How does flexible schedules affect workers' productivity? A Field Experiment

Marie Boltz[*], Bart Cockx[†], Ana Maria Diaz[‡]& Luz Magdalena Salas[§¶]

June 27, 2019

Abstract

We conducted an experiment in which we made real job offers and hired workers under different working schedules to evaluate the effects of flexibility on productivity. When evaluating these relationship, we stratified the sample considering individuals' characteristics such as gender, dependents and their ex-ante level of productivity. We offered four different contact environments varying the total number of hours worked per week - i.e. full- vs. part-time, and the working schedule - i.e. flexible vs. non-flexible. We hired workers to perform data entry tasks that didn't require a specific level of skill in the facilities of a private university in Bogota for a period of three weeks. We observed and measured workers' productivity in a natural environment. Our results show higher productivity levels for workers that perform the job under more flexible schedules than those under a regular schedule. Specifically, flexibility increases the number of correct images typed in 5 to 10 percent and reduces the time of the task in approximately 7 percent, on average. However, complete flexibility, both in terms of number of hours worked per week and in terms of working schedule, reduces productivity. These results provide evidence that flexibility could be beneficial for firms under certain conditions.

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

JEL: J21, J22, J23, J24, J33

[*]Centre d'Economie de la Sorbonne Université Paris 1 Panthéon Sorbonne Centre. marie.boltz@univ-paris1.fr

[†]Department of Social Economics, Ghent University, bart.cockx@ugent.be

[‡]Departamento de Economía, Pontificia Universidad Javeriana. a.diaze@javeriana.edu.co

[§]Departamento de Economía, Pontificia Universidad Javeriana. salas.luz@javeriana.edu.co

[¶]This paper has benefited from great research assistance by Juan David Hernandez and Juan Sebastian Ortiz. We thank Nicolas Lillo from his great work building the software and providing the data of the Chilean Census. The authors would also like to thank Pontificia Universidad Javeriana for the financial support.

1 Introduction

There is a recent debate in the literature about how different types of workers prefer jobs with more flexibility. Most of the studies show that workers, in particular women, demand more flexibility in order to balance their household responsibilities with their labor market participation Giannikis and Mihail (2011), Booth and Van Ours (2008) and Booth and Van Ours (2013), Bustelo et al. (2019). However, in some contexts, more flexibility could generate large hiring costs or lower income for employees, ocasionally transferred by firms. Therefore, offering flexible jobs may, under some circumstances, impose restrictions for firms. When flexibility is provided in terms of working schedule, it could raise operating costs, in particular for activities that require some level of coordination or monitoring. It is important, to determine how to compensate these additional costs generated to firms, when encouraging them to offer flexible work arrangements.

In order to explore this idea, we designed and conducted a field experiment in which we hired workers under different work alternatives, varying the level of flexibility. We posted ads for data entry workers in the main job-seekers' portals in Bogota, without mentioning the working schedule. Interested applicants had to fill out their demographic information in a resume format and had to perform a productivity test (similar to the task performed in the job offer). Then, we selected a subset of applicants and randomly offer jobs that varied in terms of hours worked per week (part-time or full-time) and working schedule (fixed or flexible). The workers who accepted the offers typed a Chilean Agricultural Census with 350,000 images in the computer rooms at the University of Javeriana for a period of three weeks. In this paper, we evaluate changes in the level of productivity as a result of working in one of the following four schedules: (i) full-time non-flexible schedule (i.e. Monday to Friday, from 8 AM to 5 PM), (ii) part-time non-flexible (i.e. Monday to Friday from 8 AM to 12 M, for the morning shift; and 1 PM to 5 PM, for the afternoon shift); (iii) full-time flexible (i.e. 40 hours a week, the worker could choose the entry and exit time); and (iv) part-time flexible (i.e. 20 hours a week, the worker could choose the entry and exit time).

There are different types of flexibility that we observe in the labor market. To name some, a flexible job is one in which the employee can perform the job from home or remotely, or can work a different number of hours compared with the traditional sched-

ule (8-hours-a-day or 40-hours-a-week schedule), or can decide the entry and exit hours every day. All these definitions have different implications on a variety of outcomes for firms, such as hiring costs, productivity, etc.; and for for the employers in terms of the level of effort in their activities, satisfaction, commitment with the employer, among others. For example, Kelliher and Anderson (2010) study the implications of providing flexibility in terms of number of hours worked and the workplace on the intensity of the work. They found that flexibility makes workers more satisfied and more commited to their employers, but they report working more intensively. In this case, the authors consider a reciprocity mechanism as the explanation of such behavior. In other words, under more flexible arrangements, workers exert more effort "in order to return the benefit to their employer".

In a similar study, Eaton (2003) explores the relationship between working remotely and the perception of commitment and the perception of productivity of workers using survey data of employees from different US firms. She defines flexibility as the workers' freedom to use flexibility policies in those firms. The study finds that access to policies that promote flexibility make employers more commited to the firm and make them feel more productive.

