working on the HIT are randomly assigned to one of three groups and presented the instructions in Fig. 3. Afterwards, they are shown images of scanned German text that they must transcribe for payment (like the one in Fig. 1). We include information on the number of pictures transcribed so far and the current piece rate in every step of the task. We show treatment notifications after six images have been transcribed. Subjects in the wage-decrease group are shown the treatment information illustrated in Fig. 2. A similar text is shown to subjects in the wage-increase group and the control group; the only difference is the piece rate for the additional images. The experiment ends for each subject when she decides to stop or when she transcribed 50 pictures. Online Appendix B provides more details on the instructions shown to workers during the task.

The experiment is programmed on mTurk to expire after 750 workers accept the HIT or 10 days have passed, which ever comes first. Our initial run of the experiment expired after 10 days with only 484 workers. Therefore, we initiated a second run, which expired after hitting the 750 worker threshold six days later. In total, 1,168 workers participated in the two runs. Note that the HIT is designed such that workers cannot work on the task more than once. We also excluded workers who participated in the first run from participating in the second run. Moreover, it is highly unlikely that individuals have multiple accounts to avoid these constraints due to legal requirements by Amazon.[8]

Payment. Subjects receive a participation reward of $0.10, which is paid as long as a subject accepts the HIT and completes at least one transcription. Additionally, subjects are paid a piece rate of $0.15 for each of the first six transcribed pictures, and depending on the treatment group, $0.12, $0.15 or $0.18 for each transcribed image above the first six transcriptions.

We chose this payment structure based on a small test of the real effort task that we implemented with English-speaking students in a university class before we started the field experiment. The results of this test suggested that approximately 15 pictures can be transcribed per hour, resulting in an hourly wage of about $2.35 ($= 0.1 + 6 \times 0.15 + 9 \times 0.15$) in the control group, $2.62 ($= 0.1 + 6 \times 0.15 + 9 \times 0.18$) in the increase group, and $2.08 ($= 0.1 + 6 \times 0.15 + 9 \times 0.12$) in the decrease group. In light of a median reservation wage of between $1.12 and $1.38 per hour for mTurkers, according to Horton and Chilton (2010) and Horton et al. (2011), this payment structure seemed adequate ex-ante. Although wages on mTurk may have changed since these studies have been published, we ran the experiment in summer of 2015 when wages were largely comparable.

2.3. Sample definition

Our HIT was accepted by 1,168 mTurk workers in total. We restrict the sample to those workers who completed at least one picture, and therefore received the participation fee; this leaves us with 1,158 workers (367 workers in the wage increase group, 398 in the wage decrease group, and 393 in the control group). Table 1 presents summary statistics for our sample of workers with regard to our main variables. We observe that, on average, workers transcribed 12.95 pictures over an average time span of 38.45 minutes.[9] The transcription quality was very high with an average accuracy of 97% and a standard deviation of just 0.03. This suggests that workers take the task seriously and provided high-quality transcriptions. We deliberately did not survey any demographic characteristics to avoid giving the impression that subjects are participating in an experiment.

Around two thirds of all participants (726 workers) completed at least six pictures and therefore saw the treatment notification; the remaining subjects exited the experiment before receiving treatment. Because workers did not know they were in an experiment or that the wage rate would change, the experimental groups should be balanced with respect to the characteristics that predict the pre-treatment probability of exiting the experiment. We confirm this empirically in Section 3.1 and also show that workers in the treatment groups and the control group are similar both in terms of transcription speed and accuracy.

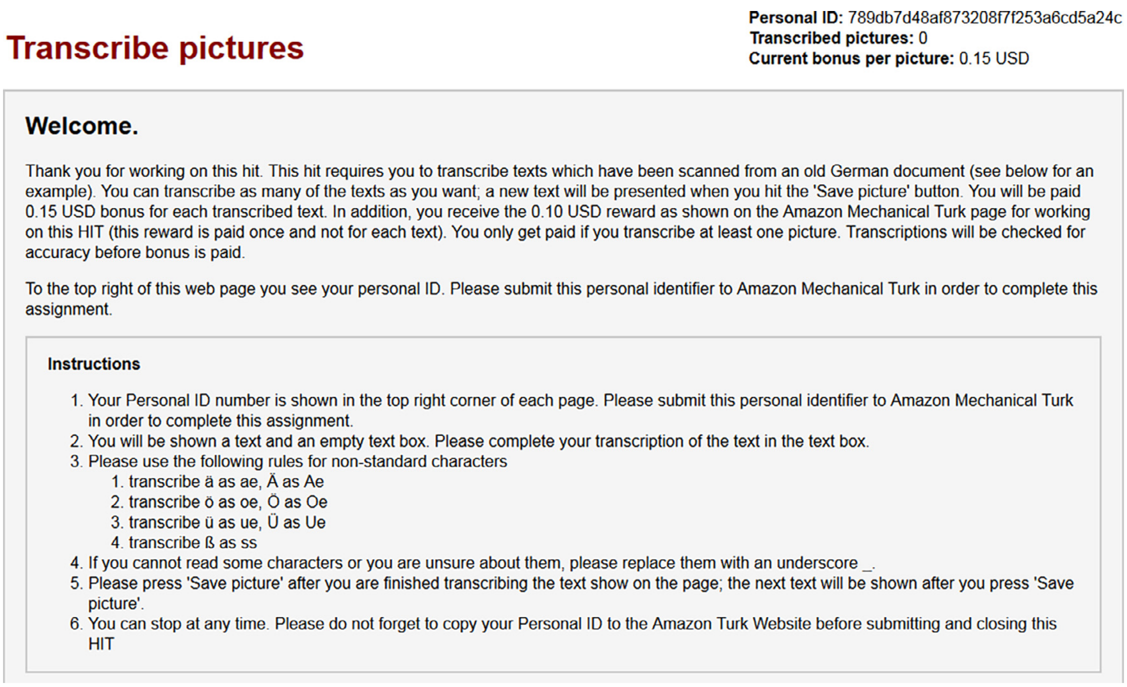
A common feature of mTurk is that workers discuss HITs on forums. This can raise issues for experimenters as those workers who have completed the experiment will unknowingly share the details of treatments with other workers who have yet to complete the experiment. We followed the forums on mTurk in order to determine if our HIT was being discussed and discovered that our HIT did in fact show up on one of

[6] The notification reads: "Thank you for transcribing these pictures. As written in the introduction, we will grant a bonus of $0.15 for each of these pictures. There are additional pictures that you can transcribe. However, the bonus payment for each additional picture will be $0.12/$0.18 from now on. You will receive $0.15 bonus for each of the six pictures you transcribed so far, though. If you want to stop and exit, just copy your Personal ID to the Amazon Turk Website and submit the HIT." Instead of the wage change, we include the following message for the control group: "There are additional pictures that you can transcribe. Just as before, the bonus for each additional picture will be $0.15."

[7] Technically we restrict the HIT to workers with a US IP address. Of course, our restriction does not preclude the possibility that German speakers participated in the task. However, any Germans who participated in our experiment are randomly distributed across our treatments and therefore should have no effect on our outcomes of interest.

