

Di Porto, Edoardo; Garibaldi, Pietro; Mastrobuoni, Giovanni; Naticchioni, Paolo

Working Paper

The Perverse Effect of Flexible Work Arrangements on Informality

CESifo Working Paper, No. 12115

Provided in Cooperation with:

Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

Suggested Citation: Di Porto, Edoardo; Garibaldi, Pietro; Mastrobuoni, Giovanni; Naticchioni, Paolo (2025) : The Perverse Effect of Flexible Work Arrangements on Informality, CESifo Working Paper, No. 12115, Munich Society for the Promotion of Economic Research - CESifo GmbH, Munich

This Version is available at:

<https://hdl.handle.net/10419/331581>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

CES ifo

**12115
2025**

September 2025

Working Papers

The Perverse Effect of Flexible Work Arrangements on Informality

Edoardo Di Porto, Pietro Garibaldi, Giovanni Mastrobuoni,
Paolo Naticchioni

CES ifo

Imprint:

CESifo Working Papers

ISSN 2364-1428 (digital)

Publisher and distributor: Munich Society for the Promotion
of Economic Research - CESifo GmbH

Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740

Email office@cesifo.de
<https://www.cesifo.org>

Editor: Clemens Fuest

An electronic version of the paper may be downloaded free of charge

- from the CESifo website: www.ifo.de/en/cesifo/publications/cesifo-working-papers
- from the SSRN website: www.ssrn.com/index.cfm/en/cesifo/
- from the RePEc website: <https://ideas.repec.org/s/ces/ceswps.html>

The Perverse Effect of Flexible Work Arrangements on Informality*

Edoardo Di Porto[†] Pietro Garibaldi[‡] Giovanni Mastrobuoni^{§¶}

Paolo Naticchioni^{||}

September 3, 2025

Abstract

Flexible work arrangements (FWAs) are often promoted as a means to regularise informal labour. Utilising unique Italian administrative data that links employer-employee records, daily voucher usage by firms, and randomly timed labour inspections (2014-2017), we demonstrate that FWAs can also hinder enforcement and increase undeclared work. We document that, upon inspection, some firms validate undeclared work with FWAs on the spot, raising the probability of FWA usage by 0.88 percentage points (18%) on average, with the largest increases occurring on the day of and the day after the inspection. A simple partial-equilibrium labour-demand model with heterogeneous tax morale rationalises these “on-the-spot” validations as an enforcement-avoidance margin. The post-inspection increase vanishes when firms are required to pre-notify the tax authority of their use of FWAs. Moreover, when FWAs are completely abolished, presumptive misusers substitute FWAs with temporary contracts.

*We thank Bruno Anastasia, Saverio Bombelli, Lorenzo Cappellari, Tullio Jappelli, Marteen Lindeboom, Roberto Nisticó, Lorenzo Pandolfi, Matteo Paradisi, Annalisa Scognamiglio, and participants at conferences and seminars at AIEL, Bank of Italy, Inps, CSEF, Mannheim, Georgian Economic Association, University of Venice. We would also like to thank the Direzione Centrale Studi e Ricerche as well as the Direzione Centrale Entrate of the Italian Social Security Administration (INPS), for their invaluable support with the data and the institutional setting, as well as the sponsors of the VisitINPS program for their financial support. The findings and conclusions expressed are solely those of the authors and do not represent the views of INPS. All errors are our own.

[†]University of Naples “Federico II” and CSEF - UCFS Uppsala University - CESifo Munich, edoardo.diporto@unina.it

[‡]Collegio Carlo Alberto, University of Torino, CEPR, IZA, pietero.garibaldi@unito.it

[§]Collegio Carlo Alberto, University of Torino, CEPR, IZA, giovanni.mastrobuoni@carloalberto.org

[¶]Corresponding author.

^{||}Roma Tre University - IZA, paolo.naticchioni@uniroma3.it

Keywords: informality, labour vouchers, flexible work arrangements, occasional work, zero-hour contracts.

JEL codes: J23, H26.

1 Introduction

In most advanced economies, the underground economy is estimated to account for 10-20% of GDP (Schneider and Enste, 2002).¹ Furthermore, undeclared work is notably persistent and, if anything, its share has been estimated to have increased over time (Ulyssea, 2018).

Partly in response to shadow work, many governments have introduced more flexible labour contracts, known as flexible work arrangements (FWAs) (ILO, 2013).² Most European countries have arrangements with no guaranteed working hours, allowing firms to swiftly adjust labour demand while offering workers more adaptable schedules.³ Boeri et al. (2020) provide a recent survey of the small but expanding economic literature on flexible arrangements.⁴

Although there is little causal evidence regarding the relationship between FWAs and undeclared work, many experts believe that this relationship should be negative.⁵ Among others, the *European Union Agency for the Improvement of Living and Working Conditions* mentions that these new forms of employment have been developed to help formalise undeclared work practices.⁶ The *European Platform Tackling Undeclared Work* has stated that FWAs, in the form of labour vouchers (government-backed payment instruments that households or firms can use to pay for casual yet legal work), should target areas where undeclared work is prevalent (see Williams, 2018).

This study challenges the notion that FWAs inherently reduce undeclared work. While they

¹In developing countries the share is closer to 50% (Schneider and Enste, 2000).

²For example, according to the ILO, Belgium, France, Finland, Denmark, and Switzerland, introduced FWAs to discourage undeclared work.

³Chen et al. (2019) estimate that Uber drivers, possibly due to selection, benefit enormously from real-time flexibility, and Chan (2018) shows that emergency department physician work schedules can distort effort allocation and patient care.

⁴Katz and Krueger (2019) cover FWAs for the US and Adams and Prassl (2018) and Datta et al. (2019) for Europe.

⁵There is also evidence that rigid employment protection legislation reduces job flows and pushes firms to hire workers with more temporary contracts or off the books (for an early and a more recent survey, see Schneider and Enste, 2000, Ulyssea, 2020). Theoretical and empirical contributions that show that labour market rigidities increase informality include Blanchard and Portugal (2001), Fugazza and Jacques (2004), Albrecht et al. (2009b), Maloney (2004), Johnson et al. (1998), and (DiPorto et al., 2017).

⁶See <https://www.eurofound.europa.eu/topic/undeclared-work>.

offer flexibility, we demonstrate that such arrangements can hinder labour inspectors' ability to identify undeclared work. When misused, FWAs, such as the Italian labour vouchers examined here, can enable firms to legitimise the physical presence of workers engaged in undeclared employment.

Similar work arrangements exist in other countries, for example, UK's "Zero Contract Hours," where workers may work more than the officially declared number of hours, or Italy's labour vouchers, where for a day of work a worker may receive a single voucher for his/her work, so as to justify his/her physical presence in the workplace, and be paid the rest under the table. The Italian Ministry of Labour in 2016 realised that some firms may strategically comply with voucher requirements only when faced with an imminent inspection, similar to "a passenger validating their ticket only when the inspector arrives."⁷ The Ministry report also provides evidence that about 10% of FWA workers were previously employed by the same firm but with regular contracts.

Consequently, we develop a model of firms' labour demand for different contract types, allowing for such "spot validation." Firms are allowed to differ based on their tax morale. When "spot validation" is prohibited, more flexible job contracts reduce hiring, increase firing costs, and lower the demand for undeclared work.⁸ Conversely, when we allow contracts to have no work schedule, nothing changes for tax-abiding firms, "while low tax morale employers" report the working hours worked by their casual employee only when they are inspected. This option allows firms with low tax morale to increase their use of undeclared work. Moreover, the model predicts that FWAs may potentially crowd out fixed-term contracts.

We bring these predictions to the data in Italy, a country with a sufficiently large underground economy and legislation that has first liberalised and later abolished FWAs. Within

⁷See Ministero del Lavoro (2016).

⁸This is in line with, among others, Albrecht et al. (2009a), Bosch and Esteban-Pretel (2012), and Ulysea (2018).

the European Union, only Eastern European countries, Spain, and Greece have a higher prevalence of undeclared work⁹ and, beginning about 15 years ago, employers in Italy could purchase 10-euro vouchers to pay for work without needing a standard labour contract. The worker would later exchange vouchers for money. This system was intended to discourage undeclared work by eliminating bureaucracy and reducing hiring and firing costs.¹⁰

Labour force surveys often lack detail to identify FWA arrangements (Katz and Krueger, 2019), and when paid sums do not contribute to future social security benefits, even administrative data can be uninformative. This is not the case in Italy, where government-backed vouchers are fully recorded.¹¹

We use a unique data set drawn from three separate Italian administrative records: i) employer-employee social security records that cover the period 2014-2017; ii) daily firm-level usage of vouchers between 2014 and 2017; and finally, iii) data on all labour inspections between 2014 and 2017. Our data reveal that almost one in four companies in 2016 was using vouchers. Exploiting documented unpredictability of the timing of labour inspections, we find clear evidence that—as soon an inspection starts—some firms tend to immediately increase their use of vouchers. The increase in the likelihood of using vouchers immediately after an inspection is 0.88 percentage points, which corresponds to a relative increase of approximately 18%. The greatest changes occur on the day of inspection and the day after, respectively, 1.5 (30%) and 1.4 percentage points (29%). Moreover, such a jump ceases to exist when the government introduces a small change in legislation, requiring firms to announce the use of vouchers with at least one hour notice.

Next, we test whether FWAs displace regular work. Dividing inspected firms into those that, upon inspection, on average, increased their use of vouchers and thus are presumably

⁹According to Williams et al. (2017), 17.2% of Italian work is undeclared (the EU average is 16.4%).

¹⁰In Europe voucher-based work is also available in Austria, Belgium, Ireland, Finland, France, Hungary, Lithuania, and Slovenia (Mandl, 2020).

¹¹Additionally, as shown by Mas and Pallais (2017), workers often select into FWA contracts, complicating the estimation of counterfactual scenarios. We discuss the identification in Section 5.1.

more likely to misuse vouchers, and those that did not, we analyse what these two sets of firms did in March 2017 when vouchers were indeed abolished.

This measure for misbehaving firms is subject to misclassification error, as changes in individual-level use of FWAs around labour inspections may reflect unobservable demand shocks. However, we show that our main estimates represent lower bounds of the true effects. In addition, we argue that varying the definition of misbehaviour, and thus the size of the effects, can be informative about the fraction of firms hiding undeclared work. We show that approximately 18% of the inspected firms that use vouchers misuse FWAs.

In line with the predictions of the model, “misusing” firms are shown, when vouchers are abolished, to revert to the next most flexible work contracts, hiring approximately two additional fixed-term workers, representing a 50% increase with respect to the pre-abolition average of misusing firms. Due to these substitution effects, the total declared wage bill, which includes the cost of vouchers, shows no changes, a possible rule-of-thumb behaviour of marginal firms.

To evaluate the impact of the abolition of vouchers on undeclared work, we compare the amount of evaded social security contributions for firms that have or have not used vouchers between January and October 2016: “treated” firms with vouchers tend to evade more when vouchers are abolished, while at the height of the use of vouchers (right before and after the introduction of the SMS requirement), evasion is about the same.

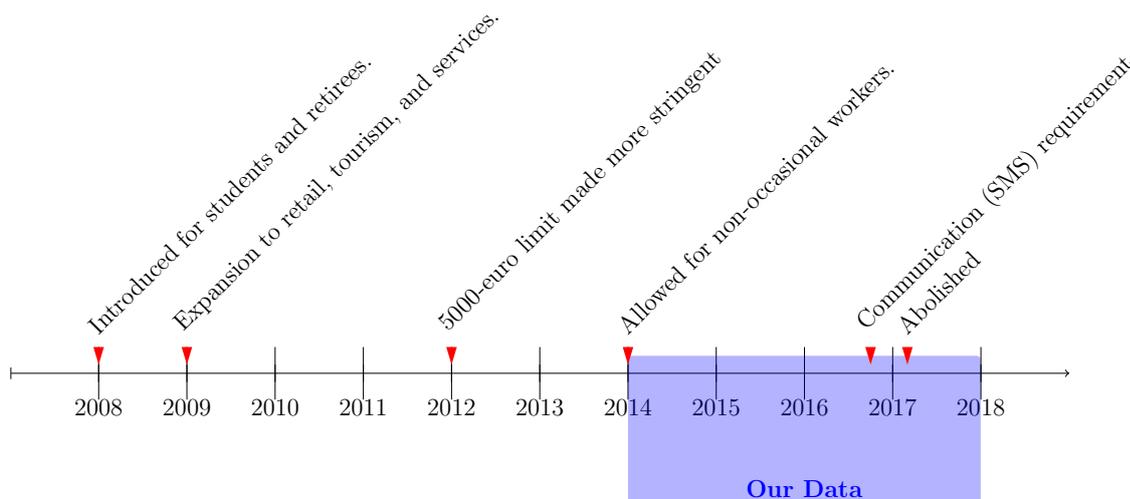
The study proceeds as follows. Section 2 describes the institutional setting. Section 3 presents and solves a simple labour demand model of labour vouchers and derives the optimal choice of contracts, with and without the option to go shadow. The section highlights the main empirical predictions. Section 4 describes the data set used in the paper. Sections 5 and 6 present the empirical models and the evidence, while Section 7 concludes.

2 The Institutional framework

In this section, we present regulatory details for the two main institutions of the analysis: Italian FWA, or vouchers, and labour inspections.

Vouchers: In 2008, the Italian legislator introduced FWAs in the extreme form of labour vouchers. Figure 1 outlines the timeline from the first time FWAs were introduced to the time they were abolished. Employers could buy vouchers from the Social Security Administration (INPS), or in tobacco stores, banks, and post offices. The vouchers looked like cheques (see Appendix Figure A6), and employers who wanted to pay someone for completing a very short temporary job would fill them with the worker’s social security number (“*Codice Fiscale*”) and the date of the job/task. Once the worker was paid with the voucher, he/she could easily exchange it for cash. For every 10 euros paid by the employer, the worker received 7.50 euros, 1.30 euros covered the social security contributions (less than the typical labour contract, where contribution rates are 1/3), 70 cents for health insurance and 50 cents for the commission fee paid to the social security administration.

Figure 1: Timeline of Voucher Legislation



Initially, vouchers had considerable restrictions: private citizens could only spend a total of €5000 vouchers for each worker who helped them in day-to-day casual work (house cleaning, gardening, etc.), and employers faced an even lower threshold (€2500). Moreover, only students and retirees were allowed to receive vouchers, and only in the agricultural sector.

Several small changes in these restrictions led to a steep increase in the use of FWAs. Initially, the centre-right government extended the use of vouchers to all workers in the agricultural sector, not just students and retirees. More limitations were lifted in the following years and, as shown in Figure 4, this led to rapid growth in the monthly number of 10-euro vouchers sold: from a few thousand in 2008 to a peak of almost 20 million in 2016.

In 2009, vouchers became available in the retail sector, tourism, and service sector, and for domestic workers. One year later, they were completely liberalised, opening up to all sectors and all workers. In 2012, the worker's €5000 limit was made more stringent, as it applied to the sum across all the employers and not to each employer separately, while a new limit was introduced: each firm could at most spend €2000 in vouchers on any single worker. In 2015, the labour reform (called the "Jobs Act") allowed vouchers not to be related to occasional work, and their annual limit increased to €7000 (leaving the amount firms could use on a single worker unchanged).

The use of vouchers reached a peak in 2016, when the pressure from labour unions to reform their use or completely abolish them intensified. In October 2016, a first reform was passed, and employers had to inform the Social Security Administration by text message at least 60 minutes before using a voucher. Several months later, as pressure increased and a new government was in power, the use of vouchers was outlawed.¹²

Labour Inspections: Inspections are organised secretly to maximise the chance of detecting any undeclared workers or any hazardous working conditions.¹³ They are divided into three

¹²A much more limited version of vouchers was reintroduced at the end of 2017.

¹³The rules can be found in the "Circolare INPS No. 76/2016."

phases. First, each inspection is secretly planned to ensure that the second phase, the subsequent access to the firm's premises, is carried out quickly by a sufficient number of inspectors, ensuring that nobody can leave the firm's premises. The second phase can last up to several days and ends with a preliminary report. Inspectors collect accounting books and any other useful information and interview employers and employees. During the last phase, which can last at most 90 days from the notification of the preliminary report (Art. 14 of Law No. 689/1981), visits to the firm are rarer, and the inspectors prepare the final report.

From the firm's standpoint, inspections occur at random and, therefore, unpredictable times, and their duration can last from a few days to several weeks. In order to deter any information leakage, corrupt inspectors who inform firms about an imminent inspection in exchange for money face up to 10 years of imprisonment (Article 319 of the Criminal Code). Depending on the outcome of the inspection, firms may have to pay a fine and/or rectify any irregularities before resuming production.