In these studies, the authors used survey data to understand how flexibility would affect firms' outcomes through the perception of workers. Our experimental setting allows us to observe real exerted effort of workers through different measures of productivity during their period of work. In another study, Cockx et al. (2019) analyze the first stages of the application process of workers that participate in this experiment in order to elicit preferences for flexibility through applicants' decisions, and estimate an empirical model to test if individuals with different characteristics i.e. productivity ex-ante, sort into different work arranegements.

Following a similar definition of flexibility than ours, Beckmann, Cornelissen, and KrÃ€kel (2017) provide a theoretical and empirical model to test the effect of self-managed working time on the level of effort that workers exert in their job. They discuss two potential channels through which this effect can be observed: instrinsic motivation and reciprocity of workers. Using the German Socio Economic Panel, the researchers estimate a moral-hazard model and find that self-managed working time workers perform better than those under non-flexible or non-autonomous job contracts, mainly as a result of more intrinsic motivation. The empirical analysis does not find

reciprocity to matter in explaining such behavior. The study uses different measures of workers' productivity, such as the difference between the contractual and actual number of working hours, absenteeism and hourly wages. However, do nor observe the actual output generated by workers to the firm.

Our study contributes to the understanding of this relationship because uses real observed productivity of workers under different work arrangements, with two measures of flexibility, under a natural environment. In contrast with other studies, we recorded and informed workers about the total number of hours worked at the end of every week, therefore, we were able to control the total number of hours worked, or, at least, workers perceived some level of monitoring. We do not use information about perception of workers or employers.

In the same line with our research, two recent field experiments evaluated the effects of flexibility on workers' productivity. Bloom et al. (2014) carried out an field experiment with individuals a call-center of a travel agency in China. The employees of a call center were asked if they would be interested in working from home. Among those who answered affirmatively, the call-center randomly offered positions working from home for a 9-month period. The study finds that working from home had a positive impact on performance and increased the self-reported level of satisfaction of employees. Based on these results, the firm extended the opportunity to work from home to the rest of its employees. However, 50 percent of those who worked from home (treated) decided to go back to the office after the 9-month period and only 35 percent of individuals in the control group chose to work from home. This result suggests that some workers may be benefited from these types of work arrangements while others not. This study is similar to ours in terms of the type of activity because in both cases, the task didn't require any specific skill nor any level of monitoring or coordination. Also, productivity was measured in terms of output per time. But it differs in, at least, two ways: first, flexibility in Bloom's paper is defined by the opportunity of working from home for a self-selected group of individuals interested in that type of flexbility. Second, the salary of workers in the call-center was not fixed. It varied according to the performance of workers, which may introduce other motivations to perform well during the period of work.

Dutcher (2012) also analyzes the effect of job flexibility on productivity in a lab setting, but he compares these effects for two types of tasks: dull and creative. Flexibility is

defined in a context in which workers (in this case, students) are free to choose to work remotely from the lab or in the lab. The results point out gains in productivity from working remotely but only for creative tasks. The opposite is observed for routine or dull activities: firms under this working climate may be harmed when they offer the possibility of working remotely because the productivity of workers is reduced. Our experimental setting resembles the dull-task from this study, however, it is not performed by students in a lab. Instead, by data entry clerks that received a salary for doing the task. Also, the hiring process was identical to the one they should normally follow to perform that kind of work. Another strand of the literature studies the preferences of workers for flexible job alternatives (Mas and Pallais (2017), Wiswall and Zafar (2017))

We constructed a panel data with observations at the worker level per quarter of the hour per day. We carried out a fixed effects model to estimate a causal relationship between job flexibility and workers' productivity. To measure productivity, we use three measures of time spent typing each image from the census and the number of correct images typed. Our estimates correct for serial correlation in order to eliminate the individual fixed effect on productivity over time. Our findings indicate that more flexibility translates into larger productivity. Specifically, individuals under a part-time flexible contract reduce the number of seconds per image (all and correct) in 11 to 14 percent and increase the number of images typed in 8.4 percent. The performance of individuals under full-time flexible contracts the effects are more modest, but still significant, pointing out an increase of 7 to 10 percent in comparison with individuals under full-time non-flexible schedules. Finally, we didn't find any significant differences in seconds per image for those under part-time-non-flexible schedule in comparison with full-time-non-flexible, implying that the gains of flexible work arrangements are negligible when too much flexibility is provided.

In general, there is a difficulty in distinguishing the mechanisms by which flexible contracts can raise workers' productivity: it could be through sorting on ex-ante productivity, or through raising ex-post productivity. In a paper about the first stage of the application process, we found no evidence of any effect on productivity through sorting. Therefore, we are able to cleanly identify the effect of the types of contracts on ex-post productivity.

