IDB WORKING PAPER SERIES Nº IDB-WP-1180

Increasing the Take-up of Public Health Services

An Experiment on Nudges and Digital Tools in Uruguay

Ana Cuesta
Lucia Delgado
Sebastián Gallegos
Benjamin Roseth
Mario Sánchez

Inter-American Development Bank
Institutions for Development Sector

July 2021

Increasing the Take-up of Public Health Services

An Experiment on Nudges and Digital Tools in Uruguay

Ana Cuesta, Inter-American Development Bank
Lucia Delgado, School of Medicine of the University of the Republic and Honorary Commission on the Fight Against Cancer
Sebastián Gallegos, Inter-American Development Bank and UAI Business School
Benjamin Roseth, Inter-American Development Bank
Mario Sánchez, Inter-American Development Bank

July 2021

Abstract[*]

In this paper, we test whether promoting digital government tools increases the take-up of an important public health prevention service: cervical cancer screening. We implemented an at-scale field experiment in Uruguay, randomly encouraging women to make medical appointments with a digital application or reminding them to do it as usual at their local clinic. Using administrative records, we found that the digital application nearly doubled attendance of a screening appointment compared to reminders and tripled the rate compared to a pure control group (3.2 percentage point increase over a base of 1.9 percent). Survey data suggests that the impacts of the intervention were mostly mediated by reduced transaction costs. Our results highlight the potential of investing in digital government to improve the take-up of public services.

JEL codes: I12, I15, I18, D90, D91

Keywords: information, transaction costs, public services, text messages, digital government, health outcomes, cancer screening, women, behavioral economics, field experiment

[*] The authors thank seminar participants at the UC San Diego Pacific Conference for Development Economics, The University of Chicago-UCEMA Joint Initiative for Latin American Experimental Economics, the Inter-American Development Bank, the PUC-Chile School of Government, and Applied Microeconomics Peru. We also thank Phil Keefer, Carlos Scartascini, and Florencia López-Boo for their contributions to study design. For their invaluable implementation support, we thank Analice Berón, Adriana Brescia, Susana Dornel, Pablo Orefice, Pablo de Santa Ana, Juan Jones, Verónica Nin, Ignacio Parietti Rosario Berterretche, Gabriela Piriz, Ana Noble, Mauricio Cuello, and, especially, Inés Caraballo and Agustín Gründel.

This experiment is listed in the American Economic Association Registry for Randomized Controlled Trials: #AEARCTR-0004716. This research was made possible through funding from the Inter-American Development Bank and support from the Uruguayan Ministry of Health and Digital Government and Information and Knowledge Society Agency. The opinions expressed in this paper are those of the authors and do not necessarily reflect the views of the Inter-American Development Bank, its Board of Directors, or the countries they represent.

1. Introduction

The low take-up of social benefits is a puzzling empirical regularity (Blanco and Vargas, 2014; Muralidharan et al., 2020) and increasing it could benefit billions of people (World Bank, 2003). Some of the main barriers to adoption documented in the literature are lack of information and different forms of transaction costs (Currie, 2004; Finkelstein and Notowidigdo, 2019).

Recent advances in technology have the potential to address these obstacles (Goldfarb and Tucker, 2019). Governments are developing the capacity to digitize, store, and use administrative data sources (Figlio et al., 2017). Combined with the widespread use of mobile phones and internet access, governments might now be better equipped to inform citizens and reduce bureaucracy and transaction costs. Yet beyond the recent work by Muralidharan et al. (2016, 2020) on India, there is still limited empirical evidence on whether combining government and new technologies ends up improving the delivery of social benefits. This paper contributes to filling this gap by presenting evidence from a large-scale experimental evaluation of state capacity investment on the take-up of an important public health service: cervical cancer screening.

Cervical cancer is one of most common causes of female mortality worldwide and one of the most treatable forms of cancer given early detection and management (WHO, 2020). Nevertheless, many women do not take the standard screening tests, such as the Papanicolaou test (Pap smear), even when offered cost-free by a health provider. This is a critical issue, particularly in developing countries, where absence of preventive health behaviors can lead to increased mortality rates (Dupas, 2011).