Before taking advantage of some of the institutional changes in our empirical analysis, we develop a labour demand model that generates potential links between labour demand, including labour vouchers, labour inspections, and the use of undeclared work.

3 A Labour Demand Model of Jobs, Temporary Jobs and FWAs

3.1 The Environment and the Institutions

Following the tradition of the search and matching literature and the Diamond-Mortensen-Pissarides model, we develop a stylised static labour demand model with heterogeneous firms characterised by a single job. There is a measure 1 of homogeneous workers and an

endogenous number of heterogeneous firms and jobs. Firm heterogeneity is defined by its technological probability of becoming unproductive. Each job (or firm) faces a probability λ of becoming unproductive, where $\lambda \in [0, 1]$ is drawn from a cumulative probability distribution $G(\lambda)$ when the worker and the firm meet, and $G(1) = 1$.¹⁴ A job produces output y for the fraction of time $1 - \lambda$, and produces 0 for the rest of the time (normalised to be one). After observing λ , the firm decides which type of contract to offer to a single worker, as contracts differ by their termination costs.

The wage ω paid to each worker is taken as given by the firm.¹⁵ The labour market is characterised by a payroll tax τ , regardless of the type of contract. The tax is paid on a flow basis by the firm and, at first, we assume that the tax cannot be evaded.

Later, in Section 3.2 we introduce the possibility of tax evasion, allowing contracts to be shadow. When labour is shadow, taxes are avoided, but firms risk being fined by the labour inspectors. Since many firms avoid the use of shadow employment, we add a further dimension of heterogeneity, a firm-specific “tax-evasion attitude” $\theta^j > 0$ associated to the status of being a tax dodger. For simplicity, we assume that the parameter θ^j can take only two values $\theta^j \in [\theta^l, \theta^h]$ with $\theta^l < \theta^h$, so that firms with $\theta = \theta^h$ are more likely to be tax dodgers. There is a proportion p of firms with high evasion attitude. The two types of firms may face different distributions of λ and $G(\lambda) = pG_1(\lambda|\theta^h) + (1 - p)G_2(\lambda|\theta^l)$, where $G_1(\cdot)$ and $G_2(\cdot)$ are the cumulative distributions for each type.

Without tax evasion, the main choice of the firm is about which contract to offer depending on λ . Broadly speaking, Italian legislation allows three different types of labour contracts: open-ended, fixed-term jobs, and FWAs/vouchers (we use the words FWAs or vouchers

¹⁴Cahuc et al. (2016) provide a matching model with heterogeneous jobs that features job heterogeneity and optimal contract selection. See Boeri and Garibaldi (2024) for a recent survey on the vast literature on fixed term contracts.

¹⁵In the model we keep the assumption that the wage ω is fixed across types of contract. A simple extension of the model would be to assume that more flexible labour contracts come with lower wages. The model would be solved with an endogenous wage and rent sharing. However, we prefer to focus on how λ influences labour demand.

interchangeably). Different contracts have different termination costs. When faced with an open-ended contract that is unproductive, the firm is better off paying a firing tax equal to $-F$. In line with Italian legislation, we assume that the tax is a multiple of the wage rate: $F = f\omega$.¹⁶ The formal modelling of labour contracts is in line with Cahuc et al. (2016). In what follows, we shall indicate with $J^{oe}(\lambda)$ the value to the firm of a λ job with an open-ended contract.

Fixed-term contracts are active for a fraction $1 - \rho^{ft}$ of the time. When a firm opens a fixed-term contract, the worker produces and is going to be paid for an expected duration equal to $1 - \rho^{ft}$, regardless of the specific value of the job λ . The advantage of a fixed-term contract is that the firm does not pay any firing costs when the productivity drop strikes (which happens with probability ρ^{ft}). However, the downside is that the firm pays the worker even if λ strikes and productivity drops to 0.¹⁷ In what follows, $J^{ft}(\lambda)$ indicates the value to the firm of a fixed-term contract.

Finally, the firm can open FWAs. FWAs do not have any dismissal costs, but are characterised by an expected duration $1 - \rho^v$ where ρ^v is considerably larger than ρ^{ft} .¹⁸ In practice, it is as if FWAs can be terminated at any time at no cost. $J^v(\lambda)$ is the value to the firm of an FWA.

There is a fixed cost of opening a job equal to K , which means that a firm will open a λ job as long as $J^i(\lambda) \geq K$ for at least one type of contract i . Conditional on the realised technological parameter λ , firms choose the type of contract to offer such that $J^*(\lambda) = \max_{\{i=[oe,ft,v]\}} \{J^i(\lambda), K\}$.

To solve this maximisation problem, we need to specify the expected value of different jobs.

¹⁶The model abstracts from the slight differences in social security contribution rates across contracts, but this should be second-order for our empirical analysis, as those differences do not vary over time with the arrival of labour inspectors.

¹⁷The data confirm that only a handful of fixed-term contracts end with a dismissal (0.5 percent in 2016), consistently with our assumption that no firing costs are paid to fixed-term contracts.

¹⁸Note that this is in line with the relatively low caps on the amounts firms are allowed to pay each worker in labour vouchers (see Section 2).

The value of an open-ended job is

$$J^{oe}(\lambda) = (1 - \lambda)(y - \tau - \omega) - \lambda F. \quad (1)$$

In contrast, the value of a fixed-term contract is

$$J^{ft}(\lambda) = (1 - \rho^{ft}) [(1 - \lambda)(y - \omega - \tau) - \lambda(\omega + \tau)] = (1 - \rho^{ft}) [(1 - \lambda)y - \omega - \tau].$$

where at rate ρ^{ft} the job is destroyed at no cost. Yet, as argued above, with probability $\lambda(1 - \rho^{ft})$ the firm is forced to pay the wage until the end of the contract.¹⁹

Finally, the value of an FWA is

$$J^v(\lambda) = (1 - \rho^v) [(1 - \lambda)(y - \omega - \tau)]. \quad (2)$$

Since all job values are monotonically decreasing in λ , the maximisation problem satisfies the reservation property. In other words, maximisation is solved by selecting particular values of λ (say λ^{ft} and λ^v) such that each contract is optimal in any particular subset of the domain of G in $[0, 1]$.

Furthermore, one can easily show that $J^{oe}(0) > J^{ft}(0) > J^v(0)$ and that $J^v(1) = 0 > J^{ft}(1) = -(\omega + \tau) > J^{oe}(1) = -F$.

Proposition 1. *Firms' optimal labour contract depends on two productivity thresholds λ^{ft} , λ^v such that $\forall \lambda < \lambda^{ft}$ firms choose an open-ended contract, $\forall \lambda \in [\lambda^{ft} - \lambda^v]$ firms choose a fixed-term contract, and $\forall \lambda > \lambda^v$ firms choose a flexible working arrangement. Moreover, firms do not open jobs $\forall \lambda : J(\lambda) > K$.*

The maximisation is thus an envelope of three downward-sloping lines. The reservation

¹⁹The value of the firing tax $F = f\omega$ has the restriction that $f < 1 + \frac{\tau}{\omega}$, so that the firing tax for open-ended contracts must be smaller than the firing tax paid to fixed term contract.

probability can be characterised as the solution to

$$J^{oe}(\lambda^{ft}) = J^{ft}(\lambda^{ft}); \quad \text{and} \quad J^{ft}(\lambda^v) = J^v(\lambda^v). \quad (3)$$

λ^{ft} is the expected duration that makes the firm indifferent between an open-ended job and a fixed-term job. Similarly, λ^v makes the firm indifferent between a FWAs and a fixed-term job.²⁰

The intuition behind this result is as follows. For a given net flow productivity $\tilde{y} = y - \omega - \tau$, a λ -type firm has a strong ordering regarding which contract to offer. On one extreme, open-ended contracts are ideal for jobs with a very long expected duration, while, on the other extreme, FWAs are ideal for jobs that are expected to last very little. Thus, introducing FWAs creates opportunities for labour demand that would otherwise be absent. Without FWAs firms may prefer not to sign labour contracts with such low expected durations. In particular, λ^{max} is the solution of $J^*(\lambda^{max}) = K$, and the firm does not sign labour contracts for any $\lambda > \lambda^{max}$. If the probability of meeting a worker is η , total employment is simply $n = \eta G(\lambda^{max})$, while $(1 - \eta G(\lambda^{max}))$ is non-employment.

²⁰The threshold λ s are:

$$\begin{cases} \lambda^{ft} &= \frac{\rho^{ft} \tilde{y}}{\rho^{ft} \tilde{y} + f\omega - (1 - \rho^{ft})(\omega + \tau)} \\ \lambda^v &= \frac{(\rho^v - \rho^{ft}) \tilde{y}}{(\rho^v - \rho^{ft}) \tilde{y} + (1 - \rho^{ft})(\omega + \tau)} \end{cases} \quad (4)$$

Note that existence of two thresholds-and thus two different contracts with fixed duration (vouchers and fixed term contracts)-requires that the duration of FWAs is sufficiently short or that

$$\rho^v > \frac{\rho^{ft} F}{F + (1 - \rho^{ft})(\omega + \tau)}.$$

3.2 Shadow Employment and the Misuse of FWA

We now introduce the possibility of evading taxes by under-reporting jobs. A shadow job allows firms to avoid paying tax τ . For each type of contract i and a given λ , let $J_j^{i,s}(\lambda, \theta^j)$ be the value of a λ shadow job. The value varies depending on the subscript j , which refers to the attitude to go shadow $\theta^j, j = l, h$. In addition, it will depend on γ , the probability of inspection, and on $C^i(\lambda)$, the fine imposed on the firm with undeclared work upon inspection. The fine is equal to a fixed term c^i and a term that depends on the duration of the productive relationship λ : $C(\lambda) = c^i + c(\lambda)$. We also assume that $c^{oe} \geq c^{ft} > 0$ and $c^v = 0$, so that FWA jobs face only variable fines. There is a natural assumption that $c(1) = 0$, since the fine on an “instant” job has to be negligible. Furthermore, we assume that $c'(\lambda) < 0$. According to common practice, inspectors will charge a higher fine to firms that appear to be in a longer-lasting employment relationship (and therefore have a job with a lower λ).

The decision to go shadow for a firm of type j with idiosyncratic probability λ is simply

$$J_j^{i,s}(\lambda, \theta^j) = (1 - \gamma)[J^i(\lambda) + \tau(1 - \lambda)(1 - \rho^i)] + \gamma[J^i(\lambda) - C^i(\lambda)] + \theta^j > J^i(\lambda), \quad i = \{oe; ft; v\}, \quad (5)$$

where $\rho^{oe} = 0$. This is the standard condition found in most of the shadow employment literature, namely that $(1 - \gamma)(1 - \rho^i)(1 - \lambda)\tau + \theta^j > \gamma C^i(\lambda)$, so that going shadow is optimal if the tax evaded is larger than the expected fine augmented by the civic cost of evading tax.²¹ To obtain simple analytical insights, we assume that the variable fine is linear in λ so that $c(\lambda) = c^i(1 - \lambda)$. For each underlying contract i there is a unique threshold $\lambda^{s,i}$ that characterises the decision to go shadow. There is a marginal shadow value $\lambda^{s,i}$

²¹The model assumes that in going shadow only the payroll tax is avoided, not the firing costs. We believe that the main reason for the existence of “implicit” contracts that make termination costs binding also when the job is shadow is that in real life workers have the option to sue the employer for having worked illegally, thus are likely to pretend to be compensated in the event of dismissal. Another example of implicit contracts are family firms. In those contexts the job held by the family member is shadow but termination costs are implicitly defined, since it is highly costly to sever the job of a family member.

that solves the equation $(1 - \gamma)(1 - \rho^i)(1 - \lambda^{s,i})\tau + \theta^j = \gamma C^i(\lambda^{s,i})$. We call this reservation value $\lambda^{s,i}(\theta^j)$, where s stands for shadow, i corresponds to the type of contract and θ^j is the tax evasion attitude. As the following proposition specifies, with a simple restriction on parameters, firms with $\theta = \theta^l$ never go shadow while firms with a high tax evasion attitude exercise the shadow option for $\lambda > \lambda^{s,v}(\theta^h)$, as the following proposition specifies.

Proposition 2. *Let the cost parameter $c > \frac{(1-\rho^{ft})(1-\gamma)\tau}{\gamma}$; let the tax evasion attitude for the low type $\theta^l < -(1 - \gamma)(1 - \rho^{ft})$ and let the tax evasion attitude for the high type $\theta^h \in [0, \gamma c - (1 - \gamma)(1 - \rho^{ft})\tau]$. It then follows that firms with $\theta = \theta^l$ never go shadow, while firms with $\theta = \theta^h$ exercise the shadow option for v and go shadow if $\lambda > \lambda^{s,v}$, where $\lambda^{s,v}(\theta^h)$ solves*

$$\lambda^{s,v}(\theta^h) = 1 - \frac{\theta^h}{\gamma c - \tau(1 - \gamma)(1 - \rho^v)} < 1 \quad (6)$$

Further, with $c^{oe} > F$ and $c^{ft} > f\omega$, only firms with a FWA job go shadow.

The proof is in the Appendix. Figure 2 reports the behaviour of the two firms. The black curve refers to firms that do not go shadow and divide their employment into open-ended, fixed-term, and FWA. The grey curve refers to firms that go shadow for $\lambda > \lambda^{s,v}(\theta^h)$. Formally, the amount of shadow employment is thus $n^s(\theta^h) = np(1 - G_1(\lambda^{s,v}|\theta^h))$. From this discussion and from Figure 2, it follows that the existence of FWAs reduces the amount of shadow employment in equilibrium. In a world with only open-ended and fixed-term contracts, any job with $\lambda > \lambda^{s,ft}(\theta^h)$ is shadow, and the fixed-term λ shadow threshold is lower than the FWA λ shadow threshold: $\lambda^{s,ft}(\theta^h) < \lambda^{s,v}$.

3.2.1 Misuse of FWA and Empirical Implications

Proposition 2 establishes the existence of shadow employment for jobs that are typically covered by FWAs. But when the institutional rules of FWAs allow for instant validation of labour contracts, firms can do even better than passively observe the arrival of a labour

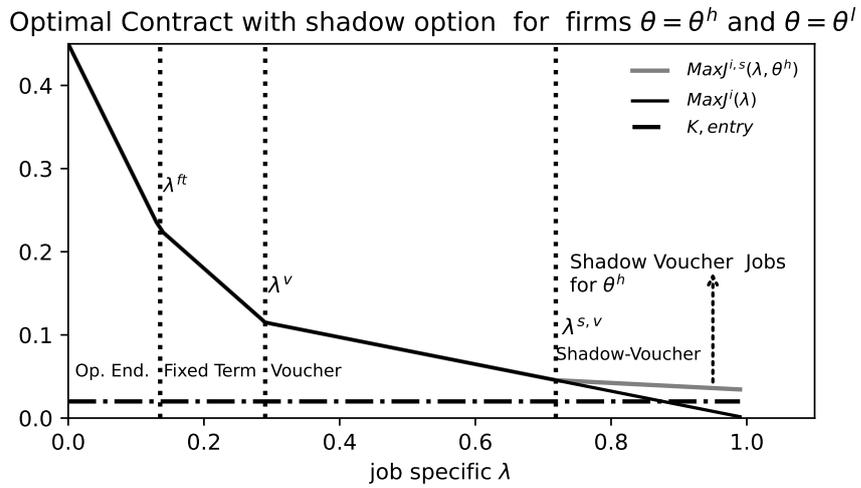


Figure 2: Optimal Contract without shadow employment and Shadow Option For Firms with high tax-evasion attitude.

Notes: The chart reports the value functions $\max[J^i(\lambda)]$ for $i = oe, ft, v$. Parameters are $y = 1; \omega = .4; \tau = .15; f = 3; \rho^{ft} = .28; \rho^v = .64; K = 0.2$. The chart considers also the shadow option with linear cost sanction $C = c^i + c\lambda$ with $c = .82; \gamma = .2, C^o = 0.03F; C^{ft} = 0.1(\omega + \tau). C^v = 0. \theta^h = 0.034$ and $\theta^l = 0.013$. The thresholds $\lambda^{ft}; \lambda^v$ and $\lambda^{s,v}$ are indicated in the chart. The grey envelope refers to the value of shadow employment.

inspector. When an inspector arrives, they can immediately sign a voucher for previously undeclared workers. We define “grey” jobs as the employment relationships that are shadow in normal times and are regularised when the inspectors arrive at the firm’s gate. In other words, “grey” jobs are shadow jobs of different contractual lengths that become flexible working arrangements when the inspector shows up.