According to the literature, reciprocity, commitment to the firm or intrinsic motivation

could be causing these differences in the effort exerted by workers. Our experimental setting does not allow us to explore these mechanisms separately. Our results therefore suggest that flexibility in terms of time and schedule raises productivity. One possible explanations for this result is that workers' autonomy to choose the number of hours enhance the intrinsic motivation of individuals because they feel less forced to work. Or, workers may also choose to take longer breaks while working which may raise the average productivity while at work. We can unambiguously conclude that full-time flexible work raises productivity. However, there could be more trade offs for part-time jobs and even an unambiguous negative effect on productivity if part-time is combined working time flexibility.

The remainder of the paper is organized as follows: In section 2, we describe the experimental design in more detail. Section 3 describes the data and discuss some descriptive statistics, section 4 presents the empirical model and results of the estimations, and section 5 concludes.

2 Experimental Design

2.1 Assignment to contract environments

We placed real job advertisements for data entry clerks in different job-posting web pages in Bogota. We selected a fraction of applicants to offer real and meaningful employment for a task that didn't require a specific level of education nor a specialized skill, and varied the contract offered in terms of number of hours (part or full time) and working schedule (flexible or non flexible). We hired a sample of 40 workers that accepted our offers in four different contract environments and measured productivity of each worker during a period of three weeks. We define four treatment groups according to the contract environment offered, as follows:

1. Full time, non-flexible: Monday to Friday from **8 AM to 5 PM** (**40 hours a week**).

2. Part time, non-flexible: Monday to Friday from **8 AM to 12 M** (morning shift, **20 hours a week**); or Monday to Friday from 1 PM to 5 PM (afternoon shift, 20 hours a week).

3. Full time, flexible: Work **40 hours a week** and the worker arranges the working schedule with her supervisor, according to worker's availability. However, they could work Monday to Friday in a time frame from 8 AM to 8 PM from only.

4. Part time, flexible: Work **20 hours a week** and the worker arranges the working schedule with her supervisor, according to worker's availability. However, they could work Monday to Friday in a time frame from 8 AM to 8 PM from only.

Our experimental setting does not include flexibility as an alternative to work remotely or from home. We stratified the assignment to the treatment groups according to characteristics of the applicant such as gender, if the applicant had dependents and an ex-ante level of productivity that we measured during the application process. Although we offered an equivalent number of positions for each treatment group, we were unable to hire a proportional number of individuals in each contract environment. For our analysis, we have information from a self-selected sample of individuals that accepted our job offers. In another paper, Cockx et al (2019) evaluated the existence of selection or sorting of job applicants into different contract schedules. The study provides evidence of sorting by analyzing the probability of continuing in the application process conditional on a given work arranegement offered.

1 shows the distribution of individuals that were actually hired to perform the data entry tasks for the experiment. We made offers to an equivalent number of individuals per stratification cell; however, not all applicants that were offered a position accepted the job. Whenever an applicant rejected our offer, we chose another applicant from the same stratification cell. We repeated this process until administrative office of the University allowed us to collect all the required documentation to sign the contracts.

Table 1: Sample by Treatment Group

	N	%
Full-Time Non-Flexible	13	32.5
Part-Time Non-Flexible	8	20.0
Full-Time Flexible	11	27.5
Part-Time Flexible	8	20.0
Total	40	100.00

We randomly offered real and meaningful work to 79 applicants and 40 were hired. However, 3 never showed up to work and 5 resigned before completing the three-week

period of work. For that reason, we use the information of the 32 workers that finished the tasks. Contracts were exactly the same for all individuals in the different treatments' groups, the only difference being the payment for part time and full time workers.[1]

2.2 Performing the job

The workers were hired to perform a data entry task in the computer labs at the University of Javeriana. We assigned one separate lab to each treatment group, located in different buildings within the university to reduce the risk of communication between individuals from different tratment groups. We hired 8 undergraduate students, that monitored each computer lab for the entire experiment. Monitors were randomly assigned to one of the 4 rooms in 2 shifts (morning from 8 AM to 2 PM, and afternoon from 2 PM to 8 PM). Monitors were present at the computer lab all the time and had to register any activity that happened inside the computer rooms in a Qualtrics survey format during their shifts, such as phone calls, interactions among participants and between participants and monitors, and other types of activities that could have affected workers' productivity. They were also asked to report the start and finish time for each worker every day of work, and to assist them with any technical difficulty they might have faced during the period of the study. Additionally, they had to verify that the computer labs were open on time each day and they couldn't leave the rooms until everyone had left.

