CRELI

Centro di ricerche per i problemi del lavoro e dell'impresa

CRELI Working Paper

Why don't you take a free shot? Free access to flu vaccination and its effects on take-up probability and health

Ylenia Brilli
Claudio Lucifora
Antonio Russo
Marco Tonello

Working Paper n. 2-2018

UNIVERSITÀ CATTOLICA del Sacro Cuore

Why don't you take a free shot? Free access to flu vaccination and its effects on take-up probability and health

Ylenia Brilli
Centre of Health Economics, Gothenburg University and Università di Verona

Claudio Lucifora
Università Cattolica del Sacro Cuore and IZA

Antonio Russo
Epidemiology Unit, Agency for Health Protection - Milan

Marco Tonello
Bank of Italy, Regional Economic Analysis - Florence and CRELI

Working Paper n. 2-2018

CRELI
Centro di Ricerche per i Problemi del Lavoro e dell'Impresa
Università Cattolica del Sacro Cuore
Largo Gemelli 1 - 20123 Milano – Italy
tel: +39.02.7234.2976 - fax: +39.02.7234.2781
e-mail: dip.economiaefinanza@unicatt.it

Why don't you take a free shot? Free access to flu vaccination and its effects on take-up probability and health*

Ylenia Brilli† Claudio Lucifora‡ Antonio Russo§ Marco Tonello¶

This version: February 21, 2018

Abstract

We study how the individual decision to get vaccinated against the seasonal flu changes according to whether the medical treatment is free of costs or not in a Regression Discontinuity framework by exploiting individual-level administrative data on the 2013 vaccination campaign in Italy, and the fact that vaccination is free for individuals aged 65 and more. Based on the exact date of birth, we compare individuals born before January 1st 1949 (eligible) with those born after (not eligible) and find that eligibility to free vaccination rises the take-up probability of about 5 percentage points. We complement our analysis assessing the subsequent short-term health consequences of getting the vaccination shot and find that vaccination decreases the use of drugs for respiratory diseases, and the probability of hospitalization, especially during the weeks in which the flu spread reached its peak. These results are precisely estimated only for some selected groups such as females and those living outside urban areas.

JEL Classification: I12, I18, J10

Keywords: vaccination, influenza, health cost-sharing, public health

*The views expressed in this paper are those of the authors and do not involve the responsibility of the Institutions they belong to. The usual disclaimers apply.

†Gothenburg University, CHEGU - Centre of Health Economics Gothenburg University, and CHILD - Collegio Carlo Alberto.

‡Catholic University Milan and IZA. **Corresponding author**: claudio.lucifora@unicatt.it

§Epidemiology Unit, Agency for Health Protection of Milan (Milan).

¶Bank of Italy, Regional Economic Analysis Division in Florence, and CRELI - Centre for the Study of Labor and Industrial Policies, Catholic University Milan.

1 Introduction

Seasonal influenza is a serious public health problem that causes severe illness and death in high risk populations.[1] Worldwide, these annual epidemics are estimated to result in about 3 to 5 million cases of severe illness, and about 290 to 650 thousands deaths, which, in industrialized countries are mostly concentrated among people age 65 or older (WHO, 2017).[2] The health consequences of flu pandemics are especially severe for populations more at risks (i.e. the children and the elderly), and may even affect fetus conditions during pregnancy, with lasting effects in later life (Almond, 2006; Kelly, 2011).

Despite the often short duration of illness, the yearly economic and health care burden of influenza is substantial (Hayward et al., 2014). Epidemics can result in high levels of worker/school absenteeism and productivity losses; clinics and hospitals can be overwhelmed during peak illness periods (WHO, 2017). In aggregate terms, the effects of the spread on the seasonal influenza might have consequences which, through health, may also affect human capital accumulation, labor force participation and economic growth (Adda, 2016).

Nevertheless, influenza is an easily preventable disease through vaccination. Safe and effective vaccines are available and have been used for more than 60 years (WHO, 2017).[3] Flu vaccination is also a peculiar type of vaccine: it is not compulsory (though recommended) and all adult population (and not only children) may decide to take it. The decision of taking the flu vaccine repeatedly happens every year and can be conceptualized as an investment in own health (Grossman, 1972). Under this regard, free access to vaccination might be a key economic variable in the decision process, which may influence individuals at the margin.

In this paper we focus on the take-up of seasonal influenza vaccination and study how free access may affect the individuals' decision to get it (so-called *take-up probability*). Moreover, we provide an assessment of the short-term health consequences of the vaccination, as measured by the use of respiratory drugs and hospitalization events.

Our work contributes to two different bodies of the literature. On the one hand, it contributes to the little knowledge of the individual-level determinants of vaccination decision (Mullahy, 1999; Schmitz and Wübker, 2011). In this case, our focus is on an economic determinant, i.e. the cost affordability of the vaccination shot. To the best of our knowledge, this work is the first to provide a causal estimate of the effect of free access to vaccination on the individual take-up probability. On the other hand, we contribute to a wider body of literature, mainly focused on the US, which evaluates individuals' propensity to make use of general health care provision depending on its cost and the subsequent effects on health. As in the tradition of those works, we exploit a discontinuity in the cost-sharing of health services determined by an age threshold (Card et al., 2008; 2009; Ponzo and Scoppa, 2016). Differently from them, we focus

[1] Seasonal influenza is a preventable infectious disease with mostly respiratory symptoms. It is caused by influenza virus and is easily transmitted, predominantly *via* the droplet and contact routes and by indirect spread from respiratory secretions (WHO, 2017).

[2] Recent estimates for Europe report about 15 to 70 thousands deaths per year (ECDPC, 2017); in the United States this figure range from a low of 12 thousands deaths (in 2011-2012) to a high of 56 thousands (in 2012-2013) (CDC, 2016).

[3] Among healthy adults, influenza vaccine provides protection, even when circulating viruses may not exactly match the vaccine viruses; among the elderly, influenza vaccination also reduces severity of disease and incidence of complications and deaths. Vaccination is especially important for people at higher risk of serious influenza complications, and for people who live with, or care for, high risk individuals (WHO, 2017).

on a specific health treatment which has not been explored yet. Only Ward (2014) studies the effects of an influenza vaccination program expanding coverage outside the typical target group in Canada. However, she is mainly interested in the estimate of health and labor market effects by exploiting aggregated data, while she does not focus on individual take-up probability and health outcomes.

We use individual-level administrative records from the largest metropolitan area in Northern Italy on the 2013-14 influenza campaign, combined with an administrative rule which assigns free vaccinations to all 65-year-old residents. Based on the exact date of birth, we estimate the variation in the vaccine take-up probability by comparing individuals individuals born just before January 1st 1949 (eligible) with those born just after (not eligible) in a Regression Discontinuity (RD) framework. Moreover, we complement our analysis by assessing the subsequent short-term health consequences of getting the vaccine shot by linking individual data on hospitalization and on the use of respiratory drugs. For the estimation, we mainly make use of the non-parametric methods developed by Calonico et al. (2014) and Calonico et al. (2017a).

We find that eligibility to free vaccination rises the take-up probability of about 5 percentage points. The effect is slightly higher for females and for residents in the main urban area. These results are robust to alternative specifications, estimation samples and methodologies. Concerning the subsequent effects on health, the results are less precisely estimated. However, we do find evidence that vaccination decreases the use of respiratory drugs and hospitalization, though the effects seem concentrated in certain population groups (females and residents outside the main urban area).

Our work bears important policy implications. Current influenza vaccination recommendations focus on immunizing the elderly and the high-risk people. However, given that the minimum thresholds of vaccination take-up are more often not reached in the targeted population, policy makers are starting to consider the expansion of flu vaccination recommendations to lower the age eligibility threshold for free access to flu vaccination in the adult population. In this regard, our estimates are useful under at least two main aspects. First, we provide a robust estimate of a structural policy-relevant parameter, which can be used, for example, in a reliable cost-benefit analysis of an expansion of free vaccination to other adult populations. Second, our results also show that the gratuity of the vaccine is not the only driver of the take-up decision in the targeted population, as a relevant share do not take the free shot. Thus, in order to increase the take-up of those who already benefit of the free access to the treatment, the policy maker should look for other leverages, excluding the price, which may, perhaps, involve more role of the family doctors or public information campaigns.

The rest of the paper is structured as follows. Section 2 reviews the related literature, while 3 provides a description of the institutional setting. Section 4 discusses the identification strategy and the main threats to identification. Section 5 describes the administrative archives used and provides descriptive evidence, as long as some tests on the validity of the RD assumptions in our setting. Section 6 describes the results and provides several robustness and sensitivity checks. Section 7 concludes and provides some policy relevance considerations.

2 Related literature

2.1 Flu vaccination determinants

Little is known on the individual-level determinants of flu vaccination decision. Mullahy (1999) and Schmitz and Wübker (2011) analyze the micro-economics determinants of getting flu vaccine using mainly *coeteris paribus* correlation models and survey data (the *National Health Interview Survey* for the US, and the *Survey of Health, Aging, and Retirement* (SHARE) on 50+ European residents, respectively). Both studies find that the most important determinants of taking flu vaccination are the risk factors, namely age and health status. Those for which vaccination is shown to have the strongest effect − the elderly and individuals with chronic conditions − are most likely to take-up vaccination. Physician quality also has a positive impact on the propensity to vaccinate, which might reflect the fact that physician's recommendation is important for a patient's decision to vaccinate against influenza. They also provide evidence that working full-time increases the likelihood to get influenza vaccination: the time costs for workers to obtain immunization seem to be much lower than the costs from being ill with the flu.

Ward (2014) estimates the overall impact of an influenza vaccination program expanding coverage outside the typical target group (i.e. children until the age of five and 65+) to all population. Using a triple-difference design, which exploits the introduction of the vaccination program in Ontario (Canada) and the variation of quality of the vaccine from year to year, higher vaccination is linked to health improvements in terms of lower hospitalization rates and fewer work absences.

The medical literature finds that individual level characteristics such as age, gender, marital status, education, ethnicity, socio-economic status, social and cultural values, place of residence, behavioral beliefs, social influences, previous vaccine experiences, sources of information, and perceived health status correlate with the decision of taking seasonal influenza vaccination (Nagata et al., 2011). Health care-system-related factors (such as accessibility, knowledge and attitudes about vaccination, and physicians advice) are also found to take part in the vaccination decision. Among them, vaccination cost is found to be an important determinant in countries where patients have to pay for the vaccine, as elderly people may consider having the influenza vaccination only if it is provided free of charge (Nagata et al., 2011; Daniels et al., 2004).