[8] First, when registering for mTurk, Amazon requires workers to confirm in the Participation Agreement that they "may not use multiple Amazon Accounts to register with Mechanical Turk". Second, the Participation Agreement further requires workers to provide "true and accurate" information on a worker's name, email address, phone number and physical address (https://www.mturk.com/mturk/conditionsofuse). Third, workers are required to provide a tax identification number (Social Security Number or Individual Tax Identification Number) after their mTurk lifetime earnings have exceeded a set threshold. Workers who fail to provide this number are not allowed to accept additional HITs on mTurk.

[9] Appendix Fig. A.1 provides the distribution of completed pictures for all workers in the sample.


Personal ID: 789db7d48af873208f7f253a6cd5a24c
Transcribed pictures: 0
Current bonus per picture: 0.15 USD

Transcribe pictures

Welcome.

Thank you for working on this hit. This hit requires you to transcribe texts which have been scanned from an old German document (see below for an example). You can transcribe as many of the texts as you want; a new text will be presented when you hit the 'Save picture' button. You will be paid 0.15 USD bonus for each transcribed text. In addition, you receive the 0.10 USD reward as shown on the Amazon Mechanical Turk page for working on this HIT (this reward is paid once and not for each text). You only get paid if you transcribe at least one picture. Transcriptions will be checked for accuracy before bonus is paid.

To the top right of this web page you see your personal ID. Please submit this personal identifier to Amazon Mechanical Turk in order to complete this assignment.

Instructions

1. Your Personal ID number is shown in the top right corner of each page. Please submit this personal identifier to Amazon Mechanical Turk in order to complete this assignment.
2. You will be shown a text and an empty text box. Please complete your transcription of the text in the text box.
3. Please use the following rules for non-standard characters
   1. transcribe ä as ae, Ä as Ae
   2. transcribe ö as oe, Ö as Oe
   3. transcribe ü as ue, Ü as Ue
   4. transcribe ß as ss
4. If you cannot read some characters or you are unsure about them, please replace them with an underscore _.
5. Please press 'Save picture' after you are finished transcribing the text show on the page; the next text will be shown after you press 'Save picture'.
6. You can stop at any time. Please do not forget to copy your Personal ID to the Amazon Turk Website before submitting and closing this HIT

Fig. 3. Instructions Shown on Our Website. *Notes:* The Figure depicts a screenshot of the external website that we set up for the purpose of the field experiment. Subjects were taken to this website once they decided on Amazon's Mechanical Turk website that they would like to work on the task. The depicted screenshots provides subjects all information relevant for the task.

Table 1
Descriptive Statistics.

	(1) Mean	(2) SD	(3) P10	(4) P25	(5) P50	(6) P75	(7) P90
Pictures Transcribed	12.95	13.37	2.00	4.00	7.00	17.00	34.00
Total Time Worked	38.62	47.05	3.12	7.77	19.40	52.58	101.20
Time Per Picture	2.90	2.18	1.37	1.73	2.30	3.22	4.85
Accuracy Ratio	0.97	0.03	0.96	0.97	0.97	0.98	0.98

Notes: This table presents summary statistics for our estimation sample. The sample consists of all 1,158 workers who started working on the task (i.e., including those who transcribed less than six pictures and did not see the treatment notification). *Pictures Transcribed* refers to the number of images that subjects transcribed. *Total Time Worked* is the time (in minutes) that subjects spent working on the labor task. *Time Per Picture* is the time (in minutes) that subjects spent working per picture. *Accuracy Ratio* denotes the share of characters that is transcribed correctly. *SD* refers to the standard deviation, *Px* indicates the x-th percentile.

the forums.[10] The first mention of our HIT occurred close to the end of the second run of the experiment. Workers discussed the fact that the wage rate changed as well as the magnitude of the changes. Few of them also discussed potential reasons for rate changes, and mostly speculated that the wage variation must be due the quality of their work. Nobody speculated that this task is an experiment; people therefore still did not know they were part of an experiment.

The forum post led to a significant spike in acceptance of our HIT; roughly half of the workers accepted the HIT after the forum discussion began. Because some of these subjects knew of a potential wage variation before accepting the HIT, self-selection might be a problem. We find no evidence that the forum discussion drives our results (see Online Appendix C).

3. Results

In this section we present the empirical results of our experiment. We start by analyzing the number of pictures workers in the three groups transcribed before leaving our labor task (Section 3.1). In a second step, we analyze the time spent working, the transcription rate, and accuracy (Section 3.2). We discuss the robustness of our results in Section 3.3.

3.1. Transcribed pictures

We begin the analysis by calculating the distribution of transcribed pictures for both treatment groups and the control group. Fig. 4 presents the resulting hazard rates, indicating the share of workers who transcribe additional pictures and did not (yet) leave our experiment (vertical axis) against the total number of transcribed pictures (horizontal axis). The dashed vertical line marks the timing of the treatment notification. Around 30–35 percent of workers leave our experiment before we show the treatment notification. Importantly, the figure gives no indication of selection by group, which would cast doubt on our randomization procedure; pre-treatment trends are very similar across groups.

While the control group continues to shrink at the same pace pre and post treatment, we observe an immediate response to both types of wage changes. Workers in the increase group are more likely to continue transcribing pictures relative to the control group. Workers in the decrease group leave our experiment disproportionally faster. The hazard rates suggest that these effects are persistent or even growing over time.

[10] See https://www.reddit.com/r/HITsWorthTurkingFor/comments/3eg39l/us_transcribe_texts_from_an_image_payment_bonus/.

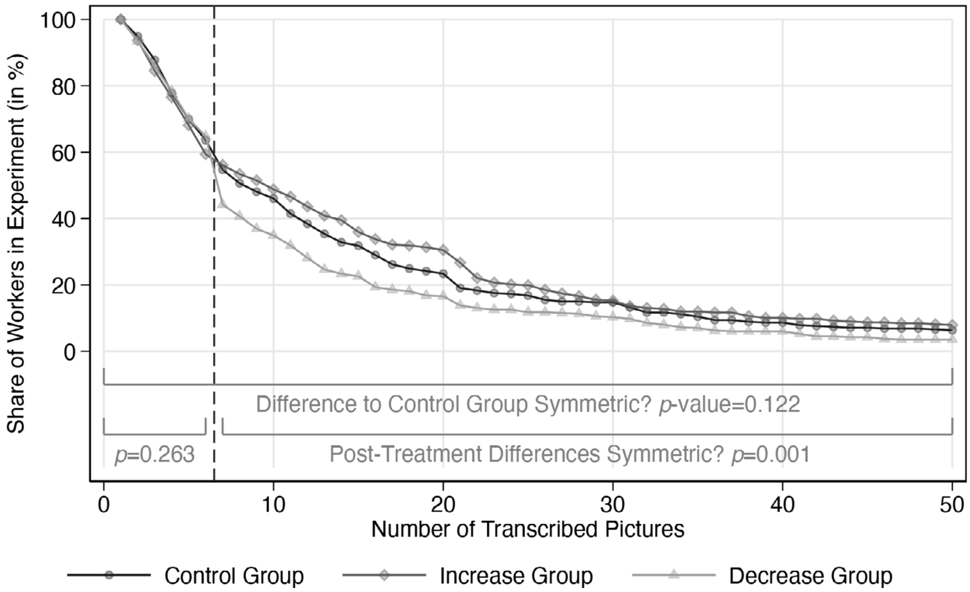


Fig. 4. How Many Pictures Do Workers Transcribe? *Notes:* This figure presents hazard rates depicting how many pictures workers in the control group and the two treatment groups transcribe in our experiment. The initial sample is based on all 1,158 workers who transcribe at least one picture p = 1. We calculate the absolute differences in the distributions between (i) increase and control group, and (ii) decrease and control group, respectively, to assess the asymmetry of the results (see Panel A of Fig. 6). Depicted p-values stem from Kolmogorov-Smirnov tests for equality of these two distributions.