Each worker was assigned an ID code to with a user name and password to log into the software created to perform the data entry process. The software displayed the image in a dialog (chart or typing space), recorded the answer, the time spent in each image, the breaks or pauses and the time at which the data clerk finished her tasks for the day. This software was created in Python and depicted images with numbers from a Chilean Agrarian Census, with 350.000 images.[2]

Each image was randomly assigned to nine workers to control for the quality of the typing. Every worker had the same probability of receiving each image. All the numbers or images from the Census was typed an average of 7 times. The following figure shows an example of what was displayed in the screen.

During the first day of work, our monitors instructed workers about how to use the

Figure 1: Task 1

software and all the details of the shedules. After the instructions, workers had to start the data entry task for the next few days. The last three days of the experiment, workers received a second task, that consisted in typing both words and numbers from electoral results in Chile. This task was more difficult than the first task because it required more attention due to the presence of spaces, tables, non-common words, and blurriness.

We recorded and informed workers the flexible work schedules, via e-mail, the total number of hours worked at the end of every week. Consequently, we told them whether they needed to work more or less hours in the coming week in order to fulfill the contract requirements. We didn't send any similar e-mail to workers from the non-flexible treatment conditions, as their working hours were fixed and they didn't have the opportunity to compensate in case they didn't meet the total number of hours. Ideally, the same messages had to be sent to all treatments, however for non-flexible employment shedules it does not make sense and would have generated suspicions from participants. In addition, sending messages for all treatments could be a treat to external validity.

3 Data and descriptive statistics

We observed and used different measures of productivity for the sample of workers who performed the data entry task during three weeks. Our measures of productivity are observed in a natural environment for 32 workers. However, we use information in blocks of 15 minutes (every quarter of the hour) for each individual every day of work. Thus, we constructed an unbalanced panel of 32 individuals over time for a total sample of 8,981 observations. Our outcome variables are: the number of seconds that each worker spends typing each image, the number of images that are correctly typed, and the number of seconds per image that is correctly typed. We also constructed a

measure of overall performance as the standard deviation of the variable seconds per correct image.

For the purpose of the analysis, we compared observable characteristics of workers including gender, if the worker has dependents, age, level of education, years of experience, marital status and different measures of ex-ante productivity, for individuals in the part-time non-flexible, full-time non-flexible, part-time flexible and full-time non-flexible contracts. In order to compare two different measures of flexibility, we aggregated the observations from treatments 2 and 4 to construct a measure for part-time work and treatments 1 and 3 to measure full-time work. This variable allows us to measure flexibility in terms of number of hours worked by individuals. 2 compares a set of observable characteristics for individuals that performed the job under part-time vs. full-time contract schedules. Although we offered an equivalent number of jobs to individuals in each job contract, not all of them accepted the job. For that reason, we have a sample of 40 workers that were actually hired.

The distribution of hired workers is not balanced across part-time and full-time groups. However, their observable characteristics are statistically the same: more than a half are women, 25 percent have dependents (both children under 5 or other family member that requires permanent care), are around 30 years of age, have between 5 and 6.8 years of work experience and almost 80 percent are single. In terms of education, most workers reached at least a vocational level. The last set of characteristics measure workers' ex-ante level of productivity, recorded during their application process. For all the set of variables we find no differences among part-time and full-time workers.

Table 2: Mean Difference of Observables, Part-Time (1) vs. Full-Time (0)

	0	1	Diff	Pvalue	N1	N2
Gender (0 male, 1 female)	.58	.69	–.10	.51	24	16
Dependents	.25	.25	.00	1.00	24	16
High Productivity	.23	.13	.10	.42	22	16
Dummy Children	.29	.25	.04	.78	24	16
Age	32.04	30.00	2.04	.46	24	16
Years of experience	6.76	5.13	1.63	.46	21	15
Educ: Graduate	.04	.00	.04	.33	24	16
Educ: Undergraduate	.29	.31	–.02	.89	24	16
Educ: Vocational	.46	.38	.08	.61	24	16
Single	.79	.81	–.02	.88	24	16
Number of correct answers - Test 1	10.92	10.69	.23	.82	24	16
Number of correct answers - Test 2	6.29	6.13	.17	.81	24	16
Percentage of correct answers - Test 1	.78	.76	.02	.82	24	16
Percentage of correct answers - Test 2	.63	.61	.02	.81	24	16
Test 3 Correct	.68	.69	–.01	.97	22	16
Test 1 Time	139.23	134.91	4.32	.80	24	16
Test 2 Time	142.30	139.67	2.63	.94	24	16
Test 3 Time	42.43	61.40	–18.96	.11	22	16
Productivity test 1 (sd)	–.07	.09	–.16	.62	24	16
Productivity test 2 (sd)	.09	–.02	.11	.73	24	16
Productivity test 3 (sd)	–.13	–.33	.20	.50	22	16

*** p<0.01, ** p<0.05, * p<0.1

Second, we constructed another measure of flexibility in terms of working schedule. As mentioned in the job offer, workers were able to decide the start time and end time of work every day within a 12-hour frame. They were instructed to work either 20, for part-time, or 40 hours a week, for full-time schedules, but could choose their working schedule. 3 shows the mean differences of the observable characteristics for individuals in the two groups, flexible vs. non-flexible schedule. The comparisons show nos statistical differences for most variables. However, there is a difference in the workers that report having children, and those with completed undergraduate and vocational education. In paprticular, 42 percent of workers in the flexible work arrangement have children, while only 14 percent in the non-flexible. Also, a larger perceentage of workers in the non-flexible work schedule completed undergradate education, while in the non-flexible, the vocational education is more frequent. All the characteristics that measure productivity ex-ante are balance statistically.

