Management Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Does the Unemployment Benefit Institution Affect the Productivity of Workers? Evidence from the Field

Mariana Blanco, Patricio S. Dalton, Juan F. Vargas

To cite this article:
Mariana Blanco, Patricio S. Dalton, Juan F. Vargas (2016) Does the Unemployment Benefit Institution Affect the Productivity of Workers? Evidence from the Field. Management Science

Published online in Articles in Advance 15 Aug 2016

. http://dx.doi.org/10.1287/mnsc.2016.2511

MANAGEMENT SCIENCE
Articles in Advance, pp. 1–18
ISSN 0025-1909 (print) | ISSN 1526-5501 (online)

http://dx.doi.org/10.1287/mnsc.2016.2511

Does the Unemployment Benefit Institution Affect the Productivity of Workers? Evidence from the Field

Mariana Blanco
Facultad de Economía, Universidad del Rosario, Bogotá, Colombia, mariana.blanco@urosario.edu.co

Patricio S. Dalton
Economics Department, CentER and CAGE, Tilburg University, 5037 AB, Tilburg, Netherlands, p.s.dalton@uvt.nl

Juan F. Vargas
CAF-Development Bank of Latin America, Bogotá, Colombia; and Facultad de Economía, Universidad del Rosario, Bogotá, Colombia, juan.vargas@urosario.edu.co

This paper studies the effects of unemployment benefit schemes on individual productivity. We created employment and unemployment in the field and compared workers' productivity under no unemployment benefits to productivity under two different unemployment schemes. In one scheme, the unemployed received an unconditional monetary transfer. In the other, the monetary transfer was obtained conditional on the unemployed spending some time on an ancillary activity. Our results challenge the standard economic theory prediction that unemployment benefits, especially unconditional compensations, hinder workers' effort. We find that workers employed under the unconditional scheme are more productive than workers under the conditional one, and both schemes make workers more productive than having no unemployment benefit. We discuss two possible explanations for our results based on reciprocity and differential psychological costs of unemployment across unemployment benefit schemes.

Data, as supplemental material, are available at http://dx.doi.org/10.1287/mnsc.2016.2511.

Keywords: unemployment benefits; productivity; psychological cost of unemployment; reciprocity
History: Received March 21, 2014; accepted January 27, 2016, by John List, behavioral economics. Published online in *Articles in Advance* August 15, 2016.

1. Introduction

This paper uses data from a novel empirical design to investigate whether and how the type of unemployment benefit system affects the productivity of workers. According to the International Labor Organization (ILO 2010), 42% of countries worldwide have some type of statutory unemployment benefit institution (UBI) that differs across countries. While most Western European nations and some developing countries like Colombia provide *unconditional* monetary benefits to the unemployed, others like Argentina, Australia, and Brazil offer benefits only *conditional* on the unemployed performing some kind of work in return. Yet a third set of countries like the United States, the United Kingdom, and Chile offer both conditional and unconditional-type programs.[1]

Despite being widely implemented through the world, little is known about the impact of these institutions on worker's productivity.[2] From a purely economic perspective, one would expect the productivity of workers under unconditional systems to be the lowest, simply because the cost of being unemployed is lower than under conditional schemes and lower than under no compensation at all. However, whether this prediction finds empirical support is still an open question that this paper aims to address.

Investigating this question using observational data is challenging for several reasons. To name only a few, the implementation of different UBIs is likely to be correlated with country-specific institutional characteristics that may affect workers' productivity through other channels. Moreover, obtaining independent and

[1] See Ochel (2005) for a survey on conditional unemployment programs (also called "workfare programs") and Gueron (1990) for a discussion on the two types of schemes.

[2] There is existing research on the effect of different variations of unconditional schemes (e.g., different sizes of the monetary benefit, different application requirements, different maturities of the benefit) on unemployment duration, search effort, and reservation wages (e.g., Abbring et al. 2005, Boone et al. 2007). However, to our knowledge, there is no empirical evidence on the effects of unemployment benefits on productivity.

objective measures of individual-level productivity is a difficult task, as is finding the right counterfactual for assessing how these measures are affected by the specific UBI in place.

For these reasons, we developed a novel empirical strategy to study this issue. In a nutshell, we created employment and unemployment in the field for a one-month period. We employed over 300 research assistants (RAs) and recorded an objective measure of their productivity under different unemployment schemes on a daily basis. The RAs were students from two major private universities in Bogotá, Colombia (University of Rosario, hereafter University A, and University of Los Andes, hereafter University B). Their job was to code into a computer spreadsheet news items on local politicians taken from the online archives of Colombian newspapers. Individual productivity was measured by counting the number of news items correctly codified. Before starting the job, the RAs were informed that the demand for coders would vary from day to day. Some days there would be a shortage of vacancies and those with lower productivity would become temporarily unemployed. Some other days there would be an increase in the demand for coding, allowing some of the unemployed to return to work.

Importantly, we did not randomize the treatment (i.e., UBIs) at the individual level. Such design was purposefully ruled out because having different RAs working under different conditions within the same university, while making the subjects more comparable across treatments, would have generated contamination inasmuch as peers or classmates are hired simultaneously to do the same job under different working conditions. This could have raised suspicion and undermined the experiment.[3]

For identification purposes, instead, we implemented the study in two separate stages that took place within a one-year difference and with different subjects, but with the task described being identical in both stages. In one stage, we ran simultaneously two identical *Comparison* interventions in each university.[4] This constituted the benchmark of our analysis and resembled a system in which no UBI is in place. Indeed, the unemployed in this stage received no monetary compensation, independent of their university. In another stage each university was randomly assigned to one UBI, with all the employees within each university receiving the same treatment. Employees in University A were assigned to the *Unconditional UBI* treatment; hence, if they became unemployed, they were given an unconditional monetary compensation, equivalent to 30% of the daily salary of their employed peers. Employees in University B were assigned to the *Conditional UBI* treatment; hence, the unemployed received an equivalent (30% of daily salary) compensation in exchange for spending 30 minutes (one third of a coding shift) doing an administrative task.

We identify the effect of unemployment benefit on productivity by comparing the productivity under no unemployment benefits to that under the two UBI schemes. We then contrast productivity under unconditional and conditional benefit payments by comparing the productivity differential of the two groups based in University A (*Unconditional UBI* and *Comparison*) with the differential productivity of the two groups based in University B (*Conditional UBI* and *Comparison*).

We show that productivity does not differ across universities in the *Comparison* treatment. However, once the different compensation schemes are in place, we find that the productivity under both unemployment benefit schemes is higher than under no unemployment benefits. Moreover, the productivity gain of the employees working under the *Unconditional UBI* scheme more than doubles that of their counterparts working under the *Conditional UBI* scheme. This suggests that an unconditional unemployment benefit stimulates rather than discourages workers to exert effort.

These results are at odds with the conventional economics perspective. Because effort is costly, when facing the possibility of being laid off, people should try to avoid unemployment more under a conditional UBI than under an unconditional one. Additionally, the highest effort should be observed when there is no unemployment benefit in place, since a laid-off worker would receive nothing. However, we observe exactly the opposite, and we show that the results are not driven by differential attrition across treatments or by the differential reaction to the different UBIs of subjects with initially unbalanced characteristics across universities.

We explore the potential mechanisms that could explain our results. One candidate explanation is that workers reciprocate with higher effort the unemployment compensation they are offered. Another possibility is that individuals facing unemployment

[3] According to the taxonomy proposed by Harrison and List (2004), our empirical design fits the characteristics of a natural field experiment in that we created a controlled environment in which subjects naturally undertook certain tasks and did so without knowing that they were taking part in an experiment. However, since the treatment was not randomly assigned at the individual level (while it was at the university level), we call our intervention a quasi-field experiment. In what follows and for simplicity we will refer to our intervention simply as the "intervention" or the "experiment."

[4] Due to the nonrandomized nature of our study across individuals, hereafter we explicitly avoid the word "control" and refer to the benchmark with no UBI as the "comparison."

experience psychological costs that differ across treatments, and workers internalize such costs. Although our data do not allow us to prove or disprove either of these two potential channels, we believe that both can play an important role in explaining the effects of unemployment benefits, and we leave the analyses of their relative importance open for future research.

Our paper contributes to several strands of the literature. First, while the literature on incentives and workers' productivity studies the effect of incentive schemes provided within the working place (Prendergast 1999; Bandiera et al. 2005, 2007), we show how productivity responds to more general institutional arrangements.

Second, we contribute to the literature on the labor market consequences of unemployment benefits (see Fredriksson and Holmlund 2006 for a review). For instance, we complement Acemoglu and Shimer (2000), who study how UBIs affect ex ante productivity by determining the type of jobs for which the unemployed apply. In contrast, we show how UBIs can affect productivity ex post, once the job has been chosen.[5]

Third, from a methodological point of view, this paper illustrates the possibility of using field or quasi-field experiments to study different questions concerning the dynamics of unemployment.[6] The closest related paper in this respect is Black et al. (2003), who use random assignment to the unemployment insurance system to show that the threat of being forced to take up reemployment training services reduces the duration of the benefit. Falk et al. (2006) run a laboratory experiment on the effect of UBIs on labor supply (instead of productivity). In their experiment, the job of "employees" is counting the number of zeros in a zeros-and-ones matrix printed on a sheet of paper. An important feature of our design is that we are able to precisely measure workers' productivity without departing from a natural job setting.

Finally, to the best of our knowledge, this is the first paper to look at the effect of different types of unemployment benefits on productivity, a problem of first-order importance that is greatly unexplored. As such, this paper helps to pave the road to answering important policy questions that have been studied very little but that are highly relevant.