We implemented our intervention in Uruguay, partnering with the Ministry of Health and the Agency for e-Government and Information Society (*Agencia de Gobierno Electrónico y Sociedad de Información*, or AGESIC) to invest in digital government capacities in two steps. The first step was to consolidate information that was digitized but disperse, bringing together data from a register of health users, an electronic clinical history of patients, and the medical appointments system.

Second, we worked with our partners to design and implement an application to schedule Pap smear appointments online. The application, compatible with smart phones and desktop computers, allowed users to make appointments choosing from a set of available dates and times at their local clinic. Before this innovation, most women made their appointments in person at their local clinics.

We then conducted a field experiment at scale randomly encouraging women to make Pap smear appointments. About 80 percent of the female population of public health users in Montevideo, the capital city, participated (N=47,600). We randomized participating women into four treatment arms and one pure control group. Those in the treatment arms received messages with either an encouragement to schedule appointments as usual with their local clinic or to schedule online using a link to the digital application. Because loss- and gain-framed messages might differentially influence health behaviors (Rivers et al., 2005), we paired each of the encouragement messages with either benefit- or risk-framed information about the PAP smear, thus generating the four treatment groups. Women in the pure control group received no messages nor additional information.

We tested effects on two main behavioral outcomes and a host of secondary measures. Our main outcome indicators were whether women scheduled and attended appointments over a 12-week period. We computed intention-to-treat (ITT) estimates and two sets of local average treatment effects (LATEs), considering both the successful delivery and receipt of the SMS messages.

Using administrative records from the national health authorities, we found that the screening appointment rate among women encouraged to use the digital application was double that of the reminded group (7.9 percent vs. 4.0 percent) and triple that of the pure control group (7.9 percent vs. 2.7 percent). There was no meaningful difference between benefit- and risk-framed messages.

These effects on appointment scheduling are informative about participant's intentions to follow the advice embedded in our treatment arms. Because it is well documented that intentions often differ from actions (e.g., Beshears, Milkman, and Schwartzstein, 2016), we used further administrative records to complement the previous intention outcome and estimate the effects on women's actions—whether they attended their appointments.

Our findings show that the digital tool intervention also tripled attendance rates, with an effect of 3.2 percentage points over the control group (5.1 percent versus 1.9 percent). Compared with women randomized to the reminders group, the digital application increased attendance rates by 1.5 times (5.1 percent versus 3.3 percent). As with scheduling, we found a precise zero difference between benefit- and risk-framed messages.

These results on attending medical appointments suggest that components of our intervention helped women overcome some obstacles to getting a Pap smear. To explore in depth which barriers were lowered, we needed additional data on secondary outcomes, which the administrative records lacked. Thus, we supplemented the administrative data with a survey of a subsample of just under 2,500 participating women, evenly distributed across the four treatment arms and the control group.

We collected data on a set of plausible mechanisms through which treatment effects might operate, including women's knowledge about the Pap smear and its costs, their beliefs about the importance of testing and the risks of not taking the test, and preferences for public services, among other factors.

The survey results show that women were already well informed about the Pap smear and its importance. Therefore, we hypothesize that part of our treatment effects worked through the salience of the encouragements. Rather than providing women new information, our messages appear to have brought the Pap smear to their minds. Our findings also indicate that the higher impacts are working through reduced transaction costs, particularly for those who received the link to make appointments online. Treated women reported that they found it easier to make appointments and were less likely to say that they would make an appointment in person at a health center. Our study complements the literature aimed at increasing preventive health services with behavioral insights (e.g., Kremer, Rao, and Schilbach, 2019), which has delivered mixed results.

Our effects were sizable compared to most of the successfully implemented behaviorally informed interventions in other countries for similar health behaviors. For example, Milkman et al. (2012) found a 1.0 percentage point increase over a control mean of 6.2 percent in the fraction of people making and sticking to appointments to be screened for colon cancer in the United States, where the treatment group received mailings, including a sticky note that prompted the recipient to write down the appointment date and the name of the doctor conducting the procedure. Using a similar intervention, Milkman et al. (2011) found an increase of 1.5 to 4.0 percentage points from using implementation intention prompts on influenza vaccination rates in the United States.