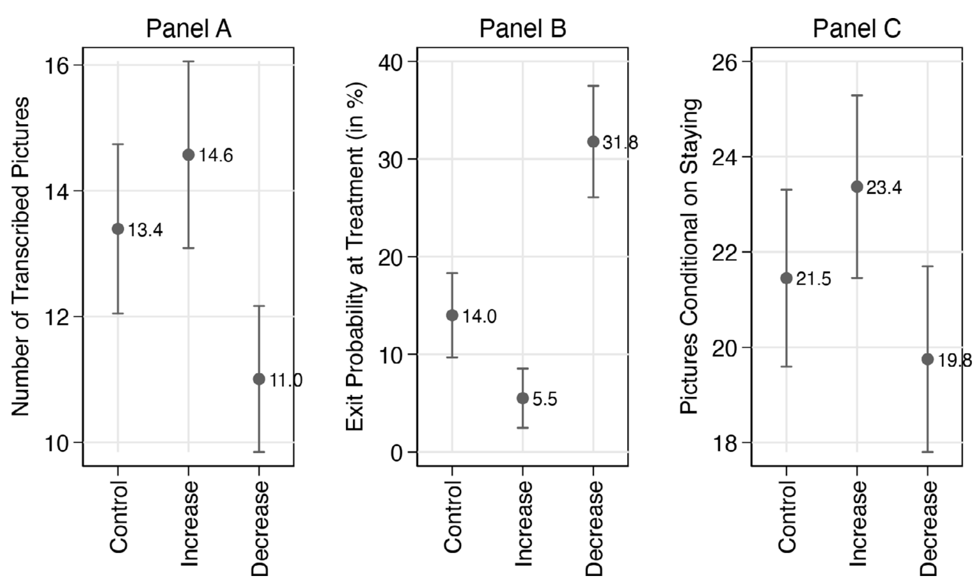


Fig. 5. Transcribed Pictures by Treatment and Control Group. *Notes:* This figure plots the average number of transcribed pictures (Panel A), the probability to quit working after being notified of treatment (Panel B), and the average number of transcribed pictures conditional on transcribing at least one pictures after the treatment (Panel C). In all panels, we present estimates for the control group and the two treatment groups. Point estimates are retrieved from an OLS regression of the respective outcome on three group indicators. The formal model reads $y_i = \beta_0 \mathbb{1}(i \in \mathcal{C}) + \beta_1 \mathbb{1}(i \in \mathcal{I}) + \beta_2 \mathbb{1}(i \in \mathcal{D}) + \epsilon_i$, where $\mathcal{C}$, $\mathcal{I}$, and $\mathcal{D}$ denote the control group, the wage increase, and the wage decrease group, respectively. Vertical bars indicate 95% confidence intervals based on robust standard errors. See columns 1–3 of Panel A in Appendix Table A.1 for detailed regression results and statistics.

Fig. 4 also allows a first, tentative check for an asymmetric response to wage changes. If the behavioral response was *symmetric* between groups, both hazard rates should decline with roughly similar distance to the control group. We observe the opposite: the (absolute) response of workers in the decrease group is larger than the reaction of workers in the increase group. In the following, we discuss each of these observations in more detail and also provide inference for the different effects.

Treatment Effect. For each of the three groups, Fig. 5 shows the average number of transcribed pictures (Panel A), the probability of leaving our experiment after receiving the treatment notification (Panel B), and the average number of transcribed pictures conditional on continuing to work after this announcement (Panel C). While the average worker transcribed 13.4 pictures in the control group, the average worker completed 14.6 and 11 pictures in the wage-increase and wage-decrease groups, respectively. The relationship between labor supply and wages is thus positive and group averages are oftentimes statistically different from each other. The treatment effects in Panel A translate into labor supply elasticities of 0.44 for the wage increase group, and 0.89 for the wage decrease group (see Appendix Table A.1).

As wage changes are only announced after having transcribed six pictures, we can decompose this effect into *extensive* and *intensive* margin responses. Panel B shows the extensive margin, which we define as the share of workers who quit the experiment immediately after the treatment notification. Our results reveal that the treatment strongly affects the exit probability. In the control group, 14 percent of workers quit after their sixth transcription (even though their wage remains constant). Workers from the increase group are 8.5 percentage points less likely to quit after being notified of the wage increase. In contrast, workers who face a wage decrease are twice as likely to quit compared to the control group.

Finally, we compare the intensive margin behavior across the three groups in Panel C. We define the intensive margin as the number of transcribed pictures conditional on completing at least one picture after seeing the treatment information. Again, our results provide evidence for a positive relationship between wages and labor supply. While workers in the control group who continue to work after the treatment notification transcribe 21.5 pictures on average, workers from the wage-increase group transcribe approximately two additional pictures. In contrast, workers from the decrease group transcribe only 19.8 pictures on average conditional on transcribing at least seven pictures.

Note that these intensive margin results may partly reflect endogenous differences in the composition of workers continuing to work in the three groups. However, we find little differences in the time spent transcribing pictures and the workers' average time per picture around the treatment notification (see Appendix Fig. A.2). We further discuss potential endogenous selection issues along this margin in Section 4.1.

Asymmetric Effects. To investigate the symmetry of these treatment effects, we calculate the absolute differences in the distributions between (i) the wage-increase and the control group and (ii) the wage-decrease and the control group, respectively. The gaps in both empirical cumulative distribution functions and thus the deviation to the control group in Fig. 4 should be similar in case the treatment responses were symmetric. We run a standard Kolmogorov-Smirnov test for equality of these gap distributions to assess the symmetry of both effects. The check results in a p-value of 0.122 (depicted in Fig. 4).

However, due to the design of our experiment, we should not expect any differences in the pre-treatment rounds where workers transcribe up to six pictures. Differences in the distribution should only appear once workers receive the treatment notification after the sixth transcribed picture. We thus apply Kolmogorov-Smirnov tests separately for the pre- and post-treatment distributions. In line with the seemingly flat trend before the treatment notification, we cannot reject that differences relative to the control group are symmetric (p-value 0.263 until the sixth picture). After seeing the treatment notification, differences in the distributions emerge compared to the control group. The respective test for equality is clearly rejected (p-value 0.001 after the sixth picture). We also employ Wilcoxon rank-sum tests to non-parametrically test for symmetry of treatment effects, finding comparable results with p-values below 1%. Employing a parametric χ^2-test on the two distributions measuring the gaps between groups provides similar evidence ($p < .001$).

We continue by testing for symmetry in means for each of the three outcome variables depicted in Fig. 5. To this end, we calculate the difference between the control group and the wage-increase or wage-decrease group, respectively. Panel A shows that, relative to the control group, workers in the increase group transcribe an average of 1.2 additional pictures while workers in the wage-decrease group transcribe about 2.4 fewer pictures. The behavioral response to wage decreases is thus twice as large as the effect of wage increases of equal absolute magnitude. Similarly, Panel B shows the differences in exit probability immediately after the treatment notification. While wage increases yield an 8.5 percentage points lower probability of exit (compared to the control group), wage decreases trigger an 17.8 percentage point increase in the exit probability. Again, the response to wage decreases is about twice as large as the wage increase effect. Finally, Panel C shows no indication of asymmetry in number of pictures transcribed conditional on continuing to work on the experiment; wage increases and decreases yield 1.9 more and 1.7 fewer pictures, respectively.