2.2 Health care provision and cost-sharing

A wider body of literature, mainly focused on the US experience, evaluates individuals' price elasticity to health care provision and subsequent health effects. Age thresholds which determine the eligibility to different health insurance cost coverages are usually exploited as source of exogenous variation.[4] Card et al. (2008) exploit the exogenous variation in the health insurance coverage of the US population at age 65 induced by *Medicare* eligibility rules. These changes lead to increases in the use of medical services, with a pattern of gains across socioeconomic groups that varies by type of service. While routine doctor visits increase more for groups that

[4]Notice that the health system in the US is such that different health insurance coverage might correspond to different health treatments, so that, in the end, health outcomes might be directly affected by the type and generosity of the insurance coverage.

previously lacked insurance, hospital admissions for relatively expensive procedures increase more for previously insured groups that are more likely to have supplementary coverage after 65, reflecting the relative generosity of their combined insurance package under Medicare. Card et al. (2009) find that the number of procedures performed in hospitals and the total list charges exhibit small but statistically significant discontinuities, implying that following the *Medicare* implementation, patients over 65 received more services. They estimate a 1 percentage point drop in 7-day mortality for patients at age 65, equivalent to a 20% reduction in deaths for this severely ill patient group. The mortality gap persists for at least 9 months after admission.

Concerning Italy, Ponzo and Scoppa (2016) evaluate the impact of cost-sharing on the use of health services exploiting the fact that in the Italian health system, individuals reaching age 65 and earning low incomes are given total exemption from cost-sharing for health services consumption. The authors exploit data from the Italian Health Interview Survey, which collects several health outcomes and behavior of a representative sample of Italian residents, though the exact information on whether the individual benefits from the cost exemption, the exact the income level or the date of birth are not available. The estimates are thus performed in Fuzzy Regression Discontinuity Design in which the age threshold is used as an instrument for exemption, as it can be only observed that the probability of exemption changes discontinuously at age 65. The authors find that drug prescription and consumption, specialist visits and diagnostic checks remarkably increase with exemption, while no effect is detected on individual health measures, such as self-reported health condition or health diseases limiting the activities in daily life.

3 Institutional setting

The Italian health care system is mainly public and managed by the Regional Governments, while minimum quality standards and accessibility criteria are defined at the State level for all the Regions. For the majority of the services (such as doctor visits or hospitalization) the public system provides free access to all Italian residents. Cost-sharing is in place for health services such as specialist visits and diagnostic checks; the level of cost-sharing might range from total exemption (i.e. free access) to a coverage of part of the costs It may also vary according to individual or family income and Regional Governments can partly customize the eligibility criteria, and decide the level of cost-sharing.[5] Every individual is assigned to a so-called *family doctor*, who follows the patient and makes prescriptions for drugs and specialist checks. The family doctor is a totally free service for individuals of all age.

The vaccination season[6] in Italy usually starts in late October and finishes by the end of December, while the flu prevalence season in Western Europe is generally concentrated from late December to March (Mereckiene, 2017). Flu vaccine is a rather peculiar medical treatment. It is recommended, though not compulsory, and free of charge for all residents aged 65 or more. It is easily accessible, as it is the same family doctor who makes the injection, either in the

[5]For example, some Regional Governments impose an equal cost-sharing, that is all residents pay the same amount which only varies depending on the service chosen; other make the level of cost-sharing for the same service vary depending on family income.

[6]In the rest of the paper, for the sake of simplicity, the general term *vaccination* will always refer to flu vaccination (and not to other vaccines).

ambulatory, or at the patient' house.

Free vaccination is also offered to (but not compulsory) to some specific groups of individuals: (*i*) those with serious or chronic disease (of all age); (*ii*) individuals in nursing homes or similar facilities (of all age); (*iii*) some specific categories of workers: those working in the health care system (e.g. doctors, nurses); workers in the public sectors strategic for public safety (e.g. policeman, fireman, kindergarten teachers); workers in contact with animals. In all cases, and crucially for our empirical exercise, the flu vaccination date is registered on the General Health Register records (see section 5).[7]

4 Identification strategy

4.1 The Regression Discontinuity design

The aim of our empirical exercise is to estimate how the individual's propensity of getting the seasonal flu vaccination changes depending on whether she is offered free access to it or not, and the possible short-run consequences on health. The individual's decision to take the flu vaccine may be interpreted under the lens of an investment in health. This decision is, however, potentially endogenous with respect to several individual-level observable and unobservable characteristics. For example, it might vary with respect to actual or self-assessed health conditions, or based on previous experience of sickness, or based on the individual's concerns on the consequences of the sickness (e.g. working absences)(Schmitz and Wübker, 2011). Moreover, individuals may react to the perceived aggregate incidence of other infectious diseases (such as colds and gastro-enteric or respiratory viral diseases), which usually precede the flu pandemic and, in turn, might influence the selection into vaccination. While by exploiting administrative individual-level data, we would be able to control for several observable characteristics that define the individual's health status, some unobserved factors, which may be correlated with individual propensity to pay or not for the treatment, would remain impossible to be traced by the researcher, so that the identification of the both the effect of free access to flu vaccination on the take-up probability and the subsequent health outcomes (net of any other confounding factor) would be not possible.

To overcome these endogeneity issues, we root on the institutional setting outlined above and make use of the administrative rule which determines an exogenous assignment to the possibility of having access to the vaccine shot without paying its price.[8] In detail, we exploit the age-rule discontinuity in the eligibility criteria to get the free access to the medical treatment by means of a *Regression Discontinuity* (RD) design based on the individual's exact date of birth. For the empirical analysis, we gained access to data of the Agency of Health Protection of Milan for the 2013 flu vaccination campaign. Thus, in our case, individuals born before January

[7] Vaccination shots are sold in pharmacies, and the medical prescription is not necessary, although only family doctors or professional nurses are allowed to make the injection, in order to control for the occurrence of adverse reactions. In this case, the individual bears the full price of both the vaccine shot and the injection. The price of the vaccine shots sold in pharmacies might vary from 12 to 30 euros or even more, depending on the year, type and producer. The price for an injection might vary between 10 and 30 euros, depending on the doctor and on whether it is done at home or in the ambulatory. The share of those who choose this opportunity is negligible.

[8] Notice that non-monetary costs, such as the time spent to go to a doctor for the injection or the dis-utility from the stress and the little pain induced by the injection remain unchanged.

1st 1949 had free access to the vaccination, while those born after could have access to the vaccination only by paying for it.[9]. We exploit the 2013 campaign for data availability reasons. The assignment to the treatment is thus given by an end-of-the-year cutoff rule. In other words, the administrative rule determines an exogenous variation in the assignment to the free access to vaccination so that the RD makes it possible to estimate the effects by comparing the vaccination behavior and health outcomes of individuals who are equal under all (observed and unobserved) characteristics, except for the fact of being born in days sufficiently close to the cutoff day, but on its two different sides.

Let d_i (*running variable*) be the distance (in days) between the individual's date of birth and the cutoff point (January 1st 1949), such that it is positive for those born before the cutoff day (*eligible*), and negative for those born after (*non eligible*). The baseline RD equation takes the following form:

$$Pr(V_i) = \alpha + ET_i(\beta + f^R(d_i)) + (1 - ET_i) + f^L(d_i) + \epsilon_i \tag{1}$$

where: V_i is a dummy indicating whether individual i got vaccinated in the 2013 campaign; ET_i is a dummy defining the eligibility for the treatment status (i.e. taking value 1 for those born before January 1st 1949 and 0 otherwise); f^R and f^L are unknown smoothing functions of the running variable d_i, on the right and left hand side of the cutoff, respectively. Given that the assignment to the treatment is deterministic and based on the date of birth, which, in a sufficiently small neighbor of the cutoff date, can be considered as-good-as random, the parameter β_{RD} is an estimate of the causal effect of free vaccination on the individual's demand for it:

$$\beta_{RD} = \lim_{d_i \to 0^+} E(V_i|d_i) - \lim_{d_i \to 0^-} E(V_i|d_i) \tag{2}$$

Similarly, defining to as Y_i the other outcome variables describing the individual health status in the weeks during and immediately after the flue disease spread, and applying the same estimation framework, we would obtain a causal estimate of the the short-run effects of vaccination on individual health.

4.2 Identification assumptions and threats

4.2.1 Manipulation of the running variable and continuity of the observables at the cutoff

For the RD design to properly capture the effects under study, two main assumptions must hold. First, there should be not manipulation in the running variable. Second, there are not other relevant discontinuities at the threshold exploited for the estimation.

The first assumption is testable by searching for discontinuities in the density of the observations close to the cutoff (McCrary, 2008). Administrative records such as the date of birth are usually difficult to manipulate. However, in our specific setting manipulation could have occurred as children registration in the Birth Registers in the '40 was still a manual task in Italy. Thus, some (intended or unintended) mis-transcription might have occurred in the days

[9]Free access to flu vaccination for individuals aged 65 or more is in place since the early Nineties in Italy

immediately before or after January 1st, as parents might have had some preferences for the baby to be registered in the new year.

Concerning the second assumption, it can be shown that some relevant observable characteristics of the individuals do not display discontinuities close to the cutoff. Then, as far as all relevant observable characteristics do not display discontinuities, the unobservables can be safely assumed to behave in the same way. We will come back on these issues with the appropriate diagnostic tests in section 5, after presenting in more detail the data.

4.2.2 Externalities in the vaccination treatment

The vaccination treatment is subject to relevant externalities, which in the medical literature are specifically labeled to as *herd immunity*. Indeed, the probability of getting sick depends not only on the individual's decision of taking the flu vaccine, but also on the consideration of the level of herd immunity of the communities to whom she is usually in contact with (e.g. colleagues at work and family members). This can be conceptualized in a simple linear-in-means model à la Manski (1993):

$$y_{ig} = \beta V_i + \gamma \overline{V}_{(-ig)} + \alpha_i X_i + \alpha_g G_g + \epsilon_{ig} \tag{3}$$

where: y_i is the outcome under study, V_i is the vaccination dummy, and $\overline{\mathrm{V}}_{(-ig)}$ indicates the share of vaccinated individuals in the individual's relevant community g (excluding herself); X_i and G_g are observable individual's and group's characteristics. In such a simplified framework, the parameters β and γ would capture, respectively, the individual's and the herd effects of vaccination. The choice of the relevant individual's community is clearly endogenous, as it depends on a variety of endogenous individual's decisions (such as housing, location, fertility and labor market choices). Moreover, the parameter γ cannot be identified in a standard linear-in-means model because of a the so-called *reflection problem*, that is the probability of vaccination is jointly determined by the individual's decision and all peers' decisions, so that, in the end, it is impossible in such a model to disentangle the contribution of each single channel on the outcome of interest (Manski, 1993).

Notice that herd immunity effects are relevant for the two different outcomes under study, i.e. the decision to get the flu vaccine and the subsequent effects on health. Concerning the first, an individual may vary its propensity to vaccinate depending on what she perceives to be the level of herd immunity of the peers she is contact with, which would change the perceived probability of getting sick. Similarly, health outcomes might vary according to the levels of herd immunity in the group of relevant peers.

The RD design illustrated helps in overcoming this problem. Indeed, we can consider herd immunity as one of the potential unobservable individual-specific characteristic. Although continuity of the unobservables cannot be directly tested, this assumption can be corroborated by showing that as many as possible observable characteristic in our data do not display discontinuities at the cutoff. Moreover, a specific attention should be devoted to those observable characteristics which may proxy for some differences in the relevant community. For example, if herd immunity varies at some geographical level, we can test whether the share of individuals resident in the main city does not change in the proximity of the cutoff, or that the distance from the municipality of residence to the main city does not change. Herd immunity might also vary according to the timing of the vaccination, so that it can be tested whether, for those who got the vaccine, the vaccination date is not systematically different close to the cut-off. Finally, herd immunity might vary substantially depending on the employment status (e.g. working individuals could be in contact with more persons as compared to retired), or on the family structure (e.g. living with children might increase the risks but also the level of immunization). While we do not have variables in our administrative data to test these latter aspects, we will complement our diagnostic tests with out-of-the-sample falsification tests, by exploiting individual-level data from the *Labor Force Survey*.

