I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 14346

Friday Morning Fever. Evidence from a Randomized Experiment on Sick Leave Monitoring in the Public Sector

Tito Boeri
Edoardo di Porto
Paolo Naticchioni
Vincenzo Scrutinio

APRIL 2021

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 14346

Friday Morning Fever. Evidence from a Randomized Experiment on Sick Leave Monitoring in the Public Sector

Tito Boeri
Bocconi University, IZA and CEP-LSE

Edoardo di Porto
CSEF, INPS, University of Naples Federico II, UCFS and Uppsala University

Paolo Naticchioni
INPS, IZA and University of Roma Tre

Vincenzo Scrutinio
LSE, IZA and University of Bologna

APRIL 2021

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Friday Morning Fever. Evidence from a Randomized Experiment on Sick Leave Monitoring in the Public Sector*

This paper provides the first analysis of a population-wide controlled field experiment for home visits checking on sick leave in the public sector. The experiment was carried out in Italy, a country with large absenteeism in the public sector, and it concerned the universe of public employees. We exploit unique administrative data from the Italian social security administration (INPS) on sick leave and work histories. We find that receiving a home visit reduces the number of days on sick leave in the following 16 months by about 12% (5.5 days). The effect is stronger for workers who are found irregularly on sick leave (-10.2 days). We interpret our findings as a deterrence effect of home visits: workers being found irregularly on sick leave experience a decline of about 2% of their wage in the following 12 months. Uncertainty aversion (there is no automatism in these sanctions) can play a role in these results. Our estimates suggest that home visits are cost-effective: every Euro spent for the visits involves up to 10 Euros reductions in sick benefits outlays. We estimate the marginal value of public funds (MVPF) spent on home visits at about 1.13, which is significantly lower than estimates of MVPF of income taxes in the US.

JEL Classification: I12, J45
Keywords: sick leave, absenteeism, randomized trial

Corresponding author:
Tito Boeri
IGIER - Università Bocconi
Via Roentgen 1
20136 Milano
Italy
E-mail: tito.boeri@unibocconi.it

* We thank Thomas Le Barbanchon, Pietro Biroli, David Yanagizawa-Drott, Nathan Hendren, Andrea Ichino, Maarten Lindeboom, Johannes Spinnewijn, Marco Pagano, Nicola Persico, participants at the AIEL, EALE-SOLE, and RES annual conferences, and seminar participants at the IEB (Barcelona), Bocconi University, Bologna University, CSEF (Naples), and Padua University.

1 Introduction

Sick leave is a key institution protecting workers' health and providing income smoothing, but its misuse can promote absenteeism, presenteeism and harm productivity. There is a huge variation across OECD countries in the number of working days lost for illness. Such differences can hardly be explained either by data sources (self-reported in surveys vs. administrative data on days of compensation), or by differences in the age structure or in the sectoral and occupational composition of the workforce. Indeed, average reported days of sickness per year and worker vary almost by a factor of 10: they range from 2.1 in the United Kingdom to 18.3 in Germany. As regulations are rather similar across the EU (OECD, 2010), and cross-country differences in epidemiological risk are second order,[1] heterogeneity in enforcement procedures is likely to play an important role in these huge cross-country differences in absenteeism rates. Several countries have introduced strict sickness monitoring policies in the last 20 years in order to contain sickness absence and reduce public expenditure on sickness benefits. Despite this high policy relevance, only a few attempts have been made to date to evaluate the effectiveness of such enforcement measures, notably in the public sector where absenteeism is generally larger than in the private sector.[2]

Our work aims at filling this gap. We investigate the effects of Home Visits (HV) to public employees on sick leave by general practitioners (GP) working for the Italian social security administration (INPS). Italy is an ideal case study as there is a national administration enforcing the regulations with the same procedures over the entire country. Absenteeism in the public sector is rather widespread: one public employee out of four benefits from sick leave at least once a year, compared with one out of ten in the private sector. Moreover, there is a large body of anectodal evidence pointing to opportunistic behaviour of public employees. For instance, on 2014 New Years's eve 764 local police officers out of 900 reported sick leave in Rome. Data from the sickness benefit register point to strategic behavior: sickness certificates are far from being uniformly distributed over week days as one would expect based on epidemiological factors. As shown in Figure 1, there are two visible peaks in the distribution, respectively on Mondays and on Fridays, de facto extending the week end.

We exploit a population-wide randomized control trial, run by the INPS between November 22, 2017 and January 5, 2018. The experiment allotted randomly HVs to ongoing medical certificates associated with sick leave spells while leaving unchanged all

[1] See for instance the country risk assessments by the European Center for Disease Prevention and Control before the Covid19 pandemic (https://www.ecdc.europa.eu/en).

[2] According to the American and European Working Conditions Surveys (AWCS and EWCS), public employees stay at home for illness reasons over a one-year span about 5.2 days in the US and 7.2 in Europe, with a peak of 8.4 days in Nordic countries. In Europe this means one and a half day more than their counterparts in the private sector (Gomes, 2018).

the procedures related to the organisation of HVs, notably the identification of eligible certificates, the assignment of the visits to doctors, and the implementation of the visits. The use of standard procedures for the implementation of the experiment reduces the possibility that doctors or workers react differently to the visits performed in the period of the experiment and improves the external validity of our results by excluding *Hawtorne effects*. Doctors belonging to the social security performed about 4,200 visits during the period of the experiment. Our analysis is based on a unique, and so-far unexploited, dataset covering sick leaves and employment histories of Italian public sector employees.

We compare access to sick leave between individuals who received a HV during the period of the experiment and workers who did not. We find that being audited leads to a reduction in sick leave in the 16 months following the experiment: workers who are audited spend about 6 days less on sick leave (over a baseline of 47 days) than workers in the control group. This corresponds to a 12% decline in the number of days on sick leave over the 16 months after the experiment. The effect is stronger for workers who are found irregularly on sick leave[3] (-10.2 days vs -4.3 for those found regularly on leave) and involves both the intensive (duration of certificates) and the extensive (number of certificates) margins.

The reduction in the number of claims for sick leave after being found irregularly absent from work is stronger on Fridays and Saturdays, in central administrations and in the health sector with respect to local administrations and schools, and for workers in the Central or Southern regions of Italy. Although no automatic (statutory) sanction is envisaged for workers irregularly on leave, we find that workers reported to their superiors to be irregularly absent from work experience a cumulative reduction in wages close to 2% of their take home pay in the 12 months following the HV. Hence, even a relatively small ex-post sanction seems to exert large deterrence effects. Uncertainty concerning the actual level of the sanctions may play a role in these relatively large deterrence effects. A simple cost benefit analysis from the standpoint of social security shows that audits are highly cost-effective: savings in sickness benefit outlays induced by HVs exceed the costs of HVs themselves by a factor of 8. Targeting individuals who are more likely to be irregularly on leave could further improve the cost effectiveness of the audits. We estimate that one Euro spent on HVs implies a 9 Euros reduction in future benefit outlays under randomized visits and 10.9 Euros if visits are targeted to workers more likely to be irregularly on leave as done in the private sector. This implies net saving of 8 and 9.9 Euros, respectively, for the two kind of HVs allocation systems. Expenditure savings do not seem to be eroded by program substitution. Actually we find

[3] A worker is considered irregularly on leave if not found at home by the doctor at the time of the inspection or if considered to be fit for work. This is discussed in greater detail in Section 2.

that workers detected to be irregularly on leave are less likely to claim disability benefits than the other workers which may also be interpreted as an extension of deterrence effect of medical controls over other health related programs.

Our results suggest that randomized control trials done by public administrations not only pay out in terms of better knowledge of the effects of policies, but also may not be a net cost for the administrations involved. We hope that this will make this kind of studies more attractive to other institutions.

Another way to assess the cost effectiveness of the audits is to estimate the marginal value of public funds (MVPF) of HVs (Finkelstein and Hendren 2020). We estimate it to be 1.13, which is significantly below estimated MVPF of the top income taxes in the US. Hendren and Sprung-Keyser (2020) compute the MVPF for several changes in top tax rate, and find an average of 3.03.

Our paper fits into two main strands of literature.

The first is the rather scant literature on sickness benefits. The focus of previous works was largely on legal rules concerning generosity (De Paola et al. 2014; Böckerman et al. 2018; Scognamiglio 2020; Marie and Vall Castelló 2020), and entitlement conditions (Markussen et al. 2012; Godøy and Dale-Olsen 2018; Hernæs 2018; Markussen et al. 2018) while enforcement was generally overlooked. Partial exceptions are Hesselius et al. (2013), Hesselius et al. (2005) and Hesselius et al. (2009), who exploit a information experiment in Gothenburg (Sweden), to assess the impact of lower monitoring on sick leave claims by treated workers and co-workers, as well as D'Amuri (2017) who studies the response of public sector employees in Italy to changing monetary incentives and monitoring probabilities. Recent studies such as Pichler and Ziebarth (2017), Pichler et al. (2020) investigated the effect of sick leave benefits on contagious presenteeism[4] by comparing sick pay mandates across US states.

The second avenue of research related to our work concerns the evaluation of enforcement mechanisms through auditing. This finds application in various settings such as school testing (Bertoni et al., 2013) and, most notably, taxation. In this context, randomization often concerns information, in terms of simplification of procedures or threats of audits (Pomeranz et al. 2014, Pomeranz 2015, and De Neve et al. 2021). A partial exception is Bergeron et al. (2021) who randomize tax rates, enforcement letters, and assignment of inspectors for property tax in the city of Kananga (Congo). To our knowledge, only two studies could randomize actual audits (Kleven et al. 2011 and Guyton et al. 2021) and assess their impact on subsequent tax payments, while De Neve et al. (2021) exploit a discontinuity in the probability of a set of enforcement actions related to an outstanding

[4] Maclean et al. (2020) provide a comprehensive assessment of the introduction of mandated sick leave in the US and its welfare implications.

tax liability.

Our contribution to this body of research is threefold. First, we are not aware of previous studies directly investigating the enforcement of sick leave regulations throughout home visits. This is a crucial component of the sick leave institutional framework which is common to most countries. As in the case of the «tax system» problem (Slemrod and Gillitzer, 2013), understanding the role played by enforcement is essential to determine if a sickness benefit system is optimal. Second, we draw on a rich experimental design which involved a population-wide randomized control trial over home visits, allowing to directly focus on the responses of the workers involved[5] to actual auditing, and to take into account of the outcomes of such visits. Our experiment takes place over the whole country, encompassing a wide range of different environments in terms of human and social capital endowments. Due to the relevant similarities in the design of sick leave benefits across countries, this experimental design strengthens the external validity of our findings (List 2020). Third, we are able to link data on sick leave to data on workers' careers, shedding light on actual workplace sanctions on workers found to be irregularly absent from work. This issue is relevant in a setting where workers do not face automatic and well defined statutory penalties.

Our findings are particularly relevant at a time in which several countries are extending paid sick leave in response to the Covid-19 pandemic. For instance, in the US, several states introduced mandated leave and the US Government is making plans to include firms and workers previously excluded from sick leave schemes.[6] Evidence on policy tools to tackle moral hazard in the use of sick leave (and potentially other health related transfers such as disability benefits) could provide useful guidance in designing more balanced and effective systems, whose costs are far from negligible (expenditures for sickness and disability cash transfers range from 1 to 4 % of GDP in Oecd countries).

The rest of the paper is structured as follows: Section 2 describes the institutional setting and the experiment; Section 3 describes the data; Section 4 provides the main results; Section 5 discusses evidence of the underlying mechanism and impact on workers' career; Section 6 provides a simple back of the envelope computation of the cost effectiveness of the system, and Section 7 concludes.

[5] As we do not have access to information to sick leave use and working careers of co-workers not involved in our experiment, we are not able to investigate potential spillovers on colleagues of the treated employees.

[6] See, for example, "President Biden Announces American Rescue Plan", link.

2 Institutional Setting and Experiment

2.1 Sick leave and Home visits

All private and public sector employees in Italy are eligible to rather generous sickness benefits. The reason for the absence from work has to be certified by a GP or a specialty doctor, and the certificate, reporting the days in which the worker will be absent from work, must be notified to both the employer and the INPS. The latter provides payments for the sick leave benefits. In the public sector, workers are entitled to up to 18 months of paid sick leave, and to an additional 18 months of unpaid leave over a 4 year period. The amount of the benefit is determined as follows: in the first 9 months of absence from work, sick leave replaces 100% of the contractual wage (excluding all variable components of pay); the replacement rate declines to 90% in the following 3 months, and to 50% from the 13th to the 18th month. Workers do not receive any accessory pay or allowance for all absences with duration lower then 10 days. Surgeries, day hospitals and treatments for chronic disease (e.g., cancer) are exempted from these reductions.

Unlike other countries, such as the Netherlands, where workers can claim disability benefits after a long period on sick leave,[7] in Italy sick leave is not automatically connected to disability benefits.[8]

Employees can be subject to monitoring to assess their health status on the days in which they are on sick leave. HVs verify whether the stated reason for the sick leave (based on a certificate issued by a GP) matches the current true health conditions of the employee. While INPS has been checking on sick leave of private sector workers since 2011, monitoring public sector employees is a new duty for the social security administration, which took up this task since November 2017 from local administrations.

There are two types of HVs: the so called "employer called" visits (ECV) ("Visite Datoriali") and the "INPS called" visits (ICV) ("Visite d'Ufficio"). Each type of visits accounts for about one half of the total HVs. The ECVs are inspections made by a INPS doctor checking whether the public employee is sick and at home. The inspector goes directly to the home residence of the employee to check her conditions, without any notice. This type of audits is performed after a request is issued by the public administration where the absent worker is employed. If INPS has resources to carry out the ECV,[9] a doctor-inspector ("medico fiscale") is sent out to verify the sickness

[7]Workers are eligible to WIA (Dutch abbreviation for the Work and Income Act) benefit if they have been ill for nearly 2 years (104 weeks) and, because of the illness or disability, only can earn 65% or less of their previous income.

[8]We explicitly test whether our treatment affects also the probability of taking up these benefits in Section 6.

[9]In the private sector the procedure is similar, but the employer pays the monetary costs associated to the visit.

certificate. The doctor assesses the health status of the worker, and checks whether it is consistent with the certified reason for the sick leave and the expected duration of the absence. Neither treatment nor therapy should be provided during the visit, as these tasks belong to the GP or other specialists chosen by the family.

ICVs follow the same procedure of the ECVs, but, in this case, it is the social security administration itself to decide whom should be subject to the inspection, and when it should take place. Certificates are selected for a visit on a daily basis, among the universe of ongoing sick leaves notified to the social security administration. Our experiment focuses only on this second type of visits that are fully under INPS control.

HVs can be carried out every day in the week and workers must be at home and available for an inspection between 9 and 12 am, and between 3 and 6 pm.[10] There are 3 possible outcomes for a HV:

1. If a HV finds that the worker is at home but fit for duty, i.e. healthy, the worker has to be at work the day after the visit.
2. If a HV finds that the public employee is not at home during the declared sickness period and the worker has no "force majeure" reason –such as hospitalization, need of life-saving therapies, or medical check ups– for this absence, then INPS reports the matter to the public employer.
3. If the employee is sick and at home, then INPS (based on the recommendation of the inspector) confirms the prognosis or may modify the duration of the sick leave.

In the first two cases, the worker is considered to be irregularly on leave. From then on, the worker involved cannot claim new certificates for the same sickness event. There is no pre-determined sanction: possible disciplinary actions are left to the discretion of the public manager, who is allowed to fire the employee in the extreme case of an unjustified absence from work. However, this event is extremely rare. For instance, in 2019 only 117 workers (.002% of public employment) were laid off for unjustified absences from the workplace. Nevertheless, even without a statutory fine, being found to be irregularly on sick leave may affect career prospects and wage dynamics: a manager could, for instance, reduce paid overtime hours or postpone promotions for employees who have been reported as irregularly on sick leave.

The selection of certificates for inspections in the private sector since 2011 is based on an algorithm denominated *Savio* ("Sistema Assegnazione Visite Ottimizzato") to better target HVs.

Savio operates according to the following steps:

[10]Due to budgetary reasons ICVs were not performed on weekends until 2020 while ECVs were performed also on non-working days.

1. A random sample of ongoing certificates is selected out of the universe of the certificates notified every day to INPS.
2. Certificates exempted from HVs (because involving some chronic disease, such as cancer) are excluded from this randomly selected sample.
3. Among the non-exempt certificates, the algorithm generates a ranking via a machine learning procedure maximising the probability of detecting irregularities.
4. A matching of visits to doctors is undertaken to minimize costs, given a target number of visits. This is implemented in order to optimize travel time of doctors (who are paid based on the distance between their office and the residence of the worker on sick leave) but also to avoid any arbitrariness in the choice of the workers to be subject to inspections.