Both the total and the extensive margin effects in Fig. 5 provide suggestive evidence for an asymmetric response, which is further strengthened as confidence bands of increase and decrease group estimates are non-overlapping (see Panels A and B). However, these comparisons of means across groups are still only indicative of stronger absolute responses in the wage-decrease group.

We provide a more formal statistical test for asymmetry in means by estimating the treatment effects in the wage-increase and the wage-decrease groups (relative to the control group) and then testing whether absolute responses in the wage-decrease group are indeed larger than those of the wage-increase group (using simple one-sided Wald tests, see Panel B of Appendix Table A.1).[11] It turns out that we cannot statistically reject the null hypothesis of symmetry in mean outcomes between both groups for the total number of transcribed pictures shown in Panel A ($p = .24$). When applying a similar test to check for asymmetry at the extensive margin, we can reject the null of symmetry at the 5% significance level.

The formal inference checks on group means thus provide support for asymmetric effects on the extensive margin only. When applying two-sided tests and also allowing for wage-decrease responses to be smaller than wage-increase effects, the corresponding p-value raises to 0.09. Despite the suggestive empirical evidence and the strong non-parametric evidence for differences in distributions of the total number of transcribed pictures, we cannot reject symmetric effects of group averages.

Treatment Dynamics. We further investigate the dynamics of the treatment effect in Fig. 6. In Panel A, we again estimate the differences in hazard rates between the two treatment groups relative to the control group as baseline. As already indicated in the descriptive Fig. 4, we find that a wage increase raises the probability that workers transcribe more pictures after the treatment notification. After 20–25 pictures transcribed, this effect vanishes and we can hardly distinguish the wage increase group from the control group. In contrast, the negative effect of wage decreases turns out rather persistent until the maximum number of pictures is reached. Comparing the effect size between both groups again yields suggestive evidence for asymmetric effects, which is confirmed by the Kolmogorov-Smirnov tests depicted in Fig. 4.

Panel B provides additional evidence by plotting the estimates from a linear probability model where we regress workers' exit decisions on a full interaction of group and picture transcribed indicators. The results show that the treatment notification affected the exit probability immediately in the decision to continue working whereas exit decisions are similar across groups both before and afterwards.

3.2. Time responses

Next we turn to time related outcome variables. We start by comparing the total time worked on our task across the three groups. Time is measured in minutes between submitting the first and the last transcribed picture.[12] Fig. 7 illustrates the average time worked on our task (Panel A) and the total time conditional on transcribing at least one more picture after the treatment stage (Panel B) across groups. In line with the results in Section 3.1, we find a positive relationship between wages and labor supply. Relative to the control group, workers in the increase group work slightly more on average while workers in the decrease group work almost eight minutes less. These effects provide suggestive evidence for asymmetry, but differences in group means are only marginally significant ($p = .11$, see Panel B of Appendix Table A.2). Results are qualitatively similar when conditioning on the subsample of workers who continue transcribing after the treatment stage (see Panel B of Fig. 7).

We also check for treatment effects on the transcription rate, i.e., the minutes worked per picture transcribed, and the work quality, which we define as the share of characters accurately transcribed. We do not

[11] We use one-sided tests because of the clear visual indication of stronger absolute responses of the wage-decrease group (see Figs. 4 and 5). Appendix Table A.1 also shows p-values for two-sided tests.

[12] This approximation may overstate the actual time worked as we ignore breaks. To limit the concern of outliers, we drop the top 1% of workers in terms of total time worked (those working more than four hours on the task). We also tested alternative cutoff values and find that results are quite stable.

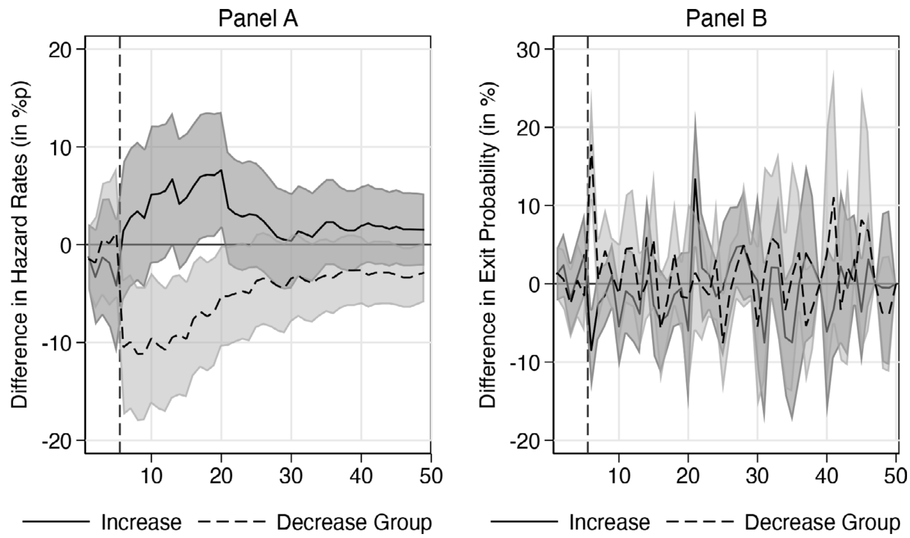


Fig. 6. Treatment Effect by Number of Pictures Transcribed. *Notes:* This figure plots the treatment effect dynamics, i.e., estimated differences in hazard rates (Panel A) and exit probabilities (Panel B) for the wage-increase and wage-decrease group relative to the control group. In both panels, we present estimates for the control group and the two treatment groups. Point estimates are retrieved from an OLS regression of the respective outcome on three group indicators fully interacted with number of picture indicators. The formal model reads $y_{ip} = \sum_{k=1}^{50} \mathbb{1}(k=p)\left[\beta_0^k \mathbb{1}(i \in \mathcal{C}) + \beta_1^k \mathbb{1}(i \in \mathcal{I}) + \beta_2^k \mathbb{1}(i \in \mathcal{D})\right] + \epsilon_{ip}$, where $\mathcal{C}$, $\mathcal{I}$, and $\mathcal{D}$ denote the control group, the wage increase, and the wage decrease group, respectively. The estimations are based on a panel data set with worker-picture observations (for all 1,158 workers). Gray shaded areas indicate 95% confidence intervals based on cluster-robust standard errors (clustering at the worker level).