The rest of this paper is organized as follows. Section 2 introduces the empirical strategy. Section 3 describes the sample and the data collected. Section 4 analyzes the data, Section 5 presents the results, and Section 6 discusses potential underlying mechanisms behind the results. Section 7 concludes.

2. Empirical Strategy

Before describing the strategy designed to study the effect of different UBIs on workers' productivity, it is useful to briefly outline the predictions of the standard labor economics model. See Online Appendix A (available as supplemental material at http://dx.doi.org/10.1287/mnsc.2016.2511) for a simplified formal version of this model, which we use to derive the hypotheses we test in this paper.

Workers in our experiment receive a fixed wage independent of their effort so, in principle, there are incentives to shirk. However, the probability of becoming unemployed depends on workers' productivity, so unemployment should act as a disciplinary device for workers. This idea, originally introduced by Shapiro and Stiglitz (1984), essentially implies that workers' exerted effort depends positively on the cost of becoming unemployed: the higher the cost, the lower the incentives to shirk and the higher the average expected productivity. This implies that workers employed under an unconditional UBI should be less productive than workers employed under a conditional UBI or under no UBI.

To test these predictions we developed an experimental design that created employment and unemployment in a controlled natural environment. We observe workers' productivity in three different settings: one featuring an unconditional UBI, another one with a conditional unemployment compensation, and a third one with no unemployment benefit scheme in place. In this section we summarize the most important aspects of our design and refer the interested reader to Online Appendix B, where we explain the details of every aspect of the intervention.

We employed over 300 RAs for a one-month period in two private universities located in Bogotá, Colombia, University of Rosario (University A) and University of Los Andes (University B).[7]

The job was to code news items on local politicians from the online archives of the two main Colombian

[5] Besley and Coate (1992) develop a theoretical model to compare unconditional- versus conditional-type schemes. Their focus is, however, different from ours, as they study the optimal design of conditional programs considering that work requirements may serve as screening and deterrence devices.

[6] See List and Rasul (2011) for an overview of the use of field experiments in labor economics.

[7] Each participant was assigned to 1 of 18 groups, according to their availability of time and their university. Subjects who indicated availability in more than one slot were randomly allocated to one of these slots. Once subjects were assigned to a group, they were not allowed to change it during the entire period of the project. Subjects could not work from home. The distribution of groups was balanced across days, time of the day, and the newspaper to code. Each group would have a work shift of 1.5 hours per day, two days per week, during four weeks. See Table B.1 in Online Appendix B.3 for a detailed description of the groups.

newspapers. We recorded daily individual productivity, measured as the number of news items correctly coded.[8]

Workers were exposed to both negative and positive employment shocks, the sequence and timing of which were unknown. Their productivity on the last day they worked determined whether they would become unemployed or remain employed (in the face of a negative shock), or be reemployed (in the face of a positive shock).[9]

Unemployed subjects were assigned to one of three treatments: *Unconditional UBI*, *Conditional UBI*, or *Comparison*. In the *Unconditional UBI* treatment, unemployed workers received one third of the daily salary in the form of an unconditional cash transfer. In the *Conditional UBI* treatment, the unemployed were required to work one third of the daily shift in order to receive the benefit. Finally, those assigned to the *Comparison* treatment did not receive any monetary benefit.[10]

For identification purposes we implemented the study in two separate stages, with the work involved in each stage being identical. In one stage, we ran simultaneously the two identical *Comparison* treatments in each university. This constituted the benchmark of our analysis and resembles a system in which no UBI is in place. In another stage we implemented the two UBI treatments with compensation, one in each university and with all the employees within each university receiving the same treatment. The *Unconditional UBI* scheme was implemented in University A and the *Conditional UBI* scheme in University B. To avoid spillover and contagion between subjects assigned to different treatments within the same university, we implemented the UBI interventions with compensation (i.e., the *Unconditional UBI* and the *Conditional UBI*) and the *Comparison* treatment with a one-year difference.[11]

This design allows us to overcome several potential confounders that are intrinsic to alternative approaches. For instance, a random allocation of the UBI at the individual level within universities, while probably making the subjects more comparable across treatments, would have generated spillovers and contamination inasmuch as peers or classmates from the same university are hired simultaneously to do the same job under different working conditions.[12] This within-school contamination is the same reason that other experiments have proposed treatment allocation at a higher unit of aggregation (e.g., Miguel and Kremer 2004).

By allocating subjects to treatments across universities, we may in principle lose comparability. However, as it can be seen in Table 1, the two populations are very similar in terms of observable characteristics, with the exception of gender and socioeconomic status.[13] Most importantly, by carrying out the same *Comparison* treatment in each university, we can control for the unobserved heterogeneity among students, to the extent that it is not correlated with the response to treatment incentives.[14] Another alternative could have been to expose the same subject to different UBIs at different moments. Although this type of design has proven to be instrumental in other labor market field experiments (e.g., Bandiera et al. 2007), in our particular setup it would have created a confusing and excessively artificial setting, thus probably introducing noise in the data.

To recruit participants, we sent an email to all social sciences students announcing the opening of temporary RA positions to work for a research group that we created for the purpose of this study. The job announcement stated that, during the one-month duration of the project, the number of vacancies per day would vary and that their daily employment status would depend on their work performance and on the shifting labor demand. These two factors would in turn determine the final pay, and applicants were told to expect a final payment varying between US$32 and US$90. The interested candidates had to fill in a preformatted CV online, which we used to gather sociodemographic information such as age, gender, residential address (from which we inferred socioeconomic status), GPA, current job, etc. Importantly, we did not mention in the open call the specific treatments (the unemployment benefit schemes) to

[8] In Online Appendix B.2 we explain how we controlled that the news items were correctly coded.

[9] Since we are interested in testing the effect of UBIs on productivity, we needed to make the probability of becoming unemployed dependent on effort. If the probability of becoming unemployed were independent on effort, there would not be a clear reason for different UBIs to affect productivity.

[10] The unemployment benefit was introduced in the most natural way possible. Employees were told that, given their commitment to work for the whole month, the unemployed were going to be offered the possibility to receive a payment for the days they remained unemployed.

[11] Students who were RAs in the first year were not eligible to work in the second year.

[12] Think of a situation in which two classmates meet and find out that they are both RAs in the same project, but one has been offered money unconditionally when unemployed while the other has to do some costly activity as a compensation for the money. Arguably, this would create an important confusion among subjects that with sufficient propagation could have jeopardized the experiment or, at least, generated all types of noise in our data.

[13] We address the concerns regarding this imbalance in Section 6.1.

[14] We compare the productivity differential of the two groups based in University A (*Unconditional UBI* and *Comparison*) with the productivity differential of the two groups based in University B (*Conditional UBI* and *Comparison*).

Table 1 Baseline Sample Differences Across Universities and Treatments

	Within University A			Within University B			Across universities	
	Unconditional UBI (1)	*Comparison* (2)	Difference (1) − (2)	*Conditional UBI* (3)	*Comparison* (4)	Difference (3) − (4)	*UBI* (1) − (3)	*Comparison* (2) − (4)
				Panel A: Individual characteristics				
Gender (1 = female)[a]	0.63	0.58	0.05	0.45	0.40	0.06	0.18**	0.19*
	(0.06)	(0.07)	(0.09)	(0.04)	(0.07)	(0.08)	(0.08)	(0.10)
Age	21.30	21.13	0.17	20.94	20.56	0.38	0.36	0.56*
	(0.25)	(0.23)	(0.34)	(0.15)	(0.23)	(0.29)	(0.28)	(0.33)
Socioeconomic status[b]	3.63	3.76	−0.13	4.50	4.34	0.16	−0.86***	−0.58**
	(0.12)	(0.12)	(0.17)	(0.09)	(0.16)	(0.18)	(0.16)	(0.20)
GPA (1–5 scale)	3.95	3.94	0.01	3.98	3.91	0.07	−0.03	0.03
	(0.07)	(0.04)	(0.09)	(0.03)	(0.04)	(0.06)	(0.07)	(0.06)
Currently working (1 = yes)[a]	0.08	0.05	0.03	0.09	0.17	−0.07	−0.01	−0.11*
	(0.04)	(0.03)	(0.05)	(0.02)	(0.05)	(0.05)	(0.04)	(0.06)
Major (= 1 econ. related)[a]	0.68	0.55	0.14	0.69	0.56	0.13	−0.01	−0.02
	(0.06)	(0.07)	(0.09)	(0.04)	(0.07)	(0.08)	(0.07)	(0.10)
				Panel B: Baseline psychological characteristics				
Self-esteem (day 0)	35.83	34.80	1.04	35.76	35.7	0.06	0.07	−0.91
	(0.41)	(0.52)	(0.65)	(0.27)	(0.45)	(0.53)	(0.49)	(0.70)
Job satisfaction (day 1)	24.19	24.20	−0.02	23.85	24.13	−0.28	0.34	0.07
	(0.46)	(0.63)	(0.77)	(0.33)	(0.51)	(0.65)	(0.59)	(0.82)
				Panel C: Initial productivity (day 1)				
All sample	29.92	18.13	11.79***	28.80	18.29	10.51***	1.11	−0.16
	(2.28)	(1.44)	(2.74)	(1.36)	(1.98)	(2.61)	(2.57)	(2.41)
Dropouts	17.22	12.12	5.10	22.96	14.50	8.46**	−5.74	−2.38
	(3.80)	(1.27)	(3.08)	(1.81)	(3.16)	(3.49)	(3.87)	(3.35)
Finalizers	32.2	23.13	9.07**	30.30	22.08	8.22**	1.9	1.05
	(2.48)	(2.02)	(3.56)	(1.62)	(2.21)	(3.67)	(2.94)	(3.00)
Difference	−14.98**	−11.01***	−3.96	−7.34**	−7.58*	0.24	7.64	3.43
(*Dropouts* − *Finalizers*)	(6.08)	(2.50)	(6.18)	(3.34)	(3.85)	(5.48)	(6.79)	(4.49)

Notes. Variables are defined as follows. *Gender*, *Age*, and *GPA* are self-explanatory. For *Socioeconomic status*, see Section 3, in particular Footnote 29. *Currently working*: Answer to the question "Are you currently working?" was scored 1 if yes and 0 if no. *Major*: Answer to the question "What is your major area of study?" was scored 1 if economics, business, or finance and 0 if any other of the social sciences. For *Self-esteem*, see Section 3, in particular Footnote 30. For *Job satisfaction*, see Section 3, in particular Footnote 31. *Finalizers* were subjects who remained in the project until the last day. *Dropouts* were subjects who dropped out of the study at some point between the first and the last working days. For *Productivity*, see Online Appendix B.2.