Doctors receive every morning a list of HVs to be completed within that day and the sequence they should follow in undertaking HVs. Doctors are fully compliant with these rules and they immediately report the outcome of the visit, first, to the social security and then to the public administration involved.

2.2 The Experiment

When INPS was given the task to perform ICVs on public sector employees, it was decided to replicate a setup similar to that operating in the private sector. This uniform application of well established procedures makes it unlikely that doctors behaved differently in the experimental setting ("*Hawtorne effect*"). The underlying procedure, which should target visits to certificates more likely to be irregular, however, needs to build on a critical mass of visits to feed the machine learning algorithm. Thus, it was decided to begin by randomly selecting the certificates potentially subject to ICVs in order to maximize the informational content of the data collected in the home visits.[11]

The decision was made at the INPS headquarters at the highest hierarchical levels and was not disclosed at lower levels.[12] Most importantly, doctors involved in the visits did not receive any information concerning the experiment and did not experience any change in their activity.

The experiment took place in the 45 days between November 22, 2017 and January 5, 2018. It assigned the HV treatment through a "typical" stratified randomized experiment,

[11]The machine learning procedure assigning visits to certificate was never fully in place due to restrictions imposed by a ruling of the Italian Privacy Authority. This ruling was largely unexpected and took place in Autumn 2018, almost one year after the experiment.

[12]The experiment design and its implementation was under the direct control of the authors as one of them, Tito Boeri, was at the time of the experiment President of Inps, while Edoardo di Porto and Paolo Naticchioni were managers at the research Directorate which was responsible for the monitoring and the assessment of the experiment.

following a five-step procedure. Visits were performed over 29 working days out of the 45 of the experiment period. The main steps, visually summarized in Figure 2, are as follows:

1. On each working day from November 22, 2017 until January 5, 2018, INPS drew daily, for each local office, a random sample of certificates from the universe of absence certificates (around 400,000 certificates).

2. Exempted certificates, i.e. those involving chronic or very serious disease, recent surgery, etc., were excluded.

3. Among the selected sample, a second random sample was drawn. Its size was determined in such a way that for each worker subject to HV there were an additional eleven workers in the control group. This sample, made up of about 60,000 certificates, was then used for the experiment.

4. Selected certificates were randomly ranked, and the resulting order defined the HV priority.

5. Visits were assigned to doctors optimizing travel time. During the period of the experiment INPS performed about 4,200 HVs, that is, about 145 visits per day.

Thus, the experiment replicated the Savio procedure except for step 3.

Two caveats are worth mentioning at this stage.

First, as in the private sector, a certificate selected in Step 1, but not assigned to a doctor for a HV in a specific day, and having a prognosis longer than one day, would re-enter the pool of certificates subject to random draw the following day. Similarly, if a worker claims multiple certificates in the period of the experiment, she could be present in the pool of eligible certificates multiple times. Hence, workers having long or multiple certificates in the period of the experiment are more likely to be assigned to a HV. This issue implies a positive correlation between the duration of the certificate and the probability of being treated.

We overcome this issue by directly controlling for the total number of days spent by an individual on sick leave in the period of the experiment.[13] Furthermore, to account for chronic illness and general propensity to use sick leave, we also control for the take-up of sick leave in the six months preceding the experiment. In practice, we include in our regression model the total number of days claimed, the total number of certificates, and their average duration in the six months before the experiment.

[13] As a worker may go back to work earlier as a consequence of a HV, we also use the time which the worker would have spent on sick leave in the period of the experiment based on the original start and end date of the certificate.

Second, certificates, rather than workers claiming sick leave, are randomly extracted due to the fact that the experiment was deliberately not revealed to the doctors carrying out the HVs to avoid changes in their behaviour as a consequence of the experimental setting. Hence, the procedures followed in the case of private sector employees, which involve daily extractions of certificates (rather than workers), were replicated in the public sector. However, as behavioural responses to HVs occur at the level of the individual, we run the analysis at the individual level for ease of interpretation. In addition, running the analysis at the certificate level might contaminate the treatment and control groups as treated individuals might send additional certificates which would be classified in the control group if not subject to visits.

Some indications about the relevance of these two problems can be obtained from Figure 3. Panel (a) reports the number of days spent on sick leave by workers in the period of the experiment. The quite wide range of values in this distribution suggests that individuals have much different treatment probabilities. Controlling for time spent on leave in the period of the experiment is therefore crucial in our setting. Panel (b) reports the number of certificates sent by individuals in the period of the experiment. In more than 75% of cases, individuals only have one active certificate, about 18% have two, 5% three, and a negligible minority more than three. Thus, collapsing the analysis at the individual level comes at little cost, as also confirmed by balancing test reported in Section 3 below.

3 Data

Our analysis uses a so far unexploited dataset, spanning the period from 2016 to 2019. It offers a record linkage of three administrative sources released by INPS for the first time for this paper.

First, our work exploits a dataset on certificates sampled in the experiment period which provides information on the sick leave and on the randomized ICV inspections. This dataset reports the start date of the disease, the beginning and the end dates of the period covered by the sick leave, the identifier of the INPS local office responsible for the inspection activity, the date of the visit, and the outcome of the visit. The diagnosis is not available due to privacy restrictions.

Second, we complement this information with data on certificates issued for individuals involved in the experiment from 2016 up to April 2019. We use this dataset to construct our dependent variables and to control for past access to sick leave. Also in this case we have information on the start date of the disease, start and end dates of each sick leave spell, and identifiers for the worker and for the INPS local office handling the certificate for the worker and managing the inspection activity.

Finally, we extend the analysis with a unique dataset on workers' careers in the public sector at monthly frequencies since 2016. The dataset includes information on wages, type of contract (part-time/full-time, permanent/temporary), occupation, subsector of activity within the public sector, and location of the worker at the municipality level.

We recover additional demographic information such as age and gender from Social Security Archives. The record linkage of these three sources generates a unique and, so far, unexploited dataset. We restrict the analysis to individuals aged 24 to 67 (workers can retire after turning 67), and having a valid employment record in the public sector at the time of the experiment as well as at least one positive monthly wage in the public sector in the six months predating the experiment.

Table 1 provides summary statistics on the workers who were involved in our experiment. About 72% of them are women and the average age is almost 53. Some 95% of the workers have a permanent contract, and 6% a part-time job. On average workers taking part in the experiment spent almost 49 days on sick leave in the 16 months after the experiment. The distribution is strongly skewed to the right, with the median (21 days) being less than half of the average. The majority of sick leaves requested (about 49%) were of short duration (between 1 and 3 days). We also observe, however, a non-negligible share (about 16%) of certificates with durations exceeding 15 days (see Figure A1 in the Appendix for additional details). Almost 10% of the workers on leave were subject to inspection, and in about four cases out of five the prognosis was confirmed, i.e. the worker was found to be regularly on leave. In the remaining cases either the worker was found *fit for duty*, while the medical certificate issued by the GP was stating the opposite, or the worker was absent without any justification, that is, the doctor did not find the employee at home when she should have been according to the existing regulation. The valid reasons for these absences are only those strictly related to the medical treatment the worker is undergoing for the certified disease (or issues of proved 'force majeure', such as the need for life-saving treatment or hospitalization). We categorize these latter two cases as "irregular" absences and the former as "regular" absences.

There is broadly the same percentage of workers reporting at least one day of sick leave in Northern and Southern regions. However, there are more public employees in the North than in the South. The incidence of sick leave is therefore higher in Southern regions.[14]

[14]For instance, Lombardia is the region that sent more certificates during the 45 days of the experiment, around 15% of the total. Sicily is close second with about 13.5% of the certificates. However, Lombardia has about 50% more public employees than Sicily. The ratio between the number of certificates and the number of public workers was in 2016 about 0.02% in Lombardia compared to 0.03% in Sicily, i.e., 50% higher in Sicily than in Lombardia.

The fact that workers in Southern regions generally display a higher propensity to send certificates may reflect either worse health conditions for the Southern employees or a higher propensity to send irregular declarations on health status. Public sector employees real earnings are higher in the South than in the North due to lower costs of living (Boeri et al., 2021). Insofar as health status and income are negatively correlated, this downplays explanations of the interregional differences in the incidence of absenteeism based on differences in the health status. Also the average age of public sector employees is lower in the South, hence the incidence of sick leave certificates should be lower in these regions, while this is not the case.[15]

All this suggests that a higher propensity to send irregular certificates might be at work in Southern regions. The decision to have a central (rather than regional) administration in charge of HVs also acknowledged the fact that there were serious problems in the enforcement of sick leave regulations in Southern regions.

In order to assess whether the randomization procedure was successful in identifying a similar treatment and control group, we compute normalized differences (Imbens and Wooldridge, 2009) for demographic and job characteristics at the individual level. The results of this exercise are reported results in Table 2. They show that differences between the treated group (i.e. those public sector workers who are subject to a HV) and the control group (i.e. those not subject to a HV) are generally small and the two groups show striking similarities. Moreover, normalized differences are always well below the threshold value of 0.25 defined by Imbens and Rubin (2015). This confirms that the randomization was successful in identifying proper treatment and control groups.

Further tests of the balancing between treatment and control groups are provided in the Appendix, where we regress the probability of receiving a HV against demographic and job characteristics both at the individual (Table B1) and certificate levels (Table B2). In both cases, regressions show only a few significant coefficients which are small in magnitude. We further test the relevance of differences in observables by including them in our regression model.

4 Results

4.1 Effects of HVs on sick leave

In this section, we present the baseline results on the effect of receiving a random HV on sick leave claims after the visit. To assess the impact of a visit, we estimate the

[15]The same type of reasoning applies to comparisons between public and private employees: as noted above, absenteeism is higher in the public than in the private sector, and yet average wages are about 25% higher in the public sector than in the private sector.

following equation at individual level:

$$Y_{ij} = \alpha + \beta HV_i + X_i\zeta + D_i\gamma + Z_i\delta + INPS_j + \eta_{ij} \tag{1}$$

where Y_{ij} is the dependent variable for individual i at INPS local office j, HV_i is a dummy equal to one if the individual received an inspection in the period of the experiment, X_i is a set of controls for demographics and job characteristics. As discussed in Section 2.2, in our experiment it is more likely that individuals with longer and repeated certificates (i.e. higher number of days in which the individual is in the experiment) are treated. Thus, we introduce in the regression the variable D_i capturing the number of days in the experiment for each worker, in order to compare treated and controls with the same ex-ante probability to be subject to a HV. We also include measures of past sick leave which control for long-term health conditions and the tendency to draw sickness benefits (vector Z_i, including the number of days on leave in the 6 months before the experiment as well as of sick leave certificates, and the average duration of the certificates). As the experiment takes place separately in each INPS office (at the provincial level), treated and controls have to be compared at the office level. To deal with this issue, we include INPS local office fixed effects ($INPS_j$). Finally, η_{ij} is a random error. Under the assumption that HVs are conditionally uncorrelated with unobservable individual traits or shocks, the parameter β allows us to identify the causal effect of HVs on the future use of sick leave benefits.

We start by considering the number of days of sick leave claimed in the 16 months following the experiment. This should capture potential deterrence effects, that is, whether receiving a random home visit discourages future claims of sick leave.

Our main results are reported in Table 3. Column (1) displays a very parsimonious specification including only the dummy indicating whether the employee received a HV in the period of the experiment. The coefficient for this treatment variable is large and positive. This is consistent with a positive bias determined by a higher probability of treatment for individuals who tend to spend more days on sick leave. Column (2) includes INPS local office fixed effects and the coefficient is still positive, but smaller. Column (3) adds the number of days spent on sick leave in the period of the experiment. This variable massively reduces the size of the coefficient of interest, and it accounts for the difference in the probability of treatment for different individuals in the period of the experiment. The coefficient is now negative and highly statistically significant indicating that receiving a HV reduces future sick leave by 4.7 days. Column (4) includes controls for the use of sick leave benefits in the six months before the experiment. The coefficient remains negative and becomes slightly larger in magnitude (-6.1). Finally, Column

(5) adds demographic and job characteristics, but, in line with our expectations from previous balancing checks, our effect of interest is marginally affected by the inclusion of these additional variables: the β coefficient remains negative, statistically significant at 1%, and of comparable magnitude. According to the last regression model, a HV reduces the number of days spent on sick leave by 5.54 days. Since the average number of days on sick leave for the control group (i.e. workers not receiving a HV during the experiment) is 47 days, the percentage reduction in the number of days on sick leave induced by a HV is of the order of 12%. For the reasons detailed above, we believe that this specification, where we control at the same time for the days of sick leave during the experiment, the health status in the six months prior to the experiment, and for possible differences in characteristics between treated and controls, provides reliable causal estimates of the effect of a HV, and we will use it as the baseline specification throughout the rest of the paper.

The controls included in Column (5) provide additional information on the pattern of sick leave use among public sectors employees. First, past use of sick leave predicts future take-up as both the time spent on leave in the period of the experiment and the use of sick leave in the six months before the experiment (the number of past certificates and the days spent on leave in the six months prior to the experiment) display positive and significant coefficients. We do not detect any difference between men and women while the number of days of sick leave increases with age: in the most extreme case, individuals between 61 and 65 spent an additional 18.4 days on leave in the following 16 months with respect to workers aged less than 35. Workers on permanent contracts are more likely to use sick leave than workers on temporary contracts (+8.1 days).[16] No difference can be detected between workers on full-time or part-time contracts. In terms of sectors, workers in schools and in the health sector take the largest number of days of sick leave, while local and central administrations display, on average, significantly lower days on sick leave (about 10 less than workers in schools). Finally, workers with higher wages are less frequently on sick leave.

Our preferred specification might suffer from two main drawbacks: first, we include the number of days spent on leave in the period of the experiment linearly which might not allow to pair precisely individuals with the same sampling probability; second the number of days spent on leave could be affected by HVs, as individuals found irregularly on leave might be sent back to work earlier. To address these issues, we report four robustness checks in Table 4: the first accounts for possible overlapping sick leave by avoiding double counting of days (Panel a); the second controls for the time spent on leave during the experiment using fixed effects (Panel b); the third includes the

[16] This is in line with the effect of higher employment protection, as shown in Ichino and Riphahn (2005), or workers on temporary contracts might overall spend less time employed in the public sector with lower possibilities to claim for sick leave benefits.

time spent on sick leave in the period of the experiment based on start and end dates of the certificate (Panel c), a duration which is not affected by the HV; finally, the fourth specification uses time in the experiment instrumented by its certificate based counterpart (Panel d), included linearly as in the main equation. Results are in line with our main specification and they tend to be a bit lower when sick leave duration in the period of the experiment based on the medical certificate is used. In addition, we also test our results for workers who only sent one certificate over the experiment period. Estimates, reported in Table B3 in the Appendix, are only marginally affected by this change in the composition of the sample.

The observed decline in the number of days on sickness benefits by audited individuals conflates two different components: the change in the number of certificates and the change in the average duration of the certificates. We assess these two elements separately in Table 5. Column (1) reproduces the main result in Column (5) of Table 3. The next two columns consider the two margins and show that both of them contribute to the observed decline in days of sick leave. The number of certificates declines by 4% relative to the control group, while average duration declines by 0.755 days per certificate (-10% with respect to the control).[17]

Next we look at the time span over which the behavioural change takes place. We compute the cumulative number of days of sick leave at monthly frequencies in the 16 months after the experiment and then we run a separate regression at each time horizon. We perform the same analysis on the number of sick leaves claimed and their average duration at each time horizon. Regression coefficients for the effect of a HV and their 95% percent confidence intervals are plotted in Figure 4. Panel (a) reports the pattern for the cumulative number of days on sick leave, Panel (b) for the cumulative number of certificates, and Panel (c) for the average duration of those certificates. The decline in the number of days builds up over time and progressively increases over the 16 months horizon. Both the extensive (Panel b) and the intensive (Panel c) margin contribute to this pattern but the cumulative decline in the number of certificates stabilizes after about 10 months since the experiment. The long-lasting effect is therefore driven by a shorter duration of certificates sent by treated workers over the whole observation period.

Finally, we analyse heterogeneous effects by demographic and job characteristics. We start by looking at the effect of HVs in different administrations within the public sector and then we move on to consider other individual and contract characteristics. Results are reported in Figures 5 and 6. Reductions in sick leave induced by HV are stronger in the Health Sector and in the Central Administration (e.g. Fiscal Agencies), while HVs seem to have little effect in local administrations and in schools. Moving to workers'

[17] If workers do not send any certificate in the 16 months considered, we assign a zero average duration.

characteristics, HVs have stronger effects for older workers, in Central regions and in metropolitan areas (Rome and Milan), while gender and type of contract do not seem to play an important role in this context. Interestingly, the negative effect on the number of days on sick leave builds up much faster for workers on temporary contracts and stabilizes after 7 months, while the reduction is more gradual but continuous for workers on permanent contracts. Estimates for workers on temporary contracts are, however, too imprecise to derive conclusive evidence. We also provide a standardized measure of these effects by normalizing the effect by the average number of cumulative days on leave at the appropriate horizon for workers in the same category but not subject to a visit. Results reported in Figure A2 in the Appendix confirm the qualitative pattern observed in Figure 6, but differences between age groups seem less relevant.