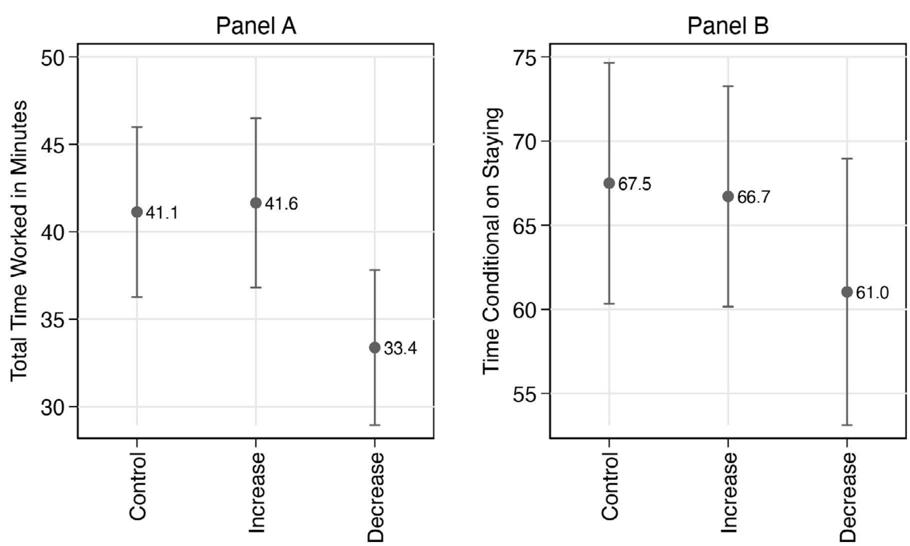


Fig. 7. Time Worked by Treatment and Control Group. *Notes:* This figure plots the average time worked on the task (Panel A), and the average time worked on the task conditional on transcribing at least one pictures after the treatment (Panel B). In both panels, we present estimates for the control group and the two treatment groups. Point estimates are retrieved from an OLS regression of the respective outcome on three group indicators. The formal model reads $y_i = \beta_0 \mathbb{1}(i \in \mathcal{C}) + \beta_1 \mathbb{1}(i \in \mathcal{I}) + \beta_2 \mathbb{1}(i \in \mathcal{D}) + \epsilon_i$, where $\mathcal{C}$, $\mathcal{I}$, and $\mathcal{D}$ denote the control group, the wage increase, and the wage decrease group, respectively. Vertical bars indicate 95% confidence intervals based on robust standard errors. See columns 4–5 of Panel A in Appendix Table A.1 for detailed regression results and statistics.

find any economically and statistically meaningful differences for both outcomes (see Appendix Fig. A.3).

3.3. Robustness

Because the workers discussed our task on the mTurk forum, it is possible that our findings are driven by selection into our experiment. We explore this by performing the analyses separately on the sample of workers who worked on our task before it was discussed online and the sample of workers who worked on it afterwards. These results, which are presented in Online Appendix C, show no evidence that our results are driven by selection among workers who participated in the post-forum period. In addition, we regress each outcome variable on a dummy variable indicating whether the subject worked on the task before or after the forum post; we do not find any significant effects of working on the task after the forum post (results not reported).

4. Discussion of results

In this section, we discuss potential explanations behind our results (Section 4.1) and then describe the implications of our findings (Section 4.2).

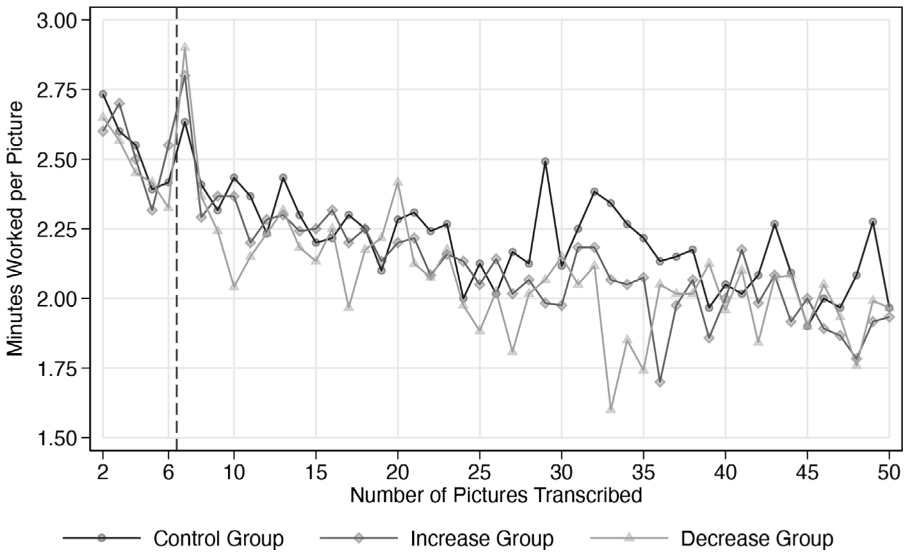


Fig. 8. Median Time Worked per Picture by Group and Picture Transcribed. *Notes:* This figure plots the median time worked by picture over the number of pictures transcribed for the control group and the two treatment groups. We measure time worked by the difference between two submitted pictures (in minutes). This also explains why the time spent for picture number seven peaks in all three groups since this time includes the time spent reading the treatment notification.

4.1. Mechanisms

Our experiment provides evidence that labor supply responses to wage changes are asymmetric. Absolute differences in distributions of pictures transcribed – compared to the control group – are statistically significant based on Kolmogorov-Smirnov tests. While estimated differences in means are suggestive of asymmetry along various dimensions, we can only reject the null of symmetry at conventional significance levels when looking at the extensive margin, i.e., the probability to quit our task immediately after the treatment notification. In the following, we discuss various theoretical explanations that help us understand the economic mechanisms behind our results.

Even a standard labor supply model implies asymmetric effects of wage increases and wage decreases if utility is concave in consumption.[13] Our evidence along the various margins is consistent with this idea. While we have no formal way to prove that (i) workers in our experiment follow the standard model and (ii) alternative models cannot explain our results, we argue that a simple labor supply framework is indeed the most plausible mechanism behind our findings. We discuss various alternative explanations in the following.

Reservation Wages. Another possible explanation for our results are differences in reservation wages. If workers are operating rationally, then the distribution of reservation wages in our experiment should be bounded between \$0 and \$0.15 per transcribed picture. Workers with reservation wages above \$0.15 will not participate in the experiment because doing so would not cover the disutility of transcribing images. We would expect workers with reservation wage between \$0.12 and \$0.15 to quit the labor task when the piece rate decreases to \$0.12 per picture. For the wage increase group, we expect a zero effect as there is no option to enter our task after the treatment notification. Such a model would thus imply asymmetric labor-supply responses.

However, this story is only partly consistent with our findings for two reasons. First, previous studies by Horton and Chilton (2010) and Horton et al. (2011) find that workers on mTurk have a median reservation wage between \$1.12 and \$1.38 per hour,[14] which is substantially lower than the implied average hourly wage of \$3.57 in our wage decrease group.[15] Although the new piece rate will fall below the reservation wage for some workers, we argue that this explanation is unlikely to account for the large exit rate in the wage decrease group given the low median reservation wage on mTurk. Second, we observe a positive extensive-margin response in the wage-increase group, which appears inconsistent with the reservation-wage argument since every worker in this group would have been paid above her reservation wage from the beginning of the experiment.

Transcription Skills and Effort Costs. The positive extensive-margin effect in the increase group may reflect imperfect information about the effort costs of the transcription before workers start working on our task. Once workers actually transcribe pictures, they update their estimate of the costs of working, which in turn influences their decision to continue working and transcribe additional pictures. This could also be one explanation for the positive quit-rates across groups before the treatment notification. Depending on the shape of workers' cost functions, this channel may also lead to asymmetric effects.

