

# **WORKING PAPER**

## **HOW DOES WORKING-TIME FLEXIBILITY AFFECT WORKERS' PRODUCTIVITY IN A ROUTINE JOB? EVIDENCE FROM A FIELD EXPERIMENT**

Marie Boltz  
Bart Cockx  
Ana Maria Diaz  
Luz Magdalena Salas

October 2020  
2020/1007

# How Does Working-Time Flexibility Affect Workers' Productivity in a Routine Job? Evidence from a Field Experiment

Marie Boltz\*, Bart Cockx<sup>†</sup>, Ana Maria Diaz<sup>‡</sup> & Luz Magdalena Salas<sup>§¶</sup>

October 16, 2020

## Abstract

We conducted an experiment in which we hired workers under different types of contracts to evaluate how flexible working time affects on-the-job productivity in a routine job. Our approach breaks down the global impact on productivity into sorting and behavioral effects. We find that all forms of working-time flexibility reduce the length of workers' breaks. For part-time work, these positive effects are globally counterbalanced. Yet arrangements that allow workers to decide when to start and stop working increase global productivity by as much as 50 percent, 40 percent of which is induced by sorting.

**JEL Codes:** J21, J22, J23, J24, J33

**Keywords:** Flexible work arrangements, part-time work, productivity, labor market flexibility, work–life balance.

---

\*BETA, University of Strasbourg, 61 avenue de la Forêt Noire, 67085 Strasbourg, France, m.boltz@unistra.fr

<sup>†</sup>Department of Economics, Ghent University; IZA, Bonn; CESifo, Munich; IRES, Université catholique de Louvain; ROA, Maastricht University, bart.cockx@ugent.be

<sup>‡</sup>Departamento de Economía, Pontificia Universidad Javeriana, a.diaze@javeriana.edu.co

<sup>§</sup>Departamento de Economía, Pontificia Universidad Javeriana, salas.luz@javeriana.edu.co

<sup>¶</sup>This paper has benefited from great research assistance by Juan David Hernandez and Juan Sebastian Ortiz. We thank Nicolas Lillo for his work building the software and providing the Chilean Census data. The authors would also like to thank Pontificia Universidad Javeriana, University Paris 1 Panthéon Sorbonne, and the University of Ghent for financial support. The research received IRB approval from Universidad Javeriana on October 3, 2018.

# 1 Introduction

Working-time and workplace flexibility play an essential role in promoting an adequate work–life balance, especially for working parents (OECD, 2016). They are also key instruments for reducing the gender gap in both labour force participation and earnings (Bertrand, 2018; Goldin, 2014). Part-time work is an option for employees who want to reduce their working time, but it comes at a price – reduced earnings. In 2019, 9.6 percent of male employment and 25.4 percent of female employment in Organisation for Economic Co-operation and Development (OECD) countries was part time.<sup>1</sup> Thus working part time may *reinforce* rather than reduce the gender pay gap, although it may reduce the gender gap in labour participation rates. Work-schedule flexibility, in which workers can determine when to start and finish work each day, does not entail such an earning penalty. However, few workers are offered this type of flexibility. In Europe, only about one-third of employees report having some form of control over their working time; far fewer employees in routine occupations have this option (Eurofound, 2015). Working from home is a third flexible arrangement, which offers flexibility in working time as well as the choice of workplace. Teleworking has boomed during the COVID-19 pandemic, but it was less widespread before. According to ILO (2020), approximately 3 percent of employees worldwide worked exclusively from home in 2019. In the US and Europe this number is larger (one in five employees),<sup>2</sup> but in developing countries, working from home is less frequent due to lower internet access rates (World-Bank, 2016).

A key question regarding such flexible working arrangements is why more firms do not offer them. If workers value such arrangements, they are willing to work for lower wages in exchange for such flexibility.<sup>3</sup> Alternatively, when keeping pay constant, the option of such autonomy may boost a worker’s productivity either by increasing their motivation or stimulating reciprocal behavior (Beckmann et al., 2017). But employers may worry that giving workers more freedom will cause them to work less or to reduce their level of effort. Some studies have demonstrated a net productivity-enhancing

---

<sup>1</sup>OECD (2020), Part-time employment rate (indicator). doi: 10.1787/f2ad596c-en (Accessed on 11 October 2020). In Colombia, where our study was conducted, the figures are very similar: 9.0 percent for men and 26.5 percent for women. They also track closely with the rates from EU28 countries: 8.0 percent for men and 26.3 percent for women.

<sup>2</sup>See the Eurofound (2015) and BLS (2019) reports.

<sup>3</sup>For recent evidence, see Mas and Pallais (2017), Chen et al. (2019), He et al. (2020), Bustelo et al. (2020), and Chen et al. (2020).

effect of work-schedule and workplace flexibility for non-routine tasks and jobs that require little coordination through interactions; yet productivity may decline for routine tasks. However, causal evidence is still scarce, especially with respect to work-schedule flexibility.<sup>4</sup>

The findings are more mixed and the causal evidence even more limited for part-time work.<sup>5</sup> Compared to other forms of flexibility, part-time workers could increase their productivity by suffering from less fatigue than their full-time colleagues or by allowing the more intensive use of capital. However, high fixed start-up costs imply that part-time workers may be less productive than their full-time counterparts, because the latter incur such costs for any number of hours worked. Moreover, part-time workers may be less committed to their career goals, and the returns on training may be lower (Garnero, 2016).

This article reports the results of a field experiment we conducted in Bogotá, Colombia's capital city, to provide causal evidence of the effects on productivity in a *routine* temporary job of two of the three flexible working arrangements described above: (1) working part time and (2) being able to decide when to start and stop working within the work week. Prior field experiments have focused on the effectiveness of working from home.<sup>6</sup> However, it is impossible to determine whether teleworking affects productivity because it provides flexibility regarding the choice of *where* to work or *when* to work. By keeping the location decision fixed and only varying the flexibility with respect to time, we can obtain more insights into this question.

Flexible work schemes can enhance productivity in two ways. They can either increase the attractiveness of the job such that intrinsically more productive workers *self-select* into it, or they incentivize workers to exert more effort or time to working (Beckmann, 2016). Our field experiment aims to capture both sources and disentangle the effect into these components. In order to capture the consequence of self-selection, the experiment began from the moment the job ads for the three-week positions as data entry operators were posted. We advertised the vacancy without reference to the contract environment. At this stage, applicants were asked to provide standard resume information and to perform an online test in which they had to carry out similar data entry tasks

---

<sup>4</sup>See Dutcher (2012), Goldin (2014), and for a survey, Beckmann (2016).

<sup>5</sup>Künn-Nelen et al. (2013) and Garnero et al. (2014) find that productivity is enhanced in pharmacies in the Netherlands and the Belgian private sector, respectively, while Specchia and Vandenberghe (2013) and Devicienti et al. (2018) find opposite results for a panel of private firms in Belgium and Italy, respectively.

<sup>6</sup>See Dutcher (2012), Bloom et al. (2015), Angelici and Profeta (2020).

as the job to which they were applying. This allowed us to construct an *ex ante* measure of productivity that can disentangle the *sorting* effects on productivity from the *ex post* motivational effects.

We then randomized the 535 applicants who provided all the required information into one of four contract environments: (i) full-time non-flexible schedule – i.e., Monday to Friday, 8 AM to 5 PM (*reference* group), (ii) part-time non-flexible – i.e., Monday to Friday, 8 AM to 12 PM for the morning shift and 1 PM to 5 PM for the afternoon shift; (iii) full-time flexible – i.e., 40 hours a week, but the worker could choose their start and finish times; and (iv) part-time flexible – i.e., 20 hours a week, and the worker could choose their start and finish times. Applicants were informed of this contractual environment, including the exact start and end dates, the location of the workplace, and the hourly wage. They were asked to express their motivation for the job by responding to a number of open-ended questions. Of the candidates who were still interested, we then randomly selected 13 individuals in each contractual environment, i.e. 52 in total, and offered them positions. However, since not all of the job offers were accepted and because we were constrained by an administrative starting date, we ended up sending out 79 job offers; 38 were accepted, and 34 stayed until the end of the contractual period.

Our overall measure of productivity is the number of correct data entries relative to the contractual time. Our point estimates suggest that an employer can increase a worker’s overall productivity by nearly 50 percent by offering a full-time flexible contract rather than a full-time non-flexible one. About 40 percent of this overall effect is attributed to attracting more productive workers. The motivational effect is almost completely driven by the fact that full-time flexible workers take fewer breaks than those with non-flexible contracts. This results in a 10-percentage-point increase in effectiveness relative to contractual working time. This is in line with the results of [Beckmann et al. \(2017\)](#), who, based on the German Socio-Economic Panel, also finds that flexible work-time scheduling arrangements increase actual working hours relative to contractual working hours.

The point estimates suggest that part-time flexible and non-flexible contracts can also enhance global productivity, but these effects are not significant at conventional levels. However, we do find that part-time workers, with or without a flexible time arrangement, spend 15 percentage points less time in breaks relative to contractual time than the reference group, and this effect is highly significant. This does not show up in the global effect on productivity, because these part-time workers make more mistakes,

or are more frequently absent from work.

Our finding that working-time flexibility can substantially enhance productivity in a routine data entry task is new. Prior studies have argued that flexibility is especially valuable in non-routine jobs, but that the impact on productivity in such jobs may be negative (Beckmann, 2016). The laboratory experiment set up by Dutcher (2012) has been highly influential on this view. In that setting, students were randomly assigned to either a “creative” or “dull” task, i.e. a data entry task as in our study. Each of these tasks was carried out either inside or outside a laboratory, to represent working from home. However, students in this study who executed the task outside the lab were not monitored; a major distinguishing feature of our study is that we did monitor workers in the flexible working-time arrangement. A further difference is that we hired real workers rather than students.

Another contribution to the literature is that our field experiment began during the recruitment process; past research has studied workers who are already employed by a firm. This allows us to jointly evaluate the total productivity effect of flexible work arrangements that is induced both by attracting more productive workers and by motivating them to exert more effort. As such, we provide a more complete picture, while at the same time propose a methodology that disentangles both components from the global effect. Moreover, we further decompose these components to explore through which mechanisms productivity is affected.

Our study was conducted in Colombia, which joined the OECD in April 2020. While our conclusions are likely to apply to any economy employing similar routine jobs, they have particularly strong echoes in emerging economies like Colombia. Indeed, we show that working-time flexibility increases productivity in routine jobs: this represents a potentially important source of economic leverage to attract workers out of the informal economy, to increase women’s labour force participation (which is typically low), *and* to increase firm productivity.<sup>7</sup> This suggests interesting research avenues to further understand how working-time flexibility can enhance women’s labour participation in economies where formal and informal labour markets coexist.

The remainder of the paper is organized as follows. The next section briefly reviews

---

<sup>7</sup>Labour force participation in Colombia is 50.5 percent in 2020, but with strong gender differences: 63.1 percent of men and only 28.3 percent of women are employed – one of the lowest figures in OECD countries. OECD (2020), Employment rate (indicator). doi: 10.1787/1de68a9b-en (Accessed on 11 October 2020).

the relevant literature. Section 3 describes the experimental design in more detail. Section 4 presents the data and associated descriptive statistics, and defines the outcome variables and their decomposition. In Section 5 we discuss whether flexible working arrangements attract more productive workers, and in Section 6 we assess how these arrangements affect on-the-job productivity overall, including both sorting and motivational effects. In each of these sections we first present the empirical strategy and then the results. In Section 6 we break down the overall effect to explore the mechanisms at work, and report the results of our robustness checks. Section 7 concludes.