Altmann and Traxler (2014) found that reminder messages (including or not additional information on the benefits of prevention) had a large and significant effect on patients that were due to schedule a dental check-up. Huf et al. (2020) found that SMS reminders improved cervical screening participation by 4.0 percentage points over a much higher control mean of 34.0 percent in the United Kingdom. This study included women in a similar age range as our study (30–64 years vs. 30–70 years) and also had different types of SMS messages: primary care physician endorsed SMS, SMS with a total or proportionate social norm, or an SMS with a gain- or loss-framed message. For Latin America, Busso, Cristia, and Humpage (2015) found that personal reminders increased demand for vaccination by 2.2 percentage points (over a control mean of 6.0 percent) in Guatemala, and Beuermann et al. (2020) found that the likelihood of prenatal care attendance increased by 9.0 percentage points (19.6 percent with respect to the control mean) when pregnant women received prenatal visit reminders in Peru.

Other similar interventions found no effects. Buchmueller and Goldzahl (2018) tested four behavioral interventions (differentiated content and presentation of an invitation letter) in France with a large-scale randomized controlled trial and found no effects on mammography use. Bronchetti, Huffman, and Magenheim (2015) reported that lower-cost nudges did not affect overall vaccine take-up for a sample of undergraduate students in Philadelphia.

Our study contributes to the literature in two important ways. First, we provide novel evidence experimenting at-scale in a region of the developing world where there is scarce evidence of scaled-up interventions (Muralidharan and Niehaus, 2017). The 47,600 participants in our experiment represented more than 80 percent of the population of women who used the public health provider in Montevideo.[1]

Randomizing at scale helped us address (i) the inability to detect effects in small-scale interventions, as was the case for some of the studies cited above, and (ii) the challenges associated with implementing a program on a larger organizational scale. What makes our study more informative is that we surveyed a subsample of participating women, which provided an estimate of the differential effects compared to the full sample. We found larger estimates on the surveyed subsample, consistent with respondents being more connected to cell phones and the internet, and more engaged with their health, among other characteristics. If we had implemented the intervention on this subsample only, we could have mistakenly expected to have higher effects once scaling.

Second, this study contributes to the literature on using new technologies in the public sector. Digital technology has reduced the cost of storage, computation, and transmission of data (Goldfarb and Tucker, 2019), and governments are developing the capacity to use it to develop and implement policy. Our results complement a recent wave of studies assessing different ways to use technology to improve state capacity in developing countries (Bossuroy, Delavallade, and Pons, 2019; Callen et al., 2020; Muralidharan et al., 2016, 2020).

Our findings suggest that promoting digital government tools is important and likely to generate the most meaningful impacts on public service delivery compared to purely behavioral interventions delivered through information or reminders, which the literature has shown to have mixed results. The prospect of investing in state digital capacity might be even more important in developing countries, where low service delivery uptake is often associated with weak institutional capacity.

[1] The population of Montevideo, the capital, represents approximately 40 percent of Uruguay's population of 3.5 million.

The remainder of this paper is organized as follows. Section 2 provides a brief description of the setting in which the intervention took place. Sections 3 and 4 describe our experimental design and the data, respectively. Section 5 outlines the empirical strategy and Section 6 discusses our results. We present our main conclusions in Section 7.

2. Background

Uruguay is a small country in South America with about 3.5 million inhabitants, of whom approximately 1.4 million live in Montevideo. With purchasing power parity–adjusted gross domestic product per capita of US$22,500 (World Bank, 2019), Uruguay ranks among the most developed Latin American countries, yet still lags well behind the OECD average (US$46,500) and the United States (US$65,300). Therefore, the country faces the development challenges of a middle-income country.

2.1. Health System

Uruguay provides health coverage for the entire population (Arbulo, Castelao, Oreggioni, et al., 2015) through National Health Insurance (*Seguro Nacional de Salud*) financed by a National Health Fund (*Fondo Nacional de Salud*, or FONASA).[2] People can enroll either with private or public health care providers.[3] The public health care system is operated by the Administration of State Health Services (*Administración de los Servicios de Salud del Estado*, or ASSE). The lower quintiles of household income use the ASSE, and about 84 percent of ASSE users are in the bottom two quintiles (Artagaveytia and Toledo, 2018).