We investigate this channel by comparing the transcription rate before and after the treatment notification (see Fig. 8). Workers' median time worked per picture decreases from 2 minutes and 40 seconds in the very beginning to 2 minutes and 18 seconds for the tenth picture transcribed. This difference of 22 seconds could be due to two distinct explanations: (i) workers becoming more efficient, or (ii) less-skilled workers quitting our task. To distinguish between these arguments, we calculate the initial transcription rate for workers who stay (at least) until the tenth period; they need 2 minutes and 33 seconds. More productive workers are thus more likely to keep transcribing pictures than less productive ones. These figures imply that two thirds of the reduction in working time are due to learning, whereas one third is due to selection, i.e., slow workers' exit.

To test whether learning and selection may drive our results and in particular the asymmetry, we regress the time worked for each transcribed picture on the full interaction of group and number of picture indicators. Appendix Fig. A.2 provides the corresponding results, showing the transcription rate for each picture (Panel A) as well as the average time per picture (Panel B). We find no systematic differences between the two treatment and the control group neither before nor after treat-

[13] For example, a CRRA utility function $\frac{1}{1-\gamma}(wL)^{1-\gamma}$ with linear costs aL yields asymmetric labor supply responses for non-marginal wage changes.

[14] The task used by Horton et al. (2011) is very similar to our environment. It also required workers to transcribe chunks of text written in a foreign language and paid them per transcribed text.

[15] This hourly wage is a lower bound estimate because our measure of the time it takes to transcribe one picture potentially overstates the actual working time as workers may have breaks in between.

ment. Therefore, learning and skill-related selection seem unrelated to the treatment and cannot explain our findings.

Loss Aversion. An alternative explanation for our finding of asymmetry could be reference dependence in consumption. Reductions in the piece rate would yield losses relative to reference consumption holding labor supply constant. To compensate for this loss, workers should increase labor supply in response to a reduction in the piece rate (Camerer et al., 1997). Our findings lend little support for this hypothesis as we find a positive relationship between wages and labor supply; workers transcribe less pictures after the piece rate decreases.[16] However, we are only one employer in a large pool of other tasks on mTurk. It is possible that workers may have simply responded by shifting to a different labor task to make up for the lost consumption. Therefore, these workers increase labor supply as predicted, but do so by working for a different firm.

While this type of response, i.e., switching to other jobs, would be consistent with loss-aversion as a behavioral explanation in the wage-decrease group, the reference point theory would also predict that wage increases lead to *lower* labor supply as income gains relative to the reference consumption bundle generate little utility gain (compared to the effect in the loss domain). Again, this theory is at odds with our finding of a positive link between wages and labor supply. Therefore we argue that loss aversion can hardly account for the effects we observe.

Reciprocity. Another potential explanation of our findings is reciprocity; workers interpret wage changes as punishment or reward, and respond accordingly. While this mechanism does not imply asymmetric responses per se, Kube et al. (2013) provide evidence that wage changes may indeed lead to asymmetric effects in a gift-exchange setting.

We argue that this is an unlikely explanation based on our experimental design which is different from a gift-exchange setup. First, subjects are paid for each completed transcription and not per unit of time. This implies that workers punish themselves in the form of lower pay-off. Workers are also likely to know that employers can easily recruit other workers to transcribe pictures and that employers therefore do not face the risk that pictures remain untranscribed. Second, workers on mTurk receive performance rating for their work, which affects their prospects of being allowed to work on other mTurk tasks. Negative reciprocity in terms of bad transcription quality risks bad performance ratings which may limit the number of tasks workers will qualify to work on in the future. In line with this, we find no meaningful differences in accuracy between groups. Third, we screened workers' discussions on the mTurk forum. While they discussed the wage changes, the posts provide little indication of work morale as the reason for quitting.

4.2. Implications

The existing labor-supply literature often identifies labor supply elasticities by exploiting panel data comprised of both wage increases and decreases. This approach becomes quantitatively problematic when one considers that wage changes are usually non-marginal. The reason is that several theoretical models—including the very standard labor supply models—predict labor supply responses to wage changes to be asymmetric. Consistent with this theoretical idea, our results provide evidence for asymmetric responses.

Our findings suggest that ignoring the direction of wage changes when estimating labor supply elasticities leads to biased own-wage labor-supply elasticities; estimates lumping together wage increases and decreases overstate the response to wage increases and understate the behavioral reaction to falling wages. We find that the asymmetry of labor supply w.r.t. wages is particularly pronounced on the extensive margin relative to the intensive margin, which is important since labor supply elasticities are mainly determined by the extensive margin response (Bargain et al., 2014; Blundell and MaCurdy, 1999; Meghir and Phillips, 2010).

Additionally, our results highlight one potential reason for the downward rigidity in wages. Prominent explanations include institutions such as minimum wages and collective bargaining. Recent evidence by Kaur (2019), however, shows that wages are downward rigid even in the absence of such institutions. Our findings of large negative extensive margin responses to wage cuts adds to this literature, suggesting that nominal wage cuts are damaging for firms—one reason why firms are reluctant to reduce nominal wages. This is likely to be even more true in the context of firm-specific labor supply where the firm has very little market power as indicated by the fairly large retention elasticity.

The results described above are obtained using an experimental design in a large online labor market. Importantly, workers did not know they participated in an experiment and thus behaved as they would in their natural occurring environment. We argue that our experiment thus generalizes to other labor markets with piece rates, flexibility, and multiple outside options. Prime examples of such labor markets are on-line crowd-sourcing labor markets, which are becoming increasingly common in the current technological age (see, e.g., Farrell and Greig, 2016; Farrell et al., 2018, and Katz and Krueger, 2019). A common feature of these labor markets is that workers tend to work for relatively low wages and have extremely high levels of flexibility. Due to randomization, our experimental design also ensures internal validity.

5. Conclusion

We estimate the effect of wage changes on labor supply using data generated in a field experiment on Amazon's Mechanical Turk. Our findings show that the labor-supply behavior of workers on mTurk is upward sloping; the relationship between wage changes and changes in labor supply is positive both for the case of wage increases and wage decreases. We further find evidence of asymmetric responses along several margins. The evidence is particularly strong for the extensive margin where the behavioral response to wage decreases is twice as large as the response to equally sized wage increases (in absolute terms). Our results are consistent with a standard model of labor supply and mostly at odds with other theories such as reciprocity or reference dependence.

Data availability

Replication data is available via doi:10.7910/DVN/YDK1ZI.

[16] This positive relationship between wages and labor supply is consistent with the study by Farber (2015) who extends the findings of Camerer et al. (1997). The results in Farber (2015) suggest that that daily labor supply of NYC taxi drivers is inconsistent with reference dependence, and it is rather explained by a standard model of labor supply.