## 2 Literature Review

Our paper contributes to a large literature based on surveys that explore how flexibility affects firms' outcomes through the job satisfaction of workers. For example, [Kelliher and Anderson \(2010\)](#) study the implications of providing working-time and place flexibility on work intensity in large multinational companies in the UK. They found that flexibility makes workers more satisfied and more committed to their employers, and therefore they work more intensively. The authors argue that this is induced by reciprocal behavior. In other words, workers exert more effort "in order to return the benefit to their employer." These findings are very similar to those of [Eaton \(2003\)](#) for US firms. However, this literature generally reveals associations rather than causal effects.

[Beckmann et al. \(2017\)](#) use a similar definition of work schedule flexibility as ours, and study the effect of self-managed working time on the level of effort that workers exert in their job. They first developed a modified moral-hazard model that shows how working-time autonomy can decrease or increase effort depending on whether extrinsic or intrinsic worker motivation dominates.<sup>8</sup> Then they used data from the German Socio-Economic Panel to test the empirical prediction of this model, paying particular attention to identifying causal effects. They find that working-time autonomy increases the total hours worked and decreases absenteeism. However, they did not directly measure productivity, as we do in our study.

Closely related to our work, two recent field experiments evaluated the effects of workplace flexibility on workers' productivity.<sup>9</sup> [Bloom et al. \(2015\)](#) carried out a field

---

<sup>8</sup>The authors consider reciprocal behavior as an alternative driver of effort, but their empirical evidence suggests that this factor is less important than intrinsic motivation.

<sup>9</sup>A third highly relevant field experiment is the one conducted by [Dutcher \(2012\)](#), which we discussed

experiment with Chinese call center employees who were asked if they would be interested in working from home. Among those interested, the call center randomly offered positions working from home for a 9-month period. The study finds that working from home increased self-reported satisfaction levels and raised productivity by about 13 percent, 9 percent of which was due to working longer hours than in the office and 4 percent of which resulted from making more telephone calls per hour. The authors also studied the effects of sorting and found that teleworking could further enhance the productivity gains to 22 percent. This study is similar to ours in that they also measured productivity in terms of output per time and the job tasks did not require any specific skills or coordination.<sup>10</sup> However, Bloom's study differs from ours in at least three respects. First, they define flexibility by place rather than by time. Second, they focus on a group of self-selected employees who expressed an interest in that type of flexibility, while we randomly assigned flexibility to job applicants. Moreover, they measured the impact of sorting by offering participants in the initial experimental group to re-select between working at home and at the office, which is different (and arguably less relevant) than measuring this during the recruitment process as we do. Third, the salary of workers in the call center was not fixed, as it was in our study. It varied according to workers' performance and may have interacted with the flexibility arrangement in its productivity-enhancing effect.

In a second closely related experiment, [Angelici and Profeta \(2020\)](#) randomly offered a "smart work" alternative to a sample of over 300 workers from an Italian multi-utility sector firm, and measured the effect on productivity, well-being, and work-life balance. Workers had the opportunity to choose the time and place of work (outside the office) one day a week during a 9-month period. Their experiment used different measures of productivity (self-reported and objective) as well as self-reported measures of well-being and work-life balance. It finds that promoting smart working is an effective way to increase productivity, and to improve well-being and work-life balance. Their study differs from ours in three important ways. First, it only considers workplace flexibility. Second, it evaluates a heterogeneous group of blue- and white-collar workers with various responsibilities, which prevented them from defining a precise productivity measure and from linking effectiveness to the nature of the jobholders' tasks. Third, they did not

---

in the Introduction.

<sup>10</sup>Nevertheless, the telephone calls may not have been as routine as the data entry tasks in our study or that of [Dutcher \(2012\)](#), which may partly explain why they did not find negative effects on productivity as in Dutcher's study.



assess the impact of sorting on productivity.

### 3 Experimental Design

We conducted a field experiment in which we posted real job ads for data entry clerks to observe job applicants' choices and performance in a natural environment. We generated random variation across job seekers in invitations to apply for jobs that differed in two dimensions of job schedule flexibility: working time (full time versus part time) and, for any given working time, the autonomy to schedule when to accomplish this work within the workweek. This resulted in randomizing job seekers into one of four contract environments: non-flexible full time, flexible full time, non-flexible part time, and flexible part time.

The experiment was designed to measure the effect of job schedule flexibility on productivity. This flexibility can affect productivity in two ways. First, inherently more productive workers may be more likely to accept jobs that provide the flexibility they prefer. Second, conditional on being hired, workers may allocate more effort and/or conscientiousness to the job. In this experiment we aim to measure the combined effect of this *ex ante* sorting and *ex post* motivationally induced behavioral reaction to flexibility on workers' productivity. To capture this combined effect, we advertised the position without referring to the contract environment and then randomly offered applicants one of the four arrangements described above. After imposing certain prerequisites (doing a pre-test, confirming interest in the contract type, and responses to open questions) designed to mimic the recruitment selection process, a random sample of remaining candidates was offered the position. The workers who accepted the position were hired for a period of 3 weeks. Their on-the-job productivity was measured throughout this time.

We also seek to disentangle the *ex ante* sorting effects from the *ex post* behavioural effects of job flexibility on productivity. Before the contract type was revealed, job applicants were asked to provide a resume and to take an online test that measured their productivity in similar tasks as the ones required in the job they applied for. This *ex ante* measure allows us to determine (i) to what extent more productive individuals self-select into more flexible work arrangements and (ii) to what extent, conditional on this *ex ante* measure of productivity, more flexible work arrangements enhance the *ex post* produc-

tivity performance between contract types. In the next two subsections, we describe the different steps in the hiring process and working environment, respectively.

### 3.1 Steps of the hiring process

#### Step 1: The job ad

We placed real job advertisements for data entry clerks every week for a month in different standard job postings in Bogota (internet and newspaper). The job offer involved real employment for a task that did not require a specific level of education or specialized skills. We advertised the position without referring to the type of contract, and asked job seekers to fill out an online CV form, which included an online productivity test. The posted job ad was phrased as follows:

**Job Advertisement** Title: Prestigious university needs data clerks to support a research project.  
Description: Contract for the provision of services.  
Duration: Three weeks.  
If you are interested in this offer, apply via the following link: LINK, or send us a message via Whats-App.

#### Step 2: Collecting resume information and the productivity test

The online CV form included standard questions about the applicants' level of education, labor market experience, and other demographic characteristics (see Appendix 8.1 for the complete questionnaire). After filling out the form, applicants were invited to complete the online test to measure their *ex ante* productivity levels.<sup>11</sup> Those who started the application but did not finish the online test were not considered in the next stages of the application process and were dropped from the experiment. The test consisted of executing three tasks that resembled those required in the job to which the job seeker applied.<sup>12</sup> Based on the answers to this test, we constructed the following *ex ante*

---

<sup>11</sup>The tests were timed, but job applicants did not know this.

<sup>12</sup>Appendix 8.2 describes the test.

productivity measure for each individual,  $i$ , and each task,  $T$ :

$$Prod_{iT} = \frac{Correct\ Answers_{iT}}{Total\ Questions_{iT}} \times \frac{1}{Time_{iT}}$$

An applicant was classified as a highly productive worker if she scored higher than the median in all three tests:

$$High_i = 1[Prod_{i1} \succeq Med(Prod_{i1}) \& Prod_{i2} \succeq Med(Prod_{i2}) \& Prod_{i3} \succeq Med(Prod_{i3})] \quad (1)$$

where  $1[A] = 1$  if  $A$  is true, and  $1[A] = 0$  if  $A$  is false.

### Step 3: Random assignment of contract types

Respondents who completed the required resume information and the productivity test were provided with more details about the contractual terms of the job: the period and place of employment, wage, and working-time regime.<sup>13</sup> The first three contractual terms were fixed: 3 weeks of employment between November 26 and December 14, 2018 in the computer labs of the university,<sup>14</sup> remunerated at a gross hourly wage of 7.000 Colombian Pesos (COP), which is about 2.33 USD.<sup>15</sup> The dimensions that varied across contract types were the four working-time regimes: non-flexible full time, flexible full time, non-flexible part time, and flexible part time.

The full-time contract required working 40 hours a week, and the part-time contract was for 20 hours. In a fixed working-time schedule the full-time worker had to work Monday to Friday, 8 AM to 5 PM with a one-hour lunch break from noon to 1 PM. Part-time workers were randomly assigned to work either in the morning (from 8 AM to 12 PM) or the afternoon (from 1 PM to 5 PM). In a flexible working-time schedule the worker could freely determine their working hours within the opening hours of the computer lab, i.e. Monday to Friday, 8 AM to 8 PM.

We stratified the sample before randomizing them to one of the four treatment conditions for two reasons. First, in order to enhance the precision of the estimates of the

<sup>13</sup>Appendix 8.3 presents the phrasing of this e-mail.

<sup>14</sup>We did not offer the possibility of teleworking. The email stated that the work had to be performed at Javeriana University in Bogota.

<sup>15</sup>This salary corresponds to approximately 1.3 times the minimum wage or 74 percent of the average wage in Colombia. The salary was based on the regular salary of a data clerk.

treatment effects, we stratified the respondents according to an attribute, *ex ante* productivity, that we expect to be a good predictor of the potential outcome, *ex post* productivity (Imbens and Rubin, 2015). We created a dichotomous stratification variable that splits the sample into high or low *ex ante* productivity as defined in equation (1) in step 1. The second reason is that since contractual flexibility is particularly relevant for women in general and caregivers in particular (defined by the presence in the household of children under 5 years of age, a disabled person or an adult over 65 years old with permanent care needs), we also stratified the sample along these two dimensions to allow us to study the heterogeneity of the impact of flexibility in these dimensions.<sup>16</sup> This resulted in eight stratification cells defined by all possible combinations of these three stratification variables.

#### **Steps 4 and 5: Mimicking the recruitment selection process**

At the end of the e-mail containing the detailed contractual information, applicants were asked whether they were still interested in the job. Those who responded affirmatively and those who did not respond received another e-mail designed to reduce the number of candidates for the position, as a recruitment exam and/or interview would do in a standard job application. This e-mail (see Appendix 8.4 for more details) inquired about applicants' motivation and asked a number of informative open-ended questions in an attempt to retain only the most motivated candidates.

In a fourth step, job offers were sent to applicants who reacted positively to this last e-mail according to a second stratified random sampling scheme based on the same stratification as the first random assignment. For this purpose, of the retaining candidates, 13 individuals were randomly selected from each contract type, so 52 (= 13 x 4) job offers were sent out. However, not all of these offers were accepted, likely due to the standard additional administrative requests involved, such as the provision of a tax identification number, an identification document, a certificate of affiliation with a health service, and a declaration that one is willing to pay the mandatory social security contributions (see Appendix 8.5 for more details). Consequently, we had to send out additional job offers. Because we were constrained by the announced starting date of the employment contract, we did not manage to hire the intended 13 workers in any of the four contract types. Eventually, after offering a job to 78 applicants, we only managed to

---

<sup>16</sup>However, as mentioned in the next subsection, since fewer individuals were hired than expected, the small sample size did not allow such heterogeneity analysis.

hire 38 workers; four of these quit shortly after the contract started. Section 4 provides additional descriptive evidence.

## 3.2 The work environment

The workers were hired to perform a data entry task in the computer labs at Universidad Javeriana in Bogota. We assigned a separate lab to each treatment group, located in different buildings, to reduce the risk of communication between individuals from different treatment groups. We hired eight undergraduate students who monitored each computer lab during the entire experiment. Monitors were randomly assigned to one of the four rooms in two shifts (morning from 8 AM to 2 PM, and afternoon from 2 PM to 8 PM), and were rotated across all rooms to avoid creating habits with monitors and thus minimize monitor-specific effects. Monitors were present at the computer lab all the time and had to register any activity that happened in the computer rooms during their shifts. They were also asked to report each worker's time of arrival and departure every working day, and to assist workers with any technical difficulties. Additionally, they had to ensure that the computer labs were open on time every day. Monitors could not leave the rooms until everyone had left. On the first working day the monitors instructed the workers on how to use the software and all the details of the working schedule. The workers then started with their data entry tasks.