A final epidemiologic argument might also help to reassure us: influenza virus are usually so easily transmitted and have a so widespread circulation that herd immunity levels should not show sharp differences in a restricted territory, as the one under study (Mereckiene, 2017).

5 Data and descriptive evidence

5.1 The General Health Register and Hospitalization records

We exploit the administrative individual-level health records of all residents aged 50+ in the territory of the Agency for Health Protection of Milan (Italy), which includes the municipality of Milan (i.e. the main city) and in the 133 municipalities of the metropolitan area of Milan[10] (about 3.2 millions inhabitants).

We exploit individual records from the *General Health Register* (GHR) which contain basic demographic characteristics (such as gender, municipality of residence and the exact date of birth linked through the Birth Register), and health information, such as whether the individual suffers from any chronic disease[11], and the date of flu vaccination (if any) in the 2013 campaign (October-December 2013). The GHR also contains information concerning the free access to health services that each individual might obtain based on chronic health conditions or low income (or both). Finally, from the GHR we retrieve the number of drugs prescriptions for respiratory diseases, which are commonly given by the doctors to combat the flu symptoms and its complications.[12]

Through an anonymous individual identifier, we merged the GHR with the *Hospitalization Records* (HR), which report all hospitalizations occurred in all hospitals in the territory of the Agency, the duration (in days) and the main cause for which the hospitalization was necessary. We focus on hospitalizations and respiratory drug prescriptions occurred between week 43-2013 and 17-2014, which is the period when the virus was circulating in Italy (based on epidemiological surveillance data from the Italian Ministry of Health, Statistics Directorate).[13] Henceforth, we will refer to these weeks as to the *observational period*, i.e. the period during which we monitor health outcomes.

[10] The metropolitan areas in Italy (10 in total) correspond to NUTS level 3.

[11] Among others, these include: cancer, diabetes, chronic renal insufficiency, heart, neurological diseases.

[12] The list includes drugs coded from R01 to R07 of the Anatomical Therapeutic Chemical (ATC) classification, i.e. nasal and throat preparations, drugs for obstructive airway diseases, cough and cold preparations, antihistamines for systemic use, other respiratory system products.

[13] These data are available on the web portal www.epicentro.it.

5.2 Variables construction

From the GHR we construct a dummy variable (V_i) which takes value 1 if an individual took the vaccination shot in the 2013 campaign, and zero for all the non-vaccinated. Then, for the analysis of the short-run effects on health, we construct a battery of variables based on the occurrence of respiratory drugs prescriptions or hospitalization events. All health outcome variables are constructed in two ways: first, considering all health events (i.e. respiratory drugs prescriptions or hospitalizations) occurred in the observational period (between week 43-2013 and 17-2014); second, considering only the health events occurred in the weeks during which the diffusion of the flu reached its peak, as defined according to the aforementioned epidemiological records (i.e. between week 3 and 9 2014).[14] Then, the health variables are constructed so to capture both the intensive margin (i.e. dummy indicating whether or not at least one event occurred) and the extensive margin (i.e. the number of respiratory drug prescriptions and the number of days of hospitalization). Finally, the hospitalization measures are calculated for all causes of hospitalization, and for the causes of hospitalization that the medical literature links flu-related diseases (Rothberg et al., 2008; Russo, 2015).

[Figure 1]

Figure 1 depicts the age profile of our main variables of interest, i.e. vaccination, drug prescription and hospitalization probability (for all causes). Vaccination probability reaches a maximum at the age of 78, with about one quarter of 78-years-old residents who took the vaccination shot. It is also apparent the sharp jump from a vaccination probability close to zero before age 65, and to about 9 percent for 65-years-old individuals. On the other hand, in this aggregate age profile figure, we do not observe sizable discontinuities for the drug prescription and hospitalization probability.

5.3 Sample selection

To perform our empirical analysis based on a RD approach, we need to focus only on the two cohorts of individuals aged 64 and 65 in the year of the vaccination campaign (2013). This corresponds to about 78,000 residents born in 1949 and 1948, respectively, that we select based on the exact date of birth extracted from the GHR and Birth Register data.[15] Moreover, we need to exclude from the sample all individuals that, independently of their age, had free access to the vaccination, which may have occurred in the three cases listed in section 3. Thanks to the precise individual-level information contained in the administrative archives, we can easily exclude from the sample all individuals who are listed in GHR as exempted from payments of sanitary treatments (including the vaccine) because of serious or chronic diseases (about 66,000), or because residing in nursing homes or similar facilities (about 334) (i.e. cases (*i*) and (*ii*)). We end up with a sample of 12,091 individuals (6,206 born in 1949 and 5,885 in 1948).

The only category of persons who have right to a free vaccination slot under age 65 that we have no way to track in our data is the specific group of workers listed in section 3, case *iii* (i.e.

[14] For all vaccinated individuals, we exclude drugs prescriptions and hospitalizations occurred in the 14 days after the vaccination, as the immunization process takes at least 14 days to produce its effects (Russo, 2015).

[15] We exclude 86 individuals who got vaccinated outside the campaign period (i.e. in January-March 2014), corresponding to less than 0.8 percent of the total number of individuals vaccinated in the 2013 campaign.

policemen, doctors, kindergarten teachers). The consistence of this group is certainly negligible in our final sample, so that our results would be unaffected from its inclusion or exclusion. Indeed, after excluding all the categories listed above, we still observe 45 individuals aged 65 who received a free vaccine (corresponding to less than 0.7 percent of those aged 65), who plausibly belong to this last category. We perform our main analysis keeping this observations in the sample, albeit this potentially introduces some fuzziness in our RD strategy: we will come back to this issue in a robustness check.

A final check we have to perform in order to define our sample is whether there is evidence of manipulation of the assignment variable (i.e. the date of birth) around the cutoff (McCrary, 2008). The individuals in our sample were born in the years immediately following World War II. At that time, it was standard to give birth at home, and in the days following the event, a parent had to go to the General Register Office of the municipality of residence to declare the birth of the baby (name, surname and date of birth). There could be an incentive for the parents not to declare as date of birth the last day of a year, and, as a matter of fact, there was no way for the Public Office to verify the declared date of birth. In our specific setting, the situation is complicated by the fact that January 1st is national holiday, so that the municipality offices were closed, and it happened to be on a Saturday in the year 1949, so that the first available working day in which a parent could have gone to declare the birth of a child born from the afternoon of December 31st 1948 to January 2nd 1949 was January 3rd.

[Figure 2]

As a first step, we perform a graphical data inspection. Indeed, Figure 2 shows an unexpectedly high number of individuals born on January 1st 1948, followed by an unexpectedly low number of individuals born on January 2nd and 3rd.[16] To provide conservative estimates of the causal effect under study, we follow the approach pioneered by and adopt a so-called RD *donut* specification (Almond et al., 2011; Almond and Doyle, 2011; Cohodes and Goodman, 2014; Barreca et al., 2016). That is, we exclude from the sample 106 individuals born exactly at the cutoff and in the two days before and after, as Figure 2 shows that manipulation of the cutoff date might have occurred until January 3rd 1949. This leaves a *donut hole* in the estimation sample, but we take these points as the first reliable points before and after the cutoff date to accurately represent the boundary. In a later section we also show that the results are robust to using different *donut* specifications (i.e. excluding only the cutoff, or the cutoff and one day before and one after) and exploiting the full sample.

[Figure 3]

As a final check before continuing in the empirical analysis, we also perform the standard diagnostic tests for the manipulation of the running variable around the cutoff. Figure 3 depicts the results of the McCrary density test (panel A) (McCrary, 2008), and a more recent test

[16] Further data inspection shows that this pattern is present for all the cohorts born in the years during or close to World War II, while it tends to reduce for younger cohorts (see Appendix Figure A.1). This seems to be in line with the interpretation of the cause of the manipulation as due to the manual filling of the birth registers. This practice slowly improved after WWII, together with an increasing number of children born in the hospitals instead that at home.

proposed by Cattaneo et al. (2017) (panel B): both tests reject the null of manipulation around the cutoff.

5.4 Descriptive statistics and falsification tests

In Table 1 we present some descriptive statistics for the outcome variables and the covariates used in the empirical analysis. These are performed on the sample of 12,091 individuals born in 1948 or 1949, obtained as described in the previous paragraph, and refer to the 2013 vaccination campaign and the 2013-14 seasonal flu. About 5 percent of the sample got the flu vaccination in the 2013 campaign, however the difference in the take-up rate between the two cohorts is sizable.

[Table 1 and Figure 4]

In Figure 4 we provide a graphical representation of the difference in the probability of getting vaccinated for individuals born close to the cutoff, which varies from about 0.05 to 0.7 percent. The discontinuity is also largely statistically significant (Calonico et al., 2014; 2015). The fact that we still observe some individuals at the left of the cutoff is compatible with the interpretation provided in the previous subsection. Moreover, the share of 64-years-old who take the vaccine is very low and it is not statistically significant different from zero close to the cutoff, which is the part of sample mainly exploited to retrieve the causal effects under study. We will come back on this issue in the robustness checks. Concerning health outcomes, about 6 percent of the sample got at least one prescription fer respiratory drugs, with an average of about 0.08 prescriptions per-capita. As expected, hospitalization, for all causes, for flu-related causes especially, is a much more rare event (0.01 and 0.004 percent, respectively).

[Figure 5]

As discussed in section 4, an important hypothesis that must be met for the estimated RD parameter to be truly capturing the effect under study, is the continuity of all observable characteristics at the threshold. We construct four main observable characteristics that describe the individuals in our sample: a gender dummy (1 if female), a dummy indicating individuals residing in the main city, a categorical variables defining the distance from the municipality of residence to the main city, the date of vaccination (only for those who got vaccinated). Testing that the vaccination date does not change significantly at the threshold is important for the assessment of the short-term health consequences of vaccination. Indeed, there is evidence that individuals with health problems tend to be vaccinated earlier (or self-select into an earlier vaccination) (Russo, 2015). In Figure 5 we show that there are not statistically significant discontinuities in the observable characteristics at the cutoff.

The data that we exploit do not contain information on some socio-economic characteristics, such as the level of education, labor force participation or family structure, which is usually difficult to retrieve from administrative archives. Labor force participation is especially important, as individuals aged 64 or 65 are close to the minimum age required by the law in 2013 to retire (i.e. age 66 and 3 months at the end of 2013), and, most importantly, health behavior might change significantly depending on the working status (Card et al., 2008; Mullahy, 1999; Schmitz and Wübker, 2011). In the Appendix B we show that, using a sample from the Italian *Labor Force Survey* (ISTAT, Italian Institute for Statistics) of individuals aged 64 and 65 and living in the same territory (Milan and its province) in the year 2013, some crucial variables constructed so to describe the individual's socio-economic characteristics do not differ significantly across the two cohorts. Finally, we point out that, to the best of our knowledge, there were not other relevant policy changes in place at the time of data collection for the discontinuity under study.