[a]Test of difference in means using χ (Pearson's χ^2) test because variable is dichotomous.

[b]Test of difference in means using Fisher's exact test because variable is categorical. The rest of the mean differences are checked using the t-test.

*Significant at 10%; **significant at 5%; ***significant at 1%.

which individuals would be assigned. This was only explained later, during the induction day. By doing this we ruled out any potential self-selection of applicants into treatments.

Coders' daily productivity was measured by the number of news items correctly coded per day. The quality of their work was controlled by our research assistants, and the workers knew this.[15] After controlling for quality, the assistants would record the true productivity of each coder and create a productivity ranking of coders per group and day. These rankings were used to determine who would become unemployed (reemployed), in case of a negative (positive) shock. Importantly, coders did not have access to these rankings until the last day of work.

2.1. Sequence of the Study and Payment

The first session was exclusively an induction day, when we explained the characteristics of the job and the contractual conditions. During the induction, coders did not work. Those who accepted the terms and conditions of the job signed a consent form and were officially hired.[16] Thereafter, each working day lasted 90 minutes in total: 70 minutes was used for

[15] In principle, coders had incentives to cheat by adding fake entries into their files. To avoid this, we assigned a team of six research assistantss to control the quality of the entries. On the induction day, coders were informed that the quality of their work was going to be monitored. Every afternoon after the working day, our assistants would randomly select 15% of each coder's entries (20% if there were 100 or fewer entries for a single coder) and control their quality following a common evaluation protocol. They would check, for example, if the entry came from real news or was invented, if the data entered were consistent with the information in the news, etc.

[16] Only one subject did not sign the informed consent and, hence, decided not to take the job.

Table 2 Sequence of Unemployment Shocks

Day	1	2	3	4	5	6	7	8
Shock		−	+	+	+	−	+	+
% unemployment rate	0	**60**	55	45	35	**60**	50	0

Note. Bold type denotes negative employment shocks.

coding and the rest of the time was used to save the file, send it to us by email (from the work account we specially created for each participant), and fill in some questionnaires.[17] Each session was supervised by two research assistants randomly allocated across groups. The research assistants would make sure that the sessions ran smoothly. They would set up the computers, record attendance, answer questions about coding, administer questionnaires, etc.

We created involuntary unemployment (shortage of vacancies) on all days except for the first and last days, when we had full employment. Unemployment increased on the second and sixth days and decreased (i.e., vacancies were opened) during the rest of the project. Table 2 shows the unemployment rate for each working day. The rate is 0 on the first and the last working days as we were interested in studying the behavior of subjects both in an environment of unemployment and in one in which there was no involuntary unemployment, and no competition for vacancies. The rest of the days featured positive unemployment rates that were rather high, varying from 35% to 60%.[18]

On the days in which there was a negative shock (i.e., an increase in the unemployment rate), the coders ranked last according to their productivity on the previous day became unemployed. In turn, in cases in which there was a positive demand shock (i.e., decrease in the unemployment rate), unemployed subjects were informed via email of the number of job vacancies to be opened the next day. If they wanted to apply to fill one such vacancy, they had to reply to that email stating so, to be considered eligible.[19] New vacancies were filled according to applicants' productivity on their last day of employment. At the beginning of every working day, two research assistants would receive the subjects, check their respective coder's ID, and inform them about their condition as employed or unemployed for that day. To be considered as part of the labor force, all subjects had to attend their working session on time, even those who were unemployed or had applied for a vacancy. If they failed to do so (without any justified cause, such as illness), they were considered to have withdrawn from the project.

The remuneration per 90 minutes of work was the market rate for research assistants in the universities where the intervention took place. We paid subjects COP$14,250 (approximately US$8).[20] This salary was over four times the minimum wage in Colombia for such a working shift. In addition, there was a bonus of COP$40,000 (approximately US$22) for the best coder of each group, i.e., the participant who codified the largest number of correct entries in any given session. The maximum possible remuneration was US$ 90, earned by subjects who came to the induction day (US$6), were employed during the eight working days (US$8 × 8 days = US$64), and won the bonus for best coder of that group (US$20). The minimum possible remuneration of someone who participated in the project over the whole period was US$32 if the person was assigned to a treatment with a UBI and US$14 if that person was assigned to the *Comparison*. To minimize attrition, we paid on the last day of the project, and subjects who withdrew from the project were entitled to request only half of the money earned by the last day they showed up to work.

2.2. Interventions

As mentioned above, we implemented three different interventions: the *Comparison* treatment, without any UBI; and two treatments designed to capture the most important aspects of the two main existing UBIs, an *Unconditional UBI* and a *Conditional UBI*.

The amount of the unemployment monetary compensation was identical in the two UBI treatments and equivalent to one third of the daily shift salary of the employed. This is, of course, an abstraction of the real-world UBIs in which unemployment compensations vary, even within countries, according to several characteristics of the unemployed like the spell of unemployment. While we do not think that abstracting from this variation is relevant for our purpose, we did make an effort to replicate other features

[17] Subjects were informed that other researchers were interested in collecting supplementary data so, as part of the work activities, they would be asked to answer a set of questionnaires. They were told that answers to these questionnaires were going to be treated anonymously and would only be used for academic purposes.

[18] Note that, by making workers experience a high risk of being laid off, we are maximizing the chance that unemployment schemes have behavioral effects. We did this on purpose for our intervention to make sense and because we needed power to make inference on the behavior of both employed and unemployed individuals. We could have obtained such power by recruiting more people into the job, but we faced both budget constraints and a limited pool of potential workers. A detailed explanation of the choice of the magnitude and the sequence of unemployment rates is provided in Online Appendix B.4.3.

[19] The unemployed were free to choose whether to apply for a new vacancy or not. Applying was not a requirement for receiving the unemployment benefits.

[20] At the time of the study, the average hourly salary of a research assistant was COP$10,300 (approximately US$5.8).

of the actual implementation of UBIs. For instance, when applying for the benefit in reality, unemployed individuals have to fulfill some requirements, including signing up, filling in paperwork, and showing up regularly at some administrative office. We had a similar procedure in our two treatments: the unemployed who wanted to receive the benefit had to show up every working day at the place of work on time. If assigned to the *Unconditional UBI*, the unemployed could leave immediately after filling in some short questionnaires, which would grant them the benefit for that day (although payable at the end of the project).[21] If assigned to the *Conditional UBI*, the unemployed could choose whether to complete the ancillary task necessary to receive the compensation for that day or to leave, which meant turning down the compensation.[22] In any case the unemployed had to show up and sign an attendance sheet. Failing to do so implied quitting the job altogether and not being considered part of the labor force for the rest of the project (i.e., being unable to apply for a position when new vacancies were opened).

For comparability, unemployed coders in the *Comparison* treatment were also required to show up every working day at the time and place of work if they wanted to be considered part of the labor force. They were invited to leave immediately after filling in some short questionnaires. This process did not entitle them to any money.[23]

3. Data

3.1. Initial Sample and Attrition

Table 3 summarizes the sample size for the different treatments at the beginning and at the end of the study (day 1 and day 8).[24] We started the project with 306 subjects: 60 assigned to the *Unconditional UBI* scheme, 143 assigned to the *Conditional UBI* scheme, and 103 assigned to the *Comparison* group (55 of which were in University A and 48 of which were in University B).[25]

Since the study lasted a month, there was some attrition. Of the starters, approximately 19% in the two UBI treatments and 48% in the *Comparison* treatment dropped out between the first and the last day of activities.[26] This difference is not surprising, as the incentives to show up for the unemployed in the *Comparison* treatment were only to ensure their continuity in the project (i.e., to be able to apply for future open vacancies), but not to get any type of unemployment compensation. In the UBI treatments, instead, showing up would not only ensure the continuity of the unemployed in the project but also the (conditional or unconditional) monetary transfer. Nevertheless, as shown in the last row of Table 3, the attrition rate is very similar between the *Unconditional UBI* and *Conditional UBI* treatments, and between the *Comparison* group in University A and the *Comparison* group in University B.[27] This is important since, for identification, we will compare the productivity differential of the two groups based in University A (*Unconditional UBI* and *Comparison*) with the productivity differential of the two groups based in University B (*Conditional UBI* and *Comparison*). If attrition were significantly higher for the group of coders assigned to the *Unconditional UBI* scheme, and the remaining employees were the most productive, differential attrition would then confound our results.[28]

[21] With the questionnaires we measure various psychological traits that we use to test the mechanisms behind our main result.

[22] The task, which lasted 30 minutes, was to place letters in envelopes and paste on them randomly assigned (real) postal addresses. We manipulated the content of the letter in two subtreatments. A *Charity* subtreatment used letters that asked for donations to a social NGO. A *Placebo* subtreatment used letters to publicize the business of a local private firm. While the existence of these subtreatments within the *Conditional UBI* treatment does not affect the incentives exploited for this paper, these data will be used in companion papers.