An ID code was assigned to each worker, together with a user name and password to log into the software created for the data entry process. The software displayed the image in a dialog box (chart or typing space), recorded the answer, the time spent on each image, the breaks reported by the worker, and the time she completed her tasks for the day. This software was created in Python and depicted images of characters (numbers) from a Chilean Agrarian Census with 350,000 images.<sup>17</sup>

Each image was randomly assigned with equal probability to approximately nine workers to determine the correct data entries, but on average all the images from the Census were typed at least seven times, because not all workers typed at the same speed and therefore did not type all the planned images. The following figure shows an example of what was displayed on the screen:

---

<sup>17</sup>The data entries were used for a research project of professor Nicolas Lillo from the Economics Department at Universidad Javeriana.

**Figure 1.** Task Sample



## 4 Data and descriptive statistics

This section describes the size of the different samples – from the pool of applicants to the final group of hired workers in our experiment – and the main variables used in the different steps of our analysis.

### 4.1 Selection process in the pool of applicants

Table 1 summarizes how the sample size decreased at each stage of the experiment. We received 686 job applications to the job ad. Only the 535 individuals who completed the three productivity tests were retained for the next steps in the application process. As mentioned above, individuals within each of the eight strata – gender, high/low *ex ante* productivity, caregiver status – were then randomly assigned to four different treatments: T1 – full-time non-flexible (control group), T2 – part-time non-flexible, T3 – full-time flexible, and T4 – part-time flexible.

In order to mimic a typical recruitment process, in step 4 another email – inquiring about motivation and additional information – was sent to job seekers who explicitly signaled an interest in the position and to those who did not click on the link (i.e.,  $438+86 = 524$  applicants). Only the 11 job seekers who were not interested in the position after the first email did not receive the second one. A total of 384 individuals responded to the second e-mail.

These 384 individuals were retained in step 5 of the application process, in which job offers were randomly sent out, again according to a stratified random sampling scheme. The initial aim was to hire 52 individuals – 13 in each type of contract. However, for reasons explained in Section 2.4, we did not attain this goal. Each applicant who did not accept the job offer in a particular treatment condition was replaced by another one randomly drawn from the same stratification cell. In the end, 79 job offers were sent out,

of which only 38 were accepted. Another four individuals never showed up or resigned before completing the contractual period.

Table 2 shows the distribution of offers and hired individuals across contract types. It shows that the standard schedule (i.e., full time, no flex) had a higher acceptance rate than the other options. The next section provides descriptive statistics for the 535 applicants who were randomly assigned to the four contract types in step 1 of the application process, and the 34 who were hired and completed the experiment. While the contracts lasted for 3 weeks, we focus only on the first 2 weeks because in the last week the workers were assigned a new task. This task was more complex and thus not comparable with the one performed during the first 2 weeks.<sup>18</sup>

**Table 1.** Samples at Each Stage of the Experiment

Sample	Experimental stages	N
0	Applicants who started the online survey	686
1	<i>First randomisation: assignment of contract types</i> Applicants who finished the survey and tests, & received 1 <sup>st</sup> email with the contract type	535
2	Applicants who stated interest after 1 <sup>st</sup> email containing the contract type	438
3	Applicants who stated interest after 2 <sup>nd</sup> email asking for new open questions	384
4	<i>Second randomisation: who gets an offer among each contract type</i> Applicants who received an offer	79
	Applicants who received an offer and accepted it, and were hired	38
	Applicants who took the job and finished it	34

**Table 2.** Job Offers and Acceptance Rate

	Offers	Acceptance Rate (%)	Workers	(%)
T1: <i>Full-time non-flex</i>	17	70.6	11	32.35
T2: <i>Part-time non-flex</i>	19	42.1	7	20.59
T3: <i>Full-time flex</i>	23	40.5	9	26.47
T4: <i>Part-time flex</i>	20	40	7	20.59
Total	79	100	34	100

<sup>18</sup>They were asked to digitize words from the census and not figures. The complexity relates to the quality of the pictures, which were more blurry, and therefore more difficult to type.

## 4.2 Descriptive statistics of applicants and hired workers

In this section, we provide descriptive statistics across treatment groups of the available explanatory variables for both the sample of applicants who took the productivity test and the hired workers who completed the experiment. We also check the extent to which the treatment groups are balanced within these samples. As the applicants are randomly assigned to the different treatments, this is a validity check of the randomization. Yet since each contract type will attract different types of workers, and as the eventual hiring of workers is the outcome of this self-selection process, it is not surprising that the treatment groups are not balanced across the sample of hired workers.

We use the standardized difference (SD) to evaluate whether the samples are balanced across treatment groups (Rosenbaum and Rubin, 1985). The advantage of the SD is that it is not sensitive to sample size. This is particularly convenient in our context, because the sample of hired workers is so small that conventional t-tests are not appropriate. We follow Rosenbaum and Rubin by considering an SD exceeding 0.20 to be large.



**Table 3.** Descriptive Statistics of Observables by Treatment Groups: Sample of Applicants

Variable	(T1)	(T2)	(T3)	(T4)	Standardized difference		
	Full time non-flexible Mean	Part time non-flexible Mean	Full time flexible Mean	Part time flexible Mean	(2)-(1)	(3)-(1)	(4)-(1)
<i>Stratification Variables</i>							
Female	0.69	0.63	0.67	0.65	-0.12	-0.04	-0.08
High Productivity	0.23	0.18	0.19	0.23	-0.12	-0.10	0.00
Dependents	0.24	0.20	0.19	0.19	-0.08	-0.12	-0.12
<i>Other Control Variables</i>							
<i>Age Groups</i>							
20- 25 yrs old	0.47	0.44	0.43	0.50	-0.06	-0.09	0.07
26- 30 yrs old	0.16	0.23	0.25	0.22	0.16	0.22	0.14
31 yrs old and more	0.36	0.33	0.32	0.27	-0.06	-0.09	-0.19
<i>Educational Levels</i>							
High School or less	0.26	0.29	0.32	0.24	0.05	0.12	-0.06
University	0.31	0.29	0.23	0.27	-0.04	-0.18	-0.09
Vocational	0.43	0.42	0.45	0.50	-0.01	0.05	0.13
N	140	132	150	113			

*Notes:* The first four columns of the Table present the mean value of the covariates in the corresponding treatment groups. The last three columns report the standardized differences defined as  $(\bar{X}_j - \bar{X}_{T1}) / \sqrt{[(Var(X_j) + Var(X_{T1})) / 2]}$ , where  $\bar{X}_j$  and  $Var(X_j)$  are the sample mean and the variance of  $X_j$  for  $j \in \{T2, T3, T4\}$ .

#### 4.2.1 The sample of applicants

In Table 3 we describe the sample of applicants. In the first four columns we report the means (and the associated SDs in parentheses) of the explanatory variables of the corresponding treatment groups. The last three columns report the SDs for treatments T2–T4 relative to treatment T1.

The profession of data clerk is a routine job that typically requires a medium level of education, as some familiarity with information and communication technologies is required. We indeed find that nearly half of the applicants have a vocational degree, which in Colombia is a technical education. Students who complete this level of education receive a diploma that is valued more than a high school degree, but less than a university degree. The temporary nature of the job explains why nearly half of the applicants are younger than 26 and about two-thirds are women (Castellani et al. 2020). We

also observe that 20 percent declared having dependents in the household, i.e. persons in need of care, and the same proportion were high-productivity types as measured by the pre-tests. As expected, these variables are balanced across treatment groups.<sup>19</sup> We only find an SD that exceeds 0.20 for the age group 26–30 when comparing T3 (full-time flexible) with the reference treatment T1: the share of applicants in T3 is 9 percentage points higher than the corresponding share in T1 (25 percent versus 16 percent). This difference is not a big concern: we will see in Section 6.4 that our findings are robust to control variables.

**Table 4.** Descriptive Statistics of Observables by Treatment Groups: Sample of Workers

Variable	(T1)	(T2)	(T3)	(T4)	Standardized difference		
	Full time non-flexible Mean	Part time non-flexible Mean	Full time flexible Mean	Part time flexible Mean	(2)-(1)	(3)-(1)	(4)-(1)
<i>Stratification Variables</i>							
Female	0.64	0.86	0.44	0.57	0.48	-0.38	-0.13
High Productivity	0.27	0.14	0.22	0.14	-0.30	-0.11	-0.30
Dependents	0.27	0.14	0.11	0.43	-0.30	-0.39	0.32
<i>Other Control Variables</i>							
<i>Age Groups</i>							
20- 25 yrs old	0.27	0.14	0.22	0.43	-0.30	-0.11	0.32
26 - 30 yrs old	0.36	0.43	0.11	0.43	0.13	-0.57	0.13
31 yrs old and more	0.36	0.43	0.67	0.14	0.13	0.59	-0.48
<i>Educational Level</i>							
High School or less	0.27	0.43	0.22	0.14	0.32	-0.11	-0.30
University	0.45	0.43	0.22	0.14	-0.05	-0.47	-0.64
Vocational	0.27	0.14	0.56	0.71	-0.30	0.56	0.86
N	11	7	9	7			

*Notes:* The first four columns of the Table present the mean value of the covariates in the corresponding treatment groups. The last three columns report the standardized differences defined as  $(\bar{X}_j - \bar{X}_{T1}) / \sqrt{[(Var(X_j) + Var(X_{T1})) / 2]}$ , where  $\bar{X}_j$  and  $Var(X_j)$  are the sample mean and the variance of  $X_j$  for  $j \in \{T2, T3, T4\}$ .

#### 4.2.2 The sample of hired workers

Table 4 describes the sample of hired workers. As expected, the variables are now imbalanced between the treatment groups, as the majority of the SDs are larger than

<sup>19</sup>It is noteworthy that the *number* of individuals assigned to the different treatments is not completely balanced. This is because we realized only after the random assignment that some individuals we retained only completed part of the productivity test; we dropped these individuals from the analysis. As demonstrated in the table, this did not result in any major selectivity in the sample.

0.20. This does not necessarily imply that these imbalances are reflected in different productivity levels. In the analysis below, we will extract the sorting that is related to productivity by controlling for an *ex ante* measure of productivity that we derive from the test implemented in the second step of our experimental design (see Section 3.1 above). As we will explain in more detail in Section 4.3, this *ex ante* productivity measure is chosen to match as closely as possible the definition of *ex post* on-the-job productivity that we use in the analysis to evaluate the productivity differences generated by the different contractual arrangements. Note that this *ex ante* productivity measure differs from the one used to construct the stratification variable “high productivity,” since we wanted to match as closely as possible the measures of *ex ante* and *ex post* productivity. Recall that the productivity tests performed during the application process did not record the time taken to perform the tests. This may partly explain why in Table 4 the differences between treatment groups in the fraction classed as “highly productive” do not fully reflect the differences that the *ex ante* productivity measure captures in the analysis reported below.

We expected *a priori* that flexible work arrangements would especially attract women and individuals with caring responsibilities. We indeed observe that the most flexible contract (T4) attracted the highest share of individuals with dependents: 0.43 is more than twice as high as the average share in the sample. However, this flexible arrangement does not contain the highest share of women hired. The share of women is highest in the part-time non-flexible arrangement (T2), which contains a below-average share with dependents. Finally, the full-time flexible arrangement (T3) seems to attract relatively more women and individuals without caring responsibilities than the other types of contracts. We do not have a clear explanation for these deviations from our expectations. It might be related to the small sample size. We do not describe the sorting patterns across the treatments for the other explanatory variables, because we do not have particular *a priori* hypotheses regarding their direction.