Appendix A

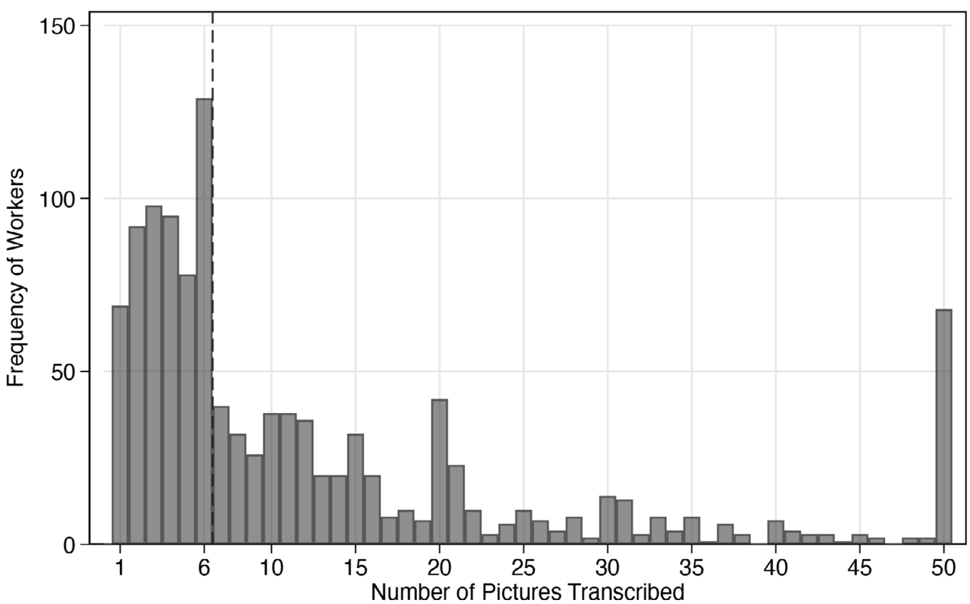


Fig. A1. Histogram of Transcribed Pictures. *Notes:* This figure plots the histogram of pictures transcribed for all 1,158 workers who worked on the task. Subjects saw the treatment notification after transcribing six pictures (indicated by the dashed vertical line).

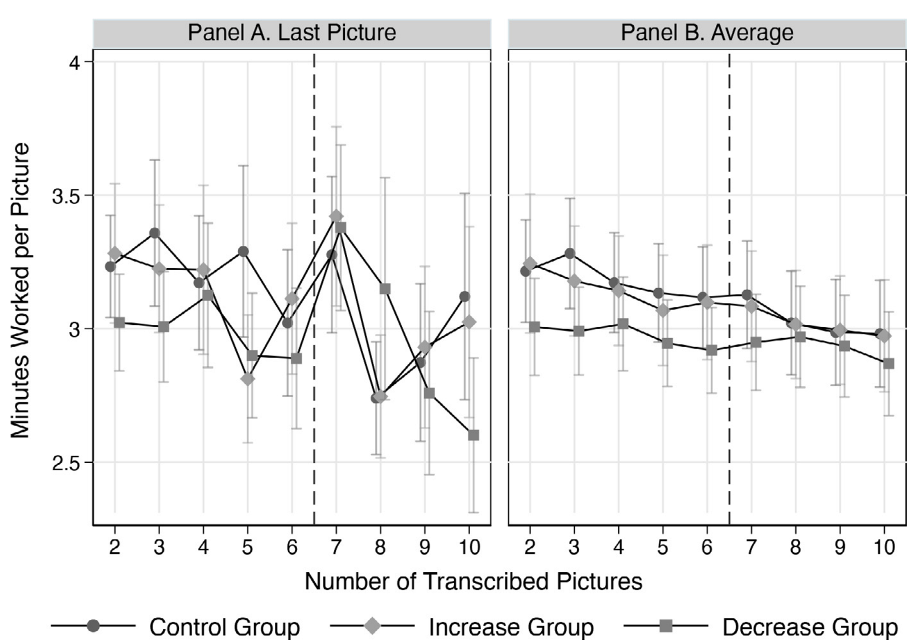


Fig. A2. Time Per Picture by Treatment and Control Group. *Notes:* This figure plots the time spent transcribing the last picture (Panel A), and the average time per picture over all transcribed pictures so far (Panel B). In both panels, we present estimates for the control group and the two treatment groups. Point estimates are retrieved from an OLS regression of the respective outcome on three group indicators fully interacted with number of picture indicators. The formal model reads $y_{ip} = \sum_{k=2}^{10} \mathbb{1}(k = p)\left[\beta_0^k \mathbb{1}(i \in \mathcal{C}) + \beta_1^k \mathbb{1}(i \in \mathcal{I}) + \beta_2^k \mathbb{1}(i \in \mathcal{D})\right] + \epsilon_{ip}$, where $\mathcal{C}$, $\mathcal{I}$, and $\mathcal{D}$ denote the control group, the wage increase, and the wage decrease group, respectively. The estimations are based on a panel data set with worker-picture observations (for all 1,158 workers). Vertical bars indicate 95% confidence intervals based on cluster-robust standard errors (clustering at the worker level).

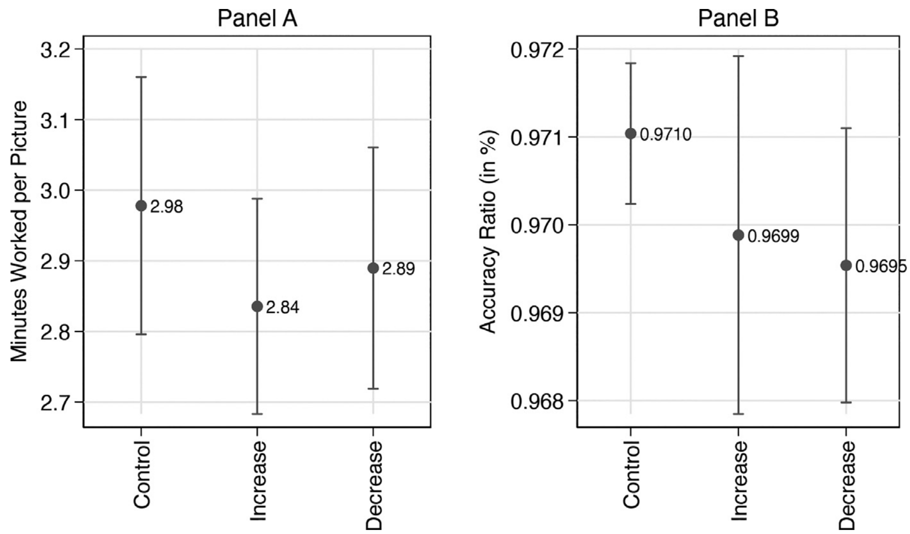


Fig. A3. Time per Picture and Accuracy by Treatment and Control Group. *Notes:* This figure plots the average time spent transcribing per picture (Panel A), and the accuracy ratio per picture (Panel B). In both panels, we present estimates for the control group and the two treatment groups. Point estimates are retrieved from an OLS regression of the respective outcome on three group indicators. The formal model reads $y_{ip} = \beta_0 \mathbb{1}(i \in \mathcal{C}) + \beta_1 \mathbb{1}(i \in \mathcal{I}) + \beta_2 \mathbb{1}(i \in \mathcal{D}) + \epsilon_{ip}$, where $\mathcal{C}$, $\mathcal{I}$, and $\mathcal{D}$ denote the control group, the wage increase, and the wage decrease group, respectively. The estimation is based on a panel data set with worker-picture observations (for all 1,158 workers). Vertical bars indicate 95% confidence intervals based on cluster-robust standard errors (clustering at the worker level). See columns 6–7 of Panel A in Appendix Table A.1 for detailed regression results and statistics.

Table A1
Regression Results.