[23] We acknowledge that asking the unemployed in the *Unconditional UBI* and *Comparison* treatments to show up every working day may be too stringent. However, we decided to impose this rule to make the three treatments comparable. If we hadn't implemented the same rule in all treatments, the additional cost of receiving compensation under the *Conditional UBI* scheme would not only have been the 30 minutes of activity, but also going to the place of work. This would have created an additional variation in the treatment. Our design solves this concern by imposing the same rule for the three treatments, at the cost of making the situation somewhat less realistic.

[24] Table B.2 in Online Appendix B.4.3 shows the daily counterpart of Table 3, reporting the number of subjects who are employed, working, or who dropped out of the study, each day and in each treatment.

[25] As already mentioned, in the *Conditional UBI* treatment we implemented two different subtreatments that had to do with the contents of the letters to be put in envelopes. This explains the oversampling in that treatment.

[26] They did so by not showing up on time to the place of work on at least one working day, regardless of their employment status.

[27] These differences are in fact not significant. The p-value of the t-test of mean differences in the attrition rate across universities is 0.38 for the UBI treatments and 0.67 for the *Comparison* group.

[28] Panel C of Table 1 confirms that attrition is negatively correlated with initial productivity. Coders who eventually dropped out through the intervention have a lower first-day productivity than coders who remained in the project until the last day, and this is true in every treatment. This rules out the possibility that the observed productivity differences across treatments are driven by the fact that dropouts in the *Conditional UBI* scheme are the most productive workers and dropouts in the *Unconditional UBI* scheme are the least productive ones. On the contrary, the observed differences in productivity suggest that attrition should mechanically increase the average productivity in the *Comparison* groups, which goes against our findings.

Table 3 Recruitment and Attrition

	UBIs		Comparison		
	University B *Conditional UBI*	University A *Unconditional UBI*	University B *No UBI*	University A *No UBI*	Total
Day 1	143	60	48	55	306
Day 8	114	51	24	30	219
Attrition rate (%)	20	15	49	45	28

Notes. The first two rows report the number of subjects in each category. The attrition rate is the share of subjects who dropped by day 8 compared to the initial labor force on day 1.

It is also worth noting that the exogenous daily unemployment rate (described in Table 2) is computed over a sample of workers net of dropouts. In other words, the higher attrition in the *Comparison* treatment does not reduce the probability of the remaining workers to become unemployed. Hence, differential attrition cannot explain the observed results by reducing differentially the effort of the subject assigned to the *Comparison* treatment relative to that of the subjects assigned to either UBI. We explore further the role of attrition in Section 5 to conclude that attrition does not drive our results.

3.2. Sample Differences

We start by looking at average baseline individual characteristics across treatments both for the entire sample as well as for the sample of dropouts. Although we show that there is balance across treatments in terms of several characteristics, a few differences remain. In the empirical analysis we deal with the remaining imbalance in several ways (see Section 6.1).

Table 1 reports the sample composition in terms of baseline individual characteristics, psychological measures, and starting productivity. The information about the individual characteristics (panel A) was gathered from the preformatted CV that applicants had to fill out online in order to apply for the job. The psychological measures of self-esteem and job satisfaction (panel B) were collected from the questionnaires that the participants filled during the induction day (called day 0 in Table 1) and during the first day of work (day 1), respectively. The initial productivity (panel C) is that recorded on the first day of work for each individual.

The students recruited in each university were quite similar in many respects. We observe initial differences between the two groups in terms of gender, age, whether the subject currently has another job, and average socioeconomic status of the coders.[29] The sample from University A displays a larger prevalence of female coders relative to University B in both the *Comparison* stage and the UBI stage, and coders are slightly older but have a lower probability of having another job in the *Comparison* stage only. In turn, coders from University B are relatively wealthier in both stages. In contrast, we see no initial differences across universities in terms of average GPA or whether the major of study is related to economics (i.e., economics, finance, or business) versus other social sciences in either treatment. Importantly, within universities there are no significant differences in any of the individual characteristics across stages, which provides support to our empirical strategy, explained in Section 2.

As indicated by panel B of Table 1, we do not observe individual differences across universities in terms of the psychological traits that we collected at the start of the study. We focus on two psychological variables: self-esteem and job satisfaction. Self-esteem was measured using a validated Spanish version of the Rosenberg (1965) self-esteem scale (RSES), one of the most widely used self-esteem measures in social science research, and it indicates subjects' overall evaluation of their own worthiness or value.[30] To measure job satisfaction, we used a modified version of the Macdonald and MacIntyre (1997) job satisfaction scale.[31]

[29] Neighborhoods in Bogotá have a score from 1 to 6 (called strata) which is used to price-discriminate the tariff charged for public services. People who live in strata 5 and 6 (and to some extent 4) subsidize the utilities of those living in strata 1 and 2. Real estate and rent prices are positively correlated with this stratification. Therefore, the income level of a household is highly correlated to the stratum of its neighborhood, hence, the use of this proxy in the absence of objective household income or consumption data. Importantly, the stratum is not self-reported. On the contrary, in the preformatted CV that applicants completed online, they reported their address of residence. We geocoded each address to determine the respective stratum.

[30] The RSES is a compound of ten items answered on a four-point scale ranging from strongly agree to strongly disagree. The Spanish version of the RSES scale was validated by Martin-Albo et al. (2007) and has been shown to have satisfactory levels of internal consistency and temporal stability in a population of Spanish university students. We only changed a few words so that the scale better matched Colombian use of the Spanish language.

[31] We shortened and adapted this scale to the specificities of the job. The resulting scale is compounded of six items, answered on a five-point scale ranging from strongly agree to strongly disagree.

In panel C of Table 1, we look at differences in day 1 productivity. Average initial productivity does show significant differences across treatments within university, but no differences between universities. The difference across treatments within each university is not surprising, since productivity is an outcome and hence the gap can be attributed to the exposure to different unemployment benefit schemes. Panel C also shows that attrition is negatively correlated with initial productivity, as already mentioned in Section 3.1. Coders who eventually dropped out through the intervention have a lower first-day productivity than coders who remained in the project until the last day, and this is true in every treatment.

The fact that the students recruited in both universities are overall very similar in terms of a large number of characteristics is not surprising since both universities are private, offer roughly the same majors within the social sciences, target students from the same backgrounds, and are among the best universities in Colombia.

Table C.1 in Online Appendix C is equivalent to Table 1, but it is limited to the sample of subjects who dropped out of the study at some point between the first and the last working days. Sample differences for dropouts are very similar to the differences shown for the whole sample. There is a relatively large balance of subjects' characteristics across treatments with the exception of age, socioeconomic status, and whether the student currently has another job. The fact that dropouts differ across some dimensions further justifies the inclusion of these as controls in the empirical analysis, to account for potential selection caused by unevenly distributed attrition.

4. Analysis

The aim of this paper is to evaluate whether and how UBIs in general, and the type of UBI in particular, affect the productivity of workers. Our measure of productivity is the number of news items correctly coded by each subject in each working session. Figures 1(a) and 1(b) show the daily average productivity by treatment stage (i.e., *Comparison* or UBI) across universities. In turn, Figures 2(a) and 2(b) show the daily average productivity by university across treatments.

The first salient feature is that workers' productivity is higher with unemployment benefits than without them. This happens in both universities, suggesting that an unemployment benefit, regardless of its type, increases workers' productivity. Second, Figure 1(a) shows that productivity in the *Comparison* stage does not differ significantly across universities. If anything, in most working days of the *Comparison* stage, the average productivity of the employees in University B is slightly higher than in University A. (This difference is, however, not significant.) But, as shown in Figure 1(b), in the *UBI* stage there is a swap in the productivity advantage on most working days, and coders of University A (who worked under the *Unconditional UBI* scheme) are significantly more productive than their peers at University B (who worked under the *Conditional UBI* scheme). Indeed, the average productivity of the entire working period in the *Comparison* stage is approximately eight news items coded less per day in University A compared to University B. In contrast, the average productivity of the entire working period in the UBI stage is approximately 16 news items coded more per day in University A (*Unconditional UBI*) compared to University B (*Conditional UBI*). This observation, which is robust to looking at the entire sample or at the sample net of dropouts, implies that the productivity gain obtained in the *Unconditional UBI* scheme is higher (24 news items or so) than that obtained in the *Conditional UBI* scheme, a finding that is more striking given the fact that the average productivity of the first day of work does not significantly differ across universities in either stage (see panel C of Table 1). The fact that the productivity difference across universities in the *Comparison* stage is small and not significant reassures us that the difference observed between the *Unconditional UBI* and *Conditional UBI* schemes is not driven by differences in characteristics between students of University A and University B.

A third noteworthy aspect of the figures is that productivity is very similar across treatments on both the first day of work (day 1) and the last day of work (day 8). Recall that on the first day nobody had experienced unemployment (i.e., the actual effect of the treatments had not been experienced) and, on the last day there was no possibility of becoming unemployed anymore. This suggests that the observed differences in productivity reported in this paper are due to treatment effects when the possibility of becoming unemployed was present and clear.

Finally, in all figures there is a spike on day 6. This spike can be explained by a pure mechanical effect. Recall that, together with day 2, day 6 is the day with the highest unemployment rate during the project: 60%. However, in contrast to day 2, on day 6 most of the attrition had already taken place. This implies that only the very best coders worked on day 6, significantly raising the mean productivity on that day. Importantly, note that the productivity spike happens in all treatments (unconditional, conditional, and *Comparison*); hence, it is not driven by a treatment-specific shock.