### 4.3 The outcome variables and their decomposition

Our main aim is to determine the impact of working-time flexibility on labor productivity. We first discuss how we measure the *ex post* productivity of hired workers and how we break this measure down into different components. This decomposition allows us to explore the mechanisms of potential productivity differences. Next, we define a

measure of *ex ante* productivity and its decomposition. In the analysis reported in Sections 5 and 6, this *ex ante* measure is used to disentangle the selection from the *ex post* motivational effects.

### 4.3.1 *Ex post* productivity

Different measures of productivity could be considered. In a private sector environment, a natural measure of productivity would be the value of production per unit of contractual working time. In the current context, the value of production is not known, but since the data entries served as an input for research purposes it is natural to value accuracy. The need to type accurately was also stressed when instructions were provided on the first working day. We therefore code each correctly typed character as 1 and each incorrectly typed character as 0. This results in the following definition of the productivity level of worker  $i$  in time period  $t$ :

$$AP_{it,0} \equiv C_{it}/T_{it} \tag{2}$$

where  $C_{it}$  is the number of correct characters that individual  $i$  typed within the contractual time  $T_{it}$  of period  $t$ . In the analysis below we consider that  $t \in \{1,2\}$  for both working weeks included in the analysis. For a full-time worker  $T_{it} = 40$  hours, while for a part-time worker  $T_{it} = 20$  hours.

This measure of productivity requires us to determine what is correct and incorrect. We determined a correctly typed character to be equal to the mode of all typed characters for each corresponding character image.<sup>20</sup> Based on a 1 percent sub-sample in which optical character recognition software was used, a 99 percent correspondence was found with our measure of correctness based on the mode.

To explore the mechanisms that generate differences in productivity between types of contracts, we define our productivity measure as follows:

$$AP_{it,0} \equiv C_{it}/T_{it} = (C_{it}/N_{it}) \times (N_{it}/D_{it}) \times (D_{it}/T_{it}) \equiv \prod_{j=1}^3 AP_{it,j} \tag{3}$$

where  $N_{it}$  is the total number of characters typed within the contractual time  $T_{it}$  and  $D_{it}$

---

<sup>20</sup>In a sensitivity analysis we used the median instead of the mode, but this did not change our findings in any way. This analysis is available from the authors upon request.

is the *actual* working time. The actual working time differs from the *contractual* working time in that workers do not work all the time. They can take *breaks* while remaining in the lab (i.e., coffee break, going to the bathroom, or just a period in which they do not type). Alternatively, they can leave the lab, which we label *absenteeism*. The actual working time is determined by discarding these periods of interruption (breaks plus absenteeism) from the contractual time. We cannot directly observe these interruptions in working time, but the software registered each moment when a character was typed. In the benchmark analysis we defined a period of interruption as any period that lasted for more than 15 seconds between two subsequent instants when characters are typed.

The above equation breaks down average productivity ( $AP_{it,0}$ ) into a measure of *precision* ( $C_{it}/N_{it}$ ), i.e. the fraction of correctly typed images, a measure of *speed* ( $N_{it}/D_{it}$ ), i.e. the number of images typed per unit of actual working time, and a measure of *actual* working time ( $D_{it}/T_{it}$ ), i.e. the fraction of contracted time that the worker actually worked.

Work interruptions can be further decomposed into breaks and absenteeism. However, this decomposition is additive, so it cannot be integrated into the multiplicative decomposition above. We therefore consider a separate additive decomposition in which we use the sum of fractions relative to contractual time ( $T_{it}$ ), actual working time ( $D_{it}$ ), breaks ( $B_{it}$ ), and periods of absenteeism ( $A_{it}$ ) is equal to 1:

$$D_{it}/T_{it} + B_{it}/T_{it} + A_{it}/T_{it} = 1 \quad (4)$$

### 4.3.2 *Ex ante* productivity

The measures defined in the previous subsection relate to *ex post* productivity, which was measured after the worker was hired. In Section 5 we evaluate whether flexible contract environments attract more productive workers, and in Section 6 we try to disentangle sorting from *ex post* motivational effects. To test this hypothesis, we use the answers from the productivity test the applicants took during the first step of the application process to construct the following average *ex ante* productivity measure:

$$AP_{i,0}^A = C_i^A / D_i^A \quad (5)$$

where the superscript  $A$  denotes that the figures refer to the *ex ante* productivity test,  $C_i^A$  is the number of correct characters that applicant  $i$  typed in all three tasks of the

productivity test, and  $D_i^A$  is the actual time that elapsed between typing the first and last characters in the test. We denote this time by  $D$  rather than  $T$ , because it is conceptually closer to  $D_{it}$  in the aforementioned decomposition formula of *ex post* productivity.

As for the *ex post* productivity measure, we can also decompose the *ex ante* measure. However, since we cannot measure the contractual time  $T$  *ex ante*, we break it down into two terms instead of three:

$$AP_{i,0}^A \equiv C_i^A / D_i^A = \left( C_i^A / N_i^A \right) \times \left( N_i^A / D_i^A \right) \equiv AP_{i,1}^A \times AP_{i,2}^A \quad (6)$$

where  $N_i^A$  is the total number of characters typed in the productivity test. This decomposes the total *ex ante* productivity into a measure of *precision* and *speed*.

## 5 *Ex ante* productivity and sorting into flexible jobs

### 5.1 Empirical strategy

One of the aims of this study is to determine the extent to which working flexibility can enhance productivity by attracting more productive workers. In order to test this we estimate the following model for  $j \in \{0, 1, 2\}$  (i.e. for the three *ex ante* productivity measures defined above):

$$\ln[AP_{i,j}^A] = \lambda + \sum_{k=2}^4 \alpha_k T_k + \eta \text{Accept}_i + \sum_{k=2}^4 \delta_k T_k * \text{Accept}_i + X_i' \beta + \epsilon_i \quad (7)$$

$T_k$  denotes the contract type  $k$  to which individual  $i$  was randomly assigned;  $T_2$  is a dummy coded 1 if individual  $i$  was assigned to a part-time non-flexible schedule,  $T_3$  to a full-time flexible schedule, and  $T_4$  to a part-time flexible schedule.  $T_1$  denotes a full-time non-flexible contract (reference contract).  $\text{Accept}_i$  equals 1 if individual  $i$  received a job offer and accepted it, and 0 otherwise. The coefficients of interest are thus  $\alpha_k$  for  $k \in \{2, 3, 4\}$  associated with the interaction terms. Whenever these coefficients are significantly different from zero they signal a differential acceptance rate of contract types  $k \in \{2, 3, 4\}$  relative to contract type 1, i.e. relative to the full-time non-flexible contract.  $X_i$  denotes the set of seven stratification variables. To enhance precision, the

other explanatory variables reported in Table 3 such as educational level and age.<sup>21</sup> Finally,  $\epsilon_i$  denotes the zero-mean error of the regression.

## 5.2 Results

Table 5 reports the main estimation results for testing whether more productive workers sort into more flexible jobs. Column (1) reveals that job applicants who sort into full-time flexible contracts are nearly 50 percent ( $\exp(.39) = 1.48$ ) more productive than those who eventually accept a full-time non-flexible contract. This effect is significant at the 5 percent level. The point estimate of  $\alpha_2$  is close to zero and far from statistical significance, suggesting that a part-time contract without time-scheduling flexibility does not attract more productive workers than the benchmark contract. By contrast, the point estimate for  $\alpha_4$  is close to that of  $\alpha_3$ , which suggests that contracts with time-scheduling flexibility attract more productive workers but the coefficient is imprecisely estimated. We therefore find firm evidence of a positive sorting effect only for full-time flexible contracts.

The decomposition analysis depicted in Columns (2) and (3) suggests that precision and speed contribute roughly equally to the effect of full-time flexible contracts, but the large standard errors of these estimates do not allow us to make a strong conclusion.

---

<sup>21</sup>This specification implicitly assumes that the treatment effect is constant across strata (see e.g., Imbens and Rubin, 2015, Section 9.6)

**Table 5.** Test of Sorting into Flexible Jobs along *Ex Ante* Productivity

<i>Ex ante</i> Productivity measures (ln)	Productivity # Correct/ Time (1)	Precision #Correct / Questions (2)	Speed # Questions / # Time (3)
$T_2$ : Part time non-flexible	-.01 (.07)	-.03 (.04)	.01 (.06)
$T_3$ : Full time flexible	-.02 (.07)	-.05 (.04)	.03 (.06)
$T_4$ : Part time flexible	.05 (.07)	.01 (.04)	.04 (.06)
<i>Accept</i>	-.21 (.15)	-.12 (.10)	
$T_2 \times Accept$	.00 (.41)	-.11 (.31)	.11 (.16)
$T_3 \times Accept$	.39** (.20)	.20 (.12)	.20 (.17)
$T_4 \times Accept$	.36 (.23)	.07 (.12)	.29 (.22)
Constant	-2.61*** (.07)	-.26*** (.04)	-2.36*** (.06)
R Squared	.16	.051	.16
N	535	535	535

*Notes:* This Table presents the coefficients from regressing the *ex-ante* measures of productivity on treatment dummies interacted with a dummy of acceptance of the job offer. Column (1) uses total measure of productivity, Column (2) uses a measure for *Precision* and Column (3) uses a measure for *Speed*. Huber-White robust standard errors in parenthesis. We add stratification dummies as controls in all the specifications. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ . In Table 10 in the Appendix the full estimation results are reported including the estimated coefficients of the control variables.

## 6 *Ex post* productivity and working-time flexibility

### 6.1 Empirical strategy

In this section we analyze how different flexible working arrangements affect on-the-job productivity. We recorded information about the same data entry task during two of the three contracted weeks using the different measures of productivity described in Section 3.3. In a first step we use the logarithm of the benchmark average productivity



measure as the outcome of interest, i.e.  $\ln(AP_{it,0}) \equiv \ln(C_{it}/T_{it})$ . In subsequent steps we take into account the decomposition in equation (3), which becomes additive after the logarithm is taken, and the decomposition in equation (4), for which no logarithms can be considered, as it is already additive. In each of those analyses, we consider both the case with and without controlling for measures of *ex ante* productivity. When we perform the analysis without controlling for *ex ante* productivity, we measure the overall impact of contract flexibility on the chosen *ex post* productivity measures, combining the selection effect and the behavioral effects. By controlling for the *ex ante* productivity measures, we seek to estimate the behavioral impact of contract flexibility. We jointly control for two measures of *ex ante* productivity: the logarithm of the overall *ex ante* productivity measure ( $AP_{i,0}^A \equiv C_i^A/D_i^A$ ) and the logarithm of the component measuring precision ( $AP_{i,1}^A \equiv C_i^A/D_i^A$ ).<sup>22</sup>

The benchmark regression without controlling for *ex ante* productivity can be written as follows:

$$\ln(Y_{it}) = \beta_1 + \beta_2 T_2 + \beta_3 T_3 + \beta_4 T_4 + \delta S_i + \lambda_2 I_{t2} + \epsilon_{it} \quad (8)$$

for  $t \in \{1,2\}$  where  $Y_{it}$  denotes one of the considered measures of productivity for individual  $i$  in week  $t$ ,  $I_{t2} = 1$  if  $t = 2$  and  $I_{t2} = 0$  if  $t = 1$ , so that  $\lambda_2$  accounts for the differential productivity in week 2 relative to week 1, the reference.  $T_j$  for  $j \in \{1, \dots, 4\}$  and  $S_i$  are, respectively, indicators of the contract type and sampling strata, and  $\epsilon_{it}$  denotes the zero-mean error of the regression.