Table 3: Mean Difference of Observables, Flexible (1) vs. Non-Flexible (0)

	0	1	Diff	Pvalue	N1	N2
Gender (0 male, 1 female)	.67	.58	.09	.58	21	19
Dependents	.24	.26	–.03	.86	21	19
High Productivity	.20	.17	.03	.80	20	18
Dummy Children	.14	.42	–.28*	.06	21	19
Age	31.86	30.53	1.33	.63	21	19
Years of experience	5.39	6.78	–1.39	.56	18	18
Educ: Graduate	.05	.00	.05	.33	21	19
Educ: Undergraduate	.43	.16	.27*	.06	21	19
Educ: Vocational	.24	.63	–.39**	.01	21	19
Single	.86	.74	.12	.36	21	19
Number of correct answers - Test 1	10.81	10.84	–.03	.97	21	19
Number of correct answers - Test 2	6.48	5.95	.53	.45	21	19
Percentage of correct answers - Test 1	.77	.77	–.00	.97	21	19
Percentage of correct answers - Test 2	.65	.59	.05	.45	21	19
Test 3 Correct	.70	.67	.03	.83	20	18
Test 1 Time	143.03	131.40	11.63	.48	21	19
Test 2 Time	136.79	146.17	–9.39	.78	21	19
Test 3 Time	42.87	58.80	–15.93	.16	20	18
Productivity test 1 (sd)	–.12	.11	–.23	.44	21	19
Productivity test 2 (sd)	.06	.02	.04	.91	21	19
Productivity test 3 (sd)	–.10	–.35	.25	.40	20	18

*** p<0.01, ** p<0.05, * p<0.1

For the purpose of our analysis, we used four measures of productivity: the number of seconds per image typed, the number of seconds per correct image typed, an overall measure of performance and the natural log of the number of correct answers. The overall performance is the standardized measured of seconds per correct image typed of each individual. 4 presents the mean differences of the outcome variables for individuals comparing part- vs. full-time, and flexible vs. non-flexible contract environments. The comparisons show that individuals in part-time and flexible work arrangements spend less seconds typing correct images relative to individuals under full-time and non-flexible contracts. Also, the number of correct images typed is larger for individuals under both types of flexibility. The means are statistically different from zero for almost all measures of productivity.

Table 4: Mean Difference of Dependent Variables

	Part-Time vs Full-Time					
	0	1	Diff	Pvalue	N1	N2
Seconds	3.02	2.97	.06	.48	6706	2275
Seconds Correct	3.00	2.87	.13**	.04	6703	2273
Overall Performance	.01	–.04	.05**	.04	6703	2273
Correct Answers (ln)	5.36	5.41	–.04*	.07	6710	2275
	Flex vs Non Flex					
	0	1	Diff	Pvalue	N1	N2
Seconds	3.19	2.80	.39***	.00	4843	4138
Seconds Correct	3.14	2.75	.39***	.00	4840	4136
Overall Performance	.07	–.08	.15***	.00	4840	4136
Correct Answers (ln)	5.25	5.51	–.26***	.00	4845	4140

*** p<0.01, ** p<0.05, * p<0.1

Notes: The first panel compares the outcome variables for the treatment groups under part-time and full-time schedules,. The panel in the bottom compares the mean differences of the outcome variables for the treatment groups under the flexible and non-flexible working schedules.

4 Empirical model and results

We recorded the information during the entry data tasks for three weeks of work and use different measures of productivity at the quarter of the hour level. In the section above we presented mean differences of the observable characteristics of workers and found no differences for the comparison of part-time and full-time schedules. However, for the flexible and non-flexible work arrangements there are differences in the number of workers that have children and their level of education. We will include these differences when estimating the effects of flexible contracts on productivity of workers.