Since “grey” jobs are black as long as the inspector does not show up, let $\tilde{J}_j^{i,s}(\lambda)$ be the value of a “grey” job that has the option to activate the voucher conditional on inspection, or the value of a shadow job that has the option to misuse vouchers. Formally, misusing vouchers adds an additional choice, generating an option value conditional on the inspector showing up at the gate. The decision to go “grey” for a firm of type j yields the following value function:

$$\begin{aligned} \tilde{J}^{i,s}(\lambda, \theta^j) = & (1 - \gamma)[J^i(\lambda) + \tau(1 - \lambda^i)(1 - \rho^{ft})] + \\ & + \theta^j + \gamma \left\{ \alpha \max \underbrace{[J^i(\lambda) - C(\lambda); J^v(\lambda)]}_{\text{option to misuse vouchers}} + (1 - \alpha)(J^i(\lambda) - C(\lambda)) \right\} > J^i(\lambda); \quad (7) \\ & i = \{o; ft; v\}, \end{aligned}$$

where the max operator inside the brackets is the option value of misusing the voucher.

Equation 7 implies that it is impossible during an inspection to pretend that a shadow worker had just been hired with an open-ended or a fixed-term contract; while it is possible to pretend that the worker was hired with a “voucher”, and such option is the economics behind the misuse of the vouchers.

Vouchers are misused when the following condition is satisfied:

$$\underbrace{(1 - \gamma)(1 - \rho^i)(1 - \lambda)\tau + \theta^j}_{\text{expected tax evaded}} > \gamma \underbrace{[(1 - \alpha)C^i(\lambda) + \alpha [J^i(\lambda) - J^v(\lambda)]]}_{\text{expected cost of pretending to have a FWA job}}. \quad (8)$$

Misuse of Voucher for firms $\theta = \theta^h$

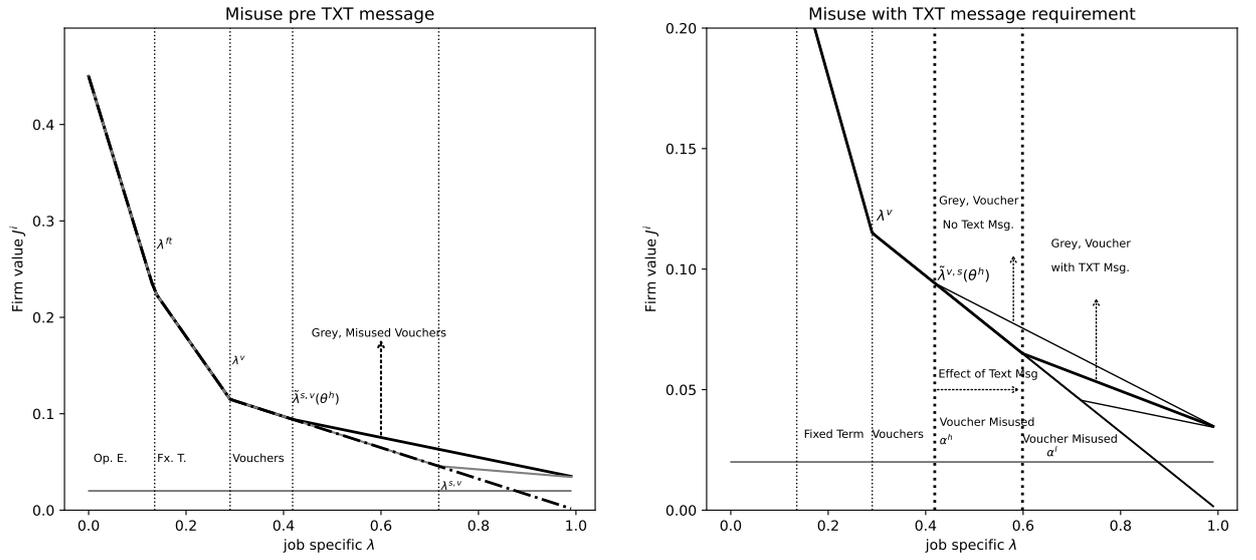


Figure 3: Optimal Contract without shadow employment and Shadow Option For Firms with high tax-evasion attitude.

Notes: The chart reports the value functions with misuse $\tilde{J}^{i,s}(\lambda, \theta^j)$ for $i = oe, ft, v$. Parameters are $y = 1; \omega = .4; \tau = .15; f = 3; \rho^{ft} = .28; \rho^v = .64; K = 0.2$. The chart considers also the shadow option with linear cost sanction $C = c^i + c\lambda$ with $c = .82; \gamma = .2, C^o = 0.03F; C^{ft} = 0.1(\omega + \tau). C^v = 0. \theta^h = 0.034$ and $\theta^l = 0.013$. The thresholds $\lambda^{ft}; \lambda^v, \lambda^{s,v}, \tilde{\lambda}^{s,v}$ are indicated in the chart. In the left hand chart $\alpha = 0.38$ while in the right hand side $\alpha = 0.24$. See text for function and parameters definitions.

Solving for λ , there is a threshold value $\tilde{\lambda}^{s,v}$ such that $\tilde{J}^{s,v}(\theta^h) = J^v$ and the firm exercises the shadow option for $\lambda > \tilde{\lambda}^{s,v}$.²²

The value of $\tilde{\lambda}^{s,v}$ is

$$\tilde{\lambda}^{s,v}(\theta^h) = 1 - \frac{\theta^h}{\gamma c(1 - \alpha) - (1 - \gamma)\tau(1 - \rho^v)}. \quad (10)$$

Comparing equation 6 with equation 10, it is easy to establish that $\tilde{\lambda}^{s,v}(\theta^h) < \lambda^{s,v}(\theta^h)$, meaning that the possibility to validate vouchers on the spot increases the amount of shadow employment, generating grey jobs. Moreover, the behaviour of the firm during an inspection is indicative of their type θ^j : only firms with $\theta = \theta^h$ misuse vouchers.

Figure 3 shows the case where $\tilde{\lambda}^{s,v}(\theta^h)$ is in the FWA region.

If, instead,

$$\lambda^{ft} \leq \tilde{\lambda}^{s,v}(\theta^h) \leq \lambda^v \iff (1 - \lambda^v)D \leq \theta^h \leq (1 - \lambda^{ft})D,$$

where $D \equiv \gamma c^i(1 - \alpha) - (1 - \gamma)\tau(1 - \rho^v)$, then even some fixed-term jobs are going to be shadow.

Thus, we have the following first key empirical implication.

EMPIRICAL IMPLICATION 1: When the inspectors show up at the gate, firms with $\theta = \theta^h$ experience a sudden increase in (misuse) of voucher relative to firms with $\theta = \theta^l$.

Note that we can also quantify the misuse of vouchers. Since the model is static, total employment misused is identical to the steady-state misused employment and is thus

²²In the case of fixed term jobs the condition is obtained by a reservation value $\tilde{\lambda}^{s,ft}$ that satisfies

$$\tilde{\lambda}^{s,ft}(\theta^h) = \frac{(1 - \gamma)(1 - \rho^{ft})\tau + \theta^h}{(1 - \gamma)\tau\rho^{ft}}. \quad (9)$$

$n^{sm}(\theta^h) = np(1 - G_1(\tilde{\lambda}^{s,v}|\theta^h)) > 0$. Conditional on being a misusing firm, employment is $n^{sm}(\theta^h) = n(1 - G_1(\tilde{\lambda}^{s,v}|\theta^h)) > 0$.

Next we consider what happens to the misuse of vouchers when the policy setting changes and it becomes more difficult to activate a voucher when the inspectors appear at the gate. In the model, such an experiment is equivalent to a reduction in the parameter α . The key comparative static is thus the derivative $\frac{\partial \tilde{\lambda}^{s,v}(\theta^h)}{\partial \alpha} < 0$, that suggests that as α falls, the marginal value shifts to the right, exactly as in Figure 3. The requirement for firms to activate a voucher through a text message 60 minutes in advance suggests that the misuse of vouchers should fall, as the following empirical implication summarises.

EMPIRICAL IMPLICATION 2: With the introduction of the text message requirement (a fall in α), at the outbreak of the inspection firms with $\theta = \theta^h$ reduce the instant validation of vouchers relative to firms with $\theta = \theta^l$.

The quantitative fall in misuse is (in modulus) equal to $|np \frac{\partial G_1(\tilde{\lambda}^{s,v})}{\partial \lambda^{s,v}} \frac{\partial \tilde{\lambda}^{s,v}(\bar{\theta})}{\partial \alpha}|$. The right panel of Figure 3 shows how the value of jobs varies when α decreases, thus when the legislation required that employers inform the administration one hour in advance about using a voucher.

The last exercise we perform considers what happens when vouchers are abolished. This policy change influences the value of jobs for both types of firms, as the margin λ^v disappears and the firms expand the amount of regular fixed-term employment. The increase in fixed-term employment is a consequence of the fact that some jobs that were previously hired as FWA jobs become fixed-term. When no vouchers are available, the margin for fixed-term jobs is defined as $J(\lambda^{ft,no-v}) = K$, where $\lambda^{ft,no-v}$ is the margin of entry post voucher elimination. Analytically, we have that $\lambda^{ft,no-v} > \lambda^{ft}$, so that firms expand fixed-term contracts.

This result implies an average increase in fixed-term employment for firms with $\theta = \theta^h$

equal to $\Delta n^{ft}(\theta^h) = n[G_1(\lambda^{ft,no-v}|\theta^h) - G_1(\lambda^{ft}|\theta^h)] > 0$ while the increase in fixed-term employment for firms with $\theta = \theta^l$ is $\Delta n^{ft}(\theta^l) = n[G_2(\lambda^{ft,no-v}|\theta^l) - G_2(\lambda^{ft}|\theta^l)] > 0$. In other words, fixed-term employment increases for both types of firms.

EMPIRICAL IMPLICATION 3: When vouchers are outlawed, fixed-term employment increases for both types of firms, and the actual change depends on the location of $\tilde{\lambda}^{s,v}$ and on the distributions $G_1(\lambda|\theta^h)$ and $G_2(\lambda|\theta^l)$. If some *ft* jobs are “grey” and/or if G_1 first-order stochastically dominates G_2 in the proper domain, meaning that high θ or low-tax morale firms have higher λ jobs than high-tax morale firms, then low-tax morale firms will see larger changes in fixed term jobs.

In the rest of the paper, we test the empirical implications against the data.

4 Data

We use data from the INPS archives, including firm-level employment data, daily firm-level voucher usage, and labour inspection records. Our model has clear predictions about the use of vouchers when firms receive a labour inspection. This is why the analysis focuses on firms that have been inspected at least once and that have used at least one voucher during our sample period.²³

The sample construction follows these steps. We start with the universe of firm-level employment data for the years 2014 to 2017. As shown by the timeline of the regulatory changes to the FWA (Figure 1) and the time series of vouchers used (Figure 4), following almost complete liberalisation, the use of vouchers peaked right before their abolition.²⁴

²³Firms that never use vouchers may require a more stable workforce (a lower λ in our model) and would not generate any empirical variation in the use of FWA.

²⁴Similar data have already been used within the VisitInps program of the Italian Social Security institute (INPS).

Later, we merge this firm-level employment data with both the universe of labour inspections and the universe of vouchers used by firms.²⁵

The INPS institute performs labour inspections to detect full or partial evasion of social security contributions. We have information on the day the inspection started and the outcome of the inspection (whether a fine was levied and its amount).

For each firm, we also know how many vouchers have been used each day. In order to maximise the time frequency of our data, we construct individual firms' daily-level use of vouchers around labour inspections. Later, when we analyse firm-level employment outcome, the finest level of aggregation we can use is at the firm and month level.

The summary statistics in Table 1 are useful for describing how labour inspections and the use of vouchers determine the type of firm that is later analysed. Each row represents monthly averages between January 2015 and September 2016 (the period before of the TXT requirement reform) of various firm-level employment measures -both regular work and vouchers- and each column represents a different sample.

In Column 1, we start with the universe of 1.8 million firms. The average firm in the INPS archives employs 7.2 workers and uses 27 vouchers per month. If each voucher covered one hour of work, the vouchers would represent only a small share of labour; however, we are going to see that firms use vouchers to hide additional undeclared work hours.²⁶ Most workers have permanent and full-time contracts.

About 23% of firms (a little more than 400,000) use at least one voucher (Column 2, Vouchers). For firms using vouchers, the average size is also close to 7, but with more temporary and part-time workers and fewer permanent and full-time ones. These firms use an average of 97.6 vouchers, or roughly 14 vouchers per full-time equivalent worker.

²⁵These two sources of data are not available within the VisitInps program and have been directly managed by two of the authors who worked for INPS, Edoardo di Porto and Paolo Naticchioni.

²⁶Unfortunately, the available data do not allow us to estimate how many hours of work tend to be underreported due to the availability of flexible work schedules.

The third column (Inspected) refers to inspected firms, which, due to the small probability of inspection, represent only 1.58% of firms.²⁷ Since larger firms are more likely to be inspected, the average workforce increases to 43.5 workers. In addition, inspected firms use a large number of vouchers, on average 300 vouchers per month, or 7 vouchers per full-time equivalent worker.

In order to analyse how the use of vouchers evolves around the onset of labour inspections, we focus on firms that have been inspected and use at least one voucher over the entire period. We are left with 3,472 firms, or 0.19% of firms (column “Both”). The number of employed workers is similar to that of the inspected firms, but with a higher share of part-time contracts and temporary contracts, and an average monthly number of vouchers of 270 (9.2 for the full-time equivalent worker). This sample provides information on how firms using vouchers, which represent at least 1/4 of the universe of Italian firms, behave upon inspection.

In the last two columns, we divide the 3,472 firms into those who presumably misused vouchers (“Misbehaving”) and those who did not (“Compliant”). The division depends on the firm’s use of vouchers when a labour inspector shows up. Later, we are going to explain in more detail how we estimate the change in the use of vouchers. The model predicts that misbehaving firms should increase such use. Right now, it is worth mentioning that “presumably” misbehaving firms are similar to compliant firms, but use more vouchers.²⁸

²⁷60% of inspections resulted in fines and the average sanction was around €20,000. Later, we are going to analyse how such fines vary depending on whether firms use vouchers.

²⁸Using an econometric terminology, we could think of “misbehaving” firms as the compliers: they activate vouchers when an inspector shows up, while the “compliant” firms are either “always takers” or “never-takers,” meaning that their behaviour does not change upon the arrival of labour inspectors.

5 Empirical Strategy

The empirical analysis is divided into four parts. First, we discuss the identification, which is based on the random timing of labour inspections (Section 5.1). Second, we discuss the empirical specification of the event study (Section 5.2) and, third, we explain how those estimates are used to identify misbehaving firms (Section 5.3). Fourth, we introduce the difference-in-difference models that compare misbehaving against compliant firms (Section 5.4). Section 6 discusses the results, whether firms are indeed more likely to use vouchers when inspectors appear at their premises (Section 6.2), and what happens to different types of regular and irregular work when vouchers are abolished (Sections 6.4 and 6.5).

5.1 Random Timing of Inspections

First, we are going to see that inspections are rare and their timing is intended to be unpredictable, allowing for a comparison of firms' behaviour before and after inspections.

A randomly selected firm has a 1 in 130 chance of being inspected in a given year and about a 1 in 50,000 chance of an inspection starting on a given day. Although some firms are more likely to be inspected than others, for example, larger firms, to be effective, the timing of inspections has to be unpredictable. Thus, from a firm's perspective, the first day that labour inspectors enter the firm's premises is as good as random.

This is testable. If the timing of inspections is as good as random, the vector of observable characteristics (X_j) of a firm j should be unable to predict the exact date (t_j) of a labour inspection. We use a balance test, the joint F-test, to determine whether all characteristics in the following cross-sectional regression have no predictive power ($\beta = 0$): $t_j = \alpha + \beta'X_j + \epsilon_j$.

Given the random timing of inspections (see Section 6.1), our model suggests a fairly simple test to determine whether an inspected firm is using vouchers to hide undeclared work: when faced with labour inspectors, firms employing undeclared workers should use

vouchers to “hedge” the risk of a fine. We compare the daily use of vouchers just before and after an inspection (the “treatment”), following the firm’s behaviour before and after the inspection.