Estimating regression equation (8) entails two important challenges with respect to the inference of the parameters of interest. First, given the small sample size, we cannot rely on the standard asymptotic inference to test our hypotheses. We therefore use Fisher's exact inference to obtain the P-values of the considered null hypotheses (Imbens and Rubin, 2015). A second potential problem of inference is that we must take into account the fact that productivity is correlated between the two weekly observations for each individual in the sample. A standard solution to this problem is to cluster the standard errors by individual. However, if there are only a small number of clusters, as is the case here, the standard errors are at risk of being biased downwards (Cameron and Miller, 2015). By relying on Fisher's exact inference, we can resolve this issue as well (Heß, 2017). Below we report the standard asymptotic standard errors for all coefficients,

---

<sup>22</sup>The logarithm of the measure of speed is, according to equation (6), linearly dependent on the two other measures, and is therefore redundant as an additional control variable.

and the exact p-values for only the coefficients of interest. The reported significance of coefficients will be based on the latter, if available.

## 6.2 Main results

Table 6 reports the estimated effects of the different forms of contract flexibility in the benchmark model, equation (8). Column (1) corresponds to the baseline regression including only the stratification variables and the week fixed effects. In the other specifications, we augment the baseline regression with the measures from the *ex ante* productivity test conducted in the selection phase.

Considering the estimated parameters of interest in column (1), we find that only the full-time flexible contract has a significant effect (exact P-value of 6 percent) on the benchmark productivity measure. Workers hired with this contract are nearly 50 percent ( $\exp(.40) = 1.49$ ) more productive than full-time non-flexible workers. For part-time workers under both flexible and non-flexible schedules, the effects on productivity are both positive, as expected, but these effects are not statistically different from that of full-time non-flexible workers.

The effects encompass both the behavioral and the selection effect of contract schedule flexibility. When controlling for selection effects, we find that there remains a role for a behavioral effect. According to the point estimate, productivity still increases relative to a full-time non-flexible contract by 33 percent ( $\exp(.28) = 1.32$ ). This effect is significant at the 10 percent level. This suggests that the behavioral effect represents about 60 percent of the total effect and the selection effect about 40 percent.

The effects of a part-time flexible and non-flexible contract are both positive, but much smaller and not significantly different from zero. We therefore conclude that there is no evidence that the flexibility of part-time work enhances productivity, even if it is combined with a flexible working schedule.

**Table 6.** Treatment Effects on Ex-Post Productivity

<i>Productivity (ln)</i>	(1)	(2)
T2: Part time non-flexible	.12 (.27) [.66]	.18 (.18) [.37]
T3: Full time flexible	.40* (.15) [.06]	.28* (.16) [.10]
T4: Part time flexible	.15 (.22) [.17]	.09 (.23) [.43]
<i>Ref: T1 Full-time non-flexible</i>		
Constant	-1.73*** (.11)	-1.64*** (.49)
R Squared	.2	.52
N	68	68

*Notes:* This Table reports the coefficients of estimating the treatment effect on the log of total productivity, measured as the number of correct digits over the contracted time – two weeks for individual level regressions, one week for week level regressions. Standard errors in parentheses are clustered at the individual level. Exact p-value for the Fisher Tests in brackets. This exact p-value is the proportion of possible treatment assignments that yield a test statistic greater than or equal to the observed test statistics. We used 3000 permutations. Note that \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$  correspond to significance levels according to Fisher p-values for T2, T3, and T4 coefficients. For the rest of coefficients, stars reflect the p-values of the standard asymptotic inference. Stratification dummies are included in all columns. Column (2) controls for the log of *ex-ante* productivity - *Precision* (number of correct answers over total questions) and for the log of *ex-ante total* productivity (number of correct answers over total time spent to answer the questions). In Table 12 in the Appendix the full estimation results are reported including the estimated coefficients of the control variables.

### 6.3 Exploring mechanisms

We explore the mechanisms of the working arrangements on productivity by breaking down the benchmark productivity measure into its components. First, we consider the logarithmic additive form of the decomposition displayed by equation (3). Next, we consider the additive decomposition displayed in equation (4).

### 6.3.1 Decomposing the effects on global productivity

Columns (1) to (4) of Table 7 present the estimates of the first decomposition of the *global* effects on productivity, which comprise both the sorting effects generated by the application process and the behavioral on-the-job effects. In columns (5) to (8) the same decomposition is displayed after controlling for the *ex ante* productivity or sorting effects. Columns (1) and (5) – denoted by "Total" – restate the total effect on productivity reported in Table 6. The subsequent columns present the coefficients of the decomposition.

**Table 7.** Decomposition of the Total Effects on Ex-Post Productivity

	Global- Productivity				Without Sorting Effects			
	Total (C/T) (1)	Precision (C/N) (2)	Speed (N/D) (3)	Effective Time (D/T) (4)	Total (C/T) (5)	Precision (C/N) (6)	Speed (N/D) (7)	Effective Time (D/T) (8)
T2: Part-time non flexible	.12 (.27) [.66]	-.08 (.10) [.52]	.00 (.14) [.98]	.20 (.12) [.13]	.18 (.18) [.37]	-.07 (.05) [.16]	-.00 (.10) [.98]	.26* (.13) [.10]
T3: Full-time flexible	.40* (.15) [.06]	.04 (.02) [.63]	.14 (.08) [.23]	.22** (.12) [.05]	.28* (.16) [.09]	-.01 (.04) [.85]	.04 (.10) [.66]	.25** (.11) [.04]
T4: Part-time flexible	.15 (.22) [.17]	.03 (.02) [.38]	.12 (.11) [.12]	.00 (.17) [1.00]	.09 (.23) [.43]	-.01 (.04) [.71]	.03 (.13) [.43]	.07 (.13) [.62]
<i>Ref: T1 Full-time non-flexible</i>								
Constant	-1.73*** (.11)	-.06*** (.02)	-.80*** (.05)	-.87*** (.08)	-1.64*** (.49)	.09 (.10)	-.25 (.33)	-1.48*** (.37)
R Squared	.2	.17	.29	.22	.52	.64	.54	.35
N	68	68	68	68	68	68	68	68

*Notes:* This table presents the treatment effects on four outcome variables. Columns (1) and (5) corresponds to the log of *Total productivity*, measured as the number of correct digits over the contracted time. *Precision*, columns (2) and (6), is the log of the number of correct digits over the total digits typed. *Speed*, columns (3) and (7), is the log of the number of total digits over working time. Finally, effective time, columns (4) and (8), is the log of working time over contracted time. Standard errors in parentheses are clustered at the individual level. Exact p-value for the Fisher Tests in brackets. This exact p-value is the proportion of possible treatment assignments that yield a test statistic greater than or equal to the observed test statistics. We used 3000 permutations. Note that \*\*\* p<0.01, \*\* p<0.05, \* p<0.1 correspond to significance levels according to Fisher p-values for T2, T3, and T4 coefficients. For the rest of coefficients, stars reflect the p-values of the standard asymptotic inference. Stratification dummies are included in all columns. Columns 5 to 8 control for the log of *ex-ante* productivity - *Precision* and for the log of *ex-ante* total productivity. Table 13 in the Appendix the full estimation results are reported including the estimated coefficients of the control variables.

We first focus on the findings for the workers on a full-time flexible contract, as we only find a significantly higher productivity for this contract type relative to full-time non-flexible contracts. The breakdown suggests that the *behavioral* component, which represents about 60 percent of the total effect on productivity, is completely explained by more effective time use in the workplace. This follows from the fact that when sorting effects are controlled for, the estimated coefficient explaining the effect on effective time use (0.25) is nearly as large as the one explaining the total productivity effect (0.28), while the coefficients that measure the reaction due to *precision* and *speed* are close to zero and not statistically significant. By contrast, the selection process seemed to have attracted workers who have the ability to type both faster and more precisely, as the coefficients associated with these components are larger for speed ( $0.14 > 0.04$ ) or statistically different from zero for precision (0.04) only when sorting is not controlled for. However, based on the findings, it is difficult to disentangle whether precision or speed matters the most, because the coefficient related to precision is statistically significant, but small, while the coefficient related to speed is large, but statistically insignificant.

Another interesting result is that workers on part-time non-flexible schedules also seem to demonstrate a more effective use of time than those in the comparison group (columns (4) and (8)), despite no observed differences in total productivity (columns (1) and (5)), and potentially at the loss of precision, though this loss is not significant (columns (2) and (6)). These effects should be interpreted as behavioral effects rather than sorting effects, as the corresponding coefficients are of similar magnitude regardless of whether sorting effects are controlled for or not.

The results from these specifications therefore provide evidence that working in more flexible environments makes individuals work more effectively, reducing their shirking time at work. The productivity of workers on full-time flexible contracts is also driven by the application process, which sorts workers who type more precisely per unit of time, but the analysis does not clearly indicate whether speed or precision dominates in this sorting process.

Our finding that the motivational effects of working-time schedules are largely driven by an increased number of working hours is in line with the findings of [Beckmann et al. \(2017\)](#), who uses the German Socio-Economic Panel. However, our results seem to contradict those of [Dutcher \(2012\)](#), who reports that workplace flexibility has negative effects on productivity for a similar “dull” data entry task as the one in our study. A potential explanation is that workplace flexibility implies not only working-time flexi-

bility, but also that the work is not monitored. For routine tasks, monitoring may be an essential device to keep workers committed to their work, as intrinsic motivation may be lacking (Beckmann et al., 2017). Yet workers may still want to reciprocate the provided working-time autonomy by increasing their effort (Ibid.; Kelliher and Anderson (2010)). Another reason why our results may differ from those of Dutcher (2012) is that we hired real workers instead of students.

Our findings suggest that part-time workers on non-flexible contracts also reciprocate by working more, but this is partly counterbalanced by a lack of precision. The latter is consistent with both the hypothesis that part-time workers are less effective than their full-time counterparts because entering data typically involves making more mistakes at the start (i.e. higher fixed start-up costs), and that part-time contracts attract less committed workers, a sorting effect for which we cannot control using our *ex ante* measure of productivity (Künn-Nelen et al. (2013); Garnero (2016)). However, this does not explain why we do not find similar or stronger positive effects on effective working time for workers on *flexible* part-time contracts. The decomposition of the effect on working time discussed in the next subsection provides more insights into this puzzle.

### 6.3.2 Decomposing the effects on time use

Table 8 presents the mean differences of the various uses of time across treatments for the global measure of productivity (columns (1) to (3)) and for *ex post* productivity after controlling for the sorting effect (columns (4) to (6)). We estimate each component from equation 8 in order to evaluate which mechanism makes workers more productive at work, as a percentage of total contracted time. The additive decomposition allows us to disentangle the global effect of productivity on effective time  $D/T$  (worked time as a percentage of contracted time), absenteeism as a percentage of contracted time  $A/T$ , and time taken for breaks as a percentage of contracted time  $B/T$ .

We observe that selection effects play no role, as the effects on *ex post* productivity are not affected when we control for *ex ante* productivity. This is in line with our expectations, because working time refers to a behavioral rather than a selection effect. The results also show that the positive productivity effects reported in the previous section are induced by the fact that workers with flexible work contracts take fewer breaks rather than being less absent.<sup>23</sup>

---

<sup>23</sup>Recall that the effects are now measured as percentage-point effects relative to the contractual time,

A new result is that part-time workers with flexible contracts (T4) are found to reduce their break times by as much as 14 percentage points of the contractual working time, a reduction that is similar to that of part-time workers on non-flexible contracts (T2), and nearly twice as large as the effect for full-time workers on flexible contracts (T3). These findings suggest that part-time contracts *do* have motivational effects, but they do not translate into higher global productivity for various reasons. For part-time workers on non-flexible contracts, this positive effect was tempered by the lower precision with which they enter data. For those on flexible contracts, this positive effect is largely absorbed by higher rates of absenteeism. A potential explanation is that workers with more caring responsibilities sort much more into these kinds of contracts than in other ones (see Section 4.2.2). As the timing of care is very uncertain, they may be forced to be absent more than other workers.