5 Behavioral Responses by HV Outcome

So far, we investigated the average impact of receiving HVs on workers' behaviour irrespective of the outcomes of the visits. The negative effect that we observe on the future use of sick leave may capture the perception of the visit as a signal of increased monitoring. The response of the worker, however, is likely to depend on the outcome of the visit. Workers who are found irregularly on sick leave may face sanctions in terms of career developments or stigma at work, and they might reassess the probability of detection for future opportunistic behaviours. Workers regularly on sick leave might as well change their behaviour due to the higher perceived probability of detection but, at the same time, feel reassured about the fairness of the system.

In order to evaluate these effects of HVs we need first to meaningfully classify HV outcomes. We grouped these outcomes in two main categories depending on the consequences for the worker: the worker could be considered regularly on sick leave (confirmed/reduced prognosis, justified absence), or irregularly on sick leave (fit for duty or unjustified absence). In the latter case, the worker may face sanctions and even be laid-off.

Irregular outcomes detected under visit randomization concern only about one fifth of the HVs, as reported in Table 1. We characterise below which workers are found to be irregularly on sick leave and when.

Figure 7, Panel (a), displays the distribution of irregular outcomes by the the day in which the HV takes place. The distribution is far from uniform: there is a spike in the share of irregular leaves detected on Friday December 22, that is, the last day of work before the Xmas break. While, on average, about one visit out of five detects an irregular behavior, on December 22, 2017 more than 40% of the workers were found to be irregularly on sick leave. Moreover, there is a concentration of irregular certificates on Fridays. This can be better appreciated by looking at Panel (b), Panel (c) and Panel (d) of Figure 7. The concentration on Fridays of irregular sick leave is more evident in Central and Southern regions than in the Northern part of Italy.

In order to characterise workers irregularly on sick leave, we restrict the sample to workers who were subject to a HV during the experiment. We then compare characteristics of workers found regularly and irregularly on leave with a linear probability model having a dependent variable equal to one if the worker is found to be irregularly on leave and zero otherwise.

Results for this set of regressions are reported in Table 6. Column (1) relates the probability of being found irregular with demographic, job and certificate characteristics, Column (2) includes local INPS office fixed effects, and Column (3) adds fixed effect for the date of the HV. Coefficients are very stable across specifications.

Six results stand out. First, women are 4 percentage points less likely to be irregularly absent from work than men, while no clear pattern emerges in terms of age groups. Second, workers found to be irregularly on leave are more likely to live in metropolitan areas and work in Southern regions. Third, visits performed on Friday are about 6 percentage points more likely to detect irregularities. Fourth, irregular conditions are less likely to be ascertained for employees in central administrations, and in the health sector relative to employees of schools (teachers and administrative staff). Fifth, workers on permanent contracts and in part-time jobs are less likely to be found irregularly on leave (by 8 and 5 percentage points respectively). Sixth, irregular behaviours are more likely to be detected among short certificates than for certificates lasting 10 days or more. Certificates of duration between 1 and 4 days, for example, are 54 percentage points more likely to be found irregular than certificates lasting more than 10 days.

Once established who are the workers irregularly absent from work and when irregular absences are more likely to occur, we assess differences in workers response to HVs depending on the outcome of the visit. We investigate the total impact of the two different types of HV outcomes over time by estimating our preferred specification (Column 5 of Table 3), with all fixed effects and controls, and decomposing our HV dummy according to the outcome of the visit. Figure 8 reports the effect of HVs by outcome of the visit. In particular, coefficients and confidence intervals for the effect of HV with regular (black) and irregular (grey) sick leaves are displayed. We consider once more the effects on the cumulative number of days on sick leave (Panel a), on the number of certificates (Panel b), and on the average duration (Panel c) in the 16 months after the experiment.

Results reported in Panel (a) indicate that the decline in the cumulative number of days of sick leave is present for both types of workers, but the effect is much stronger and significant for workers found irregularly on leave. The response is small and far from statistically significant in the early part of the observation period for workers found on regular sick leave and then it progressively builds up. After 16 months, these workers spent about 4 days less (-8%) on leave than non-inspected workers. This negative effect is much stronger for workers irregularly on leave: the difference in the *cumulative* number of days on sick leave relative to non-audited individuals is about 10 days (-21%) after 11 months and remains relatively stable thereafter. This is in line with evidence of temporary effects of auditing in other settings (see for example Bertoni et al., 2021 for the impact of monitoring during school tests). The larger reaction of individuals non-compliant with regulation is also consistent with evidence from a random tax audit in Norway by Hebous et al. (2020) where audited individuals reduced future use of tax deductions if found misreporting.[18] This suggests that a better targeting of HVs could significantly increase the effectiveness of inspections in reducing sickness benefit claims.

Panel (b) shows that the cumulative number of certificates falls for both regular and irregular outcomes: the magnitude of the effect of HVs for irregular workers is, however, three times as large as in the case of workers regularly on leave, for whom the effect is no longer significant by the end of the 16 months period. Interestingly enough, the reduction in the number of certificates does not show a slower pace by the end of the observation period for workers found irregularly on leave. Finally, Panel (c) indicates that the average certificate duration strongly decreases for irregular outcomes. In this case the decline is stronger in the short term (-3 days after 1 month) than at the end of the 16 months (-2 days) period. The reduction is substantially smaller for workers regularly on leave (-0.5 days per certificate).[19]

The average effects by outcome presented above could be affected by endogeneity as far as the current framework does not allow us to compare workers who were irregularly on leave with workers in the same situation who were not subject to audit. To address this problem, we estimate an instrumental variable model where the variable capturing detection of irregular behavior is instrumented by being subject to a HV for our three main outcomes (days on leave, number of certificates, average certificate duration), and estimate the effect of being irregular over the 16 months horizon. Results are reported in Figure A3 in the Appendix. This strategy delivers larger effects for being found irregularly on leave: indeed, these workers reduce their days of absences by 25 days, the number of certificates sent declines by 1.25, and the average duration by 3.5 days.[20] Thus the effect of HVs could be quite sizeable and strongly reduce the use of

[18]In their case, however, no effect was found on individuals compliant with the regulation.

[19]Table B4 in the Appendix reports corresponding coefficients for the effects at 16 months after the experiment for all the three outcomes.

[20]Corresponding estimates for the OLS regressions at 16 months were: -10 for the cumulative number of days on sick leave claimed; -0.6 certificates; and -1.75 days, respectively. The IV coefficients represent large declines with respect to the baseline quantities of the control group and, more specifically, about

sick leave among workers irregularly on sick leave. At the same time, we acknowledge that our instrumental variable strategy rests on the assumption that being assigned to a randomized visit (our instrument) affects (negatively) only the irregular workers (or a part of them, the so called compliers). However, we believe that, if a bias is present in our IV estimates, it is likely to be fairly limited.[21]

To sum up, workers found irregularly on leave adjust their behavior reducing both the length and the number of the sick leaves. In light of the outcome of the visit, they may update their estimated monitoring probability, and consequently reduce the duration of their certificates and the number of times they go on leave. Workers found regularly on leave, instead, show much more contained adjustments which mostly materialize in a temporary decline in the number of certificates and in slightly shorter durations in the long run. Still, this leads to a non-negligible decline in the use of sick leave.

A concern with the reduction in the days of sick leave originated by HVs is that it may induce presenteeism, causing more infections at the workplace. One way to evaluate this potentially undesirable consequence of HVs is to focus on those certificates that are still ongoing at the time of the visit and investigate the probability of sending an additional certificate while the certificate subject to the inspection is still ongoing or soon after its end date (3 days). Results are reported in Table 7. Receiving an inspection leads to different responses depending on the outcome of the visit: workers on regular leave increase their likelihood of sending a new certificate with respect to workers in the control group. The effect is rather small but positive (+4 percentage points in our preferred specification over a baseline for the control group of 40%). The opposite happens for workers found to be irregularly on leave. In this case the effect is rather sizeable: -20 percentage points, that is, 50% of the baseline probability of the control group. Rules preventing irregular workers from sending additional certificates for the same type of sickness clearly play a role in this context.[22] This suggests that workers found regularly on leave take extra time to make sure that they are fully healed before going back to work.

50% for both the cumulative number of days on sick leave and the average duration of sick leave spells, and 20% for the number of certificates.

[21]It is difficult to determine the exact magnitude of a deterrent effect for a public worker found regularly absent in our setting. According to Figure 8 Panel (a), the effect on regular workers is clearly not significant until month 11. Table B4 reports a slightly negative coefficient on cumulative days of sick leave 16 months after the experiment. This poses a threat on the validity of our instrument and suggests that there may be some compliers among those found regular. The effect for these workers is detectable only if we set up a regression on more than a year after the experiment. All this considered we believe that IV estimates comfort our view that the deterrent effect is driven mostly by irregular workers.

[22]The estimated coefficients strongly decline after the inclusion of a variable capturing past certificate use (Column 3). The positive bias present in previous columns is most likely related to the fact that more frail health conditions, in general positively associated with sick leave use in the past, lead to a higher treatment probability for the workers as they send more and/or longer certificates but are also more likely to send them in the near future.

Finally, Figure 9 displays the coefficients (and 95% confidence bands) for the number of absences by day of the week in the 16 months after the experiment according to the outcome of the visit. We estimate separate regressions for the number of days on leave by day of the week in the 16 months after January 2018 and then plot the main coefficients for being subject to HV by visit outcome together with their 95% confidence intervals. Coefficients are generally uniform over the week but increasing in magnitude, with the largest effects observed on Fridays and Saturdays (limited to irregular workers). The magnitude is, however, substantially different across outcomes and the reduction in the number of days on sick leave by day of the week is from two (on Wednesdays) to three times larger (on Saturdays) for workers irregularly on leave. A similar analysis is carried out for the start and end day of the week of certificates.[23]

Results, reported in Figure 10, show that no effects are observed for workers found regularly on leave while workers found irregularly on leave reduce the number of sick leaves claimed or ending on specific days. The largest decline is registered for certificates issued at the beginning of the week (Monday) and ending during the weekend (Saturday and Sunday). It should be noted that for some sectors, such as the health sector, Saturdays and Sundays could be normal working days, hence with potentially irregular absences from work.[24] The above is in line with a reduction in strategic behaviour in sick leave claims.

What drives the large behavioral responses of workers found to be irregularly in sick leave?

As stated above, there is no automatic sanction for these workers. Yet public managers have some leverage over the career of civil servants and may activate informal sanctions to opportunistic behavior. For instance, the number of hours with overtime pay may be reduced. Another possibility is to postpone the renewal of a temporary contracts at expiration or its upgrading to an open ended contract, as some 5% of workers involved in the experiment have temporary contracts. The presence of such *implicit* penalties could contribute to explain the decline in the number of sickness benefit claims after a HV detecting an irregular leave that we observe in our data. In other words, HVs could be a deterrent to opportunistic behaviour through actual sanctions at the workplace. This kind of informal sanctions are important from a human resource management perspective: as workers seek promotions and positive evaluations, the presence of informal sanctions would discourage co-workers from engaging in opportunistic behaviours and, at the same time, avoid conflicts with unions, which would occur in presence of more drastic

[23]Figure A4 in the Appendix shows that certificates start more frequently on Mondays and end more frequently on Fridays.

[24]Unfortunately, at the time of the experiment no ICV was performed over the week-ends due to budgetary reasons. Hence, it is not possible to assess the level of irregularity for sick leaves in those days.

measures, such as a layoff.

To assess the presence of such informal penalties to workers irregularly absent from work and receiving a HV, we look at workers' careers in the public sector in the period after the experiment. We replicate our main regression model, and estimate the effect of receiving a HV on non-employment and wages in the public sector over the 16 months following the experiment. We run a separate regression, comparing workers found regularly and irregularly on leave with the control group, for each month and then plot our main coefficients in Figure 11. Panel (a) reports results for monthly earnings in the public sector while Panel (b) looks at the probability of employment in the public sector for workers regularly and irregularly on leave. The reference group is non-inspected individuals. Workers found regularly on sick leave do not experience any change in their career while workers found irregularly on leave face penalties in terms of wage reductions (Panel a), and higher non-employment probability (Panel b), although this latter effect is too imprecisely estimated to provide conclusive evidence. Later on, differences with respect to the control groups decline and no wage gap is observable by the end of our observation period. Table 8 evaluates the cumulative consequences for these two groups of workers. On average, a HV does not have any implication for workers in terms of job outcomes, while differences emerge when decomposing the effect between workers who are found irregularly and regularly on leave. The former suffer a wage loss of about 530 Euros (about 2%), and spend 0.113 more months outside the public sector. The effect on cumulative take-home pay is relatively small in magnitude but non-negligible when we consider the strong degree of unionisation and the high level of employment protection enjoyed by public employees. Workers regularly on leave, instead, do not experience any change in their career.

Putting together the various pieces of evidence, we have that small actual sanctions induce large deterrence effects. This can be explained by risk aversion and by the uncertainty associated to having a sort of double lottery. Not only there is a positive probability of not being detected, but also there is not a well established and automatic penalty in case of misbehavior.[25] Insofar as workers do not know the probability

[25] Intuitively, suppose that workers utility is given by $u(c, a) = log(c) + a\Gamma$ where c is consumption, and a is absence from work that can take only two values: 1 if the worker is on leave (and healthy) and 0 otherwise. Γ measures the utility of the individual if he can "get away with it", that is, is not detected in his opportunistic behavior by the imperfect HV technology. Consider for simplicity that the wage is the only source of income of the worker. A regular worker would therefore enjoy $u(w, 0) = log(w)$ while the expected utility of a worker irregularly on leave will be $u(w, 1) = (1 - d)(log(w) + \Gamma)) + dlog(w^l)$ where d is the detection probability, and $w - w^l < 0$ is the sanction necessary to deter opportunistic behavior. If the premium (penalty for irregular absentees) is non-stochastic, the worker will not take an irregular leave as long as

$$log(w) - log(w^l) \geq \frac{1-d}{d}\Gamma \tag{2}$$

In other words the percentage wage increase granted to workers in order to deter misbehavior is distribution of sanctions, uncertainty aversion may play a role in this context. Moreover, workers could be facing additional sanctions in terms of reputation, hostility from colleagues who had to make up for their absence, and workload which we are not able to investigate with the current data.

6 Cost-Effectiveness of Home Visits

Our results indicate that HVs reduce the use of sick leave among audited workers, notably among those irregularly on leave. This does not necessarily imply that HVs are desirable from a public finance perspective. Indeed, sending doctors to visit workers comes at a monetary cost which may well outweigh gains in terms of lower sickness benefit expenditure. Assessing the overall impact of this work-intensive monitoring system appears therefore of paramount importance.

We perform a simple back of the envelope cost-benefit analysis taking the standpoint of the social security administration. First, we collect information on the administrative costs of a HV: the cost of a single visit is generally contained, ranging between 25 and 50 Euro depending on the distance between the residence of the worker and the office of the doctor assigned to the visit. To be conservative, we use the upper bound of the cost per visit. Under this choice, the 4,200 HV performed in the period of the experiment had a total cost of up to 210,000 Euros. Then, we move on to assess the benefits from the perspective of the social security administration organizing the HV. Our baseline estimates from Table 3, imply that a random fiscal visit reduces by 5.5 days the duration of sick leave, As a worker is paid, on average, 81.5 Euros per day[26] and the replacement rate for the first 9 months of leave is 100% of the last wage (except variable pay components), this implies a reduction in expenditure for the Social Security of about 448.3 Euros per visit. Thus each visit generates net savings of about 398.3 Euro, for a total, over all visits of the experiment, of 1,672,650 Euros. Put it another way, HVs entail a 9 Euro lower expenditure for Euro spent, or 8 Euro reduction in net expenditure. This is comforting from an institutional perspective, as it shows that gathering information through a randomized experiment did not come at a cost for social security but rather it implied a net gain.

This computation does not consider several components. First, we did not include the

decreasing in the detection probability and increasing in the utility of "getting away with it". In this context, ex-ante uncertainty as to the actual sanction in case of misbehavior is akin to uncertain detection, and the second derivative of the sanction (premium) with respect to the detection probability is increasing. This implies that a mean preserving spread of the distribution of potential sanctions will also deter misbehavior because $\frac{1}{2}(log(w^l + k) + log(w^l - k)) < log(w^l)$ for any $k > 0$. Put it another way, the expected sanction can be lower than a non-stochastic sanction.