We now investigate further and more formally the main pattern emerging from the figures, namely that the presence of any UBI fosters productivity and that

Figure 1 Daily Average Productivity by Treatment Stage Across Universities

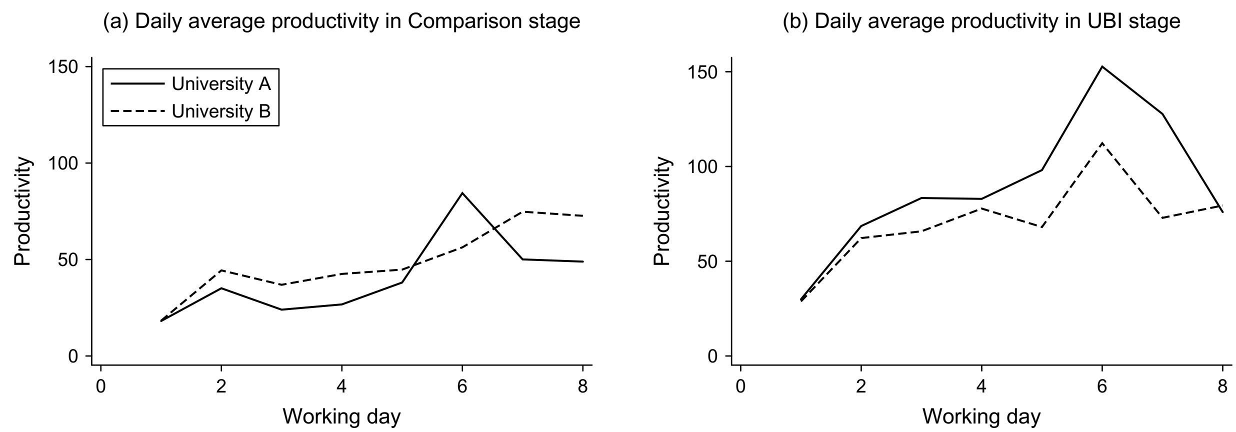
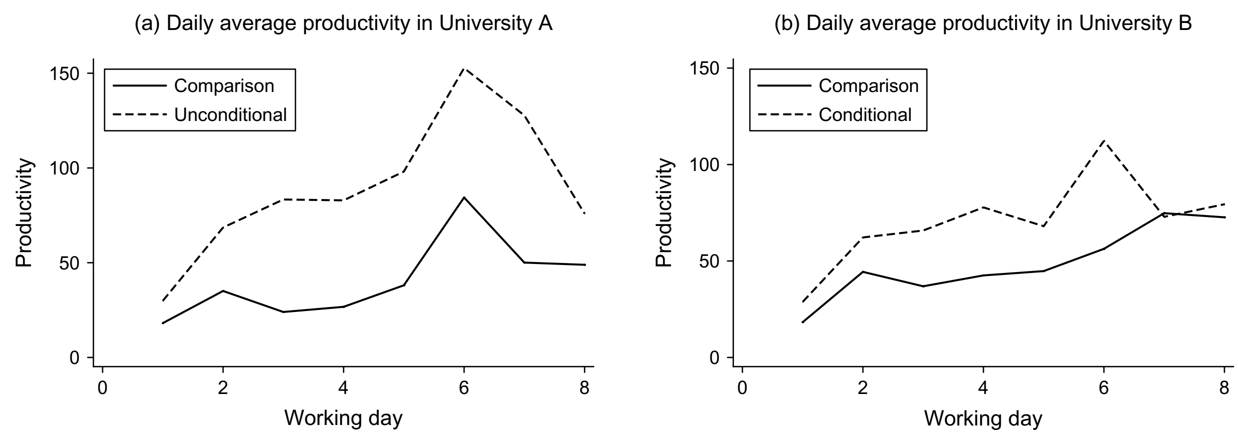


Figure 2 Daily Average Productivity by University Across Treatments



this stimulus is larger in schemes with an unconditional compensation. For this purpose, we estimate a regression model that, controlling for individual characteristics, allows us to identify both the differential productivity of each UBI relative to the *Comparison* treatment, as well as the additional productivity differential that stems from being under a particular UBI relative to the other. The latter is indeed our main interest: we want to shed light on which unemployment compensation scheme is more productivity enhancing and hence has potentially larger economic benefits.

Recall, however, that our study follows a between-subjects design; hence, each subject is either treated (by working under one of the two UBIs) or nontreated (by working under a scheme with no compensation). That is, to avoid spillovers and contagion between treated and comparison groups within each university, we do not have a panel of subjects who were first nontreated and then treated.[32] However, recall from Table 1 that the samples of coders are largely balanced both across universities and, within them, across stages. This is so in terms of important observable characteristics including, for instance, initial productivity.[33] Conditional on controlling for the characteristics that do seem to be slightly unbalanced, this ensures that any result we may find with this approach is solely driven by the effect of the UBI treatments. Pooling the data across individuals and working dates we estimate the following regression model:

$$y_{itus} = \alpha + \beta_1 UniversityA_u + \beta_2 UBI_s + \beta_3 (UniversityA \times UBI)_{us} + \delta' X_{itus} + \epsilon_{itus}, \quad (1)$$

where y_{itus} is the productivity of individual i on working session t, in university u, and experimental stage s; $UniversityA_u$ is a dummy variable that takes the value of 1 for subjects in University A (where both the

[32] Hence, while our specification resembles a difference-in-differences model, we do not claim to use a standard difference-in-differences empirical approach.

[33] In all the regressions we control for the characteristics that turned out to be significantly different across universities (i.e., gender, age, working conditions, and socioeconomic stratum).

Comparison and the *Unconditional UBI* treatments took place) and 0 for participants in University B (*Comparison* and *Conditional UBI* treatments); and UBI_s is a dummy variable that takes the value of 1 for subjects who participated in the UBI stage (where the *Unconditional UBI* scheme took place in University A and the *Conditional UBI* one in University B) and 0 for those in the *Comparison* stage. Note that the interaction between *UniversityA* and *UBI* picks up the subjects who where assigned to the *Unconditional UBI*.

The vector X_{itus} includes control variables such as gender, age, whether the subject has any other part-time job, socioeconomic stratum fixed effects, and day fixed effects (hence the t subscript). These fixed effects account, respectively, for any omitted characteristic that is common to all individuals with the same socioeconomic background, and for any aggregate shocks that may affect all subjects on any working day (including potential learning over time). One additional control we use is the size of the coding group that each subject encounters every working day. This is because the initial group sizes are unbalanced (see Table B.1), and moreover, if attrition affects groups differentially, then the resulting pool after employment shocks will differ, insomuch as the daily exogenous unemployment rate is the computed net of dropouts. In turn, the differential group size may affect the incentives for workers to make effort. Thus, controlling for the time-varying group size takes into account the tournament incentives in our design and allows us to estimate the sole behavioral effect of the UBI that the worker faces.[34]

5. Results

Table 4 reports the regression results of the estimation of Equation (1). In columns (1)–(3) the dependent variable is the productivity level (as measured by the number of correctly coded news items), and in columns (4)–(6) the dependent variable is the logarithm of the productivity. Hence, the estimated coefficients reported in the first three columns can be interpreted in terms of the number of news items coded, while those in the last three columns report percentage changes in productivity. As we do not necessarily favor one interpretation over the other, in the rest of the tables we just report absolute productivity numbers.[35]

To investigate the extent to which individual characteristics may confound the estimates, while columns (1) and (4) include no controls, columns (2), (3), (5), and (6) include the full set of individual controls in X_{itus}. The estimated coefficients are remarkably stable when the controls are included. To address any potential inference problem due to serial correlation, we report robust standard errors clustered at the group level in columns (1), (2), (4), and (5). However, as there are only 18 groups, which is a relatively small number for a cluster unit, columns (3) and (6) report bootstrap standard errors (1,000 repetitions) clustered at the group level on a regression model that includes all the controls.[36]

First, note that the estimated coefficient associated with the University A dummy (β_1 in the equation) is negative and small (and not significant when the dependent variable is the number of news items). This is equivalent to saying that the average productivity of coders across universities is very similar in the *Comparison* stage, with workers in University A being slightly less productive than their peers in University B.[37] Second, the estimated coefficient associated with the *UBI* stage dummy (β_2) is positive and significant, suggesting that, within University B, coders who worked in the UBI stage (i.e., under the *Conditional UBI* treatment) are approximately 23 news items (or 62%) more productive than their counterparts assigned to the *Comparison* treatment (19 news items, or 57%, once controls are added). Third, the increase in productivity in the UBI relative to the comparison stage is larger in University A. Indeed, workers under the *Unconditional UBI* coded approximately 24 news items (54%) in excess of the extra news items coded by workers under the *Conditional UBI* (β_3 in the equation). Note that the magnitudes and significance of these figures are robust to adding the day fixed effects and the individual controls, and significance is not compromised when standard errors are bootstrapped, even if their magnitude increases.

From a purely economic perspective, the observed productivity under the *Unconditional UBI* scheme should be the lowest among our experimental groups (see Hypothesis 1 of the formal model presented in Online Appendix A). Our results reject this hypothesis. In fact, the estimate of β_3 implies the opposite:

[34] As an alternative specification (not reported), we interact both the *UniversityA* and the *UBI* dummies, as well as their interaction, with the time-varying group size. This specification allows us to test whether the treatment effects are stronger for smaller groups, where competition is fiercer in the light of unemployment shocks. Results (available from the authors upon request) confirm that this is the case.

[35] Equivalent log-productivity tables can be provided by the authors upon request.

[36] We cluster at the group level because each subject was competing with the rest of participants within his or her group when it came to decide who would be laid off. If subjects talked to each other after a coding session, they could roughly estimate the effort-minimizing productivity level needed to keep the job.