5.2 Random Timing and Event Studies

Once we have established that inspections happen at random times, in the second part, we are going to set up the event study. We analyse the FWAs used by the firm j between 180 days before and 90 days after an inspection that occurs on day t .²⁹ We split the event studies into two periods, before and after October 2016, which is when the firms had to inform the Social Security Administration an hour in advance before using a voucher. Given that a single voucher would be sufficient to avoid the fine, our outcome variable is equal to one when on a given event day τ the firm j uses at least one voucher, and 0 otherwise ($DV_{j,\tau} = 1\{\#Vouchers_{j,\tau} > 0\}$).³⁰ Of the 180 days before the inspection date, the first 90 days will serve as the control period:³¹

$$DV_{j,\tau} = \sum_{k=-90}^{90} \beta_k D_{\tau+k} + f(t) + \epsilon_{j,\tau} . \quad (11)$$

$D_{\tau+k}$ is a dummy variable equal to one for event day $\tau + k$ and zero otherwise. Additionally, since the time series of FWAs is far from stationary (see Figure 4) it is important to control for calendar time $f(t)$. We start using several calendar-time fixed effects (e.g. year, month, and day of the week) and show that the results do not differ when using calendar-day fixed

²⁹As mentioned, the last phase of labour inspections can approximately last up to 90 days, moreover labour judges have deemed that 90 days represent the reasonable duration of labour inspections (Del Vecchio, 2019).

³⁰Here $1\{A\}$ is a dichotomous variable equal to 1 when statement A is true, and 0 otherwise. We find comparable results when using the daily number of vouchers used by the firm, though the extensive margin appears to be driving this effect. This suggests that most (declared) firms have only a few undeclared workers.

³¹Given that the days event-time effects leading to the inspections are precisely estimated to be zero, a longer baseline period is going to lead to more precisely estimated coefficients. But shorter baseline periods lead to very similar results.

effects.³²

The unpredictability of the timing of inspections is crucial in setting the correct specification for our model. This is because focusing on just treated firms that are treated at different times, conditional on firm fixed effects, calendar time, and event time of the inspection, is collinear. Fortunately, when the timing of the event is random (see Section 6.1) the timing of treatment is orthogonal to the characteristics of the firms and to their time-invariant intercept (see Borusyak et al., 2024). Our strategy is similar to the identification used in Parker et al. (2013), which exploits the random timing of the disbursement of the 2008 Economic Stimulus Payments in the United States.

It is also worth noting that, due to such randomness, we do not have to specify a two-way fixed effects model; thus, we do not have to worry about the well-known possible biases arising from this design (see Goodman-Bacon, 2021, Sun and Abraham, 2020). When, as a robustness check, we do specify a two-way fixed effects model, it is only parametrically identified. A relatively long baseline period (90 days) combined with the non-linear function of event time (for example, event-time dummy variables) breaks the perfect collinearity between event-time and calendar time, but the fix is not perfect. Such near-collinearity is measurable. Without firm fixed effects, time effects explain only 4% of the variation in the post-inspection variable (our main explanatory variable), but when we add firm fixed effects the R^2 goes up to 55%. In other words, because of near-collinearity, using both firm effects and calendar-time effects captures most of the random variation in the timing of inspection.

We are going to discuss later how we deal with these potential biases. Here, it is important to mention that with random timing of inspections, we can use the main result from Athey and Imbens (2022), which shows that the standard Difference-In-Differences estimator is an unbiased estimator of a weighted average of different causal effects, including the effect of

³²We also document how the results differ when we do not control for time or control for a simple linear time trend.

changing from never being inspected to being inspected in the first period, or changing from being inspected later to being inspected earlier in time. Since in our setup there are no reasons to believe these effects differ, apart from the collinearity issues, a simple Difference-In-Differences identifies the average causal effect.³³

Moreover, under the design-based assumption that each unit’s adoption (or inspection) date is randomly assigned, the simple before-after difference in overall means is still an unbiased estimator of a well-defined average treatment effect on the treated (the inspected firms).

The data contain information about the firms, which allows us to look for heterogeneous effects. Given our simple test for the misuse of vouchers, the heterogeneity analysis will inform us whether the misuse of vouchers is concentrated in certain parts of the country, certain sectors, and specific types of firms (large vs. small, with more or less use of temporary contracts). In order to do this, we simply test whether pre-SMS effects differ between economic sectors and firm characteristics. Given the apparently static treatment effects and the lack of pre-trends, for the heterogeneity analysis, we collapse the post coefficients over time and use the whole pre-inspection period as baseline:

$$DV_{j,\tau} = \beta D_{\tau \geq 0} + f(t) + \epsilon_{j,\tau} . \quad (12)$$

5.3 Measuring Misusing and Compliant Firms

In the third part, we use the predictions of the model to identify the misbehaving firms: those that increase their use of vouchers after the inspection.

In order to spot them, we exploit the fact that treatment effects appear to be stable over time. With firm-specific constant treatment effects, we can model the change in the

³³Athey and Imbens (2022) also show that the difference-in-difference standard errors are conservative.

probability of using a voucher to be constant over time but heterogeneous across firms:

$$DV_{j,\tau} = \beta_j D_{\tau \geq 0} + f(t) + \epsilon_{j,\tau} , \quad (13)$$

where β_j represents the firm-specific post-pre inspection differences in the likelihood of using vouchers.

Based on the estimated $\hat{\beta}_j$ s, we identify misbehaving firms and highlight how they respond to the introduction of the SMS³⁴ requirement and the abolition of vouchers. In particular, we combine the results from the previous model with the October 2016 SMS requirement and the March 2017 abolition of vouchers, separating firms with positive changes from the rest: $\widehat{M}_j^\eta = 1\{\hat{\beta}_j > \eta\}$, where η represents the cut-off above which a firm misbehaves, and in the baseline analysis $\eta = 0$. Later, we test whether the results are robust to the choice of more stringent cut-offs, $\eta > 0$.

In other words, we use behavioural changes driven by the inspections to identify misbehaving or low-tax morale firms: for each inspected firm, we compute its average use of vouchers before and after the inspections and classify firms into those that, on average, increase their use and those that do not. Our definition of misbehaving firms is subject to misclassification, both Type I and Type II, which biases the estimates towards 0. Later, we will take advantage of the possibility of changing our definition of misbehaving firms.

5.4 Difference-in-Differences between Misusing and Compliant Firms

The fourth and final part makes use of particular predictions of our labour demand model: how firms that are likely to misuse FWAs ($M_j = 1$) react when the SMS requirement is introduced and when vouchers are abolished. In this latter case, firms might: fall back to signing contracts that allow for some “grey” work, such as part-time or fixed-term contracts;

³⁴SMS are Short Messages or Text Messages in mobile telecommunication.

revert to hiding the entire work relationship; abandon the low-productivity jobs.

Empirical models are simple difference-in-differences, before and after October 2016 (SMS requirement) or March 2017 (abolition of vouchers) between firms that presumably misused FWAs and those that did not.³⁵ Since all firms share the same event date, we are not in a staggered design and do not have to worry about biases due to dynamic treatment effects (see Sun and Abraham, 2020). To assess the parallel trend assumption, we estimate event study differences with leads and lags. The number of lags is limited by the period spanned by the data, and we exclude the event time $\tau = -2$ (respectively, August 2016 and January 2017, allowing for some anticipation effect).

The outcomes available at the monthly level (m) are i) the total number of vouchers used, ii) the log of the total wage bill (including vouchers) and the total number of workers with the following contracts: iii) part-time, iv) full-time, v) fixed-term, vi) fixed-term and part-time, and vii) open-ended. The event study difference-in-difference controls for firm and year by month fixed effects:

$$Y_{j,t}^m = \sum_{k \neq -2} \gamma_k \widehat{M}_j^0 \times D_{\tau(t)+k} + \mu_j + \mu_t + \varepsilon_{j,t}, \quad (14)$$

where τ is the event period, which is 0 either in October 2016 or in March 2017, and $D_{\tau+k} = 1$ in event period $\tau + k$ and 0 otherwise.

To better understand the issue of misclassification in misusing firms \widehat{M}_j , we use a constant difference-in-difference model:

$$Y_{j,t}^m = \gamma \widehat{M}_j^\eta \times D_t + \mu_j + \mu_t + \varepsilon_{j,t}, \quad (15)$$

where D_t is a indicator that equals one after the abolition of vouchers. Defining $p(\eta)$ and $q(\eta)$

³⁵We are implicitly assuming that misbehaviour is time-invariant, and any deviation from this assumption would lead to downward biased estimates.

to be type I and II errors of the misclassified and unobserved variable M_j , it can be shown that, under constant treatment effects, $\frac{\hat{\gamma}(\eta)}{1-p(\eta)-q(\eta)}$ converges in probability to γ . This implies that, with constant treatment effects, the η^* that maximises $\hat{\gamma}(\eta)$ minimises misclassification bias. Later, we discuss whether η can tell us something about the fraction of firms that misuse FWAs.

In our final analysis, we look at how the introduction of the SMS requirement, as well as the abolition, has influenced under-reporting of work. Since under-reporting is not directly observable, we need to rely on labour inspections. Yet, since only a handful of firms are inspected more than once in our data, we cannot compare misbehaving with non-misbehaving firms, but need to rely on a coarser definition of treatment.³⁶ In particular, we compare firms based on whether they have ever used a voucher before the SMS requirement. An observation is defined as a labour inspection between 1 January 2016 and 31 December 2017, which produces 20,819 observations. The treated firms are those that have used vouchers in the pre-SMS period (4,269 observations). We use the evaded contribution as an outcome variable to measure under-reporting and define three main treatment periods: Pre-SMS requirement (1 January to 16 October 2016, which will be the omitted time period), post-SMS requirement (17 October 2016 to 17 March 2017), and post-abolition (18 March 2017 to 31 December 2017).

³⁶Based on Table 1, the number of the inspected firms is 40,121, or 2.17% of the population of firms. 2,955 firms are inspected more than once over the same period, representing 0.16% of firms. Notice that if inspections were random events the likelihood to be inspected twice would be 2.17% to the power two, or 0.047%. This implies that firms that are inspected once are more likely to be inspected again, and this might explain persistency in the effects of inspections.

6 Empirical Evidence

6.1 Random Timing of labour Inspections

We start by testing whether the timing of the labour inspection is predictable. Appendix Figure A1 shows the distribution of the day of the year inspections take place. Inspections occur throughout the year and are fairly uniformly distributed, although they are slightly more likely to occur during the summer.

The panels in Figure 5 show the linear regression coefficients of different firm characteristics on the day of inspection. In the left panel, we only control for year fixed effects, to allow labour inspectors to have different targets in different years. Only three variables are significantly different from zero, and each characteristic predicts at most a difference of a few days on the date of inspection. The largest difference is in the construction sector, where inspections tend to happen two weeks earlier. Adding a semester dummy is sufficient to drive all these differences towards zero even further. Within 180 days, all differences are, with the exception of the residual sector “Other,” smaller than 3 days. In other words, the exact day that inspections take place appears to be unpredictable.

6.2 The Misuse of FWAs: Evidence from Labour Inspections

Given random timing, we can estimate differences in the probability of using vouchers around the inspection time. Figure 6 plots the event-study differences from Equation 11 using a linear probability model. Each point represents $\hat{\beta}_{\tau+k}$, that is, the difference in voucher use between event date $\tau+k$ and the days between 90 and 180 before the inspection (the excluded event times). Upon inspection, there is a clear change in the likelihood of using vouchers. Moreover, the evidence suggests that conditional on year, month, and day of the week fixed effects, there are i) no pre-trends in the use of vouchers prior to the inspection, ii) no major

anticipation effects, and iii) fairly stable and persistent treatment effects.

With respect to point i), Appendix Figure A3 shows how important it is to control for time effects. Without such controls, the event dummies will capture part of the increasing trend in the use of vouchers (right panel). Yet, the left panel shows that adding a simple linear time trend is enough to centre the pre-period around zero (left panel).

The same is true for point ii); without time controls, the trends generate what looks like an anticipation effect. With respect to point iii) the increase in the likelihood of using vouchers immediately after an inspection is 0.88 percentage points (SE 0.16), which corresponds to a relative increase of approximately 18%. The greatest changes occur on the day of inspection and the day after, respectively, 1.5 (30%) and 1.4 percentage points (29%). If we consider that once the inspection has started, firms may also have the option to ask undeclared workers to stay home, these findings correspond to large effects. There is possibly a very slight downward trend in the effects. For the first three weeks, almost all daily effects are significantly different from zero, while there is more noise after the first 45 days. There are still several significant effects after that event date; interestingly, there is an additional spike at 90 days, which, as mentioned earlier, sometimes corresponds to the prescribed end of the last phase of inspections.³⁷

Having many pre-inspection periods allows us to perform randomisation inference. Focusing on pre-inspection data ($\tau < 0$), we sequentially generate fictitious inspection dates for $-120 \leq \tau \leq -30$ and estimate

$$DV_{j,\tau} = \beta_k D_{\tau-k \geq 0} + f(t) + \epsilon_{j,\tau} , \quad (16)$$

where $D_{\tau-k \geq 0} = 1$ when $\tau - k \geq 0$ and zero otherwise. The histogram of all placebo $\hat{\beta}_k$'s,

³⁷We know that inspections should not last more than 90 days since the preliminary report. While we do not have information on the duration of the preliminary phase, whenever such phase takes less than a day the maximal duration corresponds to our event date 90. This fuzziness in the prescribed end date (the actual date would be endogenous) makes it hard to empirically exploit the end of inspections.

shown in Figure 9, is centred around zero and is far from the vertical line, which corresponds to the estimated $\widehat{\beta}_0 = 0.88$. Sampling noise is unlikely to have generated such a large change in behaviour.

The data contain information about the firms, which allows us to look for heterogeneous effects. In particular, we test whether pre-SMS effects differ across economic sectors and firm characteristics. Given the static treatment effects and the lack of pre-trends, we collapse the treatment effects and use the whole pre-inspection period as a baseline:

$$DV_{j,\tau} = \beta D_{\tau \geq 0} + f(t) + \epsilon_{j,\tau} . \quad (17)$$

Table 2 shows that the increase is about the same in the retail sector, the tourism sector, and the manufacturing sector. For the “Other sectors” the jump is slightly lower, while it is completely absent in the construction sector. It turns out that the construction sector is highly dependent on procurement jobs and subcontracting, and since 2015, with “the Jobs Act,” vouchers cannot be used within such contracts. The aim was to reduce the number of job accidents.³⁸ This is likely to explain why, in the construction sector, firms have the lowest probability of using a voucher (0.02 per day).

In Table 3 we perform additional heterogeneity regressions. Columns 1 to 4 show that the jump is relatively constant across Italian regions, although it is slightly larger in the more productive North than in the South, with the Centre of Italy being in between. Thus, while the South is usually perceived to be more prone to tax evasion, it does not misuse vouchers more than the other parts of Italy; rather, slightly less.

Columns 5 to 7 show that medium-aged firms (those that started between 5 and 14 years earlier) are more likely to use vouchers to cover undeclared work compared to young and old

³⁸According to the Italian Work Injury Institute INAIL (“Istituto Nazionale per l’Assicurazione contro gli Infortuni sul Lavoro”) the share of deadly accidents in the construction sector is the largest. For example, in 2019 around 23% of deadly accident occurred in the construction sector, a sector that employs only 6.6% of the workforce.

ones. Firm size is highly predictive of the size of the effects, with large firms (more than 15 employees) being the ones with the largest jumps (column 10). Column 11 shows that the jump is about 40% larger for firms whose share of part-time workers is above the median. This is in line with the opinion of many labour inspectors that part-time work is sometimes used to hide what are truly full-time workers, as it lowers the social security contributions and the tax burden. Finally, according to the law, firms' use of temporary workers should not exceed 20% of the open-ended ones. Yet, in our sample (but this is true more generally) more than 50% of firms are above such limit, suggesting that enforcement of this regulation is limited. Under the assumption that such behaviour is indicative of lawless behaviour, Columns 12 and 13 show, indeed, that such firms above the threshold of 20% tend to have a more pronounced jump (1.1 p.p. against 0.85p.p.).

In the last robustness check, we add firm fixed effects that capture part of the unobserved heterogeneity in the use of vouchers. As mentioned earlier, with random timing, adding firm fixed effects is unnecessary and may do more harm than good. This can be seen in Appendix Figure A4, which shows that firm fixed effects generate pre-trends in the use of vouchers. Since firm fixed effects introduce collinearity between event time and calendar time, it appears as if some of the time effects are now captured by the event-time effects, generating a rotation in all the differences. The bias due to the rotation increases as we move further away from the date of inspection, reducing all the post-inspection changes.