## 6.4 Robustness checks

We conduct two main robustness tests on our main results: one for the inclusion of additional controls and another for the presence of outliers. In the model that tests for sorting by *ex ante* productivity (Table 5), we control for the stratification variables and for the additional control variables that were reported in Table 3. In Appendix Table 10 we report the results if these additional controls are not included in the regression; the point estimates and standard errors are hardly affected. This confirms that there need not be any concern that, conditional on the strata, the treatments in the sample of applicants is not balanced after the random assignment.

We did not perform the same sensitivity analysis for the sample of hired workers. The reason is that we do not expect this sample to be balanced across treatment conditions, because we anticipate that different work arrangements have a different level of attractiveness for different applicants. To the extent that these differences are correlated with workers' observed characteristics as well as with *ex post* productivity, introducing these characteristics as additional control variables in Equation (8) may therefore affect the results reported in Section 6.2 without invalidating them.

Second, we estimate a robust-to-outliers regression, which is a weighted least squares regression that weighs the observations differently based on absolute residuals (Verardi and Croux, 2009). Column (2) of Table 9 displays a somewhat smaller effect of full-  
whereas in the previous sections proportional effects were measured on a logarithmic scale.

**Table 8.** Decomposition of the Time Use Effects

	Global			Without Sorting Effects		
	Effective Time (D/T) (1)	Absenteeism (A/T) (2)	Breaks (B/T) (3)	Effective Time (D/T) (4)	Absenteeism (A/T) (5)	Breaks (B/T) (6)
T2: Part time non-flexible	.10 (.06) [.12]	.04 (.04) [.33]	-.14*** (.04) [.00]	.12* (.06) [.10]	.03 (.04) [.55]	-.15*** (.04) [.01]
T3: Full time flexible	.09* (.05) [.08]	-.01 (.04) [.85]	-.08** (.03) [.03]	.10* (.05) [.07]	-.01 (.03) [.73]	-.08** (.03) [.04]
T4: Part time flexible	.02 (.07) [.60]	.11 (.05) [.33]	-.14*** (.04) [.00]	.05 (.06) [.18]	.10 (.03) [.44]	-.14*** (.04) [.00]
<i>Ref: T1 Full time non-flexible</i>						
Constant	.43*** (.03)	.05* (.03)	.51*** (.02)	.22 (.15)	.20** (.10)	.58*** (.09)
R Squared	.23	.26	.46	.35	.32	.48
N	68	68	68	68	68	68

*Notes:* This table presents the treatment effects on three outcome variables. *Effective time*, columns (1) and (4), corresponds working time over contracted time. *Absenteeism*, columns (2) and (3), is the time outside the lab over the contracted time. Finally, *Breaks* is measured as the non-working time inside the lab over the contracted time. Standard errors in parentheses are clustered at the individual level. Exact p-value for the Fisher Tests in brackets. This exact p-value is the proportion of possible treatment assignments that yield a test statistic greater than or equal to the observed test statistics. We used 3000 permutations. Note that \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$  correspond to significance levels according to Fisher p-values for T2, T3, and T4 coefficients. For the rest of coefficients, stars reflect the p-values of the standard asymptotic inference. Stratification dummies are included in all columns. Columns 4 to 6 control for the log of *ex-ante* productivity. In Table 14 in the Appendix the full estimation results are reported including the estimated coefficients of the control variables.

time flexible contracts on total productivity ( $\exp(0.34) = 1.40$ ), but the effect remains statistically significant at the 5 percent level (Fisher exact p-value). When controlling for sorting effects (column (4)), the coefficient is slightly reduced and the p-value increases slightly from 10 percent to 12 percent.



**Table 9.** Robustness Checks

	Robustness			
	Global			Without Sorting
	(1)	(2)	(3)	(4)
<i>Productivity (ln)</i>				
<i>T</i> <sub>2</sub> Part-time no flexible	.12 (.27) [.66]	.29 (.15) [.19]	.17 (.19) [.37]	.18 (.15) [.40]
<i>T</i> <sub>3</sub> Full-time flexible	.40** (.15) [.06]	.34** (.14) [.05]	.28* (.16) [.10]	.28 (.15) [.12]
<i>T</i> <sub>4</sub> Part-time flexible	.15 (.22) [.17]	.08 (.15) [.57]	.09 (.23) [.61]	.07 (.16) [.70]
<i>Ref: T</i> <sub>1</sub> Full-time non-flexible				
<i>Ex ante</i> productivity - Precision			.89*** (.14)	.90*** (.22)
<i>Ex ante</i> total productivity			-0.08 (.21)	-.06 (.15)
Constant	-1.73*** (.11)	-6.89 (5.03)	-1.64*** (.49)	-6.50 (4.85)
R Squared	.2	.2	.52	.48
N	68	68	68	68

Notes: the outcome variable is the log measure of productivity, measured as the number of correct digits over the contracted time. Standard errors in parentheses are clustered at the individual level. Exact p-value for the Fisher Tests in brackets. This exact p-value is the proportion of possible treatment assignments that yield a test statistic greater than or equal to the observed test statistics. In this case, we used 3000 permutations. Note that \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$  correspond to significance levels according to Fisher p-values for *T*<sub>2</sub>, *T*<sub>3</sub>, and *T*<sub>4</sub> coefficients. For the rest of coefficients, stars reflect the p-values of the standard asymptotic inference. Strata variables are included in all columns.

## 7 Conclusion

We conducted a field experiment to test whether working-time flexibility affects the productivity of workers hired for a temporary routine job. A novelty of our study is that we randomized job applicants rather than employees into different contractual environments, and tested the productivity of these job seekers at the start of the application process before assigning them into these different environments. This allowed us to decompose the total effect of working-time flexibility into sorting and behavioral effects.

We found that both part-time work and working-time flexibility substantially reduced working time interruptions. However, for part-time workers this did not result in a statistically significant globally higher level of productivity, either because these workers

also made more mistakes or because they were more absent from work. By contrast, for full-time workers who were given the autonomy to choose when to start and stop working, this resulted in more effective working time, which in turn significantly increased their productivity by as much as 30 percent. Furthermore, the working-time autonomy enhanced the attractiveness of the job for intrinsically more productive workers: their eventual productivity was estimated to be as much as 50 percent higher than that of workers hired on a non-flexible full-time contract.

The finding that working-time flexibility can also enhance productivity in routine jobs is new, and may provide valuable advice for human resource strategies. Nevertheless, our findings cannot be generalized to all types of jobs since our research design considered a very particular job in which very specific routine tasks had to be executed. Future research could scale up the experiment and extend it to other (routine) tasks and working environments.

We identified sorting effects based on a measure of intrinsic productivity. However, we have seen that sorting effects could be related to other intrinsic characteristics of applicants, such as commitment or caring responsibilities for dependents, which in turn may generate behavioral effects. These examples reveal that sorting and behavioral effects are often very much intertwined. Disentangling these effects may therefore often be a matter of interpretation. While this suggests promising avenues of research, we believe the decompositions of the kind that we proposed here can improve our understanding of why certain flexible working arrangements are effective (or not) at enhancing workers' productivity.

## References

- Angelici, Marta and Paola Profeta**, “Smart-Working: Work Flexibility without Constraints,” Technical Report 2020.
- Beckmann, Michael**, “Working-time autonomy as a management practice,” *IZA World of Labor*, 2016.
- , **Thomas Cornelissen, and Matthias Kräkel**, “Self-managed working time and employee effort: Theory and evidence,” *Journal of Economic Behavior & Organization*, 2017, 133, 285 – 302.
- Bertrand, Marianne**, “Coase lecture—the glass ceiling,” *Economica*, 2018, 85 (338), 205–231.
- Bloom, Nicholas, James Liang, John Roberts, and Zhichun Jenny Ying**, “Does working from home work? Evidence from a Chinese experiment,” *The Quarterly Journal of Economics*, 2015, 130 (1), 165–218.
- BLS**, “Job Flexibilities and Work Schedules Summary,” Technical Report, Bureau of Labor Statistics 2019.
- Bustelo, Monserrat, Ana María Díaz Escobar, Jeanne Lafortune, Claudia Piras, Luz Magdalena Salas Bahamón, José Tessada et al.**, “What is The Price of Freedom?: Estimating Women’s Willingness to Pay for Job Schedule Flexibility,” Technical Report, Inter-American Development Bank 2020.
- Cameron, Colin and Douglas Miller**, “A Practitioner’s Guide to Cluster-Robust Inference,” *Journal of Human Resources*, 2015, 50 (2), 317–372.
- Castellani, Francesca, Giulia Lotti, and Nataly Obando**, “Fixed or open-ended? Labor contract and productivity in the Colombian manufacturing sector,” *Journal of Applied Economics*, 2020, 23 (1), 199–223.
- Chen, Kuan-Ming, Claire Ding, John A List, and Magne Mogstad**, “Reservation Wages and Workers’ Valuation of Job Flexibility: Evidence from a Natural Field Experiment,” Technical Report, National Bureau of Economic Research 2020.
- Chen, M Keith, Peter E Rossi, Judith A Chevalier, and Emily Oehlsen**, “The value of flexible work: Evidence from uber drivers,” *Journal of Political Economy*, 2019, 127 (6), 2735–2794.

- Devicienti, Francesco, Elena Grinza, and Davide Vannoni**, “The impact of part-time work on firm productivity: evidence from Italy,” *Industrial and Corporate Change*, 2018, 27 (2), 321–347.
- Dutcher, E Glenn**, “The effects of telecommuting on productivity: An experimental examination. The role of dull and creative tasks,” *Journal of Economic Behavior & Organization*, 2012, 84 (1), 355–363.
- Eaton, Susan C**, “If you can use them: Flexibility policies, organizational commitment, and perceived performance,” *Industrial Relations: A Journal of Economy and Society*, 2003, 42 (2), 145–167.
- Eurofound**, “Sixth European Working Conditions Survey. Luxembourg: Publications Office of the European Union,” 2015.
- Garnero, Andrea**, “Are part-time workers less productive and underpaid?,” *IZA World of Labor*, 2016.
- , **Stephan Kampelmann, and François Rycx**, “Part-time work, wages, and productivity: evidence from Belgian matched panel data,” *ILR Review*, 2014, 67 (3), 926–954.
- Goldin, Claudia**, “A grand gender convergence: Its last chapter,” *American Economic Review*, 2014, 104 (4), 1091–1119.
- He, Haoran, David Neumark, and Qian Weng**, “Do Workers Value Flexible Jobs? A Field Experiment on Compensating Differentials,” Technical Report, *Journal of Labor Economics - forthcoming* 2020.
- Heß, Simon**, “Randomization Inference with Stata: A Guide and Software,” *The Stata Journal*, 2017, 17 (3), 630–651.
- ILO**, “Promoting employment and decent work in a changing landscape,” Technical Report, ILO - *forthcoming* 2020.
- Imbens, Guido W. and Donald B. Rubin**, *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*, Cambridge University Press, 2015.
- Kelliher, Clare and Deirdre Anderson**, “Doing more with less? Flexible working practices and the intensification of work,” *Human relations*, 2010, 63 (1), 83–106.