[37] The constant can be interpreted as the average number of news items coded in University B during the *Comparison* stage, which in this case is equivalent to almost 42 news items (70 when controls are included). β_1 represents the additional news items coded in University A.

Table 4 The Effect of Unemployment Benefit Institutions on the Productivity of Workers

Dependent variable:	*Productivity*			*Log of productivity*		
	(1)	(2)	(3)	(4)	(5)	(6)
University A	−7.895	−3.261	−3.261	−0.374*	−0.309*	−0.309*
(*Unconditional UBI* and *Comparison*)	(5.100)	(3.093)	(4.491)	(0.211)	(0.162)	(0.180)
UBI stage	23.15***	18.71***	18.71***	0.620***	0.568***	0.568***
(*Unconditional UBI* in Uni. A and *Conditional UBI* in Uni. B)	(3.656)	(5.263)	(6.273)	(0.096)	(0.101)	(0.115)
University A × *UBI stage*	23.46***	24.10***	24.10***	0.547**	0.538***	0.538***
(*Unconditional UBI*)	(6.588)	(6.438)	(7.914)	(0.227)	(0.169)	(0.197)
Constant	42.11***	69.56**	69.56**	3.209***	3.460***	3.460***
	(1.375)	(27.63)	(28.60)	(0.0831)	(0.501)	(0.525)
Individual controls		✓	✓		✓	✓
Bootstrapped standard errors			✓			✓
Observations	1,223	1,223	1,223	1,223	1,223	1,223
R^2	0.089	0.303	0.303	0.115	0.216	0.216

Notes. Ordinary least squares regression. Robust standard errors, clustered at the group level, are in parentheses. Columns (3) and (6) use bootstrap (1,000 repetitions) to compute the standard errors. Columns (1)–(3) use as dependent variable the productivity level (i.e., the number of news items correctly coded). Columns (4)–(6) use instead the log of (1+productivity) and hence coefficients can be interpreted as percentage changes. Control variables include gender, age, whether the subject has another job, the size of the group faced by each worker every working day, socioeconomic stratum fixed effects, and day fixed effects.

*Significant at 10%; **significant at 5%; ***significant at 1%.

the productivity of workers under the *Unconditional UBI* treatment is the highest. We consider this to be the main result of the paper, namely that the productivity of workers under the *Unconditional UBI* is the highest among the three unemployment compensation schemes that we examine.

The standard economic approach also suggests that, if the productivity of those assigned to the *Conditional UBI* scheme is higher than the one observed for the *Comparison* group, it must be that the unemployment monetary compensation of the *Conditional UBI* is lower than the cost of the effort required to be entitled to the compensation (Hypothesis 2 of the model). Our estimate of β_2, which is positive and significantly different from 0, suggests that this is the case.

5.1. Multiple Hypothesis Testing

In a recent paper, List et al. (2015) suggest that multiple hypothesis testing (MHT) is likely to affect inference in experimental economics in three common scenarios: the identification of treatment effects for multiple outcomes, estimating the effect of multiple treatment conditions, and estimating heterogeneous treatment effects based on different subsamples. These three scenarios are present in our study, so we apply the statistical correction procedure suggested by List et al. (2015) to control for the fact that we use our data to test multiple hypotheses. The testing procedure asymptotically controls the probability of obtaining false rejections of the null hypothesis (that a specific treatment effect is 0). In this subsection we focus on the first two potential problems and deal with the third in Section 6.4.

Column (1) of Table D.1 in Online Appendix D reports the mean differences of all our pairwise treatment comparisons (*Conditional UBI* versus *Comparison*, *Unconditional UBI* versus *Comparison*, and *Unconditional UBI* versus *Conditional UBI*) in terms of both productivity and log productivity. Column (2) reports the p-value (and its significance) obtained from performing the standard single testing procedure on each outcome/comparison. In contrast, column (3) reports the p-value adjusted by the multiplicity of hypotheses tested, using the correction procedure of List et al. (2015).

Table D.1 shows that accounting for multiple testing does not change our substantive results, namely that workers employed under *Conditional UBI* display higher productivity than workers in the *Comparison* treatment and that this gap is even larger when comparing the *Unconditional UBI* scheme with the default *Comparison* scheme. In fact, the productivity of workers under *Unconditional UBI* is significantly higher than that of workers under *Conditional UBI*.[38] This is reassuring that our main results are not endangered by potential false positives due to the testing of multiple hypotheses.

6. Potential Explanations

In this section we investigate the potential explanations of our results. After ruling out the possibility

[38] Adjusted p-values are in fact somewhat larger only for the *Unconditional UBI* versus *Conditional UBI* comparison using productivity as outcome, but not enough to compromise statistical significance.

that they are driven by sample imbalance, by differential attrition, or by differential self-selection into working, we explore two possible explanations based on insights from behavioural economics.

6.1. Sample Imbalance

Sample imbalance is unlikely to explain our results for various reasons. First, we note that the estimated treatment effects are virtually unchanged once we control for the unbalanced characteristics (see Table 4).

Second, our design allows us to deal with the potential bias caused by unbalanced attributes. This is because, in addition to the implementation of the *Conditional UBI* and *Unconditional UBI* treatments, two identical *Comparison* treatments also took place, one in each university. We thus look at the differential productivity gain of the benefit schemes relative to the no-benefits benchmark, instead of simply comparing the outcomes of two university populations, one receiving a different active benefit scheme. While the latter comparison would likely be contaminated by the heterogeneity of subjects across universities, our approach is less vulnerable to this threat, provided that subjects of the same university look alike across treatments. This is in turn very plausible, as balance within universities is achieved on all observable characteristics (see Table 1).

Third, recall that there are no productivity differences across universities in the *Comparison* stage (these only arise in the UBI stage). In this case, unbalanced characteristics are likely to cause bias if they affect productivity differentially only after the introduction of an active unemployment benefit scheme, but are harmless when no compensation is offered. That is, any hypothetical confounder needs to drive selection only in the presence of a benefit scheme, but not so in the *Comparison* stage. This is a theoretical possibility that we acknowledge, even if it seems unlikely.

Fourth, to be sure, in Table E.1 of Online Appendix E, we reproduce the baseline results of Table 4 on randomly generated samples that are forced to be balanced, across universities and in both stages, in terms of the originally unbalanced variables, both in the entire sample (Table 1) and in the subsample of coders who dropped out during the intervention (Table C.1). As a benchmark, column (1) of Table E.1 reproduces column (2) of Table 4.[39] To compute column (2) of Table E.1 we followed the following bootstrap-like simulation procedure: (i) We randomly extracted female subjects from University A in order to get the sample perfectly balanced across universities in terms of gender, both in UBI and in the *Comparison* stages. (ii) We ran the baseline regression model on the gender-balanced sample and obtained the new estimated treatment effects. (iii) We repeated steps (i) and (ii) 100 times, each time extracting a different random subsample of female coders in order to get balance in terms of gender, and rerunning the regression model; (iv) We averaged the treatment effects across the 100 estimated coefficients and the standard errors of the 100 regression estimates. Coefficients in columns (3), (4), and (5) are each computed using this same algorithm but producing balance in terms of age, socioeconomic status, and whether the coder is currently working, respectively. Column (6) uses 100 randomly generated balanced subsamples in terms of all four originally unbalanced characteristics simultaneously (thus the smaller number of observations).[40] The results are the same, both in terms of magnitude and statistical significance, across columns (2)–(6) and compared to the benchmark column (1) (taken from Table 4). This is reassuring that our results are not driven by the fact that students from the two universities are somewhat different in terms of a few characteristics.

6.2. Attrition

The evidence shown in Section 3.1 already suggests that our main results are not likely to be explained by differential attrition patterns. We showed that baseline productivity is negatively correlated with attrition (Table 1, panel C), and that attrition rates do not differ across treatments in the UBI stage. Moreover, in our main regression estimates (Table 4) we controlled for all the individual-level characteristics that displayed significant differences across treatments for the subsample of dropouts (Table C.1). In this subsection we investigate this concern further, and we conclude that attrition is indeed very unlikely to confound our results.

We first estimate Equation (1) with daily attrition as the dependent variable. In addition to controlling for individual characteristics, we control for events that are likely to affect the individual's decision to drop out.[41] Table E.3 shows that the proportion of dropouts is significantly smaller in the UBI stage relative to the *Comparison* (coefficient of the *UBI stage* dummy) and that attrition is not significantly different across the two UBI treatments (coefficient of the interaction

[39] Recall that the estimation that produces this column includes all the individual-level controls that are unbalanced.

[40] Table E.2 in Online Appendix E reports the achieved balance in terms of all individual characteristics.

[41] These are as follows: whether subjects were unemployed on the day before each observation, the length of the unemployment spell experienced, whether subjects applied to a vacancy on the day before and were rejected, and the subject's initial productivity.

term). Hence, since attrition is negatively correlated with productivity, it cannot explain the productivity differences that we observe. Note that the only factor that significantly affects the decision of dropping out is whether the subject was unemployed on the previous day.[42] However, in Table E.4 we show that this factor does not affect attrition differentially across treatments.[43]

Second, we estimate Equation (1) for the subsample of subjects who did not drop out, who we call *finalizers*, as they stayed until the last day of the project. The results, reported in Table E.5, are remarkably similar to our benchmark results, both in magnitude and significance: (i) coders hired under a UBI scheme are on average more productive than coders who are offered no unemployment compensation; (ii) relative to the *Comparison* coders, the productivity gain of coders assigned to the *Unconditional UBI* is larger than that of coders assigned to the *Conditional UBI*; (iii) there are no differences in the productivity of coders in the *Comparison* groups in University A and University B once individual controls and day fixed effects are accounted for.