One way to reduce collinearity bias is to focus on differences around the date of inspection. In Appendix Table A1 we use 15 days (first three columns) and 30 days (last three columns) after inspection. Columns 1 and 2 show that adding fixed effects for year, month, and day of the week, or calendar date fixed effects, leads to almost identical results. With firm fixed effects (columns 3 and 6) the treatment effects drop from 0.0095 to 0.0069 and from 0.0084 to 0.0053. But firm fixed effects explain part of the variation in voucher use and thus lower the standard errors, leading to very precisely estimated coefficients.

6.3 The 60-minutes SMS/TXT Requirement

Next, we analyse the October 2016 SMS reform, which introduced the 60-minute messaging requirement and thus broke the possibility of on-the-spot insurance. On the one hand, firms would still be able to exercise the insurance option on a daily level, buying at least one voucher per worker, but this should not generate a jump on the day of the inspection. On the other hand, firms might decide to stop misusing vouchers altogether because of the additional and costly constraint of the SMS requirement.

Accordingly, Figure 7 shows that once SMS was introduced, there is no more evidence of a substantial and longer-lasting jump.³⁹ The immediate but very short-term jump (two days) would be consistent with some firms being potentially uninformed about the policy change and trying for one or two days to use vouchers on the spot, before desisting. Overall, only 10 coefficients out of 90 are significantly different from 0 (5 are positive and 5 are negative). Yet, as discussed in Section 3, firms that used vouchers to protect against inspection risk can still do so by buying at least one voucher per worker per day before possible inspections.

To test whether the total number of vouchers increased after October 2016, we use monthly data and define as “misusing” firms those who before the SMS requirement and upon inspection were more likely to use vouchers. Figure 8 shows the event-study differences from Equation 14 between “misusing” and “non-misusing” firms, where 0 refers to the timing of the reform. The estimates are noisy, possibly because the reform may have also pushed some misbehaving firms to stop using vouchers altogether, but firms appear to use 10-15 more vouchers per month, which represents a 10% increase compared to baseline, and there is evidence of some anticipation, possibly because the introduction of the October SMS requirement was known several months before implementation.⁴⁰

³⁹Given the shorter time period that is available we use -60 to -30 event days as a comparison period.

⁴⁰Newspaper articles mention this part of the future reform at least since June 2016, see, for example, *Quotidiano* (2016).

In line with the documented slight increase in the number of vouchers, Appendix Figure A2 shows that the SMS/TXT requirement had little impact on the different types of regular employment (the four panels of the figure refer to temporary, permanent, full-time, and part-time contracts).

6.4 The Effect of Vouchers on Regular Employment, Total Wage Bill, and Misclassification Error

6.4.1 The Effect of Vouchers on Regular Employment and Total Wage Bill

We have shown that FWAs can be used to hide undeclared work, even when informational requirements are designed to avoid such behaviour (the SMS). The next relevant question is whether vouchers i) crowd out regular work or ii) hide work that would otherwise be fully undeclared (black instead of grey).⁴¹

Our theoretical model predicts that, in the event that FWAs become unavailable, firms should hire some of the “grey” workers using the next most suitable flexible job, and this should occur more frequently for low-tax morale firms.

The next most FWA is arguably the combination of temporary and part-time labour market contracts.

Using monthly data, we investigate differences in employment outcomes between “misusing” and “non-misusing” firms in a event-study difference-in-difference approach. Figure 10 shows that when vouchers are abolished, contracts that are at the same time temporary and part-time increase by about 1, representing a 50% increase compared to the pre-abolition average. The difference between “misusing” and “non-misusing” firms in the total number of temporary and part-time employees is fairly flat in the months leading up to the abolition

⁴¹Given the prevalence in the misuse of vouchers, we are unable to identify whether vouchers give rise to new job opportunities. Although we can use the wage bill as a measure of declared job opportunities and a lower bound for declared and undeclared job opportunities.

of vouchers in March 2017 and then increases by approximately one additional worker.

These workers also drive the results for the total number of workers (see Figure 11), since this change is only slightly above 1 (the average number of workers is around 40). Figure 12 separates the effects for the two dimensions: temporary vs. permanent (top panel) and part-time vs. full-time (bottom panel). The largest effects are those for temporary workers. For these workers, the difference-in-differences is close to 2, an almost 50% change. There are no effects and, if anything, negative effects for open-ended/permanent contracts (upper right panel). Regarding part-time workers versus full-time workers, both groups show similar effects, indicating that firms seek flexibility with respect to the duration of the labour contracts.

In terms of total wage bill, when we include the cost of FWAs (i.e. also vouchers), there is no evidence of significant changes (see Appendix Figure A5). This suggests that there are no major changes in the job opportunities declared, possibly driven by rule-of-thumb cost calculations for the firm to undertake a strategy aimed at keeping the total wage bill constant.

6.4.2 Misclassification Error

Based on evidence on the closest substitute contract for FWAs, we use temporary and temporary part-time contracts to assess the importance of misclassification errors. Misclassification is driven by the way we identify “misusing” firms in the baseline event-study regression. When defining $\widehat{M}_j^\eta = 1\{\hat{\beta}_j > \eta\}$, with $\eta \geq 0$, type I and type II errors depend on η . Thus, we re-estimate equation 15 comparing “misusing” and “non-misusing” firms, varying η between 0, the baseline case, and 7.5%, which corresponds to, respectively, around 50% and 8% of misusers. Intuitively, when we increase η we impose a stricter requirement to be identified as a firm that is misusing vouchers. This reduces the likelihood of classification as misusing a firm that is correctly using vouchers (Type II error), but increases the misclassification

probability of misusing firms (Type I error).

Figure 13 shows that regardless of whether we use temporary contracts or temporary and part-time workers, the optimal η that minimizes the sum of Type I and II probabilities is close to 0.03, corresponding to a fraction of misbehaving firms of about 18 to 19%.⁴² When compared to the coefficients estimated with $\eta = 0$, the coefficients increase by about 60%, which implies that the relative changes in the number of temporary and temporary part-time workers increase from 50% to more than 75%.

6.5 The Effect of Vouchers on Evasion

Since the annual probability of inspection is less than 1%, between 2014 and 2017 only a handful of firms were inspected more than once. This implies that we cannot use our usual two-step identification strategy, where we first construct a measure of misbehaviour based on the use of vouchers when inspections are conducted.

We can only compare the amount of social security contributions evaded for inspected firms that have or have not used vouchers between January and October 2016, a clearly less appealing type of comparison. Separating the inspected firms according to whether they have used at least one voucher is suboptimal, as we have just seen that a minority of firms use vouchers to hide undeclared work. In our sample of inspected firms, 4,269 have used at least one voucher and 16,550 have not, with a total sample size of 20,819 firms.

During the baseline period (January to October 2016), firms that have not used vouchers evade, on average, 26,835 euros. Table 4 includes the difference-in-difference estimates of a model where time periods are interacted with the treatment variable; that is, firms using vouchers interacted with time periods (Pre-SMS; from 1 May to 16 October 2016, the omitted category; Post-SMS, from 17 October 2016 to 17 March 2017; Post-Abolition, from 18 March

⁴²It is comforting that the optimal η also corresponds to one of the two peaks in the R-squared that are close to each other in size.

to 31 December 2017). Four different specifications are reported, where column (1) does not include any additional control variables to the baseline difference-in-difference, while columns (2) to (4) include several sets of controls.

It emerges that, compared to control firms, firms that used vouchers tend to evade more (i.e. pay higher fines when inspected) when vouchers become unavailable. The results change little when we control for a monthly cubic time trend, for employment characteristics (the share of part-time workers, fixed-term workers, their interaction, and wages).

7 Conclusions

In response to the increasing demand from firms for flexible work contracts, legislators around the world are devising labour contracts that allow firms to hire workers for just a small amount of work, i.e., on-call work, zero-hour work, labour vouchers, mini-jobs, etc.

These contracts, while offering flexibility, often result in poor career development prospects, stagnant wages, and increased exposure to income risk, as workers find themselves in precarious employment situations without stable income or career growth opportunities (Boeri et al., 2020). This study documents an additional important risk: these labour contracts may complicate the job of labour inspectors whose task is to uncover undeclared work. We show that upon random inspections, whenever it is lawful to use labour vouchers on the spot, some firms use FWAs to hide undeclared work. Both the labour demand model and the evidence show that this novel mechanism can counteract the expected reduction in undeclared work resulting from lower hiring and firing costs.

There is also evidence that small changes to the bureaucratic requirements of FWAs can partially close these loopholes. When employers were forced to signal such arrangements in advance, labour inspectors would have enough time to identify undeclared workers. This policy change encouraged misbehaving firms to buy at least one voucher per day to avoid

being fined by the inspectors.

We estimate that 18% of firms used FWAs to hide undeclared work, and that these firms hired more fixed-term and part-time workers when FWA were banned. In addition, the ban increased average inspection fines for firms that used vouchers before they were banned.

In general, our results suggest that in countries with a tendency to employ undeclared work, more flexible work schedules may generate even more undeclared work. In addition, firms that use vouchers to hide undeclared work tend to use vouchers mainly when inspected. As a result, restricting the total number of vouchers used by workers or firms to be below some threshold, which has been Italy's primary policy restriction, is unlikely to successfully curb misbehaviour. Rules that restrict the use of FWA to certain categories, such as students or retirees, are probably a better way to balance the need for flexibility with legality and the protection of employment rights.

References

- Abi Adams and Jeremias Prassl. Zero-hours work in the United Kingdom. Technical Report Conditions of Work and Employment Series No. 101, International Labour Organization, 2018.
- James Albrecht, Lucas Navarro, and Susan Vroman. The Effects of Labour Market Policies in an Economy with an Informal Sector. *The Economic Journal*, 119(539):1105–1129, 06 2009a. ISSN 0013-0133.
- James Albrecht, Lucas Navarro, and Susan Vroman. The effects of labour market policies in an economy with an informal sector. *The Economic Journal*, 119(539):1105–1129, 2009b.
- Susan Athey and Guido W. Imbens. Design-based analysis in difference-in-differences settings with staggered adoption. *Journal of Econometrics*, 226(1):62–79, 2022. ISSN 0304-4076. doi: <https://doi.org/10.1016/j.jeconom.2020.10.012>. URL <https://www.sciencedirect.com/science/article/pii/S0304407621000488>. Annals Issue in Honor of Gary Chamberlain.
- Olivier Blanchard and Pedro Portugal. What hides behind an unemployment rate: Comparing portuguese and us labor markets. *American Economic Review*, 91(1):187–207, 2001.
- Tito Boeri and Pietro Garibaldi. Temporary Employment in Markets with Frictions. *Journal of Economic Literature*, 62(3):1143–1185, 2024.
- Tito Boeri, Giulia Giupponi, Alan B Krueger, and Stephen Machin. Solo self-employment and alternative work arrangements: a cross-country perspective on the changing composition of jobs. *Journal of Economic Perspectives*, 34(1):170–95, 2020.
- Kirill Borusyak, Xavier Jaravel, and Jann Spiess. Revisiting event-study designs: Robust and efficient estimation. *The Review of Economic Studies*, 91(6):3253–3285, 02 2024.

ISSN 0034-6527. doi: 10.1093/restud/rdae007. URL <https://doi.org/10.1093/restud/rdae007>.

Mariano Bosch and Julen Esteban-Pretel. Job creation and job destruction in the presence of informal markets. *Journal of Development Economics*, 98(2):270–286, 2012.

Pierre Cahuc, Olivier Carlot, and Frank Malherbet. Explaining the spread of temporary contracts and its impact on labor turnover. *The International Economic Review*, 57(2):533–572, 2016.

David C. Chan. The efficiency of slacking off: Evidence from the emergency department. *Econometrica*, 86(3):997–1030, 2018.

M Keith Chen, Peter E Rossi, Judith A Chevalier, and Emily Oehlsen. The value of flexible work: Evidence from UBER drivers. *Journal of Political Economy*, 127(6):2735–2794, 2019.

Nikhil Datta, Giulia Giupponi, and Stephen Machin. Zero-hours contracts and labour market policy. *Economic Policy*, 34(99):369–427, 2019.

Alessandro Del Vecchio. La ragionevole durata dell’ispezione in materia di lavoro e previdenza. *Rivista del Diritto della Sicurezza Sociale*, 19(4):813–832, 2019.

Edoardo DiPorto, Leandro Elia, and Cristina Tealdi. Informal work in a flexible labour market. *Oxford Economic Papers*, 69(1):143–164, 2017.

Marco Fugazza and Jean-François Jacques. Labor market institutions, taxation and the underground economy. *Journal of Public Economics*, 88(1-2):395–418, 2004.

Andrew Goodman-Bacon. Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, 2021.

- ILO. Labour Inspection and Undeclared Work in the EU . Technical report, ILO, 2013.
- Simon Johnson, Daniel Kaufmann, and Pablo Zoido-Lobaton. Corruption, Public Finances and the Unofficial Economy. *Policy Research Working Paper*, (2169), 1998.
- Lawrence F Katz and Alan B Krueger. The rise and nature of alternative work arrangements in the united states, 1995–2015. *ILR Review*, 72(2):382–416, 2019.
- William F Maloney. Informality revisited. *World development*, 32(7):1159–1178, 2004.
- Irene Mandl. New forms of employment: 2020 update. Technical report, The European Foundation for the Improvement of Living and Working Conditions, Publications Office of the European Union, Luxembourg, 2020.
- Alexandre Mas and Amanda Pallais. Valuing alternative work arrangements. *American Economic Review*, 107(12):3722–59, 2017.
- Jonathan A Parker, Nicholas S Souleles, David S Johnson, and Robert McClelland. Consumer spending and the economic stimulus payments of 2008. *American Economic Review*, 103(6):2530–53, 2013.
- Fatto Quotidiano. Voucher lavoro, renzi blocca decreto con i correttivi troppo soft. “timore di attacchi da opposizioni e sindacati”. *Il Fatto Quotidiano*, June 1 2016.
- Friedrich Schneider and Dominik H Enste. Shadow economies: Size, causes, and consequences. *Journal of economic literature*, 38(1):77–114, 2000.
- Mr Friedrich Schneider and Dominik Enste. *Hiding in the shadows: the growth of the underground economy*. International Monetary Fund, 2002.
- Liyang Sun and Sarah Abraham. Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, 2020.

Gabriel Ulyssea. Firms, informality, and development: Theory and evidence from brazil. *American Economic Review*, 108(8):2015–47, 2018.

Gabriel Ulyssea. Informality: Causes and consequences for development. *Annual Review of Economics*, 12:525–546, 2020.

Colin Williams, Predrag Bejaković, Davor Mikulić, Josip Franic, Abbi Kedir, and Ioana Alexandra Horodnic. An evaluation of the scale of undeclared work in the european union and its structural determinants: estimates using the labour input method. *European Commission, Directorate-General for Employment, Social Affairs and Inclusion, Publications Office of the European Union, Luxembourg*, 2017.

Colin W. Williams. Elements of a preventative approach towards undeclared work: an evaluation of service vouchers and awareness raising campaigns. Technical report, European Platform Undeclared Work, 2018.

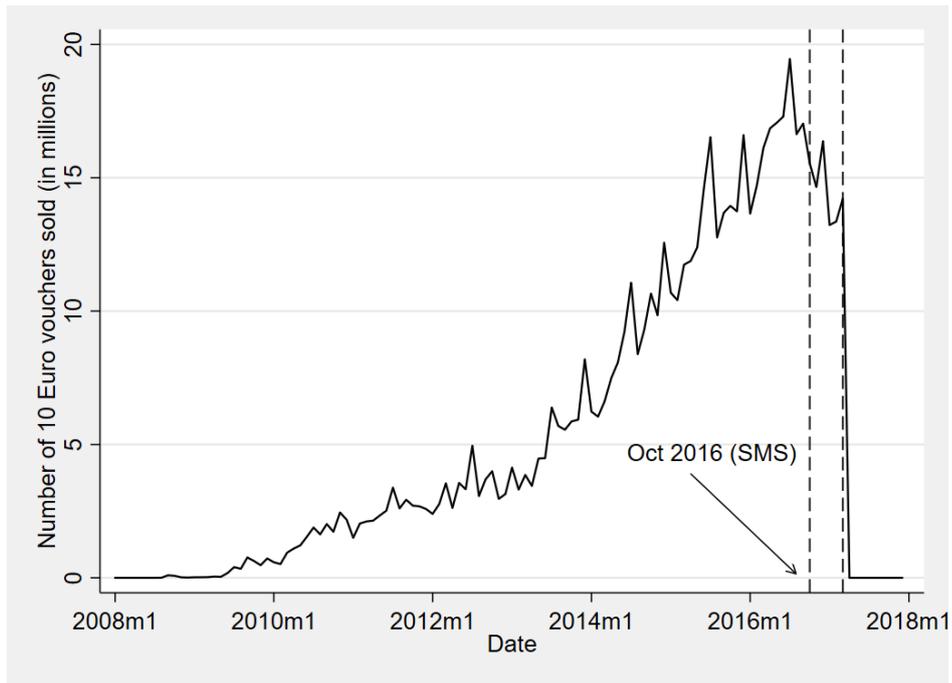


Figure 4: Vouchers Sold

Notes: The figure plots the monthly total number of 10-euro vouchers sold.