- Künn-Nelen, Annemarie, Andries De Grip, and Didier Fouarge**, “Is part-time employment beneficial for firm productivity?,” *ILR Review*, 2013, 66 (5), 1172–1191.
- Mas, Alexandre and Amanda Pallais**, “Valuing alternative work arrangements,” *American Economic Review*, 2017, 107 (12), 3722–59.
- OECD, Be Flexible**, “Background brief on how workplace flexibility can help European employees to balance work and family,” 2016.
- Rosenbaum, Paul R and Donald B Rubin**, “Constructing a control group using multivariate matched sampling methods that incorporate the propensity score,” *The American Statistician*, 1985, 39 (1), 33–38.
- Specchia, GL and Vincent Vandenberghe**, “Is part-time employment a boon or bane for firm productivity,” *Unpublished Paper, Université de Louvain*, 2013.
- Verardi, Vincenzo and Christophe Croux**, “Robust regression in Stata,” *The Stata Journal*, 2009, 9 (3), 439–453.
- World-Bank**, “World Development Report 2016: Digital Dividends,” Technical Report, World Bank 2016.

## 8 Appendix

### 8.1 Survey 1

**Greetings** Welcome, thank you for your interest in this position. We invite you to fill your data in the following application form in order to be considered in the selection process. At the end, you will do some digitizing exercises that won't take you more than 10 minutes.

**Source.** How did you hear of the position?

1. Saw the announcement in OLX
2. Saw the announcement in Elempleo.com
3. Saw the announcement in Clasificados El Tiempo
4. Other, which?

**Device.** Which device are you filling the survey on?

1. Desktop computer
2. Laptop
3. Tablet
4. Cellphone

**First Names.** Names

**Last Names.** Last Names

**E-mail.** E-mail address (check that it is spelled correctly)

**Number.** Phone number

**Date of birth.** Date of birth

**Gender.** Gender

1. Male

2. Female

**Identity document.** Identity document.

1. Colombian ID
2. Foreign ID
3. Passport

**ID number.** ID number (without any dots, commas or spaces)

**Educational attainment.** Educational attainment

1. None
2. Elementary
3. High school
4. Technical or technology degree
5. Undergraduate degree
6. Graduate degree

**Marital status.** Marital status

1. Married
2. Divorced
3. Single
4. Widow
5. Free union

**Number of children.** Number of children (if you have none, write "0")

**Pre-schoolers.** Are you responsible for any child under the age of 5 years?

1. Yes

2. No

**If answer yes.** How many children under the age of 5 live in your home?

**Older people.** Are you responsible for any adult over the age of 65?

1. Yes

2. No

**Disabled people.** Are there any disabled people living with you?

1. Yes

2. No

**Previous experience.** Do you have work experience?

1. Yes

2. No

**If answer yes.** How many years and months of work experience do you have?

1. Years

2. Months

**Currently working.** Are you currently working?

1. Yes

2. No

**If Previous experience, References.** Write the names of one professional reference, if you have it

**Reference contact.** Write the name or phone number of your professional reference

**Personal reference.** Specify the name of one personal reference, if you have it.

**Relationship.** How are you related to your personal reference?

**Personal reference contact.** Write the phone number of your personal reference



## 8.2 Pre-employment test

1. **Task 1:** Type the 14 numbers displayed on the left side column of a list displayed on the screen, and write the number in the corresponding cell on the right side column, as shown below.

**Figure 2.** Ex-ante productivity test 1

444.816,7	<input type="text"/>
—	<input type="text"/>
80,3	<input type="text"/>
128,8	<input type="text"/>
365,6	<input type="text"/>
994,7	<input type="text"/>
1.149,3	<input type="text"/>
2.408,6	<input type="text"/>
9.437,4	<input type="text"/>
13.102,4	<input type="text"/>
39.369,4	<input type="text"/>
71.055,5	<input type="text"/>
306.730,7	<input type="text"/>
—	<input type="text"/>

2. **Task 2:** Type the 8 words displayed on the left side column of a list on the screen, and write the word in the corresponding cell on the right side column. The job seekers were not able to see the words they typed and were not able to correct mistakes. The figure below shows how this task is displayed online.
3. **Task 3:** the last test aimed at checking how focused a worker was by asking how many times the letter “a” appears in the next sentence: *“Había una vez una iguana que se llamaba Mariana. Le gustaba comer moscas. Pasaba sus días y noches recorriendo manglares y selvas.”*

## 8.3 Email 1

Thank you for your interest in this position. We are sending this first response to interested applicants like you. We are trying to solve all doubts you may have about the job. Please click in the following link where you will find detailed information about the position (payment, schedule and duration of the appointment) and you will be able to continue with the application process: Link

**Figure 3.** Ex-ante productivity test 2

The image shows a screenshot of a web form titled "Ex-ante productivity test 2". The form is divided into three distinct sections, each with a header and a corresponding input field. The first section is titled "Comunas y Tamaño" and contains a text input field with three asterisks (\*\*\*) inside. The second section is titled "Total de Explotaciones" and contains a text input field with five asterisks (\*\*\*\*\*) inside. The third section is titled "Media" and contains a text input field with five asterisks (\*\*\*\*\*) inside. The form has a clean, minimalist design with horizontal lines separating the sections.

There were 4 versions of the previous link (one for each treatment condition): “Thank you very much for your expression of interest. Below you can find more information about the position:

1. The aim of the job is to support a research project at Javeriana University
  2. The length of the appointment is 3 weeks (from November 26 to December 14)
  3. The aim of the job is to support a research project at Javeriana University
  4. The work schedule is (T1, T2a, T2b , T3, T4)
  5. You would be hired under a contract for the provision of services, and the total payment will be: (T1, T2a, T2b , T3, T4)
- T1. “Monday to Friday, working hours: 8 AM – 5 PM” (Full time, non-flexible). 840.000 COP (approximately 280 USD at the time).
- T2a. “Monday to Friday, working hours: 8 AM – 12 M” (Part time, non-flexible, morning). 420.000 COP (approximately 140 USD at the time).
- T2b. “Monday to Friday, working hours: 1 PM – 5 PM” (Part time, non-flexible, afternoon). 420.000 COP (approximately 140 USD at the time).
- T3. “You must work 40 hours a week and you can agree the working hours with your supervisor, according to your availability. However, you may only work from

8 AM – 8 PM time frame Monday to Friday” (Full time, flexible). 840.000 COP (approximately 280 USD at the time).

T4. “You must work 20 hours a week and you can agree the working hours with your supervisor, according to your availability. However, you may only work in the 8 AM – 8 PM time frame between Mondays and Fridays” (Part time, flexible). 420.000 COP (approximately 140 USD at the time).

After receiving the information, we asked them: “Are you interested in this position?” Yes/No. If the applicant selected “No”, we asked him: “Why aren’t you interested?”

1. I got another job
2. The wage does not satisfy my expectations
3. The work schedule does not satisfy my expectations
4. Other

If he selected option 2 or 3, we asked them about their reservation wage and about their ideal working hours. If he applicant selected “Yes”, the survey displayed the following message: “Thank you very much for completing this survey. We will contact the selected candidates during the course of next week.”

## 8.4 EMAIL 2

“Considering the volume of applications received and the difficulty of conducting an interview to all job applicants, we would like to know more about you. Please answer the questions in the following link in order to resume the application process:

- Are you still interested in the position? o Yes (1) o No (2)
- Tell us why you are interested in this job
- How do you manage your time in order to achieve a better work-life balance?
- What is the highest number of hours you would be available to work per week?

- Why are you interested in working part time? (This question was only displayed to applicants from the part-time treatment groups)

Thank you for filling this survey, we will send the selected candidates an email during the course of next week. Please answer the previous survey as soon as possible. Best regards.

## 8.5 EMAIL 3

Congratulations! After an exhaustive selection process, we want to inform you that you have been selected for the position of Data Clerk for a study that is being carried out by Javeriana University. Remember that the job characteristics are: Duration: 3 weeks (From November 26 to December 14). Working Hours:

T1. "Monday to Friday, working hours: 8 AM – 5 PM" (Full time, non-flexible).

T2a. "Monday to Friday, working hours: 8 AM – 12 M" (Part time, non-flexible, morning).

T2b. "Monday to Friday, working hours: 1 PM – 5 PM" (Part time, non-flexible, afternoon).

T3. "You must work 40 hours a week and you can agree the working hours with your supervisor, according to your availability. However, you may only work from 8 AM – 8 PM time frame Monday to Friday" (Full time, flexible).

T4. "You must work 20 hours a week and you can agree the working hours with your supervisor, according to your needs. However, you may only work in the 8 AM – 8 PM time frame between Mondays and Fridays" (Part time, flexible).

Salary:

T1. 840.000 COP (approximately 280 USD at the time).

T2a. 420.000 COP (approximately 140 USD at the time).

T2b. 420.000 COP (approximately 140 USD at the time).

T3. 840.000 COP (approximately 280 USD at the time).

T4. 420.000 COP (approximately 140 USD at the time).

According to the new regulations, you decide whether to present in advance the corresponding social security payment of the value of this contract or, otherwise, you must sign the affidavit (attached), in which you agree to pay your social security after receiving the payment from Universidad Javeriana. No one is exempt from paying social security and the University is not responsible for these payments.

If you are still interested, please respond immediately to this offer and bring the following documents, no later than (THREE DAYS AFTER RECEIVING THE EMAIL)

1. Copy of the tax id.
2. Copy of the identification document.
3. Bank certification with date of issue no longer than 3 months.
4. Certificate of affiliation to health service.
5. Provider registration form completed (attached).
6. Declaration of information and contributions to the social security system (attached format).

Please bring these documents to the following address: Carrera 7 # 40B-36 (Jorge Hoyos Building, 7th floor). In case you cannot attend in person, please contact Carolina Franco.

If you do not bring the documentation on the date indicated, we cannot keep the job offer.

Regards,

Hiring Team



## 8.6 Tables

**Table 10.** Test of sorting into flexible jobs along *ex ante* productivity

<i>Ex ante</i> Productivity measures (ln)	Productivity # Correct/ Time (1)	Precision #Correct / Questions (2)	Speed # Questions / # Time (3)
$T_2$ : Part time non-flexible	-.01 (.07)	-.03 (.04)	.01 (.06)
$T_3$ : Full time flexible	-.02 (.07)	-.05 (.04)	.03 (.06)
$T_4$ : Part time flexible	.05 (.07)	.01 (.04)	.04 (.06)
<i>Accept</i>	-.21 (.15)	-.08 (.10)	-.12 (.10)
$T_2 \times Accept$	.00 (.41)	-.11 (.31)	.11 (.16)
$T_3 \times Accept$	.39** (.20)	.20 (.12)	.20 (.17)
$T_4 \times Accept$	.36 (.23)	.07 (.12)	.29 (.22)
Educ: High School or less	-.00 (.06)	-.00 (.03)	.00 (.05)
Educ: Vocational	-.04 (.06)	-.05 (.03)	.01 (.04)
Age 20- 25 yr old	.03 (.05)	.02 (.03)	.01 (.04)
Age 26 - 30 yr old	.01 (.07)	.01 (.04)	.00 (.05)
Men/No dep/High prod	.52*** (.07)	.14*** (.05)	.37*** (.06)
Men/Dep/Low prod	-.08 (.12)	-.03 (.06)	-.05 (.13)
Men/Dep/High prod	.52*** (.11)	-.05 (.08)	.57*** (.08)
Women/No dep/Low prod	-.04 (.07)	-.00 (.04)	-.04 (.05)
Women/No dep/High prod	.56*** (.07)	.12*** (.04)	.44*** (.05)
Women/Dep/Low prod	.04 (.09)	.03 (.05)	.01 (.08)
Women/Dep/High prod	.45*** (.07)	.15*** (.04)	.29*** (.06)
Constant	-2.61*** (.07)	-.26*** (.04)	-2.36*** (.06)
R Squared	.16	.051	.16
N	535	535	535