6.3. Self-Selection into Working

Although there is no differential attrition across treatments, there might still be differential selection into work. Workers who return to work, or those who remain always employed may be significantly different across treatments. To test this, we run Equation (1) on the subsample of the unemployed who returned to work (who we call *returners*) and on the subsample of workers who never became unemployed (*always employed*), using as dependent variable the initial (day 1) productivity. We do so to study whether workers who ended up in either subsample differ across treatments in their baseline ability to perform the task.

Results, reported in Table E.6, suggest that, while workers assigned to either UBI are more productive on day 1 than their *Comparison* counterparts, neither *returners* nor *always employed* feature a differential baseline productivity between the *Unconditional UBI* and *Conditional UBI* treatments. This suggests that, while endogenous self-selection into different subsamples may indeed bias the comparison between receiving an unemployment benefit versus receiving none, it is less likely to account for our main result, namely that unconditional unemployment schemes are more effort enhancing than conditional ones. This is because, within the subsamples of *returners* and *always employed*, and controlling for individual characteristics, workers do not seem to differ across the two UBI treatments in terms of their baseline ability.

This does not rule out, however, that selection still exists in terms of other unobserved characteristics that are correlated with effort. However, as we have argued, any existing selection is most likely driven by the intervention itself, and thus, if existing, we interpret this possibility as a treatment effect, rather than as a confounder.

6.4. Reciprocity

We have argued that results are most likely not driven by sample imbalance, differential attrition, or self-selection. One alternative candidate explanation is that workers may have reciprocated with higher effort the differential generosity of the compensations offered. While we cannot directly test this particular channel and disentangle its effect from that of alternative hypotheses, we discuss its plausibility with suggestive evidence from our data.

Reciprocal workers can reciprocate the actual materialization of the unemployment benefit (which occurs when they become unemployed and receive the compensation) or the sole existence of the benefit (i.e., the good intention).[44] One way of reciprocating the actual benefit reception is by refraining from dropping out of the labor force despite facing unemployment. However, such behavior is also consistent with the higher material gain of receiving the monetary

[42] Once we control for this, neither the length of the unemployment spell nor a rejected job application nor the baseline productivity are correlated with attrition.

[43] Table E.4 shows estimates of Equation (1) with daily attrition as the dependent variable. We limit the sample to attrition taking place between working days 3 and 8. This is because, since the first unemployment shock took place on day 2, the decision to drop out before day 3 is independent of having experienced unemployment, which is the confounder of interest. Second, in addition to controlling for the indicator of having experienced unemployment on the previous days, in columns (1)–(3) we interact this dummy with both the university and the stage indicator (respectively *University A* and *UBI stage*), as well as with their interaction. We are thus interested in the triple interaction (*University A* × *UBI stage* × *Unemployed day before*), which captures the extent to which there is differential attrition between *Unconditional UBI* and *Conditional UBI* for those who experienced unemployment on the previous day. For robustness and to facilitate the interpretation, columns (4) and (5) of Table E.4 estimate versions of Table E.3 on both the subsample of workers who, on a given working day, had experienced unemployment on the day before and the subsample of workers who had not. The results of both approaches (triple interaction in columns (1)–(3) and conditional subsamples in columns (4) and (5)) lead us to rule out the possibility that the unemployment experience led to differential attrition across UBI treatments. Indeed, neither is the triple interaction between *University A*, *UBI stage*, and *Unemployed day before* significantly different from 0, nor is the interaction between the first two terms in either the subsample of previously unemployed or that of previously not unemployed.

[44] In the model of Rabin (1993), reciprocal behavior is triggered by the other players' intentions, defined as the beliefs about how kind the other player is. Kube et al. (2012) also show that nonmonetary gifts can trigger positive reciprocity.

Table 5 The Effect of Unemployment Benefit Institutions on the Productivity of Workers (Subsamples of Returners and Always Employed)

Dependent variable: *Productivity* Sample:	Returners			Always employed		
	(1)	(2)	(3)	(4)	(5)	(6)
University A (*Unconditional UBI* and *Comparison*)	−14.48* (7.924)	−6.465 (9.177)	−6.465 (11.18)	−4.795 (9.127)	−3.096 (7.739)	−3.096 (10.92)
UBI stage (*Unconditional UBI* in Uni A and *Conditional UBI* in Uni B)	0.977 (8.206)	11.71 (8.543)	11.71 (10.59)	32.21*** (8.433)	22.85* (11.11)	22.85* (12.46)
University A × *UBI stage* (*Unconditional UBI*)	17.24* (9.690)	7.588 (10.43)	7.588 (13.04)	27.77** (11.94)	35.43*** (11.97)	35.43** (15.05)
Constant	36.23*** (6.752)	90.74** (33.89)	103.4*** (31.16)	58.16*** (4.869)	113.4 (86.62)	108.97 (92.53)
Individual controls		✓	✓		✓	✓
Bootstrapped standard errors			✓			✓
Observations	126	126	126	601	601	601
R^2	0.042	0.203	0.203	0.117	0.298	0.298

Notes. Ordinary least squares regression. Robust standard errors, clustered at the group level, are in parentheses. Columns (3) and (6) use bootstrap (1,000 repetitions) to compute the standard errors. Columns (1)–(3) focus on the sample of subjects who experienced unemployment but were rehired and look at their productivity on the day when they resume activities. Columns (4)–(6) focus on the sample of workers who never experienced unemployment. Control variables include gender, age, whether the subject has another job, the size of the group faced by each worker every working day, socioeconomic stratum fixed effects, and day fixed effects.

*Significant at 10%; **significant at 5%; ***significant at 1%.

compensation, relative to quitting the labor force and to no longer receiving it. Thus, the lower dropout rates under more generous UBIs that we do observe in our sample (see Table E.3) can neither confirm nor rule out the existence of reciprocal responses per se.

Another way of reciprocating the actual benefit reception is by exerting high effort in the case that a vacancy is opened and the unemployed worker is rehired. We compare the differential productivity of returners across treatments. Once we control for unbalanced individual characteristics, we find no evidence of returners reciprocating more generous benefit schemes with higher effort (Table 5, columns (1)–(3)). All the productivity differences we observe across treatments are caused by workers who never became unemployed (Table 5, columns (4)–(6)), which suggests that if reciprocity were indeed explaining the productivity differences, it would only be through workers reciprocating good intentions.[45]

All in all, we can neither rule out reciprocity as a potential channel nor do we have compelling evidence in its favor. Exploring the role of reciprocating both actions and intentions in the effect of different benefit schemes on productivity requires further research.

6.5. Psychological Costs of Unemployment

Besides reciprocity, another potential behavioral explanation of our results that has strong leverage in the psychology literature has to do with the psychological costs of unemployment, which may differ across unemployment schemes.

The literature distinguishes three major sources of psychological costs of unemployment: idleness, social stigma, and shame (for a review, see Darity and Goldsmith 1996). We do not expect that idleness was salient in our setup, since workers were college students engaged in all types of university activities. In contrast, precisely given our sample of well-educated students from two prestigious private schools, the feelings of social stigma or shame of becoming unemployed may well matter for behavior. In particular, becoming unemployed due to a low productivity in a relatively easy task could in principle hurt self-image.[46] Anticipating such psychological cost, workers exert higher effort to avoid unemployment. To the extent that such negative feelings are present in our setup, we can expect them to differ across UBIs, with workers under the *Unconditional UBI* anticipating more stigma or shame of losing their job and being paid for doing nothing and hence exerting the highest effort.

We test whether this conjecture is likely to explain our results by assessing whether unemployment does affect self-esteem, and whether it does so differentially

[45] Table D.2 in Online Appendix D uses the correction procedure of List et al. (2015) for the MHT described in Section 5.1 to test simultaneously the significance of all pairwise treatment comparisons on productivity, for both the subsample of *returners* and that of *always employed*. The results validate the findings in Table 5 in that there are no treatment differences for the subsample of *returners*, while the treatment effects found in the subsample of *always employed* survive the implementation of the correction procedure.

[46] Psychologists define shame as an overwhelming and unpleasant emotion associated with feelings of worthlessness, inferiority, and a damaged self-image (Ausubel 1955).

Table 6 The Effects of UBIs on the Change in Self-Esteem

Dependent variable: *Change in self-esteem*

Sample:	Always unemployed			Always employed		
	(1)	(2)	(3)	(4)	(5)	(6)
University A	3.5	3.561	3.561	0.773	1.287**	1.287
(*Unconditional UBI* and *Comparison*)	(2.161)	(2.523)	(4.182)	(0.595)	(0.449)	(0.788)
UBI stage	3.5***	4.820**	4.820**	0.079	0.207	0.207
(*Unconditional UBI* in Uni A and *Conditional UBI* in Uni B)	(0.468)	(1.901)	(1.997)	(0.552)	(0.741)	(0.928)
University A × *UBI stage*	−5.8**	−6.532**	−6.532	−0.822	−1.083	−1.083
(*Unconditional UBI*)	(2.484)	(2.998)	(4.239)	(0.723)	(0.683)	(1.037)
Constant	−4	1.577	2.182	−5	2.907	3.800
	.	(6.086)	(8.998)	(0.464)	(4.714)	(4.722)
Observations	39	39	39	76	76	76
R^2	0.119	0.276	0.276	0.014	0.066	0.066
Individual controls		✓	✓		✓	✓
Bootstrapped standard errors			✓			✓

Notes. Ordinary least squares regression. Cross-sectional regression of the change in self-esteem for the always unemployed (columns (1)–(3)) and the always employed (columns (4)–(6)). Standard errors, clustered at the group level, are in parentheses. Columns (3) and (6) use bootstrap to compute the standard errors. In columns (1), (2), (4), and (5) the standard errors are robust. Control variables include gender, age, whether the subject has another job, the size of the group faced by each worker every working day, and socioeconomic stratum fixed effects.