6 Results

We perform RD robust estimates following the non-parametric optimal-bandwidth selection procedures suggested by Calonico et al. (2014) and Calonico et al. (2017a). We show that the baseline results do not change if we adopt other sample selection criteria or implementing a parametric approach, and are unaffected by the type of kernel function used (triangular or uniform). Finally, we present the battery of results concerning the short-run effects on health outcomes.

6.1 Free access to flu vaccination and the take-up probability

Table 2 contains our baseline estimates of the effect of free access to flu vaccination on the take-up probability. The estimated parameter ranges between 5 and 6 percentage points, depending on the specification, and it corresponds to the variation in the individual's demand for vaccination generated by a binary change in the cost of the medical treatment (i.e. from costly to free). In other terms, our results suggest that a policy allowing for the free access to the medical treatment would increase the probability to get the flu vaccination by about 5 percentage points. The results remain considerably stable when covariates are included, and when the standard errors are clustered at level of the municipality of residence, to take into account potential correlations among individuals living close to one other. Moreover, the choice of the optimal bandwidth selection method (Mean Square Error or Coverage Error Rate) does not affect the statistically significance of the estimated parameter, though the CER delivers marginally smaller point estimates. To be more conservative in our results, and in light of the recent results presented in Calonico et al. (2017a), when otherwise not specified, we will adopt the CER method as our baseline, so that we can safely consider the estimates as, at least, a lower bound of the true effect.[17]

[Table 2]

In line with Card et al. (2008) and Ponzo and Scoppa (2016), we do find that individuals react to the price of the medical treatment. Concerning the magnitude, our estimates are in line to what found by Ponzo and Scoppa (2016). The authors estimate that the exemption from the payment (or a reduction in the cost-sharing) of medical treatments such as diagnostic exams and specialist doctor visits, in the Italian health care system, may determine an increase of about 10 to 16 percentage points (depending on the specification) in the demand of such services. In

[17] All the results are robust to calculations with the MSE method (available upon request to the authors).

the vaccination case, it sounds reasonable to find a smaller effect for the free access to medical treatment which is of prevention type.

Heterogeneous effects. - We further investigate some individual-level determinants by focusing on different sub-groups of the population under study, and thus distinguishing between males and females, and individuals living in or outside the main city (see Table 3). Concerning gender differences, we find that females display a slightly higher increase in the take-up probability with respect to males, although the point estimates are not statistically different. This different behavior might be explained by a combination of specific females' attitudes and behavior. As flu vaccination is certainly a prevention type of health investment, this behavior could be linked to a well-established evidence in the experimental literature which highlights a higher risk aversion for females as compared to males (Croson and Gneezy, 2009). Another possible explanation might encompass a generally higher female frequency of doctor visits as compared to males. Because women consult their general practitioners more frequently on average than men, the family doctor has a higher probability of seeing the woman in the campaign period and to offer her to take the free vaccination shot.

[Table 3]

We also observe that individuals living in the main city display a higher take-up probability (about 8 percent points), while the effect is not statistically significant for those living in the other municipalities of the urban area.

6.2 Robustness and specification checks

In the baseline results commented so far, the optimal bandwidth generally includes around 2800 to 1400 observations (depending on the specification), thus including individuals born in the last quarter of 1948 and in the first quarter of 1949 (or even closer to the cutoff). A possible pitfall of all discontinuity designs based on age cutoff rests in the fact that the quarter of birth has been shown to have some effects on later outcomes in life, including health (see among others: Carlsson et al. (2015)). Concerns about the endogenous selection of the season of birth should be less severe back in the '40, when contraceptive methods were not yet so diffused in Italy. However, in the parametric specifications in Table 4 we manually select the bandwidths to show that the results do not change once the regressions are performed on different samples. Results remain stable if obtained from the full sample of the two cohorts of born in 1949 and 1948, or in the subsamples defined by individuals born the last quarter of 1948 and in the first quarter of 1949, or, more crucially, those born in December 1948 and January 1949, under the assumption that the closer you are to the cutoff the more credible is the assumption on the random assignment to one side or the other of the discontinuity based on the birth date. In all cases, the linear models are those providing a better estimation performance, both in terms of precision and magnitude of the point estimates, which are very close to the baseline non-parametric estimates.

[Tables 4 and 5]

In the first 6 columns of Table 5, we show that altering the estimation sample using other *donut* specifications, or even including individuals born on the cutoff date leaves the results largely unchanged. A minor threat to our empirical strategy consists in the impossibility of excluding from the sample a specific group of public employees that benefit from free vaccination (group (*iii*) in the list of section 3). This group of individuals should be so small in our sample that, even in the case that it would be possible to exclude it from the estimation sample, this would hardly alter the estimates presented so far. However, at least in principle, it is possible to implement a Fuzzy RD approach, assuming that there is not perfect compliance with the rule, so that we can only observe a discontinuity in the *probability* of being assigned to the group of the eligible. In the last two columns of Table 5 we thus present the results from a non-parametric Fuzzy RD design, which do not detach from the baseline.

Finally, in the Appendix Table A.1 we also show that the results do not change if we adopt a uniform kernel, giving equal weighting to the observations, as opposed to the triangular kernel used in all other specifications, which assigns linear down-weighting to the same observations (Calonico et al., 2017b).

6.3 Effects on health

The medical literature has traditionally assessed the efficacy of flu vaccines through experiments in randomized control trials (RCT) or laboratory culture (Osterholm et al., 2012). Nevertheless, real-world evaluation of vaccine implementation is important for both policy makers and the scientific community for establishing the vaccine effectiveness (VE) and the impacts of free access or subsidization programs (King et al., 2015). Moreover, in the last decades RCT have become impractical, especially for common and repeatedly used vaccines such as the influenza one, because of the worldwide diffusion of national vaccination programs. Under these circumstances, only quasi-experimental designs offer the possibility to undertake rigorous VE evaluation (Lone et al., 2012).

Under this light, the combination of the quasi-experimental setting under study and the availability of individual-level administrative data might be exploited to provide robust evidence on the short-run effects on health of getting the flu vaccination. However, some caveats are in order. In this case, the age threshold rule plays as a source of exogenous variability in the vaccination take-up probability for individuals who are similar under all aspects, unless for the fact that some of them meet and some other do not the eligibility criteria based on the date of birth. Then, we assess short-run effects on health, which are concentrated in the weeks of the virus diffusion or in the peak weeks (as previously defined), while investigating longer-run consequences is not feasible in our context. Focusing on health outcome that are close to the vaccination campaign makes it more reliable their link to the individual's vaccination behavior.[18]

6.3.1 Drugs prescription and doctor visits

We present our results starting from the outcome variables that proxy for the less severe (and certainly more common) effects induced by the flu, that is the increase in drugs consumption. We

[18]This would be much less obvious if longer term health outcomes would be assessed, as other confounding factors might have occurred.

focus on respiratory drugs for two main reasons. First, this category of drugs is closely related to the health consequences of getting the flu (Russo, 2015), and thus its use can be considered a good predictor of having contracted the virus. Second, in the Italian health care system other types of common drugs usually taken to combat flu symptoms (such as antipyretics and drugs for gastro-enteric diseases) are usually sold in the pharmacies without need of the doctor's prescription, which makes impossible to track their individual-level consumption. Contrarily, respiratory drugs need the medical prescription and thus we can precisely track their use in our individual records. On top of that, if a respiratory drug has been taken, we can safely assume that a the individual has seen the doctor, who had made the medical prescription, thus providing a lower bound estimation of the number of medical visits done by an individual in the observation period, and, particularly, in the influenza peak weeks.

[Table 6]

Table 6 shows the results for all the sample (Panel A), and for the groups of individuals defined by gender (panels B and C), and by the residence in the main city (panel D) or in the neighboring municipalities (panel E). The outcome variables capture both the intensive margin (dummy for any respiratory drug prescription) and the extensive margin (number of drugs prescriptions), and refer both in entire the observational period, and in a period restricted to the peak weeks (see section 5). Considering the entire sample, we find a negative and statistically significant effect in the number of drugs prescription in the peak weeks: individuals who have free access to the vaccine have made use of 0.04 less respiratory drugs as compared to individuals with no access to free vaccine. The effect is large and economically meaningful as it corresponds to a decrease of about 80 percent with respect to mean of the outcome variable in the sample.

We also detect a lower probability of making usage of respiratory drugs (intensive margin) and on the number of drugs prescriptions (extensive margin) for females and individuals living in municipalities outside the main city and born just above the cutoff (panels B and E in Table 6). Females face a 6 to 7 percentage points lower probability of using respiratory drugs, respectively, in all the observational period and in the peak weeks. The lower probability of using respiratory drugs and, by extension, of seeing the doctor in the period of the influenza spread for females can be read in tandem to the higher propensity of getting the free vaccine established in the previous section.

6.3.2 Hospitalization

More severe health outcomes related to influenza and its complications might involve hospitalization. We distinguish between hospitalization for all causes and hospitalization for flu-related causes, albeit this latter group is so small in our sample that, due to the limited variability, the estimation might result problematic. As in the case of drugs prescription, the outcome variables capture both the intensive margin (dummy for any hospitalization event) and the extensive margin (days of hospitalization), and refer both in entire the observational period, and in a period restricted to the peak weeks.

[Table 7]

We start with hospitalization for all causes in Table 7. We find a lower probability of hospitalization (intensive margin) and a lower number of days of hospitalization (extensive margin) in the peak weeks (panel A). Individuals with free vaccine access face a 2 percentage points lower probability of incurring in any hospitalization event in the peak weeks, and, in case of hospitalization, they generally face a lower duration. Effects on the intensive margin are statistically significant and precisely estimated for the groups of the females (panel B).

[Table 8]

Similar results are found when looking at hospitalization for flu-related causes, as defined in section 5. We find a reduction in the intensive margin of about 0.01 percentage points, both in the entire period and in the peak weeks. The magnitude of the effect is not negligible as it is very close to the mean probability of hospitalization for flu-related causes observed in the baseline sample of the 1948 and 1949 cohorts (about 0.008 percent).

7 Concluding remarks and policy implications

Vaccines have saved an estimated 500 million lives around the world since Edward Jenner discovered how to prevent smallpox infection in 1796. But a successful vaccine campaign is about more than just medicine, as it should encompass engineering, economics, policy, government and even transport infrastructure.[19] Clearly, one key economic driver in the vaccination decision is its costs, both for the individual and for the health care public or private providers.

Current influenza vaccination recommendations focus on immunizing the elderly or high-risk people; however, influenza mortality and morbidity remain elevated (Schmier et al., 2008). Vaccination remain the safer and most effective measure to combat it and policymakers are considering expansion of flu vaccination recommendations to include adult population or school-age children (ages 518). Children are at risk for flu and propagate epidemic spread. However, flu vaccination is generally not mandatory and other policy leverages to increase its take-up should be found.