2.2. Cervical Cancer Challenges

One area of challenges in health services is related to cervical cancer testing. This type of cancer is the fourth leading cause of cancer-related deaths in the world (Ferlay, Steliarova-Foucher, Lortet-Tieulent, et al., 2013) and the fifth leading cause among women in Uruguay (Barrios and Garau, 2017). Cervical cancer is one of the most successfully treatable forms of cancer if detected early and managed effectively (WHO, 2014). Yet, at the time of this study, in Montevideo, 60 percent of the women aged 30 to 70 who used the public healthcare system (ASSE) had not had cervical cancer screening in the previous three years. Nationally, 61 percent of the women who die from cervical cancer receive care from ASSE (*Comisión honoraria de lucha contra el cáncer*, 2016).

[2] Uruguay's health system is the National Integrated Health System (*Sistema Nacional Integrado de Salud*, or SNIS).
[3] About a 56 percent of the population is enrolled with private health care providers, 34 percent with the public sector provider (ASSE), 6.5 percent with the military or the police, and 3 percent with special private insurance providers (Ministerio de Salud de Uruguay, 2018).

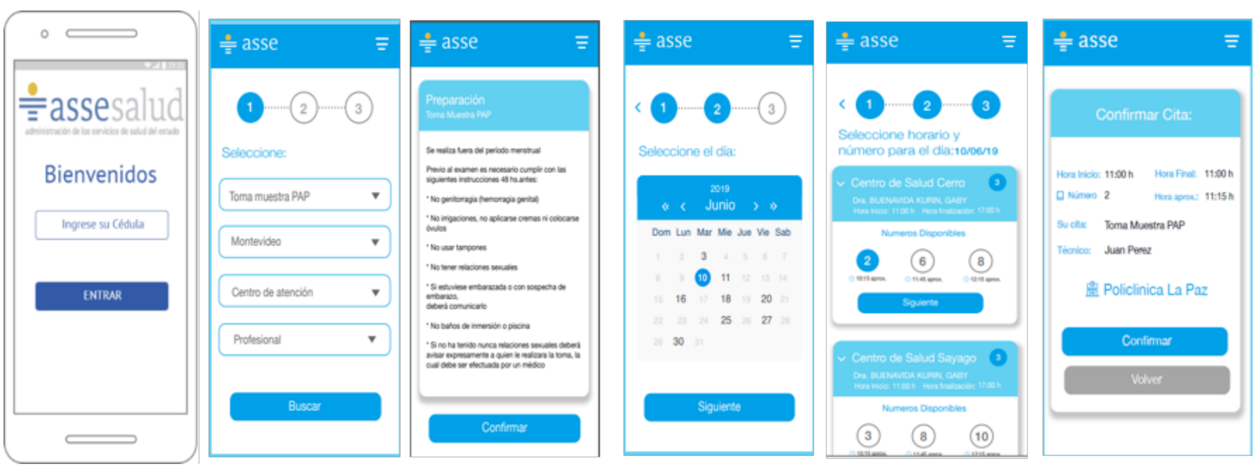
Progress and Obstacles Regarding Cervical Cancer Screening

Since 2013, in Uruguay, PAP smears have been cost-free at ASSE clinics and, since 2000, legally women have been allowed to take one paid day off from their jobs to get the test. Women report being familiar with the exam and its importance (Rodriguez et al., 2015). The most common barriers to service take-up are delays in waiting rooms and transaction costs related to appointment scheduling (Benia and Tellechea, 2000; Rodriguez et al., 2015).

Investing in State Capacity

Aiming to ease these barriers to screening, we partnered with the Government of Uruguay to design and implement an online tool to schedule Pap smear appointments with ASSE clinics.[4] The online tool is compatible with smart phones, desktop computers, and different operating systems. Figure 1 illustrates its interface when using a smartphone. Users can make their own appointments using their national identification number and selecting their health center of preference and date for the exam. The application displays consolidated information coming from different sources: the digitized register of health users in Uruguay, the electronic clinical history of patients, and the medical appointments system. It also includes a system to remind users of their appointments.

Figure 1. Online Appointment Interface Using a Smartphone



Note: Users log in with their national identification number and then follow three steps. In Step 1, they choose their local clinic and medical professional. In Step 2, they choose an available date. In Step 3, they choose an available time slot. Finally, they confirm the appointment.

[4] We designed this tool specifically for this project. Given the promising results we document in this paper, ASSE intends to broaden the use of the tool for all types of appointments for all its members.