[26]The average monthly wage in our sample is 2,120 Euros per month as reported in Table 1, and workers are paid for 26 working days per month.

gains in terms of productivity and services offered to citizens as a consequence of lower absence rates. Assuming that wages are in line with productivity, these additional benefits could amount to about 81.5 Euros per day, the average daily wage for workers in our sample. Second, a more comprehensive assessment of costs could consider the expenses related to the administrative personnel managing the assignment of certificates. This cost, however, is likely to be fairly small: indeed, most of the administrative costs come from an infrastructure which would be anyway in place to manage the ECVs paid by the private employers. Third, we did not consider costs that other workers would face to reorganize their tasks to cope with the absence of their colleague as well as spillovers on other workers, who might be deterred to claim sick leave after observing a colleague being inspected. These externalities cannot be estimated with our data.

Savings in public expenditure would be lower if workers could substitute sick leave with other kinds of benefits, e.g. by claiming disability benefits or old age pensions. To assess the scope for program substitutions, we looked at whether workers involved in the experiment started receiving these benefits between January 2018 and April 2019. Both of these substitution margins appear to be negligible in our setting. As Column (1) and Column (2) of Table 9 show, only 2.4% of workers retire over the 16 months horizon that we consider, while 1.3% claim disability benefits. Coefficients are generally small, and workers found irregularly on leave actually display a reduction in the probability of claiming disability benefits[27] (the 0.9% coefficient implies a 75% reduction in the propensity to claim disability benefits with respect to the control group). This reduction could be related to two alternative mechanisms: on the one hand, workers irregularly on leave are less likely to be actually sick with respect to the general population in the experiment and thus they might be less likely to claim disability benefits (*selection effect*); on the other hand, having been found irregularly on leave provides them with additional information about the capacity of the public sector to detect irregular behaviours, which makes them less likely to apply for these benefits (*deterrence effect*). Unfortunately, based on our data we cannot disentangle the two effects.

Finally, our back-of-the-envelope calculation does not consider the distortionary costs of taxation. Estimates of the elasticity of tax revenues to tax rates in the US are generally close to 0.3 (Finkelstein and McKnight 2008 and Olken 2007), so that an additional dollar in revenue leads to a loss for the private sector of 1.3 dollars. Given the higher marginal tax rate in Italy (up to 43% for income above 75,000 Euro and 24% for firms), this effect of taxation is possibly larger in our context. However, even if we were to consider smudge factors up to 1.5, the gains would still be substantial with 5.3 Euros

[27] The overall impact on public finance of this reduction is negligible, and hence omitted from our back of the envelope computation.

reduction in expenditure per Euro from the private sector. All balanced, our estimates offer a lower bound for the public sector gains and even large distortionary costs of taxation to finance the HV leave us with large gains.

Under the assumption that the detection technology costs, the probability of detecting irregular behavior, the treatment effects, and the number of visits per day are all constant throughout the year, we can compute the average gain for the public sector. Running the HVs has a total cost of about 2,646,250 Euros.[28] The estimated deterrent effect implies a decline in sick leave expenditure of about 23,726,278 Euros.[29] The net gain for public finance is about 21,080,028 Euros per year.[30]

All this happens under random visits that do not target individuals more likely to be absent irregularly, and have about a 20% probability of detecting irregularities. The random assignment of HVs did not aim at maximizing the effectiveness of the HVs but only at feeding the machine learning procedure. Some indications as to what could be the benefits from targeting visits also in the public sector come from the experience with the Savio algorithm in the private sector, described in Section 2. Boscarino et al. (2018) document that visits assigned by the algorithm had a irregularity detection rate close to 40%, i.e., twice as large as the detection probability in the case of random assignments. We can therefore estimate that, on average, targeted HVs could induce a decline of 6.7 days of sick leave.[31] Net benefits for sickness benefit outlays could be of the order of 26,250,000 Euros yearly.[32] This would be imply a 9.9 lower net expenditure per Euro spent on HVs.

Finally, we provide a measure of the efficiency of this policy in reducing expenditure for the government by computing the *Marginal Value of Public Funds* (MVPF) (Hendren and Sprung-Keyser 2020; Finkelstein and Hendren 2020). This measure appears particularly attractive as it relates budgetary effects to utility gains of the individuals involved and provides a consistent framework for comparisons across policies. The computation of the MVPF requires two key items. On the one hand, an evaluation should be made about

[28]Data from the experiment imply $\frac{4200}{29} = 144.8.. \approx 145$ visits per day over the course of the year with a 50 Euros cost per visit.

[29]This is obtained by multiplying the average decline in sick leave for visited workers (5.5) by the average daily wage of 81.5 Euros, the number of visits per day, and the number of days in one year.

[30]This is computed as the $\#HVperDay * 365 * 398.3 = 145 * 365 * 398.3 = 21,080,028$, where 398.3 is the lower expenditure on benefits per HV net of the cost of the visits.

[31]The effect on the number of days on sick leave can be decomposed in $S_{irr}\beta_{irr} + S_{reg}\beta_{reg} = 0.2 * 10.2 + 0.8 * 4.3 \approx 5.5$, where S_irr is the share of irregular outcome of visits, S_{reg} is the share of regular outcomes, and β_{irr} and β_{reg} are the coefficients for the effect of HVs on future days on sick leave in the following 16 months. In our experiment, this leads to a 5.5 treatment effect. By increasing the share of irregularities detected to 40%, this would lead to a $S_{irr}\beta_{irr}+S_{reg}\beta_{reg} = 0.4*10.2+0.6*4.3 \approx 6.7$ reduction in days per visit.

[32]This is computed as $145 * 365 * (81.5 * 6.7 - 50)$ where 145 is the number of visits per day, 365 is the number of days in which visits are performed, 81.5 is the average wage in the public sector (replacement rate 100%), 6.7 is the average decline in the use of sick leave, and 50 Euros is the cost of a visit.

how much the beneficiaries would be willing to pay in order to finance the extra spending, that is their willingness to pay for the policy (numerator); on the other hand, the net cost of the policy for the government (denominator) should be ascertained. We start by considering the denominator of the MVPF, that is, the net cost of the government for one Euro expenditure on HVs. As discussed above, spending an additional Euro on HVs involves a reduction of 9 Euros of expenditure on sickness benefits. Hence, the net government saving is 8 Euros per Euro spent. Next, we turn to the monetary value that inspected workers assign to the public expenditure on HV. HVs reduce sickness benefit outlays by 9 Euros per Euro spent. We showed in the previous section that the policy has a very small average effect on workers' wage and so we can consider the implicit sanction negligible in this context. Thus, we can assume that workers would be willing to pay the entire reduction in gross expenditure in sickness benefits associated to one Euro of HVs, that is, 9 Euros. Thus the implied MVPF is -9/-8 or about 1.13. This appears well below many of the estimates reported for the MVPF of taxes in the US reported by Hendren and Sprung-Keyser (2020). The average MVPF for the top income tax is about 3 and available estimates range between 1.16 and 44.23. Hence, HVs appear to have a relatively little cost in terms of efficiency with respect to classical revenue raising policy tools.[33] The dimension of the policy, however, makes it unlikely that it could become a major source of fiscal gains for the government. This computation, however, shows that the reduction in expenditure appears efficient from a welfare analysis standpoint.

7 Conclusions

In this paper we analyse the results of a randomized control trial for home visits checking on sickness benefit claims in the public sector.

Our experiment concerns the universe of public employees in Italy, and draws on unique administrative data on sick leave and work histories in the public sector. We find that receiving a home visit decreases by about 6 days the duration of sick leave over the 16 months following the experiment. We document that this decline in sick leave is driven by workers who are found to be irregularly on sick leave, and involves both the extensive (number of certificates) and the intensive (duration of certificates) margins.

Although there is no statutory sanction for those found to be absent irregularly from work, we observe that the workers involved are informally sanctioned for their behavior in their career developments. In particular, workers found to be irregularly on sickness experience wage losses close to 2% of their overall take-home pay over the 16 months

[33] Little evidence is available for the MVPF of taxation in Italy. Recent estimates by Cerqua and Galli (2020), who exploit differential regional tax rates, report elasticities up to 5% which imply a MVPF of 1.05. Thus, in the Italian context, HVs appear to be relatively less efficient than taxation in raising revenue.

after the HV. This is consistent with the presence of an implicit sanction, thereby public managers punish irregular claiming, e.g. by reducing hours in overtime pay, in order to deter misbehavior also by other workers.

The profile of reductions in sickness benefit claims is also consistent with HVs being a deterrent for opportunistic behavior. Reductions are indeed stronger on Fridays and Saturdays, and involve mostly certificates sent on Monday and ending during the weekend, which would be consistent with strategic behaviour to extend week-end holidays.

Our results confirm that the way sick leave regulations are enforced is extremely important. Given that our experiment takes place nationwide, covering a variety of different institutional, cultural and labour market conditions, we believe that it has a validity that goes beyond the Italian case. Our results highlight how HVs can contribute to repress opportunistic behaviour in sick benefit claims and could represent an efficient monitoring tool in many countries. Due to the size of these programs, and the potential spillovers of deterrence effects to disability claims, net savings in public expenditure can be rather substantial.

References

Bergeron, A., Tourek, G., and Weigel, J. (2021). The State Capacity Ceiling on Tax Rates: Evidence from Randomized Tax Abatements in the DRC. *Job Market Paper.*

Bertoni, M., Brunello, G., De Benedetto, M. A., and De Paola, M. (2021). Does Monitoring Deter Future Cheating? The Case of External Examiners in Italian Schools. *Economics Letters*, 201:109742.

Bertoni, M., Brunello, G., and Rocco, L. (2013). When the Cat is Near, the Mice won't Play: The Effect of External Examiners in Italian Schools. *Journal of Public Economics*, 104:65–77.

Böckerman, P., Kanninen, O., and Suoniemi, I. (2018). A Kink that Makes You Sick: The Effect of Sick Pay on Absence. *Journal of Applied Econometrics*, 33(4):568–579.

Boeri, T., Ichino, A., Moretti, E., and Posch, J. (2021). Wage Equalization and Regional Misallocation: Evidence from Italian and German Provinces. *The Journal of the European Economic Association*, forthcoming.

Boscarino, R., di Porto, E., and Naticchioni, P. (2018). SAVIO Shut Down: Effetti sulle Visite Mediche di Controllo. *INPS DCSR Studi e Analisi, Nota n. 2.*

Cerqua, A. and Galli, E. (2020). Income Tax Rate Increases and Heterogeneous Taxpayers' Reactions: a Spatial Regression Discontinuity Design. *Sapienza University of Rome, Working Paper N. 17/2020.*

Correia, S. (2019). REGHDFE: Stata Module to Perform Linear or Instrumental-Variable Regression Absorbing any Number of High-Dimensional Fixed Effects. *Boston College Department of Economics.*

De Neve, J.-E., Imbert, C., Spinnewijn, J., Tsankova, T., and Luts, M. (2021). How to Improve Tax Compliance? Evidence from Population-Wide Experiments in Belgium. *Journal of Political Economy*, forthcoming.

De Paola, M., Scoppa, V., and Pupo, V. (2014). Absenteeism in the Italian Public Sector: The Effects of Changes in Sick Leave Policy. *Journal of Labor Economics*, 32(2):337–360.

D'Amuri, F. (2017). Monitoring and Disincentives in Containing Paid Sick Leave. *Labour Economics*, 49:74–83.

Finkelstein, A. and Hendren, N. (2020). Welfare Analysis Meets Causal Inference. *Journal of Economic Perspectives*, 34(4):146–67.

Finkelstein, A. and McKnight, R. (2008). What did Medicare do? The Initial Impact of Medicare on Mortality and out of Pocket Medical Spending. *Journal of Public Economics*, 92(7):1644–1668.

Godøy, A. and Dale-Olsen, H. (2018). Spillovers from Gatekeeping–Peer Effects in Absenteeism. *Journal of Public Economics*, 167:190–204.

Gomes, P. (2018). Heterogeneity and the Public Sector Wage Policy. *International Economic Review*, 59(3):1469–1489.

Guyton, J., Langetieg, P., Reck, D., Risch, M., and Zucman, G. (2021). Tax Evasion at the Top of the Income Distribution: Theory and Evidence. *National Bureau of Economic Research, Working Paper N. 28542*.

Hebous, S., Jia, Z., Løyland, K., Thoresen, T. O., and Øvrum, A. (2020). Do Audits Improve Future Tax Compliance in the Absence of Penalties? Evidence from Random Audits in Norway. *CESifo Working Paper N. 8480*.

Hendren, N. and Sprung-Keyser, B. (2020). A Unified Welfare Analysis of Government Policies. *The Quarterly Journal of Economics*, 135(3):1209–1318.

Hernæs, Ø. (2018). Activation against Absenteeism–Evidence from a Sickness Insurance Reform in Norway. *Journal of Health Economics*, 62:60–68.

Hesselius, P., Johansson, P., and Larsson, L. (2005). Monitoring Sickness Insurance Claimants: Evidence from a Social Experiment. *Institute for Labour Market Policy Evaluation, Working Paper N. 2005:15*.

Hesselius, P., Johansson, P., and Nilsson, J. P. (2009). Sick of Your Colleagues' Absence? *Journal of the European Economic Association*, 7(2/3):583–594.

Hesselius, P., Johansson, P., and Vikström, J. (2013). Social Behaviour in Work Absence. *The Scandinavian Journal of Economics*, 115(4):995–1019.

Ichino, A. and Riphahn, R. T. (2005). The Effect of Employment Protection on Worker Effort: Absenteeism during and after Probation. *Journal of the European Economic Association*, 3(1):120–143.

Imbens, G. W. and Rubin, D. B. (2015). *Causal Inference in Statistics, Social, and Biomedical Sciences*. Cambridge University Press.

Imbens, G. W. and Wooldridge, J. M. (2009). Recent Developments in the Econometrics of Program Evaluation. *Journal of Economic Literature*, 47(1):5–86.

Kleven, H. J., Knudsen, M. B., Kreiner, C. T., Pedersen, S., and Saez, E. (2011). Unwilling or Unable to Cheat? Evidence from a Tax Audit Experiment in Denmark. *Econometrica*, 79(3):651–692.

List, J. A. (2020). Non est Disputandum de Generalizability? A Glimpse into The External Validity Trial. *NBER Working Paper N. 27535*.

Maclean, J. C., Pichler, S., and Ziebarth, N. R. (2020). Mandated Sick Pay: Coverage, Utilization, and Welfare Effects. *NBER Working Paper N. 26832*.

Marie, O. and Vall Castelló, J. (2020). If Sick-Leave Becomes More Costly, Will I Go Back to Work? Could It be too Soon? *IZA Discussion Paper Series N. 13379*.

Markussen, S., Mykletun, A., and Røed, K. (2012). The Case for Presenteeism—Evidence from Norway's Sickness Insurance Program. *Journal of Public Economics*, 96(11-12):959–972.

Markussen, S., Røed, K., and Schreiner, R. C. (2018). Can Compulsory Dialogues Nudge Sick-Listed Workers back to Work? *The Economic Journal*, 128(610):1276–1303.

OECD (2010). *Sickness, Disability and Work: Breaking the Barriers*. OECD publications.

Olken, B. A. (2007). Monitoring Corruption: Evidence from a Field Experiment in Indonesia. *Journal of Political Economy*, 115(2):200–249.

Pichler, S., Wen, K., and Ziebarth, N. R. (2020). Positive Health Externalities of Mandating Paid Sick Leave. *IZA Discussion Paper Series N. 13530*.

Pichler, S. and Ziebarth, N. R. (2017). The Pros and Cons of Sick Pay Schemes: Testing for Contagious Presenteeism and Noncontagious Absenteeism Behavior. *Journal of Public Economics*, 156:14–33.

Pomeranz, D. (2015). No Taxation without Information: Deterrence and Self-Enforcement in the Value Added Tax. *American Economic Review*, 105(8):2539–69.