In this paper, we focus on one potential candidate, i.e. the price. By exploiting newly available and rich administrative individual-level records, we estimate that giving free access to the treatment would increase take-up by almost 5 percentage points, with peaks of about 8 percentage points in the female population. Moreover, we detect some positive effects on subsequent health: vaccination decreases the use of respiratory drugs and the probability of hospitalization, though the effects seem concentrated in certain population groups (females and residents outside the main urban area). We obtain our results by exploiting a RD design and, based on the exact date of birth, by comparing the vaccination probability or the health outcomes of individuals born just above or below the implicit eligibility threshold settled by the rule of the free access to vaccination to the 65-year-olds. We also show that our estimates are robust to

[19] See, among others, the intervention by Professor Jeffrey Almond speech at *Great Health Challenges of the 21st Century* seminar series at Oxford University (Oxford Martin School) in December 2014 (available at http://www.oxfordmartin.ox.ac.uk/videos/view/459). Professor Jeffrey Almond is an Oxford Martin Visiting Fellow with the Oxford Martin Programme on Vaccines.

several falsification tests and do not depend on the non-parametric or parametric specification chosen.

Some possible caveats on the policy application and external validity of our results should be discussed. While the causal effect estimated would be easily applicable to asses a policy change targeted to a part of population close to the age threshold used (e.g. extending free vaccination to 60+ or 50+), a potential limit in the interpretation of our results lies in the fact that it would be less advisable to extend it to other targeted populations (e.g. children). However, increasing by about 5 percentage points the take-up of some adult population for whom the usual take-up is around 1-2 percent could be extremely powerful in terms of increasing herd immunity for the other parts of the population. Our results are also useful from a policy perspective in order to understand the causes that limit the voluntary vaccination of those who have free access to the service, as its monetary cost does not seem to be the reason for reaching the full (or targeted) coverage. In order to increase the take-up of those who already benefit of the free access to the treatment the policy maker should look for other leverages, excluding the price, which may perhaps involve the doctors or the public information campaigns.

While no medical evidence is available on the negative effects of the vaccination, cost-effectiveness considerations of the extension of the free eligibility to other adult cohorts (such as the 50+ or the 60+) should take into account that the limited effects that we find on the subsequent health consequences have to be considered as a lower bound of the plausible true effects on health. Indeed, it is not possible to measure important milder consequences of flu, such as the consumption of antipyretics and gastro-enteric drugs, or the illness-induced absenteeism at the workplace or at school (Ward, 2014; Schmier et al., 2008).

References

Adda, J. (2016). Economic activity and the spread of viral diseases: Evidence from high frequency data. *The Quarterly Journal of Economics 131*(2), 891–941.

Almond, D. (2006). Is the 1918 influenza pandemic over? long-term effects of in utero influenza exposure in the post-1940 u.s. population. *Journal of Political Economy 114*(4), 672–712.

Almond, D. and J. J. Doyle (2011). After midnight: A regression discontinuity design in length of postpartum hospital stays. *American Economic Journal: Economic Policy 3*(3), 1–34.

Almond, D., J. J. Doyle, Jr., A. E. Kowalski, and H. Williams (2011). The role of hospital heterogeneity in measuring marginal returns to medical care: A reply to barreca, guldi, lindo, and waddell. *The Quarterly Journal of Economics 126*(4), 2125–2131.

Barreca, A. I., J. M. Lindo, and G. R. Waddell (2016, 1). Heaping-Induced Bias In Regression-Discontinuity Designs. *Economic Inquiry 54*(1), 268–293.

Calonico, S., M. D. Cattaneo, and M. Farrell (2017a). On the effects of bias estimation on coverage accuracy in nonparametric inference. *Journal of the American Statistical Association*. forthcoming.

Calonico, S., M. D. Cattaneo, and M. Farrell (2017b). rdrobust: Software for regression-discontinuity designs. *The Stata Journal 17*(2), 372–404.

Calonico, S., M. D. Cattaneo, and R. Titiunik (2014). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica 82*(6), 2295–2326.

Calonico, S., M. D. Cattaneo, and R. Titiunik (2015). Optimal data-driven regression discontinuity plots. *Journal of the American Statistical Association 110*(512), 1753–1769.

Card, D., C. Dobkin, and N. Maestas (2008, December). The impact of nearly universal insurance coverage on health care utilization: Evidence from medicare. *American Economic Review 98*(5), 2242–58.

Card, D., C. Dobkin, and N. Maestas (2009). Does medicare save lives? *The Quarterly Journal of Economics 124*(2), 597–636.

Carlsson, M., G. B. Dahl, B. Öckert, and D.-O. Rooth (2015). The Effect of Schooling on Cognitive Skills. *The Review of Economics and Statistics 97*(3), 533–547.

Cattaneo, M. D., M. Jansson, and X. Ma (2017). Simple Local Polynomial Density Estimators. Working paper, University of Michigan.

CDC (2016). Estimated Influenza Illnesses, Medical Visits, Hospitalizations, and Deaths Averted by Vaccination in the United States. Technical report, Centers for Disease Control and Prevention (CDC).

Cohodes, S. R. and J. S. Goodman (2014). Merit aid, college quality, and college completion: Massachusetts' adams scholarship as an in-kind subsidy. *American Economic Journal: Applied Economics 6*(4), 251–85.

Croson, R. and U. Gneezy (2009). Gender differences in preferences. *Journal of Economic Literature 47*(2), 448–74.

Daniels, N. A., T. Juarbe, M. Rangel-Lugo, G. Moreno-John, , and E. J. PÃ©rez-Stable (2004). Focus group interviews on racial and ethnic attitudes regarding adult vaccinations. *Journal of the National Medical Association 96*(11), 1455â1461.

ECDPC (2017). Seasonal Influenza. Fact scheet, European Centre for Disease Prevention and Control (ECDPC).

Grossman, M. (1972). On the concept of health capital and the demand for health. *Journal of Political Economy 80*(2), 223–255.

Hayward, A. C., E. B. Fragaszy, A. Bermingham, L. Wang, A. Copas, W. J. Edmunds, N. Ferguson, N. Goonetilleke, G. Harvey, J. Kovar, M. S. C. Lim, A. McMichael, E. R. C. Millett, J. S. Nguyen-Van-Tam, I. Nazareth, R. Pebody, F. Tabassum, J. M. Watson, F. B. Wurie, A. M. Johnson, and M. Zambon (2014). Comparative community burden and severity of seasonal and pandemic influenza: results of the flu watch cohort study. *The Lancet Respiratory Medicine 2*(6), 445 – 454.

Kelly, E. (2011). The scourge of asian flu: In utero exposure to pandemic influenza and the development of a cohort of british children. *The Journal of Human Resources 46*(4), 669–694.

King, C., J. Beard, A. Crampin, A. Costello, C. Mwansambo, N. Cunliffe, R. Heyderman, N. French, and N. Bar-Zeev (2015). Methodological challenges in measuring vaccine effectiveness using population cohorts in low resource settings. *Vaccine 33*(38), 4748 – 4755.

Lone, N. I., C. Simpson, K. Kavanagh, C. Robertson, J. McMenamin, L. Ritchie, and A. Sheikh (2012). Seasonal influenza vaccine effectiveness in the community (sive): protocol for a cohort study exploiting a unique national linked data set. *BMJ Open 2*(2).

Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *The Review of Economic Studies 60*(3), 531–542.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics 142*(2), 698–714.

Mereckiene, J. (2017). Seasonal influenza vaccination in Europe. Technical report, European Centre for Disease Prevention and Control (ECDC.

Mullahy, J. (1999). It'll only hurt a second? microeconomic determinants of who gets flu shots. *Health Economics 8*(1), 9–24.

Nagata, J., I. Hernandez-Ramos, A. Sivasankara-Kurup, D. Albrecht, C. Vivas-Torrealba, and C. Franco-Paredes (2011). Social determinants of health and seasonal influenza vaccination in adults â65 years: a systematic review of qualitative and quantitative data. *BMC Public Health 13*(338), 2â25.

Osterholm, M., N. Kelley, A. Sommer, and E. Belongia (2012, 1). Efficacy and effectiveness of influenza vaccines: A systematic review and meta-analysis. *The Lancet Infectious Diseases 12*(1), 36–44.

Ponzo, M. and V. Scoppa (2016, February). Cost-Sharing and Use of Health Services in Italy: Evidence from a Fuzzy Regression Discontinuity Design. IZA Discussion Papers 9772, Institute for the Study of Labor (IZA).

Rothberg, M. B., S. D. Haessler, and R. B. Brown (2008). Complications of viral influenza. *The American Journal of Medicine 121*(4), 258 – 264.

Russo, A. (2015). A population study to evaluate the efficacy of influenza vaccination, 2014-2015. *Epidemiologia Prevenzione 39*(4), 234–242.

Schmier, J., S. Li, J. C. K. Jr., K. Nichol, and P. J. Mahadevia (2008). Benefits and costs of immunizing children against influenza at school: An economic analysis based on a large-cluster controlled clinical trial. *Health Affairs 27*(2), 96 – 104.

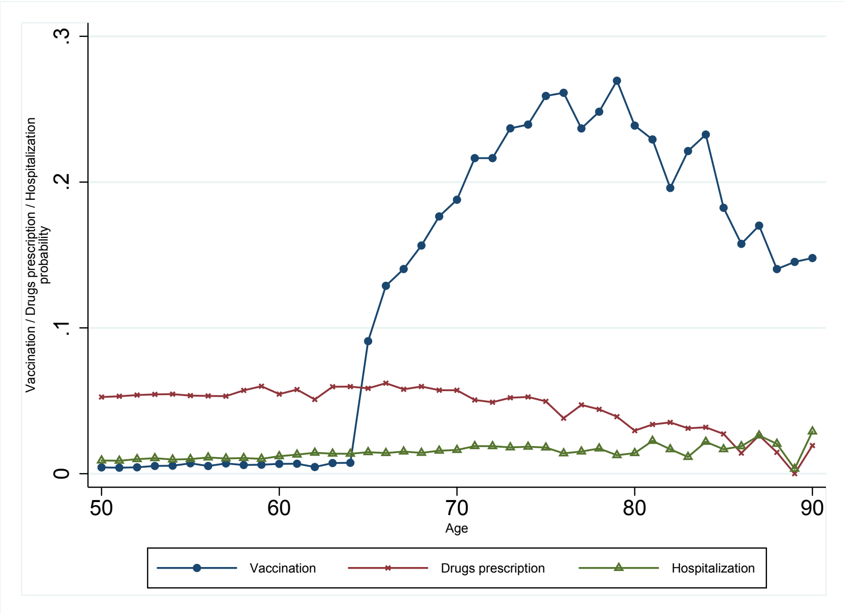
Schmitz, H. and A. Wübker (2011). What determines influenza vaccination take-up of elderly europeans? *Health Economics 20*(11), 1281–1297.

Ward, C. J. (2014). Influenza vaccination campaigns: Is an ounce of prevention worth a pound of cure? *American Economic Journal: Applied Economics 6*(1), 38–72.

WHO (2017). Seasonal Influenza. Fact scheet, World Health Organization (WHO).