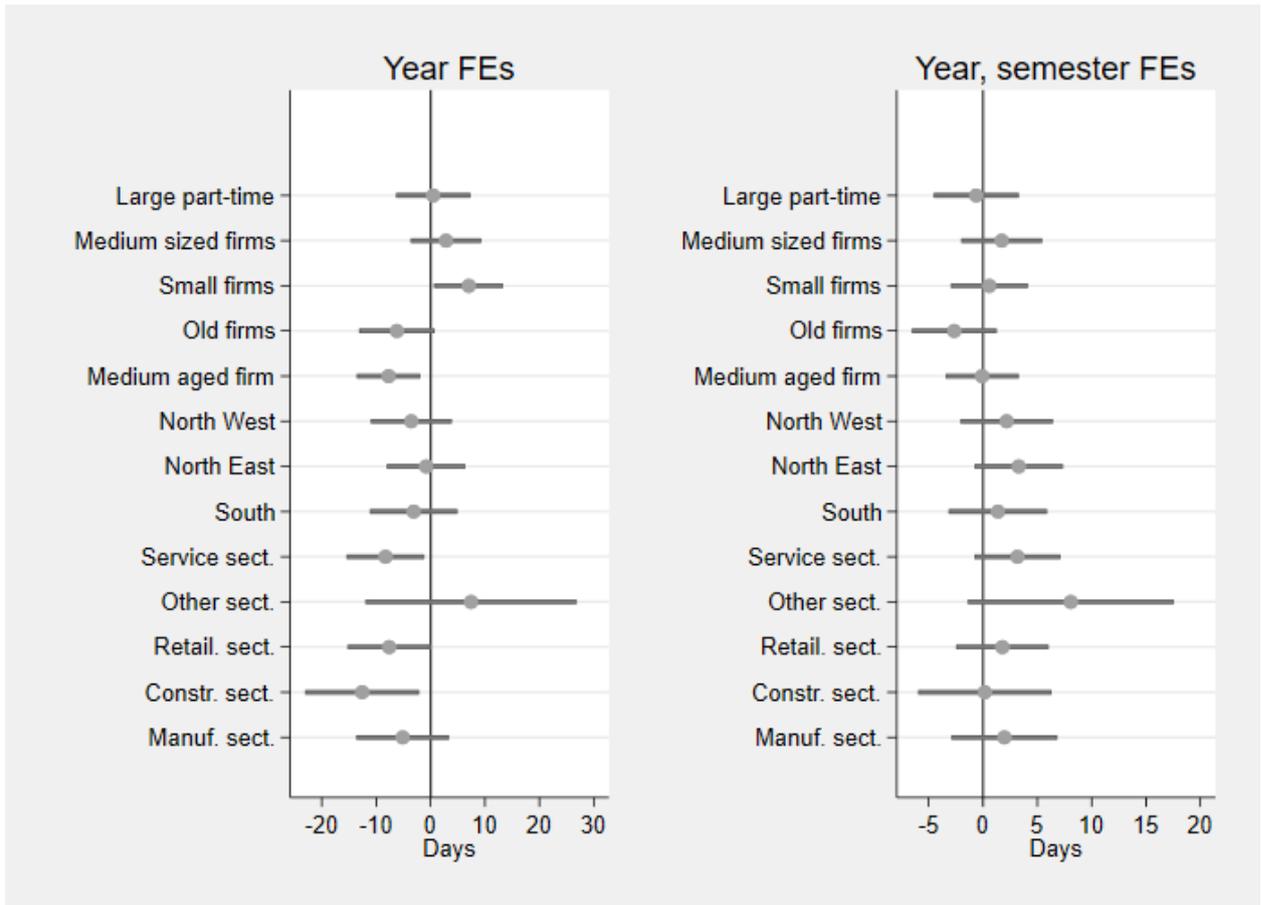


Figure 5: Balance Test

Notes: The figure plots the coefficients of day of inspection on firm characteristics. The regression results shown on the left control for year fixed effect, the ones in the second add also semester fixed effects. The vertical bars represent the 95% confidence intervals based on robust standard errors. The p-values for the F-test that sets all coefficients equal to zero are respectively 0 and 72%.

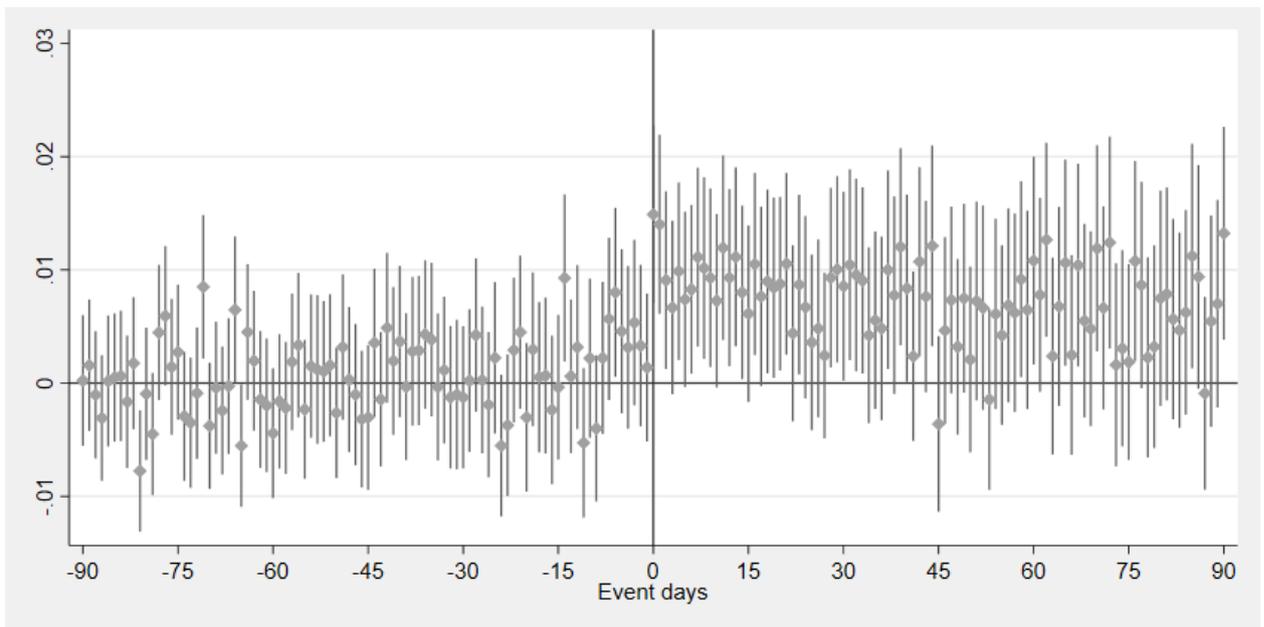


Figure 6: Event Study pre-SMS

Notes: The figure plots event study coefficients, where the event is a labour inspection. The excluded time period is between 180 and 90 days prior to the inspection. The regression controls for year, month, and day of the week fixed effects. Standard errors are clustered at the firm level.

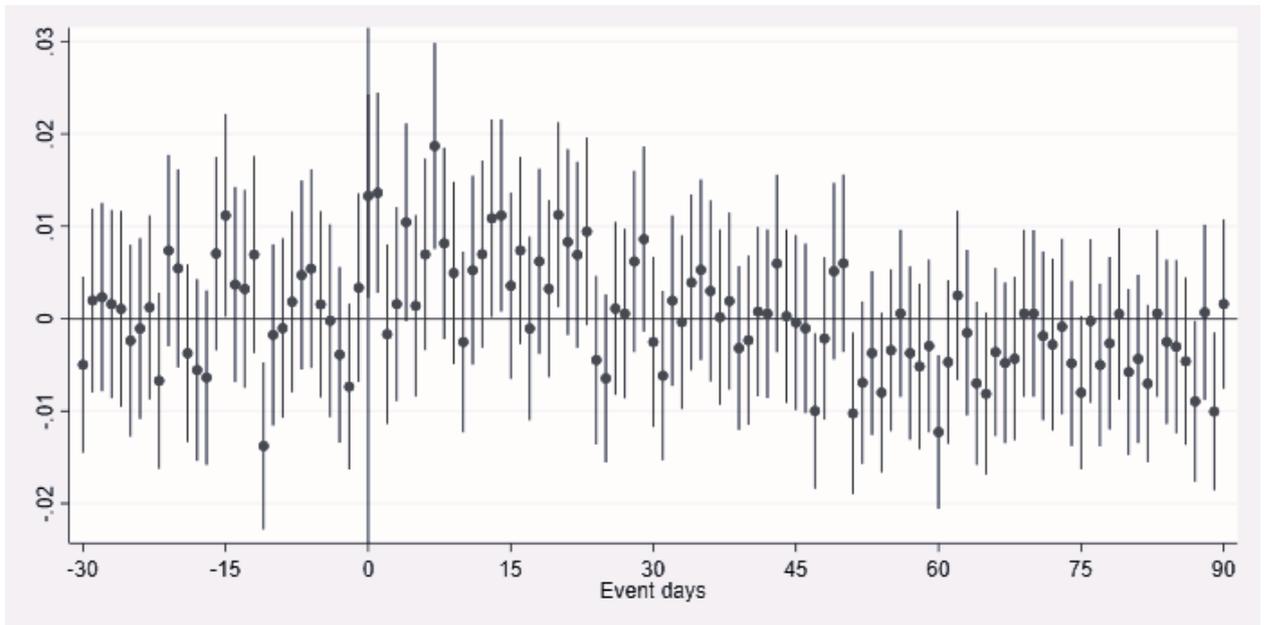


Figure 7: Event Study post-SMS

Notes: The figure plots event study coefficients, where the event is a labour inspection. The excluded time period is between 60 and 30 days prior to the inspection. Standard errors are clustered at the firm level.

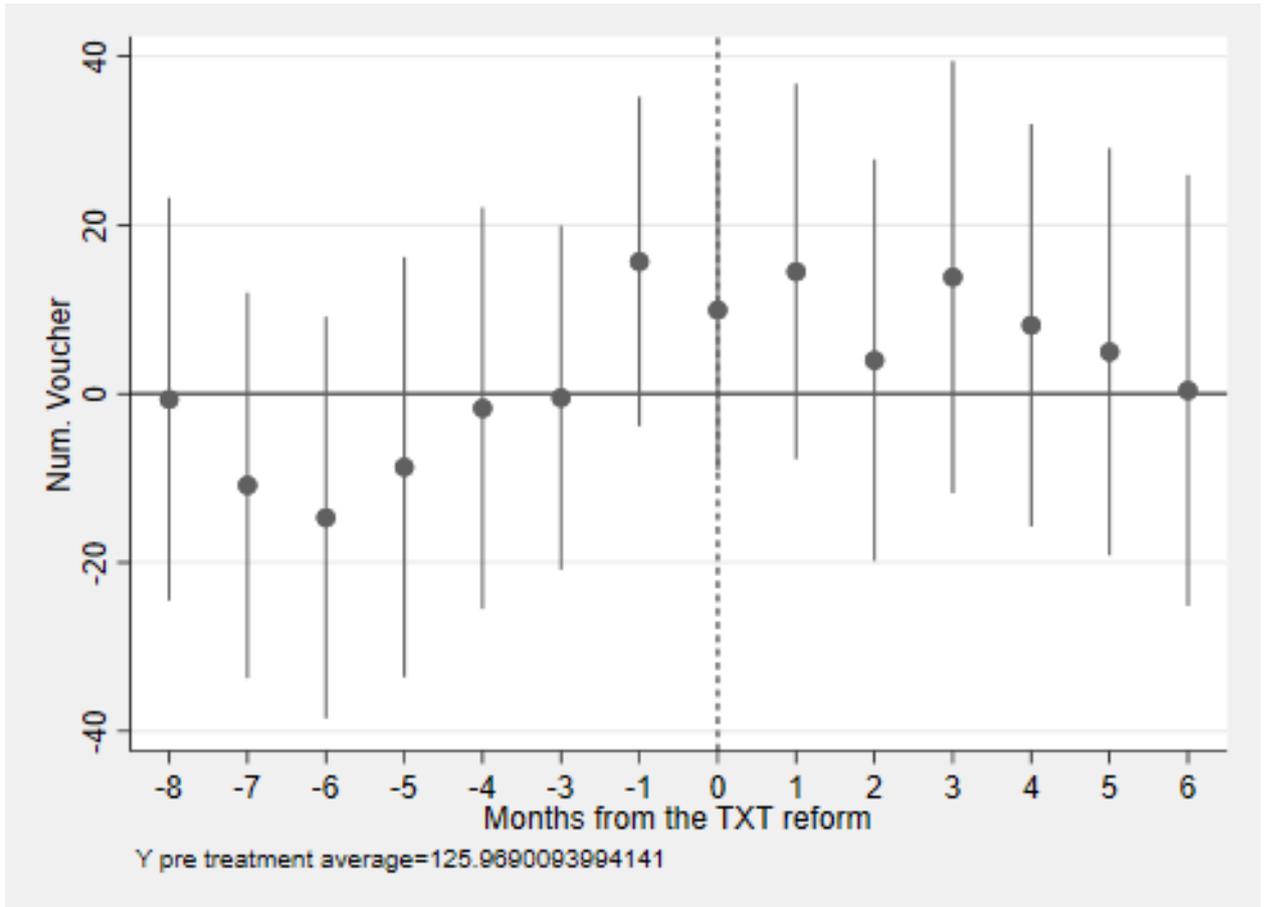


Figure 8: Event Study post-SMS

Notes: The figure plots differences in the number of vouchers used at firms that on average “mis-used” vouchers and those that did not, 8 months before and 6 months after the abolition of vouchers. Standard errors are clustered at the firm level.

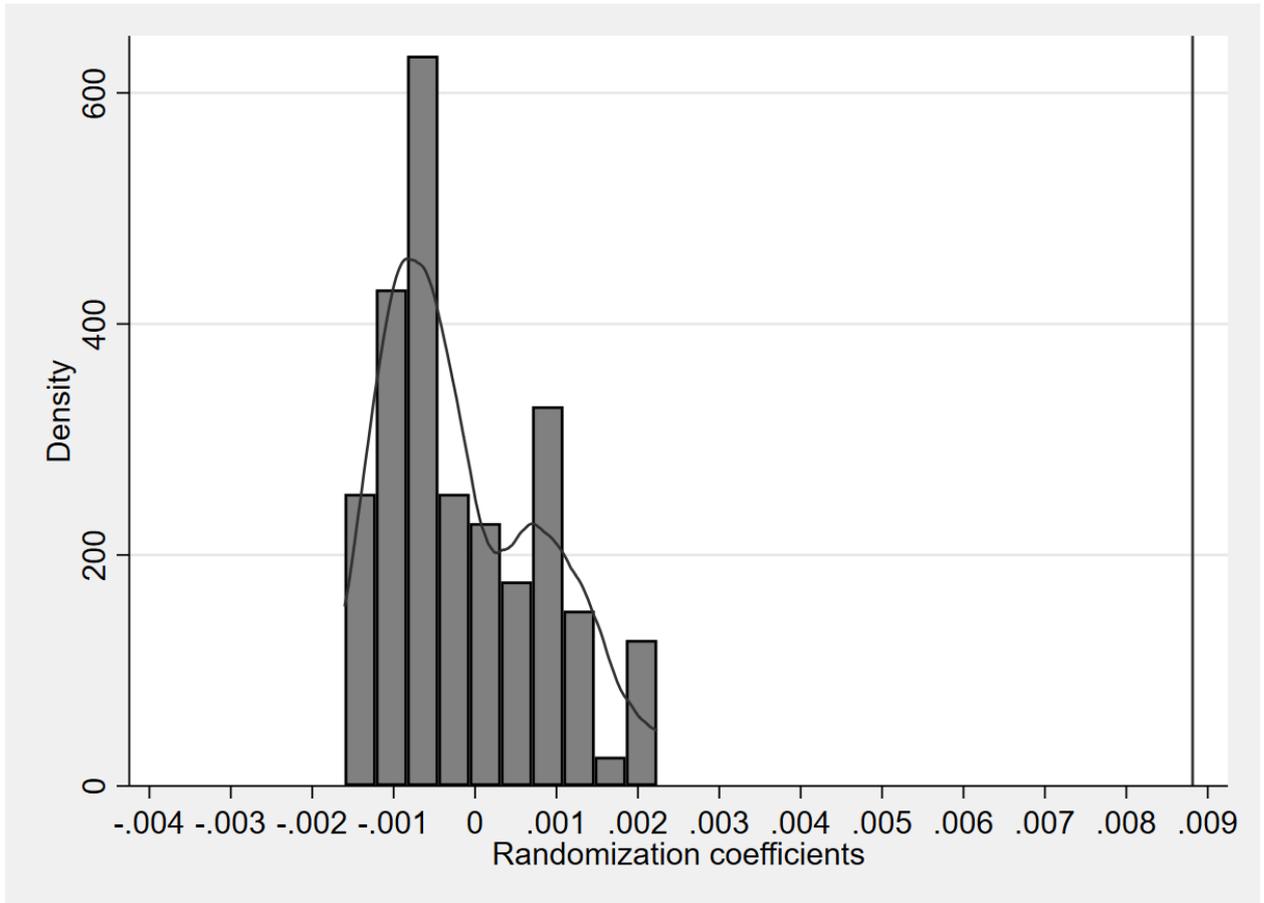


Figure 9: Randomization Test

Notes: The figure plots the density of the randomization coefficients in a simple difference model: $DV_{j,\tau} = \alpha + \beta D_{\tau-t} + \epsilon_{j,\tau}$ where $t < 0$ measures the event time of inspection (it's zero the day of the inspection), while $D_{\tau-t}$ takes value one t days before the inspection. The vertical line on the right corresponds to the coefficient when the timing of the labour inspection is correct ($t = 0$) and the full sample is used ($t \leq 0$).