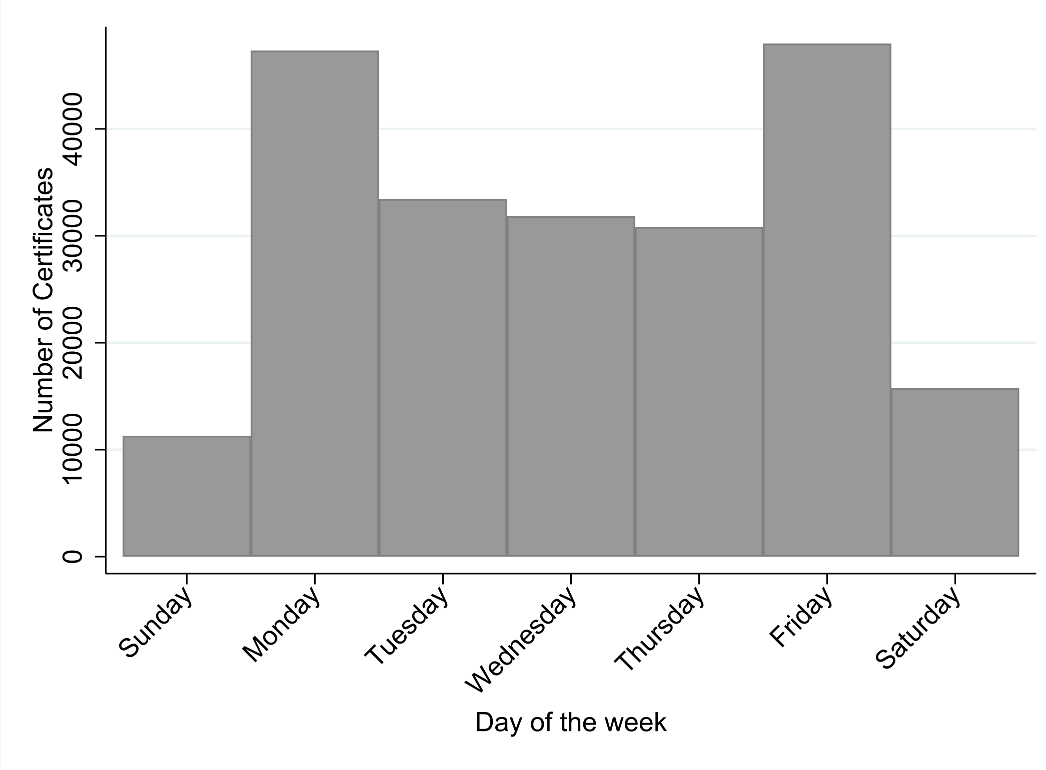
Pomeranz, D., Marshall, C., and Castellon, P. (2014). Randomized Tax Enforcement Messages: a Policy Tool for Improving Audit Strategies. *Tax Administration Review*, (36):1–21.

Scognamiglio, A. (2020). Paid Sick Leave and Employee Absences. *LABOUR*, 34(3):305–322.

Slemrod, J. and Gillitzer, C. (2013). *Tax Systems*. MIT Press.

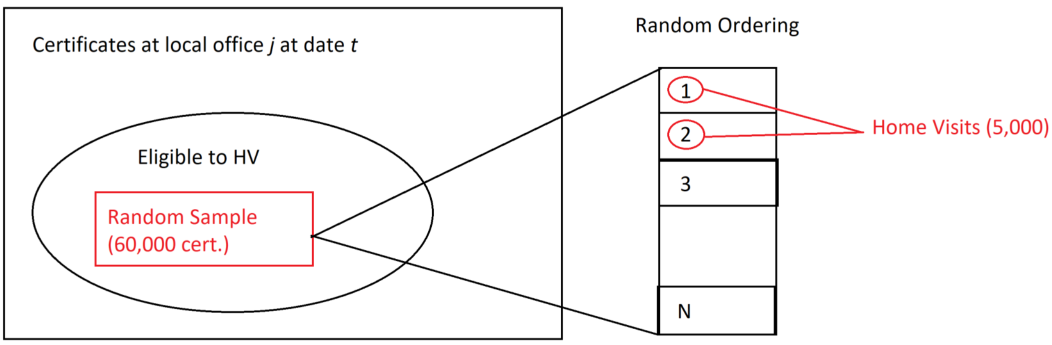
Graphs

Figure 1: Certificates by Start and End Day of the Week



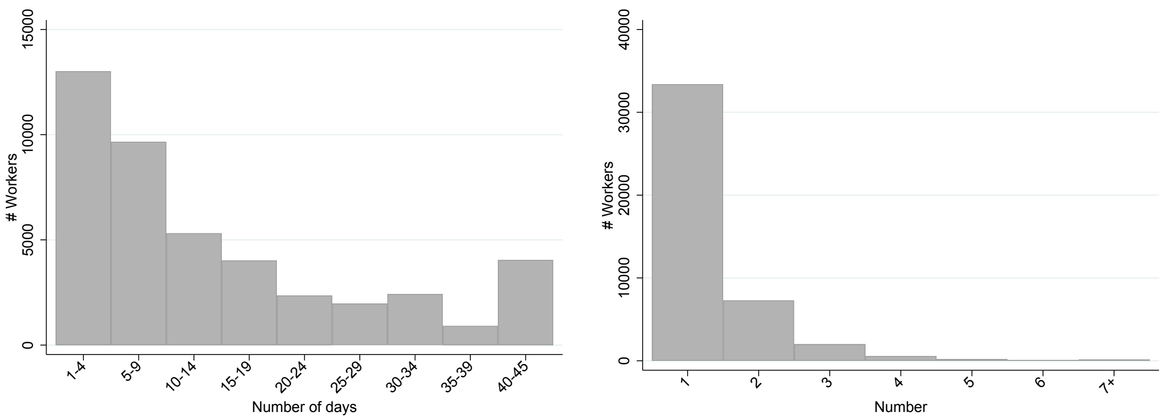
Note: Statistics are based on data for public sector workers involved in the experiment. We consider sick leave spells which started between May and October 2017. Figure reports the number of certificates by start and end day of the week. Each certificate is counted both for the day of the week in which it starts and the day of the week in which it ends.

Figure 2: Structure of the Experiment



Note: Description of the structure of the experiment. A random set of medical certificates is drawn at INPS local office (*sede*) j and date t level. Each certificate corresponds to a sick leave spell. Then, exempt certificates are removed (e.g. cancer), and a second random sample is drawn among certificates eligible to HV. Among these certificates, a random order, which determines which certificates will be inspected, is assigned for HV, and, finally, certificates are assigned to doctors to minimize mobility costs.

Figure 3: Number of Workers by Days Spent on Sick Leave and Number of Certificates during the Experiment.

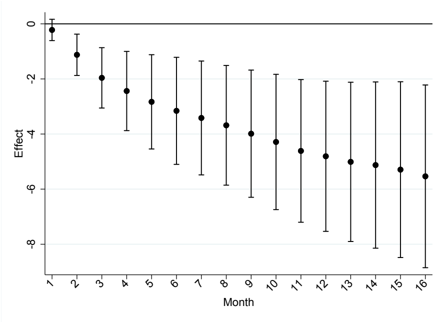


(a) Number of days

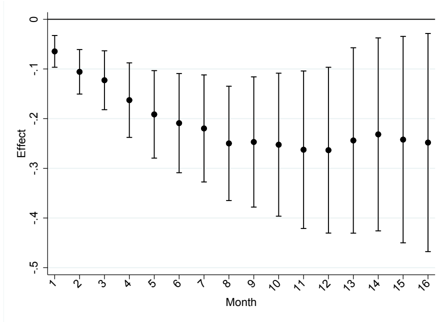
(b) Number of certificates

Note: Figure reports the number of individuals by number of days spent on sick leave in the period of the experiment (11/22/2017-1/5/2018) in Panel (a) and by number of certificates sent in the period of the experiment in Panel (b). The number of certificates corresponds to the number of sick leave spells by the worker.

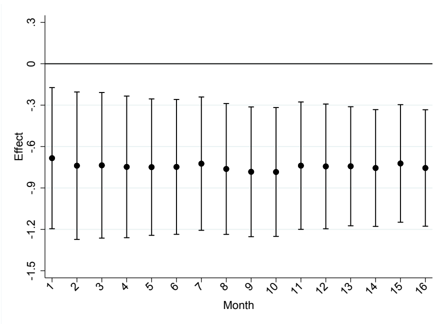
Figure 4: Effect of HV on Sick Leave over 16 Months after the Experiment



(a) Cumulative Days on Sick Leave



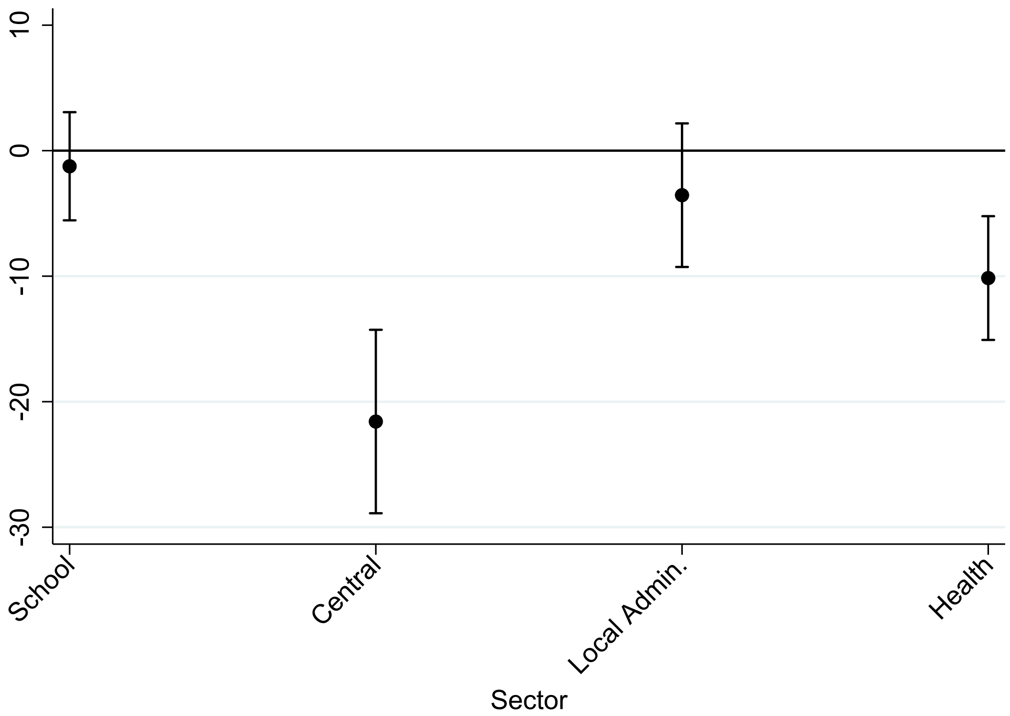
(b) Cumulative Number of Certificates



(c) Average Certificate Duration

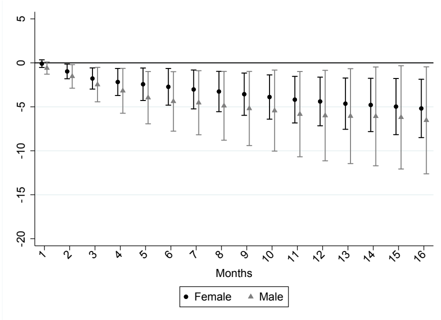
Note: Figure reports estimates for the effect of HV on cumulative number of days spent on sick leave (Panel a), cumulative number of certificates (Panel b), and average duration of certificates (Panel c) in the 16 months after the end of the experiment. The cumulative number of certificates corresponds to the cumulative number of sick leave spells and the average certificate duration corresponds to the average duration of sick leave spells. Regression run separately for each month, and they include a dummy for receiving HV, demographic characteristics (gender, age category dummies), job characteristics (sector dummies, average salary in the 6 months before the experiment, part time dummy, dummy for permanent contract), fixed effects at INPS local office level, number of days spent on sick leave in the period of the experiment (11/22/2017-01/05/2018), number of certificates, average duration, and number of days spent on sick leave in the 6 months before the experiment. Standard errors clustered at local office level. Coefficient and 95% confidence interval of HV dummy reported.

Figure 5: Effect of HV on Cumulative Days on Sick Leave by Sector

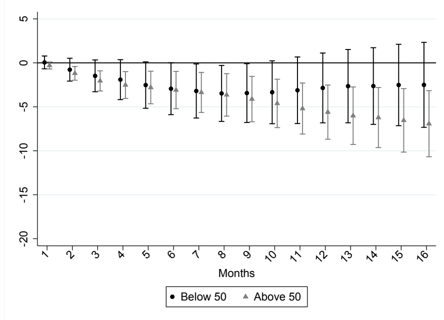


Note: Figure reports estimates for the effect of HV on cumulative number of days on sick leave in the 16 months after the end of the experiment. Regressions include a dummy for receiving HV its interaction with sectors dummies, demographic characteristics (gender, age), job characteristics (sector dummies, average salary in the 6 months before the experiment, part time dummy, dummy for permanent contract), fixed effects at INPS local office level, number of days spent on sick leave in the period of the experiment (11/22/2017-01/05/2018), number of certificates, average duration, and number of days spent on sick leave in the 6 months before the experiment. Standard errors clustered at local office level. Coefficients and 95% confidence intervals for the sum of the HV dummy and the appropriate interaction term reported.

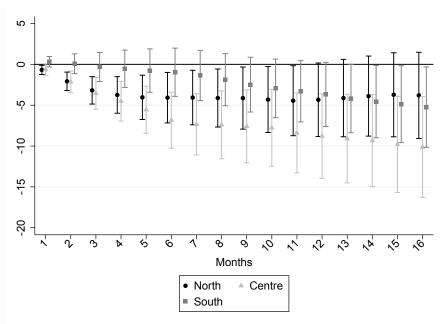
Figure 6: Effect of HV on Cumulative Days on Sick Leave by Worker/Job Characteristics



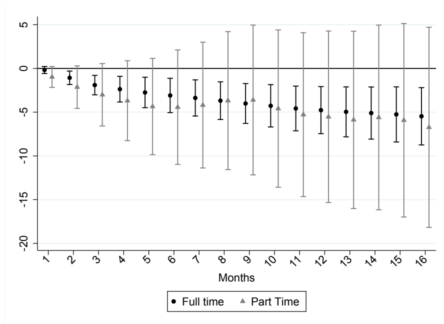
(a) Gender



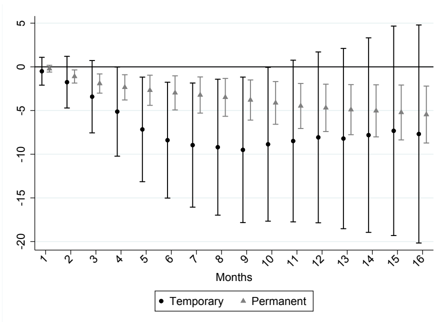
(b) Age



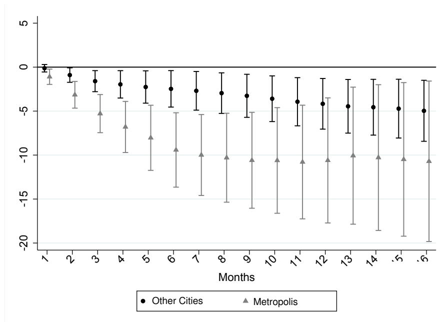
(c) Geographic Area



(d) Full-Time/Part Time



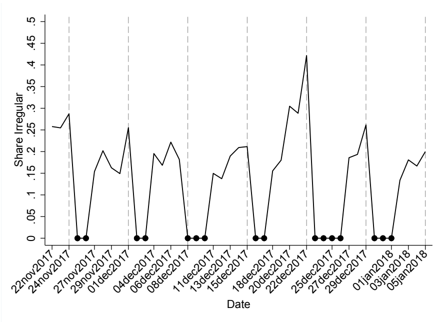
(e) Permanent/Temporary



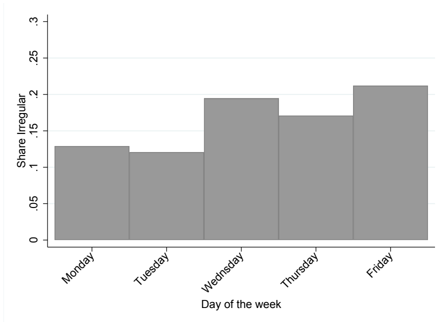
(f) Municipality Size

Note: Figure reports estimates for the effect of HV on cumulative number of days on sick leave use in the 16 months following the experiment. Regressions include a dummy for receiving HV and its interaction with relevant dummies, demographic characteristics (gender, age), job characteristics (sector dummies, average salary in the 6 months before the experiment, part time dummy, dummy for permanent contract), fixed effects at INPS local office level, number of days spent on sick leave in the period of the experiment (11/22/2017-01/05/2018), number of certificates, average duration, and number of days spent on sick leave in the 6 months before the experiment. Metropolis are the two main administrative centre in Italy: Rome (the capital of the country) and Milan (the main economic centre). Standard errors clustered at local office level. Coefficients and 95% confidence intervals for the sum of the HV dummy and the appropriate interaction term reported.