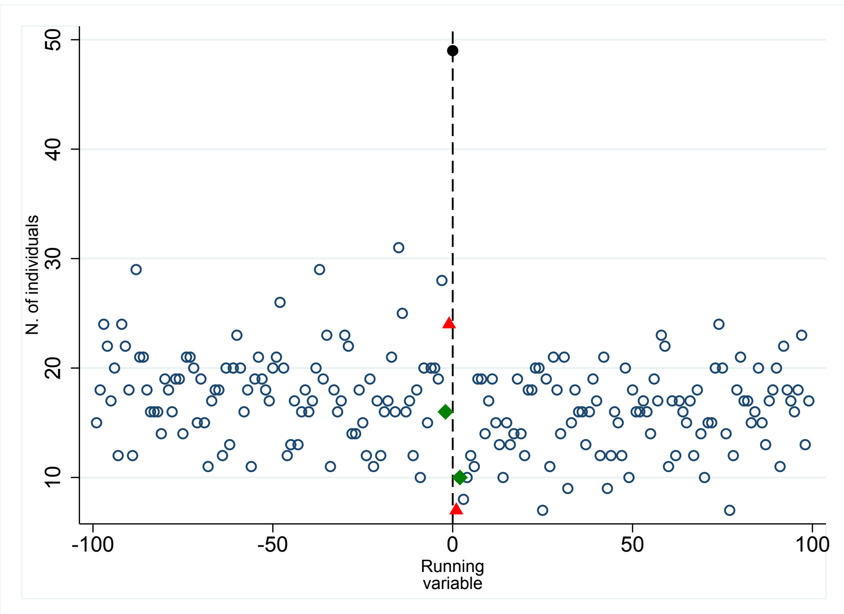
Figures

Figure 1
Vaccination, drugs prescription and hospitalization: the age profile



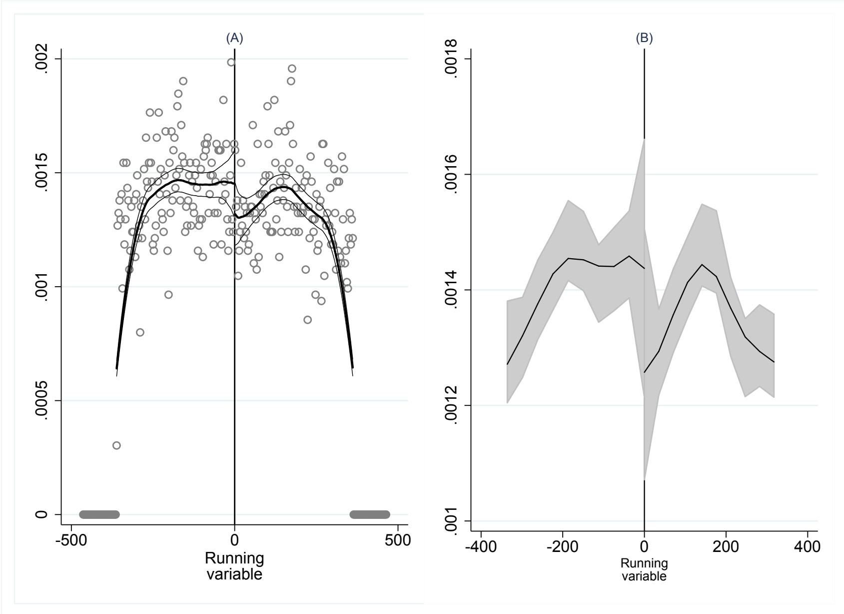
Notes: the lines show the average value by age of the dummy variables indicating whether an individual got vaccinated (*Vaccination*), got at least one drug prescription for respiratory diseases in the observational period (*Drugs prescription*), got hospitalized at least one day in the observational period (*Hospitalization*) (see also Table 1 for variables definitions). The sample is truncated at age 90 due to the very small number of individuals olther than 90. **Source**: based on General Health Register, ATS-Milan.

Figure 2
Number of individuals born per calendar day close to the cutoff



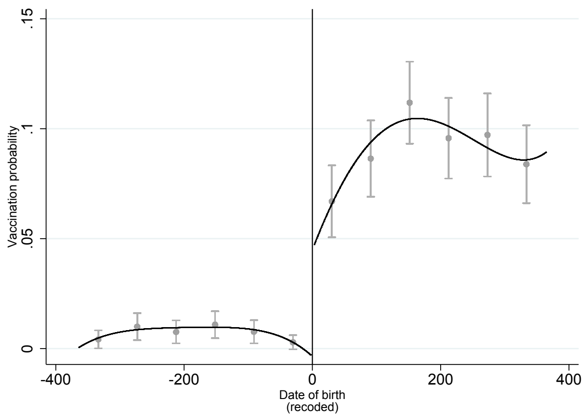
Notes: the figure shows on the vertical axis the number of individuals born in each calendar day (running variable) close to the cutoff date, indicated on the horizontal axis (recoded so that the value 0 corresponds to the cutoff date of January 1st 1949); the black circle indicates individuals born on January 1st 1949, the triangles individuals born on January 2nd 1949 and December 31st 1948, the diamonds individuals born on January 3rd 1949 and December 30th 1948. **Source**: based on General Health Register, ATS-Milan.

Figure 3
Tests for the manipulation of the running variable



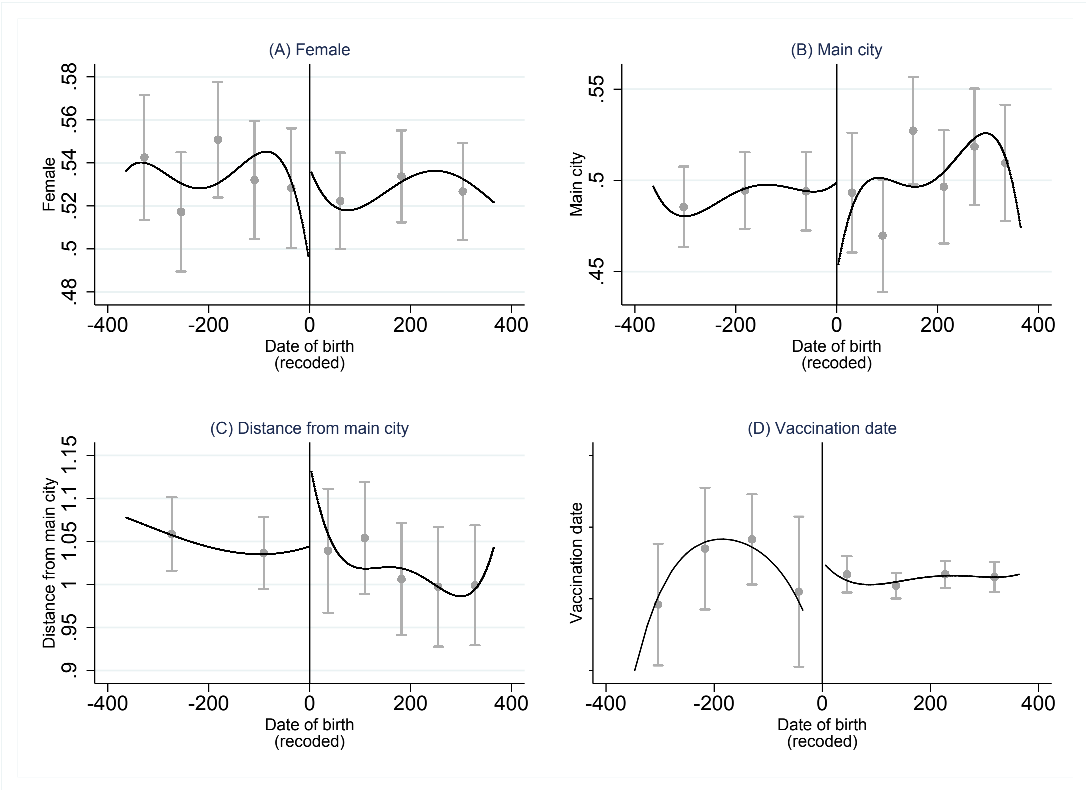
Notes: the figures depict standard tests for the density of the running variable around the cutoff. The running variable is indicated on the horizontal axis (recoded so that the value 0 corresponds to the cutoff date of January 1st 1949, i.e. function d_i as specified for eq. 1). Figure in panel (A) depicts the McCrary test (McCrary, 2008), while figure in panel (B) depicts the Cattaneo et al. (2017) test. The tests are performed on the *baseline donut sample*, which excludes observations at the cutoff, and in the two days before and after. **Source**: based on General Health Register, ATS-Milan and McCrary (2008); Cattaneo et al. (2017).

Figure 4
Discontinuity of the vaccination probability



Notes: the figure shows IMSE (Integrated Mean Square Error) optimal evenly spaced RD plot with 95% CIs (Calonico et al., 2014; 2015). The horizontal axis indicates the running variable (recoded so that the value 0 corresponds to the cutoff date of January 1st 1949). **Source**: based on General Health Register, ATS-Milan.

Figure 5
Falsification test: continuity of the observable characteristics



Notes: the figures show IMSE optimal evenly spaced RD plot with 95% CIs (Calonico et al., 2014; 2015). The horizontal axis indicates the running variable (recoded so that the value 0 corresponds to the cutoff date of January 1st 1949). See Table 1 for definitions of variables in Panels (A) and (C); in Panel (B) *Main city* is a dummy equal to 1 for individuals residing in the main city (Milan); in Panel (D) *Vaccination date* indicates, for the subsample of those who got vaccinated, the date of vaccination. **Source**: based on General Health Register, ATS-Milan.

Tables

Table 1
Descriptive statistics

	mean	sd	max	min	N
Dependent variables:					
Vaccination (dummy)	0.048	0.214	1	0	12091
Respiratory drugs (dummy)	0.059	0.236	1	0	12091
Respiratory drugs (dummy, peak weeks)	0.020	0.140	1	0	12091
Respiratory drugs (number of prescriptions)	0.077	0.348	6	0	12091
Respiratory drugs (number of prescriptions, peak weeks)	0.055	0.274	5	0	12091
Hospitalization (dummy)	0.014	0.118	1	0	12091
Hospitalization (dummy, flu-related causes)	0.000	0.009	1	0	12091
Hospitalization (dummy, peak weeks)	0.003	0.059	1	0	12091
Hospitalization (dummy, flu-related causes in peak weeks)	0.000	0.009	1	0	12091
Hospitalization (days)	0.066	0.829	26	0	12091
Hospitalization (days, flu-related causes)	0.001	0.136	26	0	12091
Hospitalization (days, peak weeks)	0.017	0.470	15	0	12091
Hospitalization (days, flu-related causes in peak weeks)	0.001	0.136	15	0	12091
Covariates:					
Female	0.531	0.499	1	0	12091
Resident in Milan city	0.497	0.500	1	0	12091
Distance from Milan (categorical)	1.034	1.197	3	0	12091

Notes: descriptive statistics performed on the entire sample of individuals born in 1948 or 1949. *Respiratory drugs (dummy)* indicates whether the individual got at least one doctor prescription for respiratory drugs; *Respiratory drugs (number)* indicates the number of doctor prescriptions for respiratory drugs; *Hospitalization (dummy)* indicates whether the individual got hospitalized at least once; *Hospitalization (days)* indicates the number of days of hospitalization; the hospitalization variables are calculated for all causes of hospitalization and for flu-related causes only. Based on the official epidemiological reports, the health outcomes variables are calculated both in the overall observational period (which corresponds to the weeks of diffusion of the influenza virus from 2013 week 42 to 2014 week 17), and in the weeks in which the influenza reached its peak (from week 3 to week 9 2014). *Female* is a dummy equal to 1 for females; *Distance from Milan* is a categorical variable for the distance from the municipality of residence to the main city, where 0 corresponds to the main city (i.e. resident in Milan), 1 to the neighboring municipalities (so-called *first circle* municipalities), 2 to the municipalities neighboring the neighboring municipalities (so-called *second circle* municipalities), and 3 to the more distant municipalities. **Source**: based on General Health Register, ATS-Milan, and Epicentro (Ministry of Health).