Conditions to Promote Cervical Cancer Screening

We leveraged three key factors to promote cervical cancer screening using the online tool.

- The significant government experience in providing digital public services and their commitment to implementing the online tool. The Government of Uruguay is explicitly committed to the digitalization of public services (United Nations, 2018). Accordingly, AGESIC has taken important steps toward digitizing public services and internal administrative processes (Roseth et al., 2018).[5]
- ASSE's good reputation among its enrollees (Berterreche and Sollazo, 2012). Comparatively, Uruguay's health system ranks among the best rated in terms of perception in Latin American countries (OECD/World Bank, 2020).
- Uruguayan citizens report having access to and being frequent users of the internet. About 86 percent of adults use the internet daily, and 9 of 10 connect using a cell phone (AGESIC, 2019).

This backdrop of longstanding commitment to providing e-government services, the good reputation of the health authority, and high rates of internet usage provided an ideal setting for our intervention based on text messages and internet links.

3. Experimental Design

This section presents our experimental design. We describe our intervention, eligibility, randomization procedures, the implementation and timeline, and our outcome measures.

3.1. Intervention

Our intervention consisted of messages reminding women to schedule a cervical screening appointment either using the online tool (with the internet link provided) or by contacting their local clinic. The messages also contained either benefit or risk information. All of the messages were designed to increase scheduling and attendance of cervical cancer screening. We randomized the intervention in four treatment arms and one pure control group. Women in the control group continued with the health center's standard procedures regarding appointment scheduling and did not receive any communications or additional information regarding PAP smears.

[5] Uruguay is well established as the regional leader in digital government and is recognized as a leading country on the world stage. In the 2018 United Nations ranking of e-government development, Uruguay placed #34 in the world and #1 in Latin America. It is also institutionally strong compared to its regional peers: it ranks first in Latin America and the Caribbean on the Democracy Index (EIU, https://www.eiu.com/n/campaigns/democracy-index-2020), the Corruption Perceptions Index (Transparency International, https://www.transparency.org/en/cpi/2020/index/nzl), and the Rule of Law Index (World Justice Project, https://worldjusticeproject.org/rule-of-law-index/).

Table 1 classifies our four treatment arms.[6] Women in treatment arms 1 and 2 (first row) were sent informational messages with a link to use the online scheduling tool. We designed this component as an effort to decrease transaction costs in appointment scheduling. Women that made appointments through the online channel were also sent one reminder message 48 hours prior to their appointment. Women in treatment arms 3 and 4 (second row) received messages recommending they schedule appointments at their health center as usual. We label this component status quo.

We also provided benefit or risk information in the messages because it is well documented that under-adoption of preventive health behaviors might be related to underestimation of the benefits and risks (Kremer, Rao, and Schilbach, 2019).[7] Messages for women in column 1 (treatment arms 1 and 3) emphasized the benefits and importance of taking the Pap smear. Messages for women in column 2 (treatment arms 2 and 4) highlighted the importance of the Pap smear considering the risks associated with cervical cancer.

Table 1. Treatment Arms

	Information on:	
Appointment encouraged to be made:	**Benefits**	**Risks**
Online (lower transaction cost)	Arm 1	Arm 2
At health center (status quo)	Arm 3	Arm 4

Notes: Table 1 shows the content of the behaviorally informed messages in our treatment arms. We sent information on benefits taking and risk of taking the PAP smeal (in columns), and encouraged woment to make appointments using the online system or at their local health center (in rows).

All messages were personalized with the first name of the recipient and signed by the ASSE, the Ministry of Health, and the municipal government (*Intendencia Municipal de Montevideo*):

- **Treatment Arm 1:** *Hi [Name]. The Pap smear can prevent cervical cancer. It is important to get screened. Here is a link to schedule an appointment.*

[6] Note that this is not a factorial (or cross-cutting) design as defined in Duflo, Glennerster, and Kremer (2007) and discussed in Muralidharan, Romero, and Wüthrich (2020). In a cross-cutting design, the interaction of two interventions, A and B, would result in four groups: no interventions (*pure control*), A only, B only, and A and B together. In our case, each of the four treatment arms and the pure control group are individually defined ex ante and are not the result of interacting two interventions. This table demonstrates a cross-cutting design.

		Intervention A (information)	
		No	Yes
Intervention B (encouragement)	No	