We estimated a dynamic data panel model represented in the following equation:

$$Y_{it} = \beta_1 + \beta_2 T_2 + \beta_3 T_3 + \beta_4 T_4 + \delta X_{it} + \lambda_t + \epsilon_{it}$$

with $i = 1, ...N$ and $t = 1, ...T$. Our unit of time is quarter of the hour from 8 AM to 8 PM for 13 days.Y_{it} is the dependent variable that captures different measures of productivity on-the-job, T_2 is a dummy that measures if individual i worked in part-

time non-flexible schedule, T_3 is a dummy that measures if individual i worked in full-time flexible schedule, T_4 is a dummy that measures if individual i worked in part-time flexible schedule is the treatment group, λ_t measures time fixed effects. ϵ_{it} is the disturbance term with $\epsilon_{it} = u_i + v_{it}$, where $u_i \sim IID(0, \sigma_u^2)$ and $v_{it} \sim IID(0, \sigma_v^2)$. X_{it} is the set of controls such as the stratification variables (gender, dependents and the ex-ante level of productivity) and other observable characteristics (age, years of experience, level of education and marital status) that could affect the observed productivity of workers.

The performance of workers during their tasks could be correlated over time, therefore ϵ_{it} suffers from serial correlation. Given that N is small and T is large, we correct the possible serial correlation by estimating a fixed effects model with a lagged dependent variable as follows:

$$Y_{it} = \alpha_0 + \alpha_1 Y_{it-1} + \alpha_2 T_2 + \alpha_3 T_3 + \alpha_4 T_4 + \delta X_{it} + \lambda_t + \mu_{it}$$

were Y_{it-1} is the observed level of productivity in the period $t-1$ and μ_{it} is the disturbance term.

5 presents the results of regressing the outcome variables on the lagged dependent variable, the treatment dummies and all other controls. By including a lag, we get rid of individual specific effects. We find that all lagged variables are statistically significant in each specification, confirming the strong correlation of the dependent variables over time. The first column shows the treatment effects on the number of seconds that a worker spends typing an image. We find that workers in the full-time flexible and part-time flexible schedules spend less minutes than under full-time no flexible schedule. In particular, 7.7 and 11.4 percent less, respectively. Similarily, the direction and significance of the effects hold for the number of seconds spend in each correct image abd the overall performance of the workers (see columns 2 and 3). The last column shows that individuals in treatment groups with any type of flexibility are more efficient that those in the full-time non-flexible schedule. They type more correct images on average than the reference group. Other regressors such as if workers have dependents, the level of education, productivity ex-ante and gender are statistically significant in explaining the productivity measures.[3]

Table 5: Treatment Effects on Productivity

	Seconds	Seconds Correct	Overall Performance	Correct (ln)
L.Seconds	.273***			
	(.018)			
Part-time no flex	−.039	−.085	−.034	.143***
	(.104)	(.097)	(.038)	(.034)
Full-time flex	−.249***	−.184**	−.073**	.376***
	(.082)	(.077)	(.030)	(.026)
Part-time flex	−.367***	−.347***	−.137***	.318***
	(.103)	(.096)	(.038)	(.033)
L.Seconds Correct		.404***		
		(.020)		
L.Overall Performance			.404***	
			(.020)	
L.Correct Answers (ln)				.236***
				(.014)
Constant	3.236***	2.487***	.285***	3.781***
	(.298)	(.279)	(.107)	(.116)
F	41.2	58	58	72.3
R Squared	.354	.32	.32	.388
N	7496	7491	7491	7499

*** p<0.01, ** p<0.05, * p<0.1

Notes: The coefficients are the result of estimating a time effects model for different measures of productivity on its lag and treatment dummies. This estimation includes as controls the observable characteristics of individuals such as gender, level of education, age, marital status and the ex-ante level of productivity.

Next, we estimated the treatment effects including interactions of our measures of flexibility, in terms of number of hours worked and schedule. We estimated a fixed effects lagged dependent variable model for the productivity measures and find no effects of part-time nor part-time flexible for the first three outcome variables. Results are presented in 6. In this model, flexibility matters because it raises all measures of productivity. To be more precise, having the freedom to choose the entry and exit time of work reduces the time spent typing numbers and increases the total number of images typed. In fact, it reduces the number of seconds per image typed (correct or all) in around 7 percent, and increases the number of correct images in 10 percent. However, when we interact this variable with part-time job schedules, thus, providing more flexibility with the total number of hours worked per week, the coefficients for the measures of productivity in terms of time are not statistically different from zero, and the effect disappears. Moreover, the effect of the interaction on the number of images correctly typed is negative, implying that too much flexibility harms the firm by lowering the level of productivity of workers in 5.3 percent. In sum, under full

flexibility (total hours and freedom to choose the working schedule), workers are less productive than workers in a more traditional working schedule. Both part-time and flexible-time schedules increase productivity in terms of number of correct entries, on their own, but the interaction is negative, suggesting that a part-time-flexible contract will be detrimental for the employer.[4] In other words, flexibility is beneficial up to a certain point.