**Significant at 5%; ***significant at 1%.

under different UBI schemes. We do so by estimating the model specified in Equation (1), but using as outcome the change in self-esteem between the beginning (day 0) and the end of the intervention.[47] We focus on the sample of workers who remained unemployed since the first unemployment shock on day 2 and throughout the experiment.

Results from this estimation are described in columns (1)–(3) of Table 6. The estimated coefficient associated with the interaction term suggests that unemployed coders under the *Unconditional UBI* experienced a differential drop is self-esteem relative to their peers in the *Comparison* and the *Conditional UBI* treatments, both without (column (1)) and with (column (2)) individual controls.[48] Because of the cross-sectional nature of the regression and the focus on the subsample of the subjects who remained unemployed throughout the intervention, we only have 39 observations. Not surprisingly, when group-clustered standard errors are bootstrapped, significance is lost (column (3)).

As a placebo, we run the same regressions for the sample of *always employed*. The results (columns (4)–(6)) confirm the intuition that self-esteem is not harmed for those who do not experience unemployment. Although this suggestive evidence is consistent with the psychological costs hypothesis, it is far from conclusive, and more research aiming at understanding the psychological costs of unemployment across UBIs is needed.

7. Conclusion

This paper uses a novel empirical design to study, for the first time, the effects of UBIs on individual productivity. Our design allows us to measure workers' productivity under the two most widely implemented unemployment benefit schemes (conditional and unconditional) and under a situation without unemployment benefit.

Contrary to the conventional economic perspective, we observe that an unconditional unemployment benefit makes workers more productive than a conditional compensation scheme, and that both schemes make workers more productive than having no unemployment benefit. After ruling out several candidate explanations related to potential threats to internal validity, we discuss two plausible behavioral mechanisms, namely positive reciprocity and the psychological costs of unemployment. While both explanations could in principle explain our results, pinning down the relative contribution of each mechanism is outside the scope of this paper. We leave this for future research.

The paper sheds light on a topic that is of foremost importance for policy purposes, but of which very little is known. The lack of evidence has to do, precisely, with the challenges of identifying causal

[47] We measured self-esteem on day 6 instead of on day 8 (the last day) because day 6 was the last unemployment shock and hence the last day on which we could measure self-esteem for a large fraction of the unemployed.

[48] Arguably, in our experiment the unemployed had the chance to avoid the negative psychological feeling by choosing not to receive the benefit. However, nobody chose that option. One possible explanation is that, once the self-esteem was damaged, rejecting or accepting the benefit would not restore it.

effects, either by using observational data or by conducting a randomized field experiment with random treatment assignment at the individual level. Our intervention, coupled with the empirical strategy, was designed to overcome these obstacles in the best possible way, while at the same time providing a fair balance between internal and external validity. Since individual random assignment to treatments was not feasible, we picked similar schools, with similar student populations, to do our intervention on balanced samples, both across and within universities. In addition, and in spite of this balance, to account for potential persisting unobserved differences across samples, instead of making a simple comparison of conditional and unconditional UBIs, we looked at the differential productivity gain of [*Unconditional UBI* − *Comparison*] relative to [*Conditional UBI* − *Comparison*]. Further, we implemented identical treatments in all respects except for the type of unemployment compensation offered. Of course, while no alternative is a perfect substitute for individual-level randomization, we believe that our approach minimizes potential biases and thus we have confidence in our results.

Although using a real job in a natural environment gives higher external validity than an artificial task in a lab setting, we acknowledge that our population is not representative of the labor force as a whole. Having said that, our sample is arguably representative of an important segment of the world's population who currently suffers from relatively high unemployment rates, namely well-educated youth. According to the European Commission (2013), the youth unemployment rate in Europe is more than twice as high as the adult one. The chances for a young unemployed person to find a job are low, and, when young people do work, their jobs tend to be less stable. Moreover, the unemployment among young people with academic degrees is an increasing problem both in developed and the developing countries. In Africa, for example, young people with a university education have the highest unemployment rates (African Economic Outlook 2013).

Supplemental Material

Supplemental material to this paper is available at http://dx.doi.org/10.1287/mnsc.2016.2511.

Acknowledgments

The authors are grateful to Lina Díaz for excellent research assistance and fieldwork leadership. For assistance in the field, the authors also thank Darío Romero, Viviana Garcia, Julián Hidalgo, Laura Hincapié, Natalia Lemus, Carlos Salamanca, Diana Salazar, and Mauricio Vela. The authors thank Ximena Cadena, Juan C. Cárdenas, Leopoldo Fergusson, and Ana M. Ibáñez for support with the logistics of the experiment at University of Los Andes. The authors are grateful to Dan Houser, Tobias Klein, Steven Poelhekke, participants at the 2012 International Economic Science Association Conference, the Fourth Development Economics Workshop at Tilburg, the 2013 Florence Workshop on Behavioral and Experimental Economics, the Sixth Maastricht Behavioral and Experimental Economics Symposium, the WZB Conference on Fertile Fields in Development, the Monash-Warwick Workshop in Development Economics, and the economic seminars at Tilburg, University of Rosario, University of Los Andes, University of Buenos Aires, and University of Maastricht for useful comments and suggestions. Financial support from University of Rosario is gratefully acknowledged.

References

Abbring JH, van den Berg GJ, van Ours JC (2005) The effect of unemployment insurance sanctions on the transition rate from unemployment to employment. *Econom. J.* 115(505):602–630.

Acemoglu D, Shimer R (2000) Productivity gains from unemployment insurance. *Eur. Econom. Rev.* 44(7):1195–1224.

African Economic Outlook (2013) *African Economic Outlook 2012* (African Economic Outlook, Paris), 116–123.

Ausubel DP (1955) Relationships between shame and guilt in the socializing process. *Psych. Rev.* 62(5):378–390.

Bandiera O, Barankay I, Rasul I (2005) Social preferences and the response to incentives: Evidence from personnel data. *Quart. J. Econom.* 120(3):917–962.

Bandiera O, Barankay I, Rasul I (2007) Incentives for managers and inequality among workers: Evidence from a firm-level experiment. *Quart. J. Econom.* 122(2):729–773.

Besley T, Coate S (1992) Workfare versus welfare: Incentive arguments for work requirements in poverty-alleviation programs. *Amer. Econom. Rev.* 82(1):249–261.

Black DA, Smith JA, Berger MC, Noel BJ (2003) Is the threat of reemployment services more effective than the services themselves? Evidence from random assignment in the UI system. *Amer. Econom. Rev.* 93(4):1313–1327.

Boone J, Fredriksson P, Holmlund B, van Ours JC (2007) Optimal unemployment insurance with monitoring and sanctions. *Econom. J.* 117(518):399–421.

Darity W Jr, Goldsmith AH (1996) Social psychology, unemployment and macroeconomics. *J. Econom. Perspectives* 10(1): 121–140.

European Commission (2013) Employment, social affairs and inclusion: EU measures to tackle youth unemployment. Last accessed July 26, 2016, http://ec.europa.eu/social/main.jsp?langId=en&catId=1036.

Falk A, Huffman D, Mierendorff K (2006) Incentive properties and political acceptability of workfare: Evidence from real effort experiments. Working paper, Center for Economics and Neuroscience, Bonn, Germany.

Fredriksson P, Holmlund B (2006) Improving incentives in unemployment insurance: A review of recent research. *J. Econom. Surveys* 20(3):357–386.

Gueron J (1990) Work and welfare: Lessons on employment programs. *J. Econom. Perspectives* 4(1):79–98.

Harrison GW, List JA (2004) Field experiments. *J. Econom. Literature* 42(4):1009–1055.

International Labour Office (2010) *World Social Security Report 2010/11: Providing Coverage in Times of Crisis and Beyond* (International Labour Office, Geneva)

Kube S, Marchal MA, Puppe C (2012) The currency of reciprocity: Gift exchange in the workplace. *Amer. Econom. Rev.* 102(4):1644–1662.

List JA, Rasul I (2011) Field experiments in labor economics. Ashenfelter O, Card D, eds. *Handbook of Labor Economics* (Elsevier, Amsterdam), 103–228.

List JA, Shaikh AM, Xu Y (2015) Multiple hypothesis testing in experimental economics. NBER Working Paper 21875, National Bureau of Economic Research, Cambridge, MA.

Macdonald S, MacIntyre P (1997) The generic job satisfaction scale: Scale development and its correlates. *Employee Assistance Quart.* 13(2):1–16.

Martin-Albo J, Nuniez JL, Navarro JG, Grijalvo F (2007) The Rosenberg self-esteem scale: Translation and validation in university students. *Spanish J. Psych.* 10(2):458–467.

Miguel E, Kremer M (2004) Worms: Identifying impacts on education and health in the presence of treatment externalities. *Econometrica* 72(1):159–217.

Ochel W (2005) Unconditional-to-work experiences with specific work-first programmes in selected countries. *Internat. Soc. Security Rev.* 58(4):67–93.

Prendergast C (1999) The provision of incentives in firms. *J. Econom. Literature* 37(1):7–63.

Rabin M (1993) Incorporating fairness into game theory and economics. *Amer. Econom. Rev.* 83(5):1281–1302.

Rosenberg M (1965) *Society and the Adolescent Self-Image* (Princeton University Press, Princeton, NJ).

Shapiro C, Stiglitz J (1984) Equilibrium unemployment as a worker discipline device. *Amer. Econom. Rev.* 74(3):433–444.