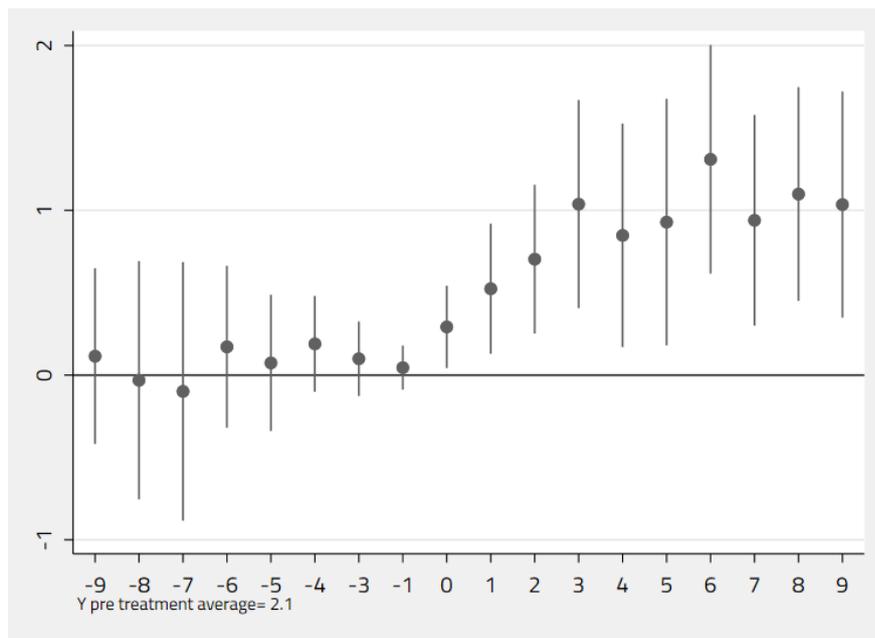


Figure 10: Part-time and Temporary Contracts Around the Abolition of Vouchers

Notes: The figure plots differences in the total number of part-time and temporary workers employed at firms that on average “mis-used” vouchers and those that did not, 9 months before and 10 months after the abolition of vouchers. Standard errors are clustered at the firm level.

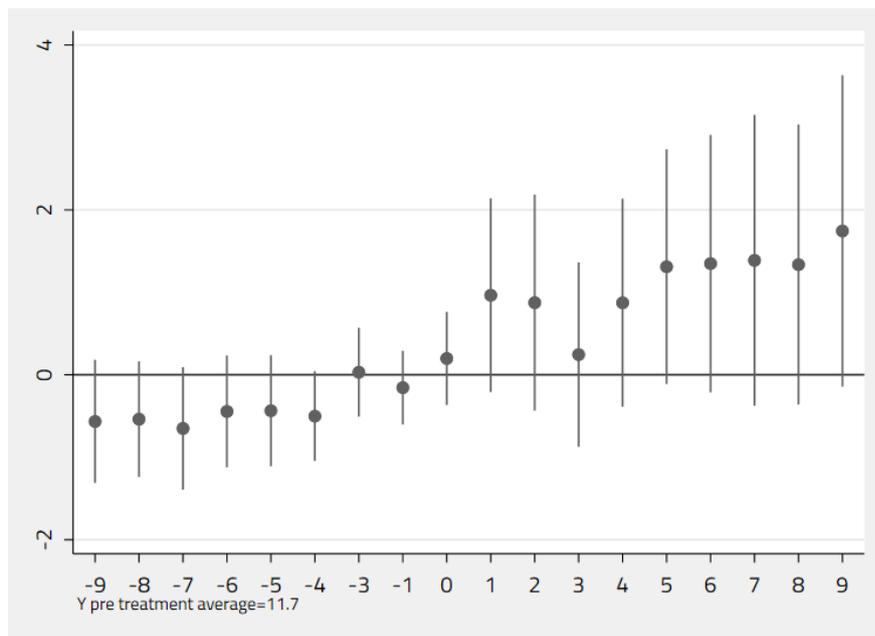


Figure 11: Workers Employed around the Abolition of Vouchers

Notes: The figure plots differences in the total number of workers employed at firms that on average “mis-used” vouchers and those that did not, 9 months before and 10 months after the abolition of vouchers. Standard errors are clustered at the firm level.

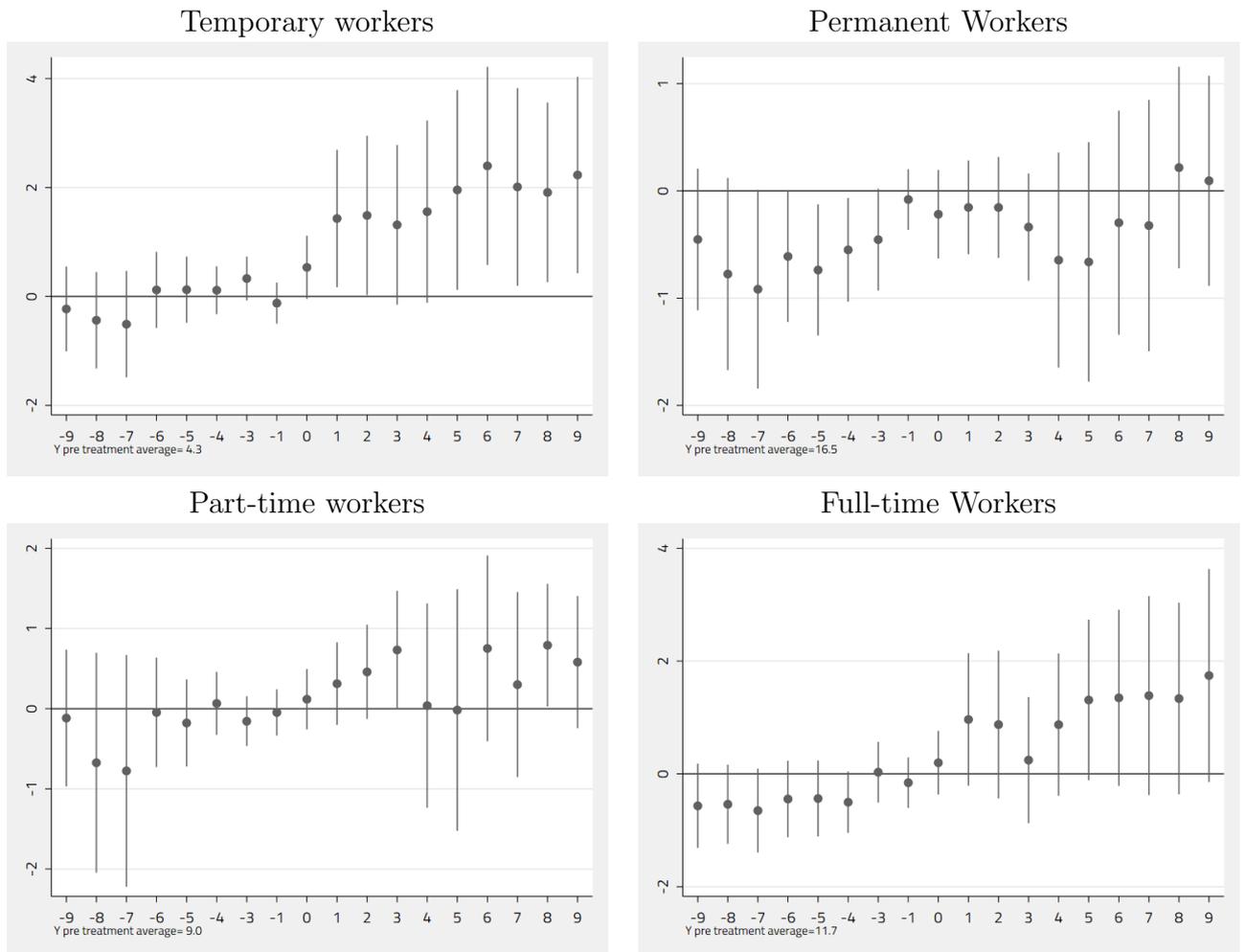


Figure 12: Event Study around the Abolition of Vouchers

Notes: The figure plots differences in the number of workers employed at firms that on average “mis-used” vouchers and those that did not, 9 months before and 10 months after the abolition of vouchers. Standard errors are clustered at the firm level.

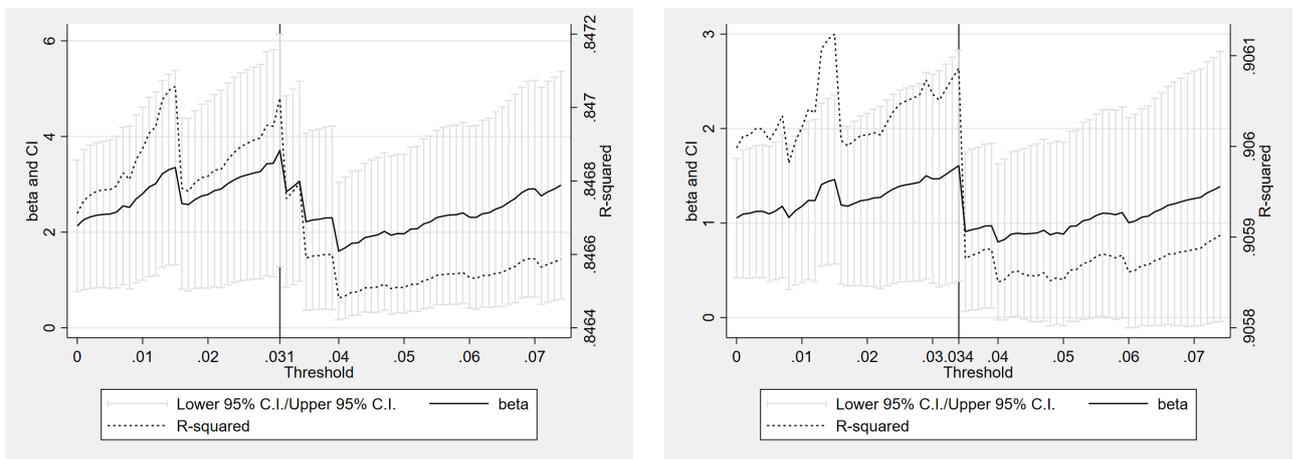


Figure 13: Sensitivity analysis on the misclassification bias

Notes: The figures plot differences in the total number of temporary workers (left panel) as well as temporary and part-time workers (right panel) employed at firms that on average “mis-used” vouchers and those that did not, and the corresponding 95% confidence intervals. The definition of misusing firm \widehat{M}_j^η varies based on the η threshold: $\widehat{M}_j^\eta = 1\{\hat{\beta}_j > \eta\}$. Standard errors are clustered at the firm level.

Table 1: Summary Statistics

Firms:	All	Vouchers	Inspected	Both	Misbehaving	Compliant
Workforce	7.171	7.212	43.511	40.818	42.737	39.411
Monthly Wage bill	14,329	11,968	93,170	63,042	61,664	64,051
Average wage	1,346	1,271	1,316	1,247	1,231	1,258
Monthly Wage Bill (FTE)	15,894	13,737	102,458	75,886	78,431	74,021
Workforce (FTE)	6.3	6.1	38.5	33.7	33.7	33.7
Temporary workers	1.1	1.8	9.4	16.5	14.4	18.0
Part-time workers	2.1	2.7	11.8	17.0	20.8	14.1
Permanent workers	6.1	5.4	34.1	24.3	28.3	21.4
Full-time workers	5.0	4.5	31.7	23.9	21.9	25.3
Monthly number of vouchers	27.9	97.9	306.8	252.9	271.3	239.4
Voucher per workforce (FTE)	4.461	15.928	7.979	7.500	8.050	7.097
Number of firms	1,840,339	416,942	40,121	3,472	1,469	2,003
Fraction	100%	22.66%	2.18%	0.19%	0.08%	0.11%

Notes: This table shows averages for different samples using 2015 and 2016 data. The samples are the following: “Full” is the universe of private firms in 2015 and 2016; “Vouchers” are firms that have used at least one voucher in those years; “Inspected” are firms that have been inspected between January 2015 and October 2016 (almost all are inspected only once); “Both” are firms that use at least on voucher and have been inspected. Presumably “Misbehaving” and “Compliant” firms either increase or decrease the use of vouchers upon inspection. “FTE” stands for full-time-equivalent.

Table 2: Vouchers and labour Inspections by Sector

Sector	(1) Manufacturing	(2) Construction	(3) Retail	(4) Tourism	(5) Other Services
Post-Inspection	0.013*** (0.003)	-0.002 (0.003)	0.012** (0.006)	0.011*** (0.002)	0.006* (0.003)
Constant	0.032*** (0.003)	0.021*** (0.003)	0.030*** (0.002)	0.054*** (0.002)	0.051*** (0.004)
Observations	157,329	98,087	208,194	614,758	255,718
R-squared	0.014	0.013	0.010	0.022	0.016
Mean dep. var.	0.0381	0.0201	0.0352	0.0592	0.0541

Notes: Linear probability model of using at least one voucher, daily data with year, month and day of the week fixed effects. Clustered standard errors (by firm) in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 3: Vouchers and labour Inspections Heterogeneity

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
	South	Center	North-East	North-West	Young Firms	Medium-aged F.	Old Firms	Small Firms	Medium-sized F.	Large Firms	High PT Use	Above 20% Temporary	Below 20% Temporary
Subsample	0.009** (0.004)	0.008** (0.003)	0.010*** (0.004)	0.011*** (0.003)	0.009*** (0.002)	0.011*** (0.004)	0.008** (0.003)	0.006*** (0.002)	0.008*** (0.003)	0.012*** (0.004)	0.014*** (0.003)	0.011*** (0.003)	0.085*** (0.002)
Post-Inspection	0.047*** (0.003)	0.046*** (0.003)	0.047*** (0.002)	0.043*** (0.003)	0.045*** (0.002)	0.048*** (0.002)	0.044*** (0.003)	0.035*** (0.001)	0.044*** (0.002)	0.058*** (0.003)	0.057*** (0.002)	0.052*** (0.002)	0.039*** (0.002)
Constant	255.262 (0.015)	256.735 (0.015)	409.649 (0.014)	347.459 (0.014)	581.120 (0.014)	414.932 (0.013)	273.053 (0.017)	461.643 (0.012)	373.953 (0.016)	433.509 (0.017)	446.903 (0.005)	644.001 (0.015)	534.543 (0.013)
Observations													
R-squared	0.015	0.015	0.014	0.014	0.014	0.013	0.017	0.012	0.016	0.017	0.005	0.015	0.013
Mean dep. var.	0.0495	0.0477	0.0504	0.0461	0.0477	0.0498	0.0481	0.0368	0.0466	0.0609	0.0574	0.0548	0.0421

Notes: Linear probability model of using at least one voucher, daily data with year, month and day of the week fixed effects. Clustered standard errors (by firm) in parentheses: *** p<0.01, ** p<0.05, * p<0.1. Large firms are those with more than 15 employees: old firms are those with more than 14 years of age, medium are between 5 and 14 years of age, and young firms are below 5 years of age.