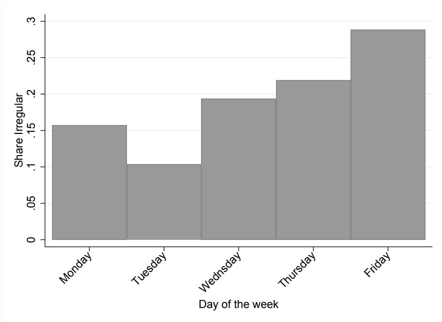
Figure 7: Share of Visit with Irregular Outcome by Day of the Week, Date, and Geographic Area



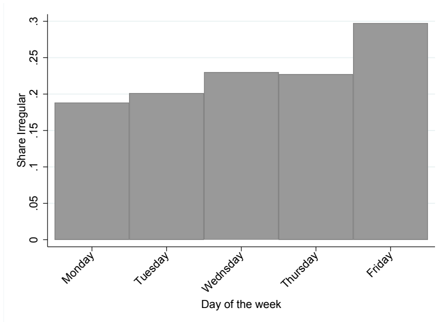
(a) Date



(b) Day of the Week: North



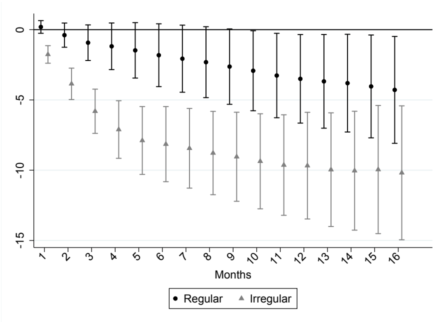
(c) Day of the Week: Centre



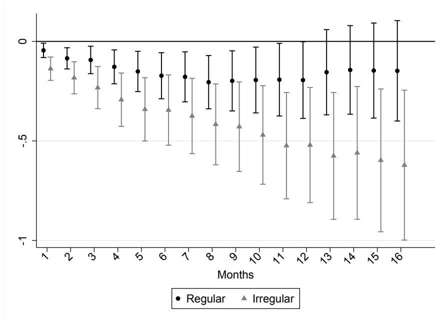
(d) Day of the Week: South

Note: Figures report the share of visits with Irregular outcome in the period of the experiment by date in Panel (a) and by day of the week and geographic area in Panels from (b) to (d). During non-working days (mostly Saturday and Sunday), corresponding to the zeros in Panel (a), no INPS called HV was performed due to budgetary reasons.

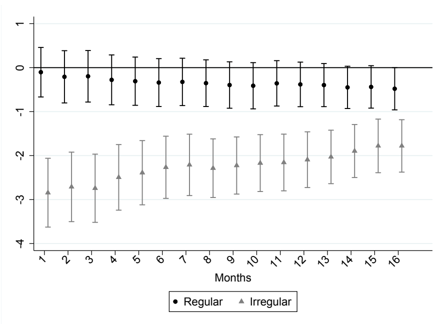
Figure 8: Effect of HV on Sick Leave over 16 Months after the Experiment by Visit Outcome



(a) Cumulative days on sick leave



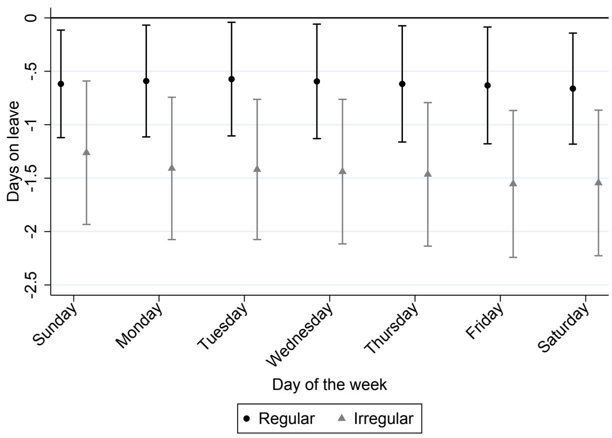
(b) Cumulative certificates



(c) Average Certificate Duration

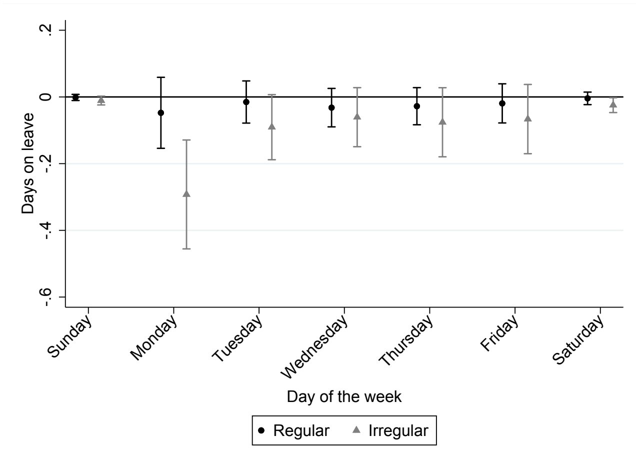
Note: Figure reports estimates for the effect of HV on cumulative number of days spent on sick leave (Panel a), cumulative number of certificates (Panel b), and average duration of certificates (Panel c) in the 16 months after the end of the experiment by outcome of the visit. The cumulative number of certificates corresponds to the cumulative number of sick leave spells and the average certificate duration corresponds to the average duration of sick leave spells. Regression are estimated separately for each month, and they include dummies for receiving HV by outcome of the visit, demographic characteristics (gender, age), job characteristics (sector dummies, average salary in the 6 months before the experiment, part time dummy, dummy for permanent contract), fixed effects at INPS local office level, number of days spent on sick leave in the period of the experiment (11/22/2017-01/05/2018), number of certificates, average duration, and number of days spent on sick leave in the 6 months before the experiment. Standard errors clustered at local office level. Coefficients and 95% confidence interval for HV by outcome reported.

Figure 9: Effect of HV on Sick Leave by Day of the Week over 16 Months after the Experiment by Visit Outcome.

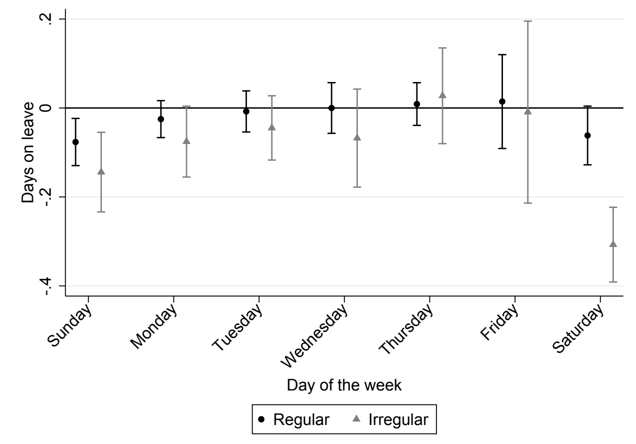


Note: Figure reports coefficients for effect of HV on the number of absences by day of the week according to the outcome of the visit. Coefficients obtained by separate regressions by day of the week. Coefficients obtained from OLS regressions including indicators for workers subject to HV with regular and irregular outcome, demographic characteristics (gender, age), job characteristics (sector dummies, average salary in the 6 months before the experiment, part time dummy, dummy for permanent contract), fixed effects at INPS local office level, number of days spent on sick leave in the period of the experiment (11/22/2017-01/05/2018), number of certificates, average duration, and number of days spent on sick leave in the 6 months before the experiment. Figure includes confidence interval at 95% with clustered standard errors at INPS local office level.

Figure 10: Effect of HV on Number of Certificates by Start and End day of the week.



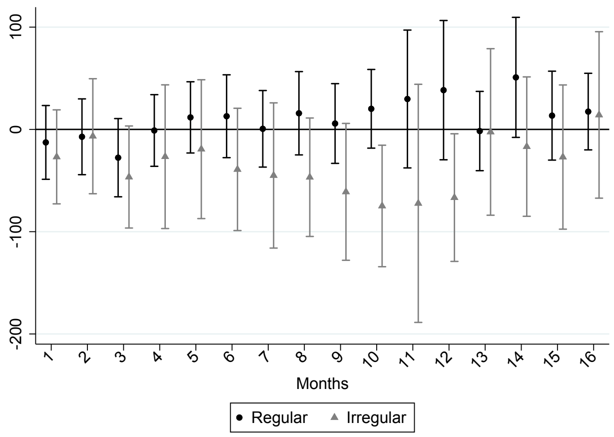
(a) # Certificate by Start day of the week



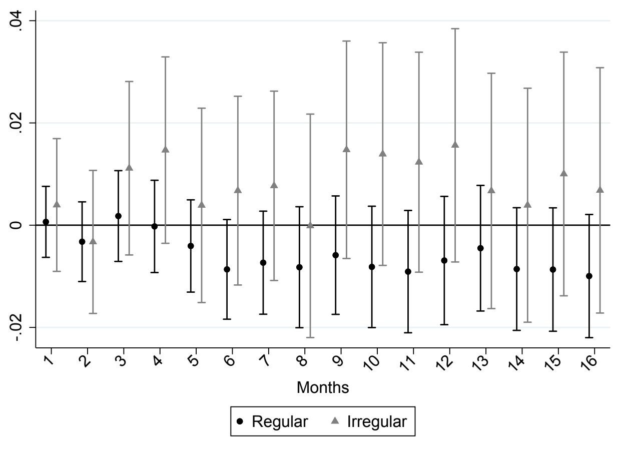
(b) # Certificate by end day of the week

Note: Figure reports the effect of HV by outcome of the visit on the number of certificates by day of the week start and end. Coefficients obtained by separate regressions by day of the week. Coefficients obtained from OLS regressions including indicators for workers subject to HV with regular and irregular outcome, demographic characteristics (gender, age), job characteristics (sector dummies, average salary in the 6 months before the experiment, part time dummy, dummy for permanent contract), fixed effects at INPS local office level, number of days spent on sick leave in the period of the experiment (11/22/2017-01/05/2018), number of certificates, average duration, and number of days spent on sick leave in the 6 months before the experiment. Figures include confidence interval at 95% with clustered standard errors at INPS local office level.

Figure 11: Effect of HV on Workers' career in the Public Sector in the 16 months after the Experiment



(a) Earnings



(b) Non-Employed in Public sector

Note: Effect of HV on workers' career in the public sector by visit outcome. Panel (a) reports coefficients for the effect of HV on regular and irregular worker by month on total take-home pay in the 16 months after the experiment. Panel (b) reports results for a linear probability model for the probability of being not employed in the public sector per month. Coefficients estimated by running a separate regression for each month with indicators for workers subject to HV with regular and irregular outcome, demographic characteristics (gender, age), job characteristics (sector dummies, average salary in the 6 months before the experiment, part time dummy, dummy for permanent contract), fixed effects at INPS local office level, number of days spent on sick leave in the period of the experiment (11/22/2017-01/05/2018), number of certificates, average duration, and number of days spent on sick leave in the 6 months before the experiment. Figures include confidence interval at 95% with clustered standard errors at local office level.

Tables

Table 1: Summary Statistics at Individual Level

Variables	Average	Se	Minimum	Median	Maximum
Female	0.724	0.447	0	1	1
Age	53.423	8.463	24	55	67
North	0.394	0.489	0	0	1
Center	0.177	0.381	0	0	1
South	0.429	0.495	0	0	1
School and University	0.396	0.489	0	0	1
Central Administration	0.061	0.239	0	0	1
Local Administration	0.234	0.423	0	0	1
Health Sector	0.310	0.462	0	0	1
Permanent Contract	0.948	0.222	0	1	1
Part Time	0.060	0.238	0	0	1
(log) Mean Monthly Earnings	7.658	0.338	0	8	10
Days on sick leave in following 16 months	48.859	70.354	0	21	551
Certificates in following 16 months	6.180	7.439	0	4	190
Average Certificate duration in following 16 months	7.509	8.980	0	4	92
Number of Certificates (bef. exp.)	2.332	3.107	0	1	57
Number of Days (bef. exp.)	21.827	35.898	0	5	315
Mean Duration Certificate (bef. exp.)	6.754	10.217	0	3	92
Home Visits and outcome: individual					
Individual subject to Home Visit	0.096	0.294	0	0	1
Outcome Home Visit: Regular	0.076	0.265	0	0	1
Outcome Home Visit: Irregular	0.020	0.138	0	0	1
Home Visits and outcome: certificate					
Certificates subject to Home Visit	0.073	0.260	0	0	1
Outcome Home Visit: Regular	0.058	0.234	0	0	1
Outcome Home Visit: Irregular	0.014	0.119	0	0	1
# Workers	43718				

Note: Summary statistics at individual level for public sector employees who had at least one ongoing sick leave certificate in the period of the experiment (11/22/2017-01/05/2018) and were randomly selected in the experiment. (log) Mean Monthly Earnings is the log of average earnings in the public sector for the worker from May to October 2017. Number of Certificates (bef. exp.), Number of Days (bef. exp.), and Mean Duration Certificate (bef. exp.) are the number of certificates, the total number of days on sick leave, and the average duration of certificates in the 6 months before the experiment.

Table 2: Normalized Differences for Individual Characteristics of Treated and Control Workers

Variable	Avg Treatment	Avg Contol	Se Treatment	Se Control	Normalized Difference
Female	0.740	0.722	0.438	0.448	0.029
Age: 36-40	0.051	0.055	0.219	0.228	-0.014
Age: 41-45	0.091	0.095	0.287	0.293	-0.010
Age: 46-50	0.139	0.141	0.346	0.349	-0.005
Age: 51-55	0.192	0.204	0.394	0.403	-0.022
Age: 56-60	0.235	0.243	0.424	0.429	-0.013
Age: 61-65	0.226	0.204	0.419	0.403	0.039
Age: 66-67	0.039	0.022	0.193	0.147	0.070
Central Admin.	0.066	0.060	0.249	0.237	0.018
Local Admin.	0.196	0.238	0.397	0.426	-0.072
School	0.447	0.391	0.497	0.488	0.081
Health Sector	0.291	0.312	0.454	0.463	-0.032
Permanent	0.968	0.946	0.176	0.226	0.078
Part Time	0.050	0.061	0.217	0.240	-0.037
(log) Mean Monthly Earnings	7.676	7.656	0.350	0.337	0.041

Note: The Table reports normalized differences for demographic and job characteristics for treated and control individuals. Normalized differences computed as $\Delta = \frac{\bar{X}_T - \bar{X}_C}{(S_T^2 + S_C^2)^{\frac{1}{2}}}$ and reference value is 0.25. (log) Mean Monthly Earnings is the log of average earnings in the public sector for the worker from May to October 2017.

Table 3: Regression for the Effect of HV on Cumulative Days on Sick Leave in the 16 Months after the Experiment

Variables	(1) # Days	(2) # Days	(3) # Days	(4) # Days	(5) # Days
HV	18.405***	13.297***	-4.466**	-6.062***	-5.535***
	(1.889)	(1.977)	(1.912)	(1.695)	(1.692)
Duration sick leave in experiment			1.373***	1.034***	1.000***
			(0.043)	(0.038)	(0.036)
Mean Duration Certificate (bef. exp.)				0.056	0.025
				(0.066)	(0.065)
Number of Certificates (bef. exp.)				5.354***	5.277***
				(0.198)	(0.197)
Number of Days (bef. exp.)				0.167***	0.161***
				(0.030)	(0.029)
Female					-0.396
					(0.855)
Age: 36-40					2.389
					(1.458)
Age: 41-45					5.756***
					(1.417)
Age: 46-50					6.280***
					(1.473)
Age: 51-55					8.711***
					(1.448)
Age: 56-60					11.599***
					(1.403)
Age: 61-65					18.347***
					(1.606)
Age: 66-67					0.882
					(2.569)
Central Admin.					-10.562***
					(1.692)
Local Admin.					-9.695***
					(1.113)
Health					-3.877***
					(1.055)
Permanent					8.089***
					(1.240)
Part Time					-2.549
					(1.564)
(log) Mean Monthly Earnings					-14.074***
					(1.230)
Observations	43,742	43,739	43,739	43,739	43,092
Mean Dep	47.097	47.097	47.097	47.097	47.097
Sede FE	NO	YES	YES	YES	YES

Note: The Table reports estimates for the effect of HV on cumulative days on sick leave in the 16 months after the experiment (# Days). Regressions are estimated with OLS with the *reghdfe* stata command developed by (Correia, 2019). HV is a dummy equal to one if the worker was subject to HV in the period of the experiment. Days in the experiment is the number of days spent on sick leave in the period of the experiment (11/22/2017-05/01/2018). (log) Mean Monthly Earnings is the log of average earnings in the public sector for the worker from May to October 2017. Number of Certificates (bef. exp.), Number of Days (bef. exp.), and Mean Duration Certificate (bef. exp.) are the number of certificates, the total number of days on sick leave, and the average duration of certificates in the 6 months before the experiment. Mean Dep is the average for the dependent variable for individuals who did not receive a HV (i.e. the control group). Sede FE are local INPS office fixed effects. Sample size excluding singletons reported. Standard errors clustered at local office level. Level of significance: 0.1 *, 0.05 **, 0.01 ***.