	(1) Total Pictures		(2) Exit Probability		(3) Pictures Conditional		(4) Total Time		(5) Time Conditional		(6) Time Per Picture		(7) Accuracy Ratio	
Panel A – Averages by Treatment and Control Group														
Control Group	13.394	***	14.000	***	21.451	***	41.132	***	67.495	***	2.978	***	0.971	***
	(0.686)		(2.199)		(0.946)		(2.476)		(3.643)		(0.093)		(0.000)	
Increase Group	14.572	***	5.505	***	23.369	***	41.649	***	66.710	***	2.836	***	0.970	***
	(0.756)		(1.548)		(0.975)		(2.466)		(3.336)		(0.078)		(0.001)	
Decrease Group	11.008	***	31.783	***	19.750	***	33.377	***	61.039	***	2.890	***	0.970	***
	(0.591)		(2.905)		(0.991)		(2.261)		(4.030)		(0.087)		(0.001)	
Observations	1,158		726		597		1,147		586		13,696		14,993	
Adjusted R^2	0.489		0.242		0.714		0.405		0.620		0.639		0.999	
Panel B – Treatment Effects Relative to Control Group														
Increase Group	1.178		-8.495	***	1.918		0.517		-0.785		-0.143		-0.001	
	(1.021)		(2.689)		(1.358)		(3.494)		(4.940)		(0.121)		(0.001)	
Decrease Group	-2.387	***	17.783	***	-1.701		-7.754	**	-6.456		-0.088		-0.001	*
	(0.905)		(3.643)		(1.370)		(3.353)		(5.433)		(0.127)		(0.001)	
Constant	13.394	***	14.000	***	21.451	***	41.132	***	67.495	***	2.978	***	0.971	***
	(0.686)		(2.199)		(0.946)		(2.476)		(3.643)		(0.093)		(0.000)	
Observations	1,158		726		597		1,147		586		13,696		14,993	
Adjusted R^2	0.011		0.079		0.008		0.005		-0.000		0.001		0.000	
p-val. Symmetry (one-sided)	0.235		0.046		0.463		0.113		0.210		0.146		0.043	
p-val. Symmetry (two-sided)	0.470		0.091		0.927		0.226		0.420		0.293		0.085	
Panel C – Implied Elasticities														
Wage Increases	0.440		-3.034				0.063							
Wage Decreases	0.891		-6.351				0.943							

Notes: This table presents the estimation results for our empirical analysis. Panel A presents group averages for the control group and the two treatment groups. Panel B shows the treatment effects for the wage-increase and the wage-decrease group, respectively, relativ to the control group (see constant). In Panel B we also formally test for asymmetric effects. The presented p-values test the null hypothesis that treatment effects are symmetric, i.e., they sum up to zero. Panel C shows implied elasticities for the key outcomes. Columns 1–5 are estimated on a cross section of workers, columns 6 and 7 are based on a panel data set with worker-picture observations. Standard errors are robust to clustering at the worker level. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.labeco.2022.102305.

References

Bargain, O., Orsini, K., Peichl, A., 2014. Comparing labor supply elasticities in Europe and the United States – New results. J. Human Resourc. 49 (3), 723–838.

Benzarti, Y., Carlon, D., Harju, J., Kosonen, T., 2020. What goes up may not come down: Asymmetric incidence of value added taxes. J. Polit. Econ. 128 (12).

Blundell, R., MaCurdy, T.E., 1999. Labor Supply: A Review of Alternative Approaches. In: Ashenfelter, O., Card, D. (Eds.), Handbook of Labor Economics. North-Holland, Amsterdam.

Buhrmester, M., Kwang, T., Gosling, S.D., 2011. Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality, data? Perspect. Psychol. Sci. 6 (1), 3–5.

Camerer, C., Babcock, L., Loewenstein, G., Thaler, R., 1997. Labor supply of new york city cabdrivers: One day at a time. Q. J. Econ. 112 (2), 407–441.

Charness, G., Kuhn, P., 2011. Lab Labor: What Can Labor Economists Learn from the Lab? In: Ashenfelter, O., Card, D. (Eds.) Handbook of Labor Economics Vol 4A. North Holland, Amsterdam, pp. 229–330.

Cohn, A., Fehr, E., Goette, L., 2015. Fair wages and effort provision: Combining evidence from a choice experiment and a field experiment. Manage. Sci. 61 (8), 1777–1794.

DellaVigna, S., Pope, D., 2018. What motivates effort? Evidence and expert forecasts. Rev. Econ. Stud. 85 (2), 1029–1069.

Evers, M., De Mooij, R., Van Vuuren, D., 2008. The wage elasticity of labour supply: A Synthesis of empirical estimates. Economist (Leiden) 156 (1), 25–43.

Falk, A., Fehr, E., Zehnder, C., 2006. Fairness perceptions and reservation wages – The behavioral effects of minimum wage laws. Q. J. Econ. 121 (4), 1347–1381.

Farber, H.S., 2015. Why you can't find a taxi in the rain and other labor supply lessons from cab drivers. Q. J. Econ. 130 (4), 1975–2026.

Farrell, D., Greig, F., 2016. Paychecks, paydays and the online platform economy. JPMorgan Chase Co. Inst..

Farrell, D., Greig, F., Hamoudi, A., 2018. The online platform economy in 2018: Drivers, workers, sellers, and lessors. JPMorgan Chase Co. Inst..

Fehr, E., Goette, L., 2007. Do workers work more if wages are high? Evidence from a randomized field experiment. Am. Econ. Rev. 97 (1), 298–317.

Gneezy, U., List, J.A., 2006. Putting behavioral economics to work: Testing for gift exchange in labor markets using field experiments. Econometrica 74 (5), 1365–1384.

Hennig-Schmidt, H., Sadrieh, A., Rockenbach, B., 2010. In search of workers' real effort reciprocity – A field and a laboratory experiment. J. Eur. Econ. Assoc. 8 (4), 817–837.

Horton, J.J., Chilton, L.B., 2010. The labor economics of paid crowdsourcing. Proceedings of the 11th ACM Conference on Electronic Commerce.

Horton, J.J., Rand, D.G., Zeckhauser, R.J., 2011. The online laboratory: Conducting experiments in a real labor market. Exp. Econ. 14, 399–425.

Kahn, S., 1997. Evidence of nominal wage stickiness from microdata. Am. Econ. Rev. 87 (5), 993–1008.

Katz, L.F., Krueger, A.B., 2019. Understanding trends in alternative work arrangements in the united states. RSF: Russell Sage Found. J. Soc. Sci. 5 (5), 132–146.

Kaur, S., 2019. Nominal wage rigidity in village labor markets. Am. Econ. Rev. 109 (10), 3585–3616.

Keane, M.P., 2011. Labor supply and taxes: A Survey. J. Econ. Lit. 49 (4), 961–1075.

Kube, S., Marechal, M.A., Puppe, C., 2013. Do wage cuts damage work morale? Evidence from a natural field experiment. J. Eur. Econ. Assoc. 11 (4), 853–870.

Mason, W., Suri, S., 2011. Conducting behavioral research on Amazon's Mechanical Turk. Behav. Res. 44, 1–23.

Meghir, C., Phillips, D., 2010. Labour Supply and Taxes. In: Mirrless, J., Adam, S., Besley, T., Blundell, R., Bond, S., Chote, R., Gammie, M., Johnson, P., Myles, G., Poterba, J. (Eds.), Dimensions of Tax Design: The Mirrlees Review. Oxford University Press, pp. 202–274.

Paolacci, G., Chandler, J., Ipeirotis, P.G., 2010. Running experiments on Amazon Mechanical Turk. Judgm. Decis. Mak. 5 (5).