Table 2
Baseline results: free access to flu vaccination and take-up probability

	(1)	(2)	(3)	(4)	(5)	(6)
RD estimate	0.059***	0.060***	0.056***	0.046**	0.047**	0.050***
	(0.018)	(0.018)	(0.017)	(0.020)	(0.020)	(0.016)
Kernel Type	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular
BW Type	MSE	MSE	MSE	CER	CER	CER
N.Obs.: total	12091	12091	12091	12091	12091	12091
N.Obs.: effective left	1417	1368	1350	877	840	1019
N.Obs.: effective right	1244	1189	1171	741	707	885
Order Loc. Poly. (p)	1	1	1	1	1	1
Order Bias (q)	2	2	2	2	2	2
Covariates included		✓	✓		✓	✓
Clustered std. error			✓			✓

Notes: MSE and CER indicate, respectively, the Mean Square Error and the Coverage Error Rate optimal bandwidth selector for the RD treatment effect estimator, robust estimates are presented (Calonico et al., 2014; 2017a); the list of covariates includes a dummy for female and the distance from the main city (see Table 1 for definitions). Significance level: *** p<0.01, ** p<0.05, * p<0.1. **Source**: based on General Health Register, ATS-Milan.

Table 3
Heterogeneous effects by gender and residence in the main city

	(1) Females	(2) Males	(3) Main city	(4) Other municipalities
RD estimate	0.060***	0.053**	0.081**	0.037
	(0.023)	(0.021)	(0.035)	(0.025)
Kernel Type	Triangular	Triangular	Triangular	Triangular
BW Type	CER	CER	CER	CER
N.Obs.: total	6421	5670	6009	6082
N.Obs.: effective left	517	636	483	674
N.Obs.: effective right	453	574	404	579
Order Loc. Poly. (p)	1	1	1	1
Order Bias (q)	2	2	2	2

Notes: CER indicates the Coverage Error Rate optimal bandwidth selector for the RD treatment effect estimator, robust RD estimates are reported (Calonico et al., 2014; 2017a); standard errors clustered at the municipality level. Significance level: *** p<0.01, ** p<0.05, * p<0.1.**Source**: based on General Health Register, ATS-Milan.

Table 4
Robustness check: parametric RD estimates

	(1)	(2)	(3)	(4)	(5)	(6)
	Dep. Var.: Vaccination Probability					
	Linear		Quadratic		Cubic	
Panel A: full sample						
Treated	0.078***	0.078***	0.078***	0.057***	0.065***	0.043**
	(0.010)	(0.010)	(0.010)	(0.013)	(0.012)	(0.019)
R2	0.04	0.04	0.04	0.04	0.04	0.04
N	12091	12091	12091	12091	12091	12091
Panel B: Jan-Mar 1949 and Dec-Oct1948						
Treated	0.055***	0.053***	0.054***	0.064***	0.065***	0.026
	(0.015)	(0.015)	(0.015)	(0.023)	(0.021)	(0.017)
R2	0.04	0.04	0.04	0.04	0.04	0.04
N	2949	2949	2949	2949	2949	2949
Panel C: Jan-Feb 1949 and Dec-Nov1948						
Treated	0.061***	0.060***	0.061***	0.035	0.044**	0.021
	(0.016)	(0.016)	(0.016)	(0.023)	(0.021)	(0.023)
R2	0.03	0.03	0.03	0.04	0.04	0.04
N	1912	1912	1912	1912	1912	1912
Panel D: Jan 1949 and Dec 1948						
Treated	0.052**	0.051**	0.051**	-0.000	0.025	-0.078**
	(0.021)	(0.022)	(0.022)	(0.024)	(0.015)	(0.038)
R2	0.04	0.04	0.04	0.04	0.04	0.04
N	939	939	939	939	939	939
Covariates included	✓	✓	✓	✓	✓	✓
Interacted model		✓		✓		✓

Notes: *Treated* is a dummy equal to 1 for all eligible individuals (i.e. the 1948 cohort); the list of the covariates includes a dummy for female, and fixed effects for the distance to the main city variable (see Table 1 for definitions). The *Linear model* includes the running variable linearly; *Quadratic model* adds the square of the running variable; *Cubic model* adds the cubic of the running variable. The *Interacted models* interact each running variable (and its powers) with the *Treated* dummy. Estimates in Panel (A) are performed on the full sample of all individuals born in 1948-1949; estimates in Panel (B), (C) and (D) are performed on the restricted samples of individuals born between October 1948 and March 1949, November 1948 and February 1949, December 1948 and January 1949, respectively. Robust standard errors clustered at the municipality level. Significance level: *** p<0.01, ** p<0.05, * p<0.1. **Source**: based on General Health Register, ATS-Milan.

Table 5
Robustness checks: alternative specifications

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
RD estimate	0.045***	0.047***	0.049***	0.042***	0.041***	0.042***	0.052***	0.050***
	(0.010)	(0.015)	(0.015)	(0.010)	(0.014)	(0.014)	(0.018)	(0.016)
Kernel Type	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular
BW Type	MSE	MSE	MSE	CER	CER	CER	MSE	CER
N.Obs.: total	12197	12148	12117	12197	12148	12117	12091	12091
N.Obs.: effective left	1959	1528	1449	1473	1165	1111	1019	1019
N.Obs.: effective right	1806	1327	1269	1325	990	952	885	885
Order Loc. Poly. (p)	1	1	1	1	1	1	1	1
Order Bias (q)	2	2	2	2	2	2	2	2
Specifications:								
(*i*) No Donut	✓			✓				
(*ii*) Donut Zero		✓			✓			
(*iii*) Donut One			✓			✓		
(*iv*) Fuzzy							✓	✓

Notes: MSE and CER indicate, respectively, the Mean Square Error and the Coverage Error Rate optimal bandwidth selector for the RD treatment effect estimator, robust RD estimates are reported (Calonico et al., 2014; 2017a); all specifications include covariates (see Table 2 for definitions). Specification (*i*) No Donut is performed including all individuals (also those born between December 30th and January 3rd); specification (*ii*) Donut Zero is performed excluding only individuals born on January 1st (i.e. individuals for whom the running variable is equal to zero); specification (*iii*) Donut One is performed excluding only individuals born between December 31st and January 2nd (i.e. individuals for whom the running variable is equal to zero or $+/-1$). Standard errors clustered at the municipality level; specification (*iv*) Fuzzy performs a fuzzy RDD estimate in the baseline sample, in which the treatment variable is a dummy for the 1948 cohort: estimates in columns (7) and (8) are, respectively, with and without the covariates included. The specification Significance level: *** p<0.01, ** p<0.05, * p<0.1. **Source**: based on General Health Register, ATS-Milan.

Table 6
Effects on health: drugs prescriptions for respiratory diseases

	(1)	(2)	(3)	(4)
	Dep. Var.: Drugs prescription			
	All period		*Peak weeks*	
	Dummy	Prescriptions No.	Dummy	Prescriptions No.
Panel A: all sample				
RD estimate	-0.020	0.014	-0.020	-0.043*
	(0.031)	(0.019)	(0.037)	(0.025)
N.Obs.: total	12091	12091	12091	12091
N.Obs.: effective left	1205	1125	1401	1626
N.Obs.: effective right	1044	973	1227	1429
Panel B: females				
RD estimate	-0.073*	0.004	-0.060*	-0.065**
	(0.041)	(0.018)	(0.036)	(0.028)
N.Obs.: total	6421	6421	6421	6421
N.Obs.: effective left	484	859	821	933
N.Obs.: effective right	409	724	682	785
Panel C: males				
RD estimate	0.021	0.028	0.022	-0.018
	(0.032)	(0.020)	(0.045)	(0.034)
N.Obs.: total	5670	5670	5670	5670
N.Obs.: effective left	731	650	670	757
N.Obs.: effective right	665	602	627	695
Panel D: main city				
RD estimate	0.033	0.037	0.069	0.020
	(0.039)	(0.034)	(0.055)	(0.038)
N.Obs.: total	6009	6009	6009	6009
N.Obs.: effective left	665	554	587	660
N.Obs.: effective right	587	482	514	580
Panel E: other municipalities				
RD estimate	-0.066**	0.002	-0.079**	-0.079**
	(0.034)	(0.018)	(0.036)	(0.031)
N.Obs.: total	6082	6082	6082	6082
N.Obs.: effective left	576	814	693	701
N.Obs.: effective right	500	701	589	600
Kernel Type	Triangular	Triangular	Triangular	Triangular
BW Type	CER	CER	CER	CER
Order Loc. Poly. (p)	1	1	1	1
Order Bias (q)	2	2	2	2

Notes: for the definition of the dependent variables see Table 1; CER indicates the Coverage Error Rate optimal bandwidth selector for the RD treatment effect estimator, robust RD estimates are reported (Calonico et al., 2014; 2017a); specification in Panel (A) includes the list of covariates as specified for Table 2; standard errors clustered at the municipality level. Significance level: *** p<0.01, ** p<0.05, * p<0.1. **Source**: based on General Health Register, ATS-Milan.

Table 7
Effects on health: hospitalization for all causes

	(1)	(2)	(3)	(4)
	Dep. Var.: Hospitalization			
	All period		*Peak weeks*	
	Dummy	Days No.	Dummy	Days No.
Panel A: all sample				
RD estimate	-0.020	-0.041	-0.021**	-0.055*
	(0.012)	(0.049)	(0.010)	(0.033)
N.Obs.: total	12091	12091	12091	12091
N.Obs.: effective left	1488	954	1107	515
N.Obs.: effective right	1310	804	958	441
Panel B: females				
RD estimate	-0.028*	-0.064	-0.020**	-0.103
	(0.014)	(0.077)	(0.009)	(0.065)
N.Obs.: total	6421	6421	6421	6421
N.Obs.: effective left	787	494	767	217
N.Obs.: effective right	659	420	642	193
Panel C: males				
RD estimate	-0.013	-0.020	-0.005	-0.013
	(0.018)	(0.050)	(0.010)	(0.018)
N.Obs.: total	5670	5670	5670	5670
N.Obs.: effective left	794	425	451	482
N.Obs.: effective right	731	369	381	410
Panel D: main city				
RD estimate	-0.012	0.040	-0.012	-0.072
	(0.017)	(0.151)	(0.010)	(0.054)
N.Obs.: total	6009	6009	6009	6009
N.Obs.: effective left	737	498	700	390
N.Obs.: effective right	655	429	625	323
Panel E: other municipalities				
RD estimate	-0.026	-0.098	-0.016	-0.045
	(0.021)	(0.061)	(0.016)	(0.034)
N.Obs.: total	6082	6082	6082	6082
N.Obs.: effective left	814	528	721	653
N.Obs.: effective right	701	454	628	570
Kernel Type	Triangular	Triangular	Triangular	Triangular
BW Type	CER	CER	CER	CER
Order Loc. Poly. (p)	1	1	1	1
Order Bias (q)	2	2	2	2

Notes: for the definition of the dependent variables see Table 1; CER indicates the Coverage Error Rate optimal bandwidth selector for the RD treatment effect estimator, robust RD estimates are reported (Calonico et al., 2014; 2017a); specification in Panel (A) includes the list of covariates as specified for Table 2; standard errors clustered at the municipality level. Significance level: *** p<0.01, ** p<0.05, * p<0.1. **Source**: based on General Health Register, ATS-Milan.