Table 4: Regression for the Effect of HV on Cumulative Days on Sick Leave in the 16 Months after the Experiment: Alternative Specifications

Variables	(1) # Days	(2) # Days	(3) # Days	(4) # Days	(5) # Days
	Panel (a): Correction for Overlapping Sick Leave Spells				
HV	18.405***	12.917***	-4.474**	-6.026***	-5.521***
	(1.889)	(1.918)	(1.858)	(1.644)	(1.640)
	Panel (b): Fixed effects for Effective Duration				
HV	18.405***	13.297***	-3.942**	-5.647***	-5.095***
	(1.889)	(1.977)	(1.876)	(1.679)	(1.674)
	Panel (c): Theoretical Time in Experiment				
HV	18.405***	13.297***	-2.897	-4.190**	-3.660**
	(1.889)	(1.977)	(1.996)	(1.717)	(1.718)
	Panel (d): IV, Effective Duration with Theoretical Duration				
HV	18.405***	13.297***	-2.258	-3.715**	-3.198*
	(1.889)	(1.977)	(1.959)	(1.695)	(1.691)
Cragg-Donald F-test			172,716.237	146,675.391	142,925.837
Sede FE	NO	YES	YES	YES	YES
Past Cert.	NO	NO	YES	YES	YES
Controls	NO	NO	NO	NO	YES

Note: The Table reports estimates for the effect of HV on cumulative days on sick leave in the 16 months after the experiment. Regressions are estimated with OLS with *reghdfe* stata command developed by (Correia, 2019). HV is a dummy equal to one if the worker was visited in the period of the experiment. Sede FE are local INPS office fixed effects. Past Cert. includes: the number of certificates, the total number of days on sick leave, and the average duration of certificates in the 6 months before the experiment. Controls include: demographic characteristics (gender, and age dummies), and job characteristics (sector dummies, average salary in the 6 months before the experiment, part time dummy, dummy for permanent contract). Mean dep is the average for the dependent variable for individuals who did not receive a HV (i.e. the control group). Effective duration is the number of days spent on sick leave in the period of the experiment (11/22/2017-05/01/2018). Panel (a) reports effects of HV with a correction to the dependent variable to avoid double counting days in sick leave present in two separate sick leave claims (overlapping certificates). Panel (b) reports effects of HV with fixed effects for time spent on sick leave in the experiment. Panel (c) reports results with theoretical time spent on leave in the period of the experiment. This is computed based on reported certificate start and end date. Panel (d) reports effects of HV with (linear) effective time spent on leave in the period of the experiment instrumented with theoretical time spent on leave in the period of the experiment, based on reported certificate duration. Standard errors clustered at local office level. Level of significance: 0.1 *, 0.05 **, 0.01 ***.

Table 5: Extensive vs Intensive Margin for Sick Leave Use

Variables	(1) # Days	(2) # Certificates	(3) Mean Days in 16 months
HV	-5.535***	-0.248**	-0.755***
	(1.692)	(0.112)	(0.215)
Observations	43,092	43,092	43,092
Mean Dep	47.097	6.169	7.211
Controls	YES	YES	YES
Past Cert.	YES	YES	YES
Sede FE	YES	YES	YES

Note: The Table reports estimates for the effect of HV on cumulative days on sick leave in the 16 months after the experiment (# Days in 16 Months), cumulative number of certificates (# Cert in 16 months), and average certificate duration (Mean days in 16 months). The cumulative number of certificates corresponds to the number of sick leave spells and the average certificate duration corresponds to the average duration of sick leave spells. Regressions are estimated with OLS with *reghdfe* stata command developed by (Correia, 2019). HV is a dummy equal to one if the worker was visited in the period of the experiment. Sede FE are local INPS office fixed effects. Past Cert. includes: the number of certificates, the total number of days on sick leave, and the average duration of certificates in the 6 months before the experiment. Controls include: demographic characteristics (gender, and age dummies), and job characteristics (sector dummies, average salary in the 6 months before the experiment, part time dummy, dummy for permanent contract). Mean dep is the average for the dependent variable for individuals who did not receive a HV (i.e. the control group). Sample size excluding singletons reported. Standard errors clustered at local office level. Level of significance: 0.1 *, 0.05 **, 0.01 ***.

Table 6: Differences between Regular and Irregular Workers in terms of Observable Characteristics

Variables	(1) Irregular Outcome	(2) Irregular Outcome	(3) Irregular Outcome
Female	-0.043**	-0.037*	-0.040**
	(0.019)	(0.019)	(0.018)
Age: 36-40	-0.022	-0.025	-0.024
	(0.048)	(0.053)	(0.053)
Age: 41-45	-0.019	-0.019	-0.015
	(0.050)	(0.057)	(0.057)
Age: 46-50	-0.051	-0.058	-0.056
	(0.047)	(0.053)	(0.053)
Age: 51-55	-0.051	-0.051	-0.053
	(0.049)	(0.054)	(0.054)
Age: 56-60	-0.034	-0.034	-0.034
	(0.043)	(0.050)	(0.051)
Age: 61-65	-0.024	-0.027	-0.025
	(0.048)	(0.055)	(0.055)
Age: 66-67	-0.007	-0.011	-0.001
	(0.064)	(0.068)	(0.069)
Metropolis	0.063**		
	(0.028)		
Center	0.007		
	(0.024)		
South	0.058***		
	(0.020)		
Friday	0.058***	0.065***	
	(0.017)	(0.019)	
Central Admin.	-0.068***	-0.057**	-0.052**
	(0.023)	(0.022)	(0.023)
Local Admin.	-0.003	0.007	0.005
	(0.015)	(0.014)	(0.015)
Health Sector	-0.031*	-0.030*	-0.025
	(0.016)	(0.017)	(0.016)
Permanent	-0.083**	-0.083**	-0.085**
	(0.041)	(0.040)	(0.041)
Part Time	-0.050**	-0.045*	-0.047*
	(0.022)	(0.024)	(0.026)
(log) Mean Monthly Earnings	-0.028	-0.025	-0.031
	(0.023)	(0.022)	(0.021)
Number of Certificates (bef. exp.)	0.004	0.003	0.002
	(0.004)	(0.003)	(0.003)
Number of Days (bef. exp.)	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)
Mean Duration Certificate (bef. exp.)	0.000	-0.000	-0.000
	(0.001)	(0.001)	(0.001)
Duration Certificate: 1-4	0.540***	0.554***	0.555***
	(0.050)	(0.061)	(0.064)
Duration Certificate: 5-7	0.220***	0.243***	0.240***
	(0.045)	(0.043)	(0.035)
Duration Certificate: 8-9	0.090***	0.103***	0.105***
	(0.031)	(0.030)	(0.031)
Observations	4,287	4,282	4,251
Mean Dep	.199	.199	.199
Sede FE	NO	YES	YES
Date FE	NO	NO	YES

Note: The Table reports estimates of a linear probability model with dependent variable one if the worker is found irregularly on sick leave and zero otherwise. Sample restricted to workers subject to home visit. Regressions are estimated by OLS with *reghdfe* stata command developed by (Correia, 2019). Metropolis is a dummy taking value one for Rome (the capital of the country) and Milan (the major economic centre of the country). (log) Mean Monthly Earnings is the log of average earnings in the public sector for the worker from May to October 2017. Number of Certificates (bef. exp.), Number of Days (bef. exp.), and Mean Duration Certificate (bef. exp.) are the number of certificates, the total number of days on sick leave, and the average duration of certificates in the 6 months before the experiment. Sede fixed effects are INPS local office fixed effects, while Date fixed effects are fixed effects for the day in which the HV was performed. Sample size excluding singletons reported. Standard errors clustered at local office level. Level of significance: 0.1 *, 0.05 **, 0.01 ***.

Table 7: Probability of Sending a New Certificate within three days from end previous certificate: Regular vs Irregular

Variables	(1) Any renewal	(2) Any renewal	(3) Any renewal	(4) Any renewal
HV: Regular Outcome	0.201***	0.168***	0.041***	0.039***
	(0.019)	(0.017)	(0.010)	(0.010)
HV: Irregular Outcom	-0.114***	-0.144***	-0.211***	-0.209***
	(0.023)	(0.023)	(0.018)	(0.018)
Observations	59,416	59,334	59,334	58,447
Mean Dep	.407	.407	.407	.407
Demographics	NO	NO	NO	YES
Past Cert.	NO	NO	YES	YES
Sede FE	NO	YES	YES	YES
Date FE	NO	YES	YES	YES

Note: The Table reports estimates of a linear probability model with dependent variable one if the worker sends another certificate while the certificate is active or within 3 days from its end. Regression estimated at certificate level. Date fixed effects are fixed effect for the start date of the certificate. Controls include: demographic characteristics (gender, and age dummies), job characteristics (sector dummies, average salary in the 6 months before the experiment, part time dummy, dummy for permanent contract). Past Cert. include: number of days spent on sick leave in the period of the experiment (11/22/2017-01/05/2018), number of certificates, average duration, and number of days spent on sick leave in the 6 months before the experiment, certificate duration and active time of the certificate in the period of the experiment. Mean dep is the average for the dependent variable for certificates which did not receive a HV (i.e. the control group). Regressions are estimated with *reghdfe* stata command developed by (Correia, 2019). Sample size excluding singletons reported. Standard errors clustered at local office level. Level of significance: 0.1 *, 0.05 **, 0.01 ***.

Table 8: Regression for the Effect of HV on Career in the Public Sector over 16 months by Visit Outcome

	(1)	(2)	(3)	(4)	(5)
VARIABLES	M. not Public	M. not Public	Tot Earnings 12 Months	Tot Earnings 16 Months	Tot Earnings 16 Months
HV	-0.048			11.917	
	(0.061)			(222.296)	
HV: Regular Outcome		-0.091	86.514		166.699
		(0.066)	(176.623)		(241.712)
HV: Irregular Outcome		0.113	-531.833**		-563.974
		(0.128)	(268.100)		(378.831)
Observations	43,092	43,092	43,092	43,092	43,092
Mean Dep	0.272	0.272	27423.88	35439.403	35439.403
Demographics	YES	YES	YES	YES	YES
Past Cert.	YES	YES	YES	YES	YES
Sede FE	YES	YES	YES	YES	YES
Date FE	YES	YES	YES	YES	YES

Note: Estimates for the effect of HV on career outcomes in the 16 months after the experiment by visit outcome. Dependent variable are defined as follows: Months not in the public sector (Column 1 and Column 2); total earnings in the public sector in the 12 months after the experiment (Column 3); total earnings in the public sector in the 16 months after the experiment (Column 4 and Column 5). Regressions are estimated by OLS with *reghdfe* stata command developed by (Correia, 2019). Sede FE are local INPS office fixed effects. Past Cert. include: the number of certificates, the total number of days on sick leave, and the average duration of certificates in the 6 months before the experiment. Controls include: demographic characteristics (gender, and age dummies), and job characteristics (sector dummies, average salary in the 6 months before the experiment, part time dummy, dummy for permanent contract). Mean dep is the average for the dependent variable for individuals who did not receive a HV (i.e. the control group). Standard errors clustered at local office level. Level of significance: 0.1 *, 0.05 **, 0.01 ***.

Table 9: Take-up of Other Benefits in the 16 Months after the Experiment.

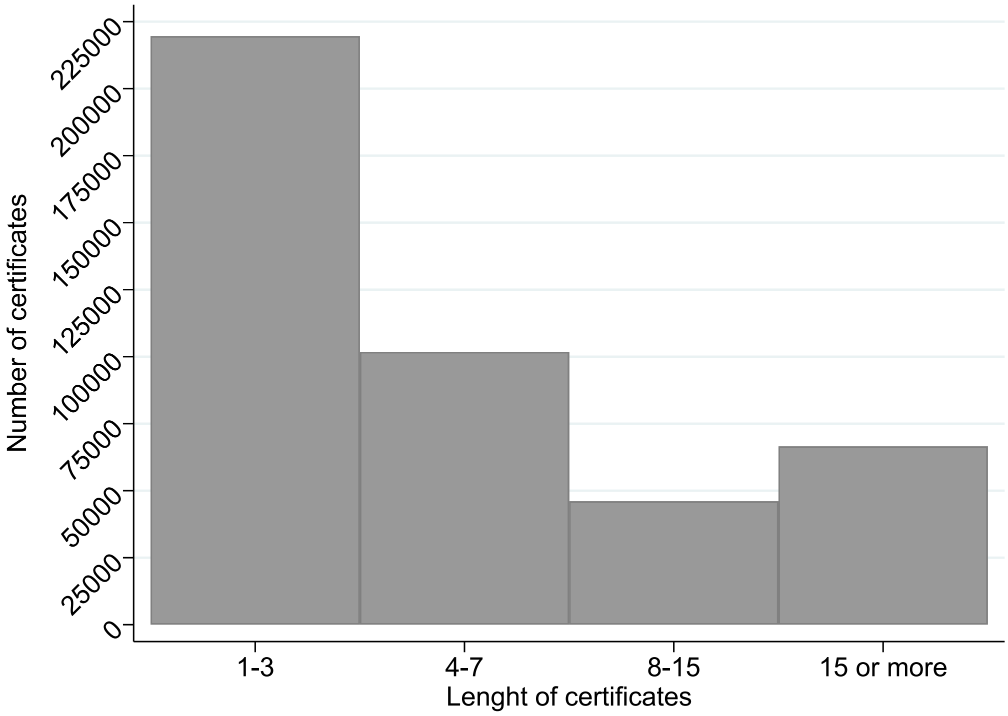
Variables	(1) Old Age Pension	(2) Disability Benefit
HV: Regular Outcome	-0.005	-0.001
	(0.003)	(0.003)
HV: Irregular Outcome	-0.005	-0.009**
	(0.006)	(0.005)
Observations	43,092	43,092
Mean Dep	.024	.013
Demographics	YES	YES
Past Cert.	YES	YES
Sede FE	YES	YES
Date FE	YES	YES

Note: Estimates for the effect of HV on take up probability of other benefits in the 16 months after the experiment by visit outcome. Dependent variables are defined as follows: Dummy for take-up of old- age pension benefits (Column 1); Dummy for take-up of disability benefits (Column 2). Regressions are estimated by OLS with *reghdfe* stata command developed by (Correia, 2019). Sede FE are local INPS office fixed effects. Past Cert. includes: the number of certificates, the total number of days on sick leave, and the average duration of certificates in the 6 months before the experiment. Controls include: demographic characteristics (gender, and age dummies), and job characteristics (sector dummies, average salary in the 6 months before the experiment, part time dummy, dummy for permanent contract). Mean dep is the average for the dependent variable for individuals who did not receive a HV (i.e. the control group). Standard errors clustered at local office level. Level of significance: 0.1 *, 0.05 **, 0.01 ***.

APPENDIX

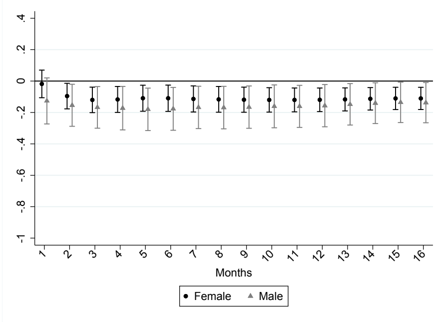
A Figures

Figure A1: Distribution of Certificates Length after the Experiment.

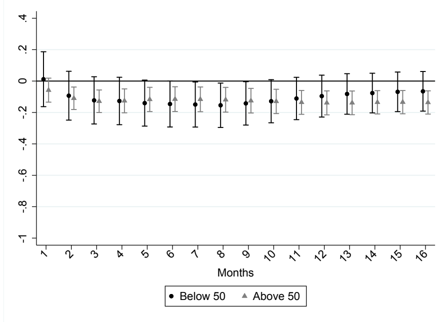


Note: Figure reports the number of certificates claimed by workers sampled in the experiment in the 16 months after the experiment (January 2018-April 2019).