Notes: This Table presents the coefficients from regressing the *ex-ante* measures of productivity on treatment dummies interacted with a dummy of acceptance to the job offer. Column (1) uses total measure of productivity, Column (2) uses a measure for *Precision* and Column (3) uses a measure for *Speed*. Huber-White robust standard errors in parenthesis. We add stratification dummies as controls in all the specifications and additional controls such as age group and the level of education. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

**Table 11.** Test of sorting into Flexible Jobs along *Ex Ante* Productivity

<i>Ex ante</i> Productivity measures (ln)	Productivity # Correct/ Time (1)	Precision #Correct / Questions (2)	Speed # Questions / # Time (3)
Part time no Flex	-.01 (.07)	-.03 (.04)	.02 (.06)
Full time flex	-.02 (.07)	-.05 (.04)	.03 (.05)
Part time flex	.05 (.07)	.01 (.04)	.04 (.06)
Accepted Job Offer=1	-.21 (.15)	-.08 (.10)	-.13 (.10)
Part time no Flex × Accepted Job Offer=1	.01 (.40)	-.10 (.30)	.11 (.16)
Full time flex × Accepted Job Offer=1	.38* (.20)	.18 (.12)	.20 (.17)
Part time flex × Accepted Job Offer=1	.35 (.23)	.05 (.13)	.30 (.22)
Men/No dep/High prod	.52*** (.07)	.15*** (.05)	.37*** (.06)
Men/Dep/Low prod	-.09 (.13)	-.04 (.06)	-.05 (.13)
Men/Dep/High prod	.51*** (.10)	-.06 (.08)	.57*** (.08)
Women/No dep/Low prod	-.04 (.07)	-.00 (.04)	-.04 (.05)
Women/No dep/High prod	.56*** (.07)	.11*** (.04)	.44*** (.05)
Women/Dep/Low prod	.03 (.09)	.03 (.05)	.01 (.08)
Women/Dep/High prod	.45*** (.07)	.15*** (.04)	.30*** (.06)
Constant	-2.61*** (.07)	-.27*** (.04)	-2.35*** (.05)
R Squared	.16	.043	.16
N	535	535	535

*Notes:* This Table presents the coefficients from regressing the *ex-ante* measures of productivity on treatment dummies interacted with a dummy of acceptance to the job offer. Column (1) uses total measure of productivity, Column (2) uses a measure for *Precision* and Column (3) uses a measure for *Speed*. Huber-White robust standard errors in parenthesis. We add stratification dummies as controls in all the specifications and additional controls such as age group and the level of education. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.



**Table 12.** Treatment Effects on Ex-Post Productivity

<i>Productivity (ln)</i>	(1)	(2)
T2: Part time non-flexible	.12 (.27) [.66]	.18 (.18) [.37]
T3: Full time flexible	.40* (.15) [.06]	.28* (.16) [.10]
T4: Part time flexible	.15 (.22) [.17]	.09 (.23) [.43]
<i>Ref: T1 Full time non-flexible</i>		
<i>Ex ante productivity - Precision (ln)</i>		.89*** (.14)
<i>Ex ante total productivity (ln)</i>		-.08 (.21)
Men/No dep/High prod		-.11 (.17)
Men/Dep/Low prod		.39* (.19)
Women/No dep/Low prod		.07 (.20)
Women/No dep/High prod		-.04 (.14)
Women/Dep/Low prod		.46* (.25)
Women/Dep/High prod		.07 (.15)
Constant	-1.73*** (.11)	-1.64*** (.49)
R Squared	.2	.52
N	68	68

*Notes:* This Table reports the coefficients of estimating the treatment effect on the log of total productivity, measured as the number of correct digits over the contracted time – two weeks for individual level regressions, one week for week level regressions. Standard errors in parentheses are clustered at the individual level. Exact p-value for the Fisher Tests in brackets. This exact p-value is the proportion of possible treatment assignments that yield a test statistic greater than or equal to the observed test statistics. We used 3000 permutations. Note that \*\*\* p<0.01, \*\* p<0.05, \* p<0.1 correspond to significance levels according to Fisher p-values for T2, T3, and T4 coefficients. For the rest of coefficients, stars reflect the p-values of the standard asymptotic inference. Stratification dummies are included in all columns. Column (2) also controls for the log of *ex-ante* productivity - *Precision* (number of correct answers over total questions) and for the log of *ex-ante total* productivity (number of correct answers over total time spent to answer the questions).

**Table 13.** Decomposition of the Total Effect on Ex-Post Productivity

	Global- Productivity				Without Sorting Effects			
	Total (C/T) (1)	Precision (C/N) (2)	Speed (N/D) (3)	Effective Time (D/T) (4)	Total (C/T) (5)	Precision (C/N) (6)	Speed (N/D) (7)	Effective Time (D/T) (8)
T2: Part-time non flexible	.12 (.27) [.66]	-.08 (.10) [.52]	.00 (.14) [.98]	.20 (.12) [.13]	.18 (.18) [.37]	-.07 (.05) [.16]	-.00 (.10) [.98]	.26* (.13) [.10]
T3: Full-time flexible	.40* (.15) [.06]	.04 (.02) [.63]	.14 (.08) [.23]	.22** (.12) [.05]	.28* (.16) [.09]	-.01 (.04) [.85]	.04 (.10) [.66]	.25** (.11) [.04]
T4: Part-time flexible	.15 (.22) [.17]	.03 (.02) [.38]	.12 (.11) [.12]	.00 (.17) [1.00]	.09 (.23) [.43]	-.01 (.04) [.71]	.03 (.13) [.43]	.07 (.13) [.62]
<i>Ex ante</i> productivity - <i>Preci- sion</i>					.89*** (.26)	.24** (.09)	.15 (.19)	.50** (.20)
<i>Ex ante total</i> productivity ( <i>ln</i> )					-.08 (.21)	.02 (.03)	.18 (.13)	-.29* (.16)
<i>Ref: T1 Full-time non-flexible</i>								
Men/No dep/High prod	-.00 (.14)	-.00 (.02)	.05 (.09)	-.05 (.11)	-.13 (.23)	-.05 (.04)	-.04 (.14)	-.03 (.09)
Men/Dep/Low prod	.21 (.23)	.00 (.03)	-.07 (.10)	.28 (.26)	.18 (.32)	.05 (.07)	-.05 (.11)	.40** (.17)
Women/No dep/Low prod	-.08 (.22)	-.05 (.06)	-.16 (.12)	.12 (.15)	.06 (.24)	.02 (.04)	-.06 (.11)	.11 (.13)
Women/No dep/High prod	.12 (.19)	.03 (.04)	.11 (.09)	-.02 (.16)	-.10 (.24)	-.04 (.04)	-.04 (.10)	.04 (.14)
Women/Dep/Low prod	.54 (.34)	.05 (.05)	.26 (.19)	.23 (.17)	.48 (.31)	.02 (.04)	.21 (.15)	.22* (.13)
Women/Dep/High prod	.13 (.24)	-.04 (.03)	-.04 (.17)	.21 (.18)	.07 (.25)	-.06 (.08)	-.09 (.14)	.22** (.10)
Constant	-1.73*** (.11)	-.06*** (.02)	-.80*** (.05)	-.87*** (.08)	-1.64*** (.49)	.09 (.10)	-.25 (.33)	-1.48*** (.37)
R Squared	.2	.17	.29	.22	.52	.64	.54	.35
N	68	68	68	68	68	68	68	68

Notes: This table presents the treatment effects on four outcome variables. Columns (1) and (5), corresponds to the log of *Total productivity*, measured as the number of correct digits over the contracted time. *Precision*, columns (2) and (6), is the log of the number of correct digits over the total digits typed. *Speed*, columns (3) and (7), is the log of the number of total digits over working time. Finally, effective time, columns (4) and (8), is the log of working time over contracted time. Standard errors in parentheses are clustered at the individual level. Exact p-value for the Fisher Tests in brackets. This exact p-value is the proportion of possible treatment assignments that yield a test statistic greater than or equal to the observed test statistics. Ee used 3000 permutations. Note that \*\*\* p<0.01, \*\* p<0.05, \* p<0.1 correspond to significance levels according to Fisher p-values for T2, T3, and T4 coefficients. For the rest of coefficients, stars reflect the p-values of the standard asymptotic inference. Columns 5 to 8 controls for *ex-ante* productivity - *Precision* and for the log of *ex-ante total* productivity. Stratification dummies are included in all columns.

**Table 14.** Decomposition of Time Use Effects

	Global			Without Sorting Effects		
	Effective Time (D/T)	Absenteeism (A/T)	Breaks (B/T)	Effective Time (D/T)	Absenteeism (A/T)	Breaks (B/T)
T2: Part time non-flexible	.10 (.06) [.12]	.04 (.04) [.33]	-.14*** (.04) [.00]	.12* (.06) [.10]	.03 (.04) [.55]	-.15*** (.04) [.01]
T3: Full time flexible	.09* (.05) [.08]	-.01 (.04) [.85]	-.08** (.03) [.03]	.10* (.05) [.07]	-.01 (.03) [.73]	-.08** (.03) [.04]
T4: Part time flexible	.02 (.07) [.60]	.11 (.05) [.33]	-.14*** (.04) [.00]	.05 (.06) [.18]	.10 (.03) [.44]	-.14*** (.04) [.00]
<i>Ex ante productivity</i> - Precision (ln)				.20** (.08)	-.12** (.06)	-.08* (.05)
<i>Ex ante total productivity</i> (ln)				-.11 (.07)	.07 (.05)	.04 (.04)
<i>Ref: T1 Full time non-flexible</i>						
Men/No dep/High prod	-.04 (.04)	-.01 (.02)	.04 (.04)	-.03 (.04)	-.01 (.03)	.04 (.04)
Men/Dep/Low prod	.12 (.12)	.01 (.08)	-.13*** (.04)	.17* (.08)	-.02 (.06)	-.15*** (.03)
Women/No dep/Low prod	.05 (.06)	.04 (.04)	-.09* (.04)	.05 (.06)	.04 (.03)	-.09* (.05)
Women/No dep/High prod	-.03 (.07)	.03 (.05)	.00 (.06)	-.01 (.06)	.02 (.05)	-.00 (.05)
Women/Dep/Low prod	.09 (.08)	.04 (.04)	-.13* (.07)	.08 (.07)	.04 (.03)	-.13* (.07)
Women/Dep/High prod	.07 (.07)	.02 (.04)	-.09** (.04)	.07 (.04)	.02 (.02)	-.09** (.03)
Constant	.43*** (.03)	.05* (.03)	.51*** (.02)	.22 (.15)	.20** (.10)	.58*** (.09)
R Squared	.23	.26	.46	.35	.32	.48
N	68	68	68	68	68	68

*Notes:* This table presents the treatment effects on three outcome variables. *Effective time*, columns (1) and (4), corresponds working time over contracted time. *Absenteeism*, columns (2) and (3), is the time outside the lab over the contracted time. Finally, *Breaks* is measured as the non-working time inside the lab over the contracted time. Standard errors in parentheses are clustered at the individual level. Exact p-value for the Fisher Tests in brackets. This exact p-value is the proportion of possible treatment assignments that yield a test statistic greater than or equal to the observed test statistics. We used 3000 permutations. Note that \*\*\* p<0.01, \*\* p<0.05, \* p<0.1 correspond to significance levels according to Fisher p-values for T2, T3, and T4 coefficients. For the rest of coefficients, stars reflect the p-values of the standard asymptotic inference. Stratification dummies are included in all columns. Columns 4 to 6 control for *ex-ante* productivity.