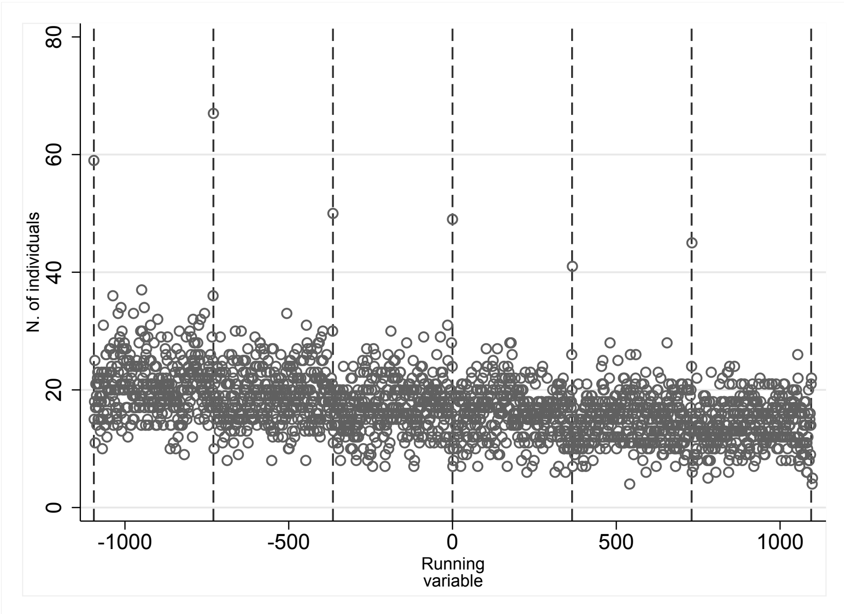
Table 8
Effects on health: hospitalization for flu-related causes

	(1)	(2)	(3)	(4)
	Dep. Var.: Hospitalization for flu-related causes			
	All period		*Peak weeks*	
	Dummy	Days No.	Dummy	Days No.
RD estimate	-0.0001*	1.6956	-0.0001*	-0.0010*
	(0.0000)	(1.1782)	(0.0000)	(0.0006)
Kernel Type	Triangular	Triangular	Triangular	Triangular
BW Type	CER	CER	CER	CER
N.Obs.: total	12091	171	12091	12091
N.Obs.: effective left	380	39	380	380
N.Obs.: effective right	310	44	310	310
Order Loc. Poly. (p)	1	1	1	1
Order Bias (q)	2	2	2	2

Notes: for the definition of the dependent variables see Table 1; CER indicates the Coverage Error Rate optimal bandwidth selector for the RD treatment effect estimator, robust RD estimates are reported (Calonico et al., 2014; 2017a); the specification includes the list of covariates as specified for Table 2; standard errors clustered at the municipality level. Significance level: *** p<0.01, ** p<0.05, * p<0.1. **Source**: based on General Health Register, ATS-Milan.

A Appendix: Additional Figures and Tables

Figure A.1
Descriptive evidence: manipulation of the running variable for younger and older cohorts



Notes: the figure shows on the vertical axis the number of individuals born in each calendar day close to the cutoff dates (i.e. January 1st of each calendar year) indicated on the horizontal axis with dashed lines, and recoded so that the value 0 corresponds to the cutoff date of January 1st 1949. Younger cohorts thus appear on the right hand side of the value 0, while older cohorts on the left hand side. **Source**: based on General Health Register, ATS-Milan.

Table A.1
Additional results: uniform kernel specification

	(1)	(2)	(3)	(4)	(5)	(6)
Type of RD estimate:						
Conventional	0.057***	0.057***	0.066***	0.061***	0.061***	0.055***
	(0.013)	(0.013)	(0.016)	(0.017)	(0.017)	(0.016)
Bias-corrected	0.055***	0.055***	0.062***	0.060***	0.061***	0.052***
	(0.013)	(0.013)	(0.016)	(0.017)	(0.017)	(0.016)
Robust	0.055***	0.055***	0.062***	0.060***	0.061***	0.052***
	(0.015)	(0.015)	(0.016)	(0.017)	(0.017)	(0.016)
Kernel Type	Uniform	Uniform	Uniform	Uniform	Uniform	Uniform
BW Type	MSE	MSE	MSE	CER	CER	CER
N.Obs.: total	12091	12091	12091	12091	12091	12091
N.Obs.: effective left	1727	1727	1296	1062	1062	965
N.Obs.: effective right	1516	1516	1138	913	913	823
Order Loc. Poly. (p)	1	1	1	1	1	1
Order Bias (q)	2	2	2	2	2	2
Covariates included		✓	✓		✓	✓
Clustered std. error			✓			✓

Notes: MSE and CER indicate, respectively, the Mean Square Error and the Coverage Error Rate optimal bandwidth selector for the RD treatment effect estimator (Calonico et al., 2014; 2017a); for the the list of covariates included see Table **??**; standard errors are clustered by municipality of residence. Significance level: *** p<0.01, ** p<0.05, * p<0.1. **Source**: based on General Health Register, ATS-Milan.

B Appendix: Out-of-sample falsification tests

As discussed in section 4, one important assumption for our identification strategy consists in the absence of discontinuities at the cutoff in all observable and unobservable characteristics, but the outcomes of interest. While continuity at the cutoff of the unobservable characteristics cannot be tested, in Figure 5 we show that the condition is met for the observable characteristics that we have in the data. However, the administrative records that we exploit might have the drawback of not containing additional information on some individual characteristics, such as marital status, education level or labor force participation, which previous literature has shown to be correlated to the decision to take the flu vaccine (Nagata et al., 2011). Labor force participation is especially important, as individuals aged 64 or 65 are close to the minimum age required by the law in 2013 to retire (i.e. age 66 and 3 months at the end of 2013), and, most importantly, health behavior might change significantly depending on the working status (Card et al., 2008; Mullahy, 1999; Schmitz and Wübker, 2011). To complement our falsification tests we thus perform some additional analysis using a sample of individuals taken from the Italian *Labor Force Survey* (LFS).

From the quarterly cross-sectional releases of the LFS micro-data in the year 2013, we select all individuals aged 64 and 65 and resident in the province of Milan, a territory that corresponds to the one of the main analysis. We end up with a sample of 483 individuals. From the LFS, we can construct some dummies indicating whether an individual is retired, married, has at least a high school education level, lives alone, or lives with at least one son, or works in the education or health sectors. Of course, LFS do not have any information on the health status nor they contain the precise date of birth (i.e. the running variable of the RD analysis). We can only distinguish individuals aged 64 in 2013, for whom the dummy *Eligible* takes value 0, and those aged 65 in 2013, for whom the dummy *Eligible* takes value 1.

Table B.1
Mean differences: regression results

	Dependent variables:					
	Retired		Married		High education	
Eligible	0.012	0.012	0.010	0.002	-0.006	-0.001
	(0.038)	(0.038)	(0.045)	(0.045)	(0.047)	(0.047)
R2	0.00	0.00	0.00	0.01	0.00	0.01
N	483	483	483	483	483	483
Mean	0.816		0.720		0.443	
	Living alone		Living with son(s)		Education or health sectors	
Eligible	-0.011	-0.001	0.048	0.043	-0.017	-0.017
	(0.042)	(0.042)	(0.041)	(0.042)	(0.014)	(0.013)
R2	0.00	0.01	0.00	0.01	0.00	0.01
N	483	483	483	483	483	483
Mean	0.207		0.251		0.023	
Quarter FE		✓		✓		✓

Notes: *Eligible* takes value 1 for all individuals aged 65 in 2013 and 0 for those aged 64; *Retired*, *Married*,*High education*, *Living alone*, *Living with son(s)*, *Education or health sectors* are a dummy variables indicating, respectively, individuals who are retired, married, have at least a high school education level (ISCED 3 or more), live alone, or live with at least one son, work in the education, defense or health sectors. Quarter fixed effects (FE) account for differences in the four quarterly cross-sectional releases of the 2013 LFS used. Robust standard errors, survey weights applied. Significance level: *** p<0.01, ** p<0.05, * p<0.1. **Source**: based on Labor Force Survey (ISTAT, Italian Institute for Statistics).

To test whether the above characteristics are statistically different for the subsamples of the *Eligible* and *Not Eligible*, in Table B.1 we perform a simple regression analysis running

OLS specifications in which the dependent variables are the observed characteristics, and the parameter of interest is the coefficient estimate for the variable *Eligible*. The results show that there are not statistically significant mean differences in the characteristics across the two groups, even when accounting for unobserved differences in the quarterly waves in the LFS (as captured by the quarter fixed effects). Similar results hold if a simple test on the mean difference of these characteristics across the two groups is performed (Table B.2). Albeit out-of-sample falsification tests should be taken with proper caution, these results further corroborate the hypothesis that individuals aged 65 or 64 in 2013 were not different in some important aspects, such as marital status, education, family structure and labor force participation, and working in sectors (education and health) which might have access to free vaccination independently from the age rule.

Table B.2
Mean differences: test statistics

	Retired		Married		High education	
	Eligible	Not Eligible	Eligible	Not Eligible	Eligible	Not Eligible
Mean	0.817	0.815	0.725	0.715	0.434	0.453
N	251	232	251	232	251	232
Difference	0.002		0.01		-0.019	
	Living alone		**Living with son(s)**		**Education or health sectors**	
	Eligible	Not Eligible	Eligible	Not Eligible	Eligible	Not Eligible
Mean	0.817	0.815	0.725	0.715	0.016	0.030
N	251	232	251	232	251	232
Difference	0.002		0.01		0.014	

Notes: *Eligible* takes value 1 for all individuals aged 65 in 2013 and 0 for those aged 64; *Retired*, *Married*,*High education*, *Living alone*, *Living with son(s)*, *Education or health sectors* are a dummy variables indicating, respectively, individuals who are retired, married, have at least a high school education level (ISCED 3 or more), live alone, or live with at least one son, work in the education, defense or health sectors. Significance levels if the mean difference between the two groups is statistically significant: *** p<0.01, ** p<0.05, * p<0.1. **Source**: based on Labor Force Survey (ISTAT).