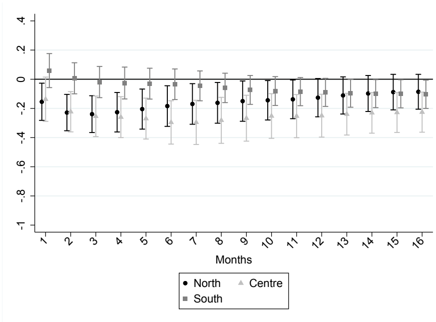
Figure A2: Effect of HV on Cumulative Days on Sick Leave Worker/Job Characteristics: Standardized



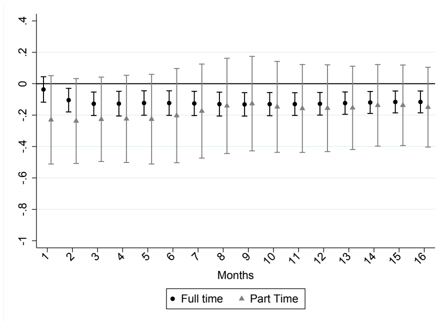
(a) Gender



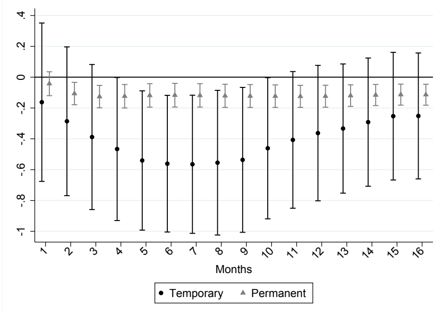
(b) Age



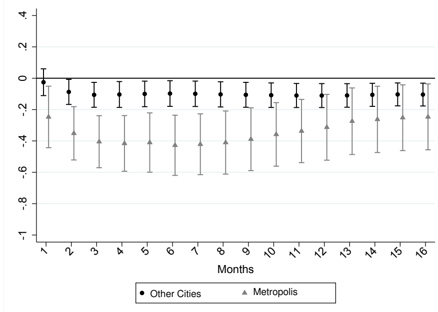
(c) Geographic Area



(d) Full-Time/Part Time



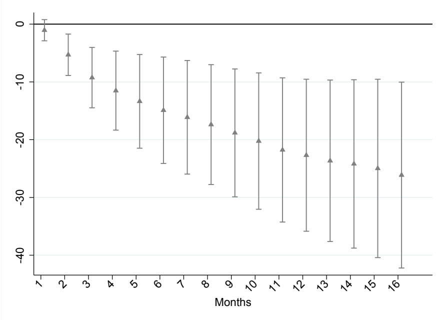
(e) Permanent/Temporary



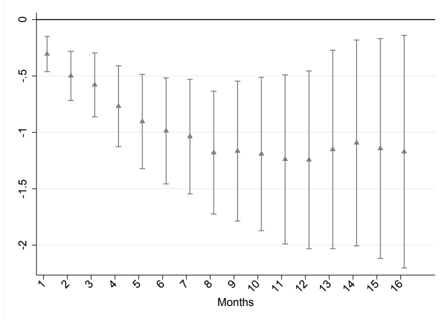
(f) Municipality Size

Note: Figure reports estimates for the effect of HV on cumulative number of days on sick leave use in the 16 months following the experiment. Regressions include a dummy for receiving HV outcome and its interaction with relevant dummies, demographic characteristics (gender, age), job characteristics (sector dummies, average salary in the 6 months before the experiment, part time dummy, dummy for permanent contract), fixed effects at INPS local office level, number of days spent on sick leave in the period of the experiment (11/22/2017-01/05/2018), number of certificates, average duration, and number of days spent on sick leave in the 6 months before the experiment. Standard errors clustered at local office level. Metropolis are the two main administrative centre in Italy: Rome (the capital of the country) and Milan (the main economic centre). Coefficients and 95% confidence intervals for the sum of the HV dummy and the appropriate interaction term reported. Point estimate and confidence interval normalized by the average cumulative number of days on sick leave for the corresponding group of workers not subject to HV at the same time horizon.

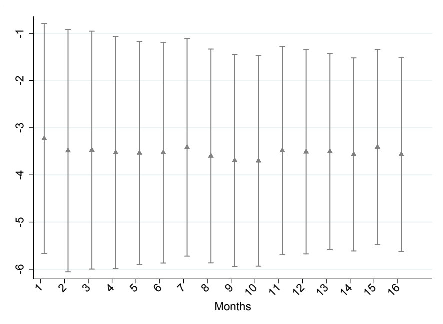
Figure A3: Effect of HV on Sick Leave over 16 Months after the Experiment for Irregular outcome: IV



(a) Cumulative days on sick leave



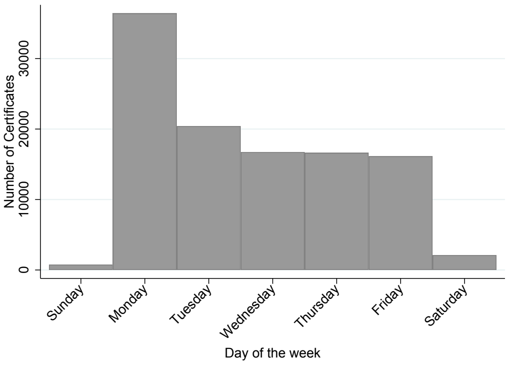
(b) Cumulative certificates



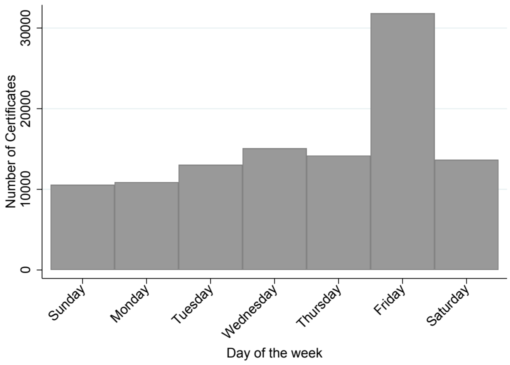
(c) Average Certificate Duration

Note: Figure reports estimates for the effect of HV on cumulative number of days spent on sick leave (Panel a), cumulative number of certificates (Panel b), and average duration of certificates (Panel c) in the 16 months after the end of the experiment for workers found irregularly on leave. The cumulative number of certificates corresponds to the number of sick leave spells and the average certificate duration corresponds to the average duration of sick leave spells. Regression are estimated separately for each month, and they include a dummy for being found irregularly on leave instrumented by a dummy for being subject to the inspection, demographic characteristics (gender, age), job characteristics (sector dummies, average salary in the 6 months before the experiment, part time dummy, dummy for permanent contract), fixed effects at INPS local office level, number of days spent on sick leave in the period of the experiment (11/22/2017-01/05/2018), number of certificates, average duration, and number of days spent on sick leave in the 6 months before the experiment. Standard errors clustered at local office level. Coefficients and 95% confidence interval for HV by outcome reported.

Figure A4: Start and End Day of the Week for Sick Leave Spells.



(a) Day of the Week: Start



(b) Day of the Week: End

Note: Statistics are based on data for public sector workers involved in the experiment. We consider sick leave spells which started between May and October 2017. Panel (a) reports the day of the week in which sick leave spells start while Panel (b) reports the day of the week in which sick leave spells end.

B Tables

Table B1: Balancing Regressions for the Probability of Receiving HV at Individual Level

Variables	(1) Visit	(2) Visit	(3) Visit	(4) Visit	(5) Visit	(6) Visit
Female	0.007**	0.007**	0.010***	0.011***	0.010***	0.009***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Age: 36-40	0.006	-0.000	-0.002	-0.003	-0.002	-0.002
	(0.009)	(0.009)	(0.008)	(0.009)	(0.009)	(0.009)
Age: 41-45	0.008	0.002	-0.001	-0.002	-0.002	-0.001
	(0.008)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Age: 46-50	0.009	0.002	-0.004	-0.005	-0.004	-0.003
	(0.009)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Age: 51-55	0.005	-0.000	-0.009	-0.011	-0.010	-0.009
	(0.008)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Age: 56-60	0.006	-0.001	-0.013*	-0.016**	-0.015**	-0.013*
	(0.009)	(0.008)	(0.007)	(0.007)	(0.007)	(0.007)
Age: 61-65	0.016*	0.010	-0.008	-0.011	-0.010	-0.009
	(0.009)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Age: 66-67	0.067***	0.050***	0.033***	0.028**	0.030**	0.032***
	(0.014)	(0.012)	(0.012)	(0.012)	(0.012)	(0.011)
Central Admin.	-0.005	-0.009	-0.006	-0.007	-0.005	-0.003
	(0.009)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Local Admin.	-0.025***	-0.016***	-0.003	-0.003	-0.001	-0.000
	(0.006)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Health Sector	-0.020***	-0.017***	-0.010**	-0.010***	-0.008**	-0.008**
	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Permanent	0.032***	0.022***	0.006	0.003	0.003	0.002
	(0.006)	(0.005)	(0.005)	(0.005)	(0.004)	(0.004)
Part Time	0.005	0.006	0.002	0.002	0.001	0.002
	(0.007)	(0.006)	(0.005)	(0.005)	(0.005)	(0.005)
(log) Mean Monthly Earnings	0.012**	0.015***	0.014***	0.017***	0.015***	0.015***
	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Duration sick leave in experiment			0.006***	0.006***		
			(0.000)	(0.000)		
Number of Certificates (bef. exp.)				0.000	-0.000	0.000
				(0.001)	(0.001)	(0.001)
Number of Days (bef. exp.)				0.000*	0.000**	0.000
				(0.000)	(0.000)	(0.000)
Mean Duration Certificate (bef. exp.)				0.001**	0.001**	0.000
				(0.000)	(0.000)	(0.000)
Observations	43,269	43,266	43,266	43,266	43,266	43,266
Mean Dep	.096	.096	.096	.096	.096	.096
Sede FE	NO	YES	YES	YES	YES	YES
Time in Exp. FE	NO	NO	NO	NO	YES	NO
Theoretical Time in Exp. FE	NO	NO	NO	NO	NO	YES

Note: The table reports results for a linear probability model at individual level with dependent variable a dummy equal to one if the worker is subject to a HV in the period of the experiment (11/22/2017-01/05/2018) and zero otherwise. All regressions are estimated by OLS. with the *reghdfe* stata command developed by (Correia, 2019). From Column 2, the regression includes fixed effects for the INPS local office (*Sede*). Column 5 includes dummies for the number of days on sick leave in the period of the experiment (Duration sick leave in exp.). Column 6 includes dummies for the number of days spent on sick leave in the period of the experiment based on stated certificate duration. (log) Mean Monthly Earnings is the log of average earnings in the public sector for the worker from May to October 2017. Number of Certificates (bef. exp.), Number of Days (bef. exp.), and Mean Duration Certificate (bef. exp.) are the number of certificates, the total number of days on sick leave, and the average duration of certificates in the 6 months before the experiment. Sample size excluding singletons reported. Standard errors clustered at local office level. Level of significance: 0.1 *, 0.05 **, 0.01 ***.

Table B2: Balancing Regressions for the Probability of Receiving HV at Certificate Level

Variables	(1) Visit	(2) Visit	(3) Visit	(4) Visit
Female	0.007***	0.004*	0.006***	0.007***
	(0.003)	(0.002)	(0.002)	(0.002)
Age: 36-40	0.003	0.002	0.001	-0.000
	(0.008)	(0.007)	(0.007)	(0.007)
Age: 41-45	0.006	0.004	0.001	0.001
	(0.007)	(0.006)	(0.006)	(0.006)
Age: 46-50	0.004	0.002	-0.001	-0.002
	(0.007)	(0.006)	(0.006)	(0.006)
Age: 51-55	0.001	0.001	-0.003	-0.004
	(0.006)	(0.006)	(0.005)	(0.006)
Age: 56-60	-0.001	-0.001	-0.007	-0.008
	(0.007)	(0.006)	(0.006)	(0.006)
Age: 61-65	0.005	0.008	-0.001	-0.002
	(0.007)	(0.006)	(0.006)	(0.006)
Age: 66-67	0.013	0.017*	0.010	0.007
	(0.010)	(0.009)	(0.009)	(0.009)
Central Admin.	-0.003	-0.000	0.001	0.001
	(0.007)	(0.006)	(0.005)	(0.005)
Local Admin.	-0.017***	-0.004	0.003	0.003
	(0.004)	(0.003)	(0.003)	(0.003)
Health Sector	-0.015***	-0.004	-0.000	-0.001
	(0.004)	(0.003)	(0.003)	(0.003)
Permanent	0.021***	0.013***	0.004	0.003
	(0.005)	(0.004)	(0.004)	(0.004)
Part Time	0.004	0.003	-0.000	0.000
	(0.005)	(0.004)	(0.004)	(0.004)
(log) Mean Monthly Earnings	0.014***	0.013***	0.009***	0.011***
	(0.004)	(0.003)	(0.003)	(0.003)
Duration sick leave in experiment			0.005***	0.004***
			(0.000)	(0.000)
Number of Certificates (bef. exp.)				-0.000
				(0.000)
Number of Days (bef. exp.)				0.000
				(0.000)
Mean Duration Certificate (bef. exp.)				0.000**
				(0.000)
Observations	58,728	58,647	58,647	58,647
Mean Dep	.073	.073	.073	.073
Sede FE	NO	YES	YES	YES
Date start FE	NO	YES	YES	YES

Note: The table reports results for a linear probability model at certificate level with dependent variable a dummy equal to one if the worker is subject to a HV in the period of the experiment (11/22/2017-01/05/2018) and zero otherwise. All regressions are estimated by OLS. with the *reghdfe* stata command developed by (Correia, 2019). From Column 2, the regression includes fixed effects for the INPS local office (*Sede*). Column 5 includes dummies for the number of days on sick leave in the period of the experiment (Duration sick leave in exp.). Column 6 includes dummies for the number of days spent on sick leave in the period of the experiment based on stated certificate duration. (log) Mean Monthly Earnings is the log of average earnings in the public sector for the worker from May to October 2017. Number of Certificates (bef. exp.), Number of Days (bef. exp.), and Mean Duration Certificate (bef. exp.) are the number of certificates, the total number of days on sick leave, and the average duration of certificates in the 6 months before the experiment. Sample size excluding singletons reported. Standard errors clustered at local office level. Level of significance: 0.1 *, 0.05 **, 0.01 ***.

Table B3: Regression for the Effect of HV on Cumulative Days on Sick Leave in 16 Months after the Experiment: Workers with only one Certificate

Variables	(1)	(2)	(3)	(4)	(5)
HV	13.139***	5.650***	-4.644**	-6.475***	-6.003***
	(2.016)	(2.140)	(2.099)	(1.865)	(1.834)
Days in experiment			1.164***	0.945***	0.910***
			(0.057)	(0.051)	(0.051)
Mean Duration Certificate (bef. exp.)				0.118	0.098
				(0.083)	(0.082)
Number of Certificates (bef. exp.)				5.422***	5.309***
				(0.221)	(0.221)
Number of Days (bef. exp.)				0.161***	0.154***
				(0.036)	(0.035)
Observations	33,385	33,382	33,382	33,382	32,885
Mean Dep	39.972	39.972	39.972	39.972	39.972
Controls	NO	NO	NO	NO	YES
Past Cert.	NO	NO	YES	YES	YES
Sede FE	NO	YES	YES	YES	YES

Note: The Table reports estimates for the effect of HV on cumulative days on sick leave in the 16 months after the experiment (# Days). Regressions are estimated with OLS with the *reghdfe* stata command developed by (Correia, 2019). The sample is restricted to workers with only one ongoing certificate in the period of the experiment. HV is a dummy equal to one if the worker was subject to HV in the period of the experiment. Days in the experiment is the number of days spent on sick leave in the period of the experiment (11/22/2017-05/01/2018). (log) Mean Monthly Earnings is the log of average earnings in the public sector for the worker from May to October 2017. Number of Certificates (bef. exp.), Number of Days (bef. exp.), and Mean Duration Certificate (bef. exp.) are the number of certificates, the total number of days on sick leave, and the average duration of certificates in the 6 months before the experiment. Mean Dep is the average for the dependent variable for individuals who did not receive a HV (i.e. the control group). Sede FE are local INPS office fixed effects. Sample size excluding singletons reported. Standard errors clustered at local office level. Level of significance: 0.1 *, 0.05 **, 0.01 ***.

Table B4: Extensive vs Intensive Margin for Sick Leave Use by Visit Outcome

Variables	(1) # Days	(2) # Certificates	(3) Mean Days in 16 months
HV: Regular	-4.285**	-0.148	-0.480*
	(1.940)	(0.129)	(0.245)
HV: Irregular	-10.185***	-0.621***	-1.780***
	(2.431)	(0.192)	(0.304)
Observations	43,092	43,092	43,092
Mean Dep	47.097	6.169	7.211
Controls	YES	YES	YES
Past Cert.	YES	YES	YES
Sede FE	YES	YES	YES

Note: The Table reports estimates for the effect of HV, by outcome of the inspection, on cumulative days on sick leave in the 16 months after the experiment (# Days in 16 Months), cumulative number of certificates (# Cert in 16 months), and average certificate duration (Mean days in 16 months). The cumulative number of certificates corresponds to the number of sick leave spells and the average certificate duration corresponds to the average duration of sick leave spells. Regressions are estimated with OLS with *reghdfe* stata command developed by (Correia, 2019). HV is a dummy equal to one if the worker was visited in the period of the experiment. Sede FE are local INPS office fixed effects. Past Cert. includes: the number of certificates, the total number of days on sick leave, and the average duration of certificates in the 16 months before the experiment. Controls include: demographic characteristics (gender, and age dummies), and job characteristics (sector dummies, average salary in the 6 months before the experiment, part time dummy, dummy for permanent contract). Mean dep is the average for the dependent variable for individuals who did not receive a HV (i.e. the control group). Sample size excluding singletons reported. Standard errors clustered at local office level. Level of significance: 0.1 *, 0.05 **, 0.01 ***.