Table 6: Treatment Effects on Productivity with Interactions of Treatments

	Seconds	Seconds Correct	Overall Performance	Correct (ln)
L.Seconds	.273***			
	(.018)			
Part time (1) Full Time (0)	−.039	−.085	−.034	.143***
	(.104)	(.097)	(.038)	(.034)
Non Flexible (0) vs Flexible (1)	−.249***	−.184**	−.073**	.376***
	(.082)	(.077)	(.030)	(.026)
Part x Flex	−.079	−.078	−.031	−.202***
	(.154)	(.143)	(.057)	(.050)
L.Seconds Correct		.404***		
		(.020)		
L.Overall Performance			.404***	
			(.020)	
L.Correct Answers (ln)				.236***
				(.014)
Constant	3.236***	2.487***	.285***	3.781***
	(.298)	(.279)	(.107)	(.116)
F	41.2	58	58	72.3
R Squared	.354	.32	.32	.388
N	7496	7491	7491	7499

*** p<0.01, ** p<0.05, * p<0.1

Notes: The coefficients are the result of estimating a time effects model for different measures of productivity on its lag and treatment dummies. This estimation includes as controls the observable characteristics of individuals such as gender, level of education, age, marital status and the ex-ante level of productivity.

5 Conclusion

We conducted a field experiment to test if working under different job schedules affect the observed level of productivity. We hired workers to perform a data entry task for a period of three weeks in computer labs at a private university in Bogota. We recruited workers under a regular application job process. We posted entry data clerk job ads in the main job-search web pages in Bogota without making any reference to the work schedule. We received over 600 applications in which individuals had to

complete an application format with information similar to a resume, and also had to perform a productivity test. We stratified the sample according to gender, if the applicant had dependents and the level of ex-ante productivity, and randomly offered positions under four different working schedules to 80 individuals, and 32 accepted and showed-up to work for the total period of work. We observed and measured the productivity of workers for the whole period of time and use that information to evaluate the relationship between flexibility and productivity.

Our results show that workers under flexible work arrangements exert more effort and increase the level of productivity in comparison with individuals that work under a regular schedule (full-time non-flexible). However, the level of productivity does not increase always with flexibility. Part-time workers under flexible schedules, are not statstically more productive than the regular full-time schedule. This study contributes to understanding how job flexibility affects firm's outcomes, and how workers respond to different types of contract environments, by exerting different levels of effort, controlling all other factors that may influence worker's productivity. As Eaton (2003) discusses, flexibility is constrained by the nature of the task. Therefore, flexibility cannot be a generalized practice. Many activities in the labor market with some degree of coordination, monitoring, etc., would be harmed by providing flexibility even with an increase in workers' productivity, because the gains from self-motivation, commitment or reciprocity of workers cannot compensate the lerger hiring costs.

With our setting we identify some limitations: (i) the mechanisms that generate the raise in productivity cannot be clearly identified (although we can say that it is not induced by sorting); (ii) we considered "dull" tasks only; it would be interesting to see what happens for more creative tasks, as in Dutcher (2012).

References

Beckmann, Michael, Thomas Cornelissen, and Matthias KrÃ€kel (2017). "Self-managed working time and employee effort: Theory and evidence". In: *Journal of Economic Behavior & Organization* 133, pp. 285–302. ISSN: 0167-2681. DOI: https://doi.org/10.1016/j.jebo.2016.11.013. URL: http://www.sciencedirect.com/science/article/pii/S0167268116302694.

Bloom, Nicholas et al. (2014). "Does working from home work? Evidence from a Chinese experiment". In: *The Quarterly Journal of Economics* 130.1, pp. 165–218.

Booth, Alison L and Jan C Van Ours (2008). "Job satisfaction and family happiness: the part-time work puzzle". In: *The Economic Journal* 118.526, F77–F99.

– (2013). "Part-time jobs: What women want?" In: *Journal of Population Economics* 26.1, pp. 263–283.

Bustelo, Monserrat et al. (2019). *What's the price of freedom? Estimating women's willingness to pay for job flexibility*. Tech. rep.

Dutcher, E Glenn (2012). "The effects of telecommuting on productivity: An experimental examination. The role of dull and creative tasks". In: *Journal of Economic Behavior & Organization* 84.1, pp. 355–363.

Eaton, Susan C (2003). "If you can use them: Flexibility policies, organizational commitment, and perceived performance". In: *Industrial Relations: A Journal of Economy and Society* 42.2, pp. 145–167.

Giannikis, Stefanos K and Dimitrios M Mihail (2011). "Flexible work arrangements in Greece: a study of employee perceptions". In: *The International Journal of Human Resource Management* 22.02, pp. 417–432.

Kelliher, Clare and Deirdre Anderson (2010). "Doing more with less? Flexible working practices and the intensification of work". In: *Human relations* 63.1, pp. 83–106.

Mas, Alexandre and Amanda Pallais (2017). "Valuing alternative work arrangements". In: *American Economic Review* 107.12, pp. 3722–59.

Wiswall, Matthew and Basit Zafar (2017). "Preference for the workplace, investment in human capital, and gender". In: *The Quarterly Journal of Economics* 133.1, pp. 457–507.