Table 4: Vouchers and Evaded Contributions

	(1)	(2)	(3)	(4)
Post-Abolition*Treated	30,562** (14,835)	32,615** (14,654)	34,465** (16,503)	32,876** (16,566)
Post-Sms*Treated	-7,076 (6,726)	-6,732 (6,714)	-6,272 (7,564)	-7,628 (7,559)
Treated	-3,640 (3,635)	-3,737 (3,621)	-4,386 (4,058)	-4,032 (4,070)
Post-Abolition	12,426** (5,806)	10,202* (5,573)	11,474* (6,085)	10,756* (6,001)
Post-Sms	7,334 (4,834)	3,021 (5,655)	737 (6,361)	1,134 (5,219)
Observations	20,817	20,817	18,108	18,108
Controls for Employment	No	No	No	Yes
Controls for Wages	No	No	Yes	Yes
Cubic Monthly Trend	No	Yes	Yes	Yes

Notes: The dependent variable is the amount of evaded contributions for each inspected firm. Treated firms have used at least one voucher during the sample period, control firms have never used vouchers. The baseline period is from January to October 17, 2016, while the Post-SMS period goes from October 17, 2016 to March 17, 2017 (TXT introduction), and the Post-Abolition period starts in March 17, 2017 (Abolition date) to the end of 2017. Standard errors are clustered at the firm level. *** p<0.01, ** p<0.05, * p<0.1.

A Appendix

A.1 Proof of Proposition 2

The decision to go shadow for *FWA* jobs requires that for $\gamma c > (1 - \gamma)(1 - \rho^{FWA})\tau + \theta^h$ which implies $\theta^h < \gamma c - ((1 - \gamma)(1 - \rho^{FWA})\tau)$ or $c > \frac{((1 - \gamma)(1 - \rho^{FWA})\tau)}{\gamma}$. Further, $\theta^h < \gamma c - (1 - \rho^{FWA})(1 - \gamma)\tau > 0$. With these restrictions, the marginal shadow value is

$$\tilde{\lambda}_{s,FWA}(\theta^h) = 1 - \frac{\theta^h}{D}, \quad D \equiv \gamma c(1 - \alpha) - (1 - \gamma)\tau(1 - \rho^{FWA}) > 0, \quad (18)$$

which is lower than λ^s as long as θ^h satisfies the condition above.

A.2 Tables and Figures

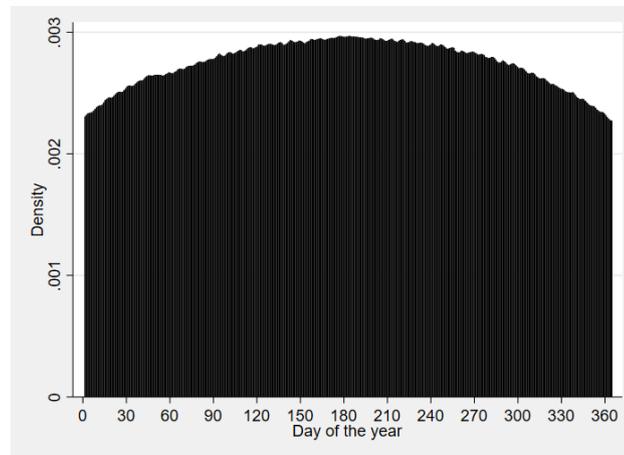


Figure A1: Distribution of Labour Inspections

Notes: The figure plots the histogram of the exact day of inspection.

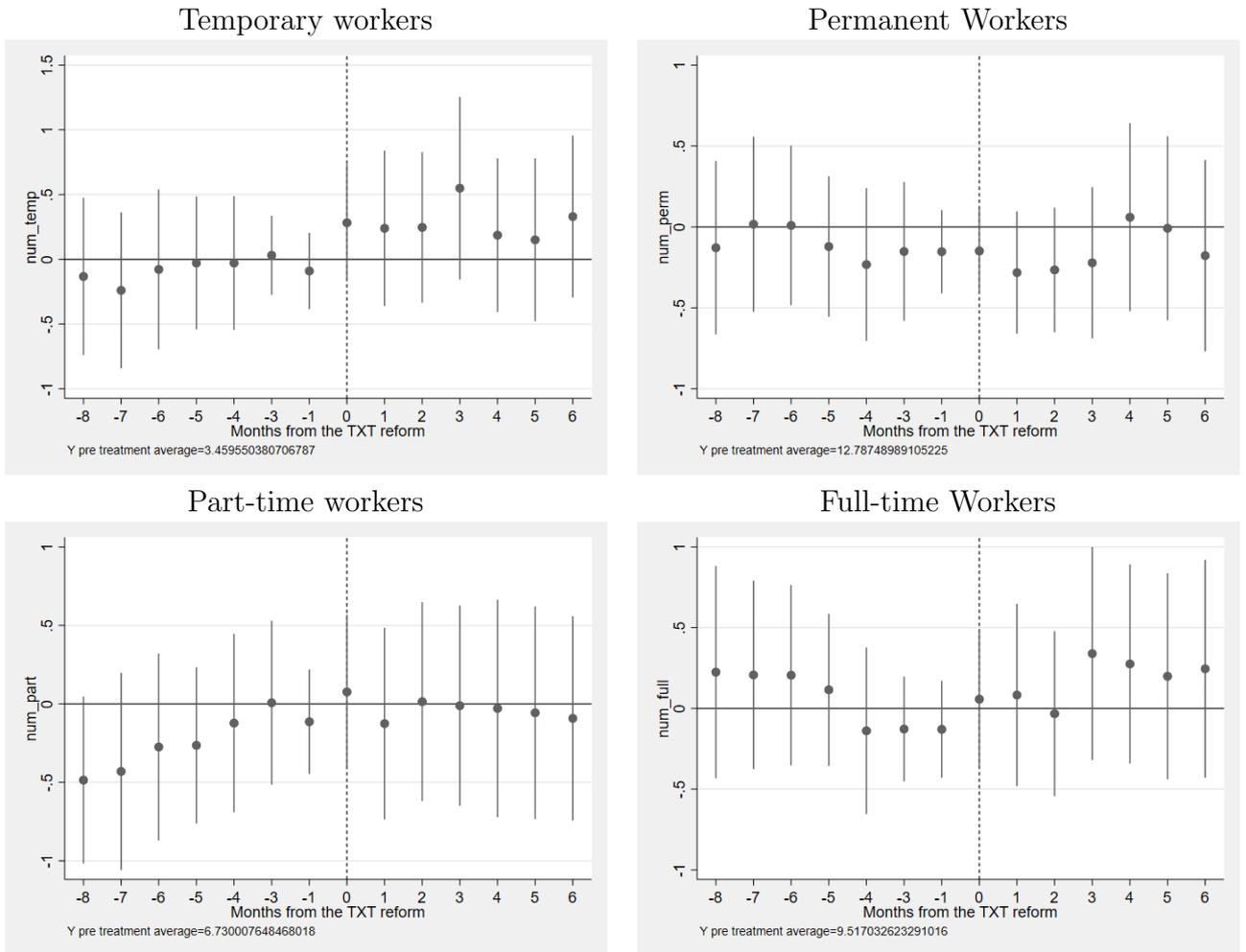


Figure A2: Event Study Around the Introduction of the SMS/TXT Requirement

Notes: The figure plots differences in the number of workers employed at firms that on average “mis-used” vouchers and those that did not, 9 months before and 6 months after the Introduction of the SMS/TXT (we have to limit the analysis to six months after as in March 2017 vouchers were abolished). Standard errors are clustered at the firm level.

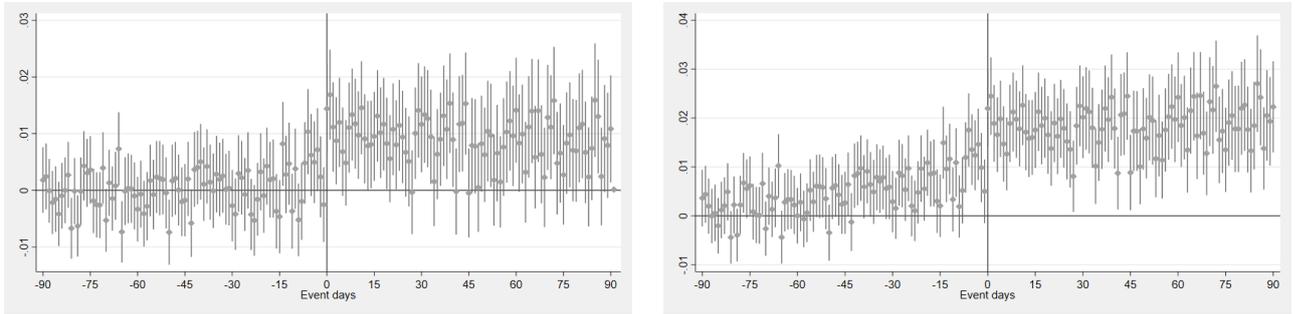


Figure A3: Event Study pre-SMS

Notes: The figure plots event study coefficients, where the event is a labour inspection. The excluded time period is between 180 and 90 days prior to the inspection. Standard errors are clustered at the firm level. The figure on the left controls for calendar time t (in Eq. 11, $f(t) = \alpha t$), the figure on the right has no additional controls (in Eq. 11, $f(t) = 0$).

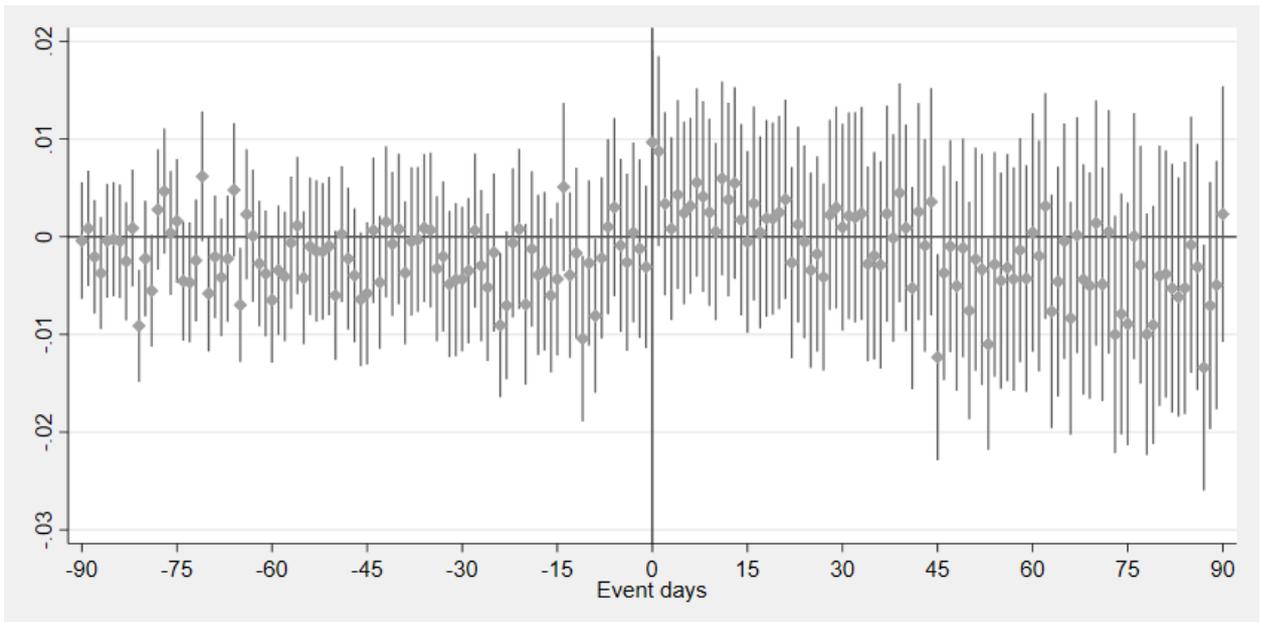


Figure A4: Event Study pre-SMS with Firm Fixed Effects

Notes: The figure plots event study coefficients, where the event is a labour inspection. The excluded time period is between 180 and 90 days prior to the inspection. The regression controls for date and firm fixed effects. Standard errors are clustered at the firm level.

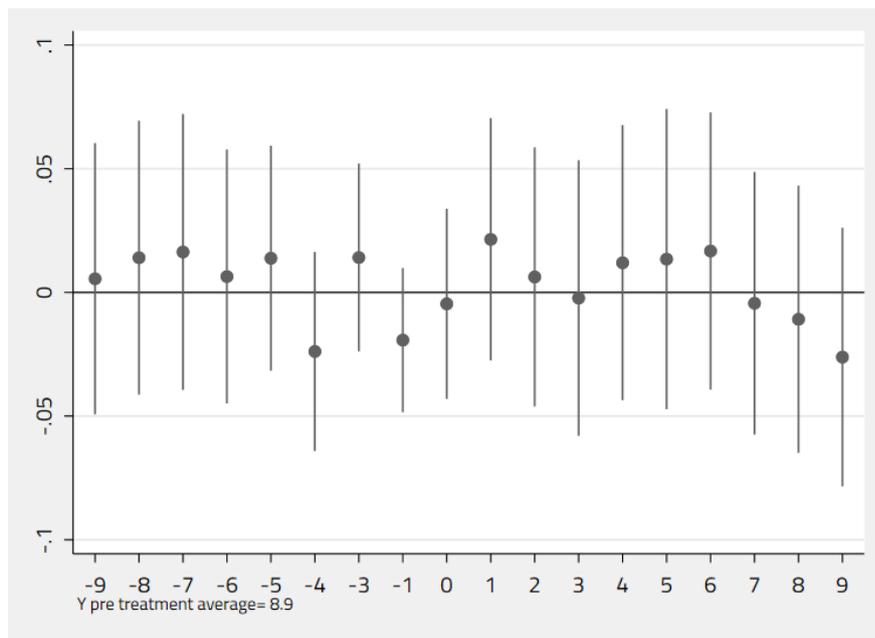


Figure A5: Log-Total Wage Bill Around the Abolition of Vouchers

Notes: The figure plots differences in the total wage bill between firms that on average “mis-used” vouchers and those that did not, 10 months before and 9 months after the abolition of vouchers. Standard errors are clustered at the firm level.

Table A1: Difference-in-Differences Models of Vouchers and labour Inspections

	(1)	(2)	(3)	(4)	(5)	(6)
Post period	15 days			30 days		
Post-Inspection	0.0094*** (0.002)	0.0095*** (0.002)	0.0069*** (0.001)	0.0083*** (0.002)	0.0084*** (0.002)	0.0053*** (0.001)
Firm Fixed Effects	No	No	Yes	No	No	Yes
Time Fixed Effects	No	Yes	Yes	No	Yes	Yes
Year, month, dow Fes	Yes	No	No	Yes	No	No
Observations	713,849	713,779	713,773	761,100	761,015	761,012
R-squared	0.023	0.025	0.165	0.023	0.025	0.164

Notes: Linear probability model of using at least one voucher with daily data. Clustered standard errors (by firm) in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

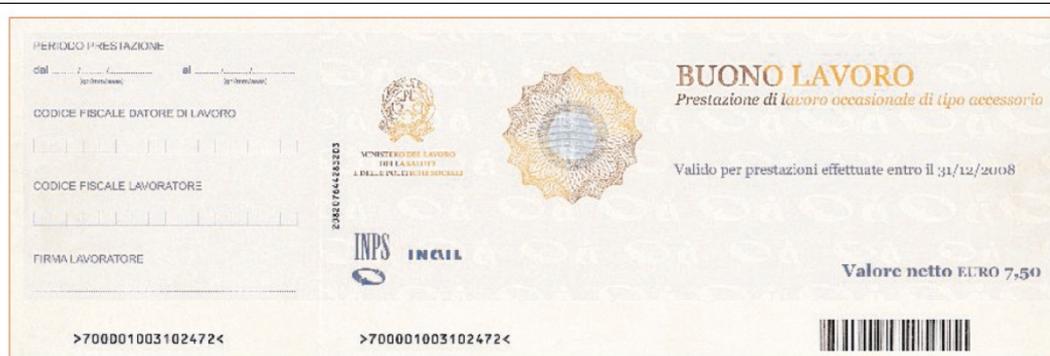


Figure A6: The Photograph of an Italian voucher