Appendix

Table 7: Treatment Effects on Productivity with Controls

	Seconds	Seconds Correct	Overall Performance	Correct (ln)
L.Seconds	.273***			
	(.018)			
Part-time no flex	−.039	−.085	−.034	.143***
	(.104)	(.097)	(.038)	(.034)
Full-time flex	−.249***	−.184**	−.073**	.376***
	(.082)	(.077)	(.030)	(.026)
Part-time flex	−.367***	−.347***	−.137***	.318***
	(.103)	(.096)	(.038)	(.033)
Gender (0 male, 1 female)	−.248***	−.177***	−.070***	.108***
	(.062)	(.058)	(.023)	(.020)
Dependents	−.462***	−.358***	−.141***	.190***
	(.081)	(.076)	(.030)	(.026)
High Productivity	−.381***	−.302***	−.119***	.038
	(.086)	(.080)	(.032)	(.028)
Dummy Children	.171**	.108	.043	−.081***
	(.072)	(.067)	(.027)	(.023)
Age	−.040***	−.030***	−.012***	.011***
	(.009)	(.008)	(.003)	(.003)
Years of experience	.015*	.009	.003	−.005*
	(.008)	(.008)	(.003)	(.003)
Educ: Graduate	1.010***	.811***	.320***	−.389***
	(.196)	(.183)	(.072)	(.063)
Educ: Undergraduate	.488***	.434***	.171***	−.258***
	(.080)	(.075)	(.030)	(.026)
Educ: Vocational	.431***	.398***	.157***	−.212***
	(.082)	(.077)	(.030)	(.026)
Single	.072	.112	.044	−.052*
	(.095)	(.089)	(.035)	(.031)
L.Seconds Correct		.404***		
		(.020)		
L.Overall Performance			.404***	
			(.020)	
L.Correct Answers (ln)				.236***
				(.014)
Constant	3.236***	2.487***	.285***	3.781***
	(.298)	(.279)	(.107)	(.116)
F	41.2	58	58	72.3
R Squared	.354	.32	.32	.388
N	7496	7491	7491	7499

*** p<0.01, ** p<0.05, * p<0.1

Notes: The coefficients are the result of estimating a time effects model for different measures of productivity on its lag and treatment dummies. This estimation includes as controls the observable characteristics of individuals such as gender, level of education, age, marital status and the ex-ante level of productivity.

Table 8: Treatment Effects on Productivity with Interactions of Treatments and Controls

	Seconds	Seconds Correct	Overall Performance	Correct (ln)
L.Seconds	.273***			
	(.018)			
Part time (1) Full Time (0)	−.039	−.085	−.034	.143***
	(.104)	(.097)	(.038)	(.034)
Non Flexible (0) vs Flexible (1)	−.249***	−.184**	−.073**	.376***
	(.082)	(.077)	(.030)	(.026)
Part x Flex	−.079	−.078	−.031	−.202***
	(.154)	(.143)	(.057)	(.050)
Gender (0 male, 1 female)	−.248***	−.177***	−.070***	.108***
	(.062)	(.058)	(.023)	(.020)
Dependents	−.462***	−.358***	−.141***	.190***
	(.081)	(.076)	(.030)	(.026)
High Productivity	−.381***	−.302***	−.119***	.038
	(.086)	(.080)	(.032)	(.028)
Dummy Children	.171**	.108	.043	−.081***
	(.072)	(.067)	(.027)	(.023)
Age	−.040***	−.030***	−.012***	.011***
	(.009)	(.008)	(.003)	(.003)
Years of experience	.015*	.009	.003	−.005*
	(.008)	(.008)	(.003)	(.003)
Educ: Graduate	1.010***	.811***	.320***	−.389***
	(.196)	(.183)	(.072)	(.063)
Educ: Undergraduate	.488***	.434***	.171***	−.258***
	(.080)	(.075)	(.030)	(.026)
Educ: Vocational	.431***	.398***	.157***	−.212***
	(.082)	(.077)	(.030)	(.026)
Single	.072	.112	.044	−.052*
	(.095)	(.089)	(.035)	(.031)
L.Seconds Correct		.404***		
		(.020)		
L.Overall Performance			.404***	
			(.020)	
L.Correct Answers (ln)				.236***
				(.014)
Constant	3.236***	2.487***	.285***	3.781***
	(.298)	(.279)	(.107)	(.116)
F	41.2	58	58	72.3
R Squared	.354	.32	.32	.388
N	7496	7491	7491	7499

*** p<0.01, ** p<0.05, * p<0.1

Notes: The coefficients are the result of estimating a time effects model for different measures of productivity on its lag and treatment dummies. This estimation includes as controls the observable characteristics of individuals such as gender, level of education, age, marital status and the ex-ante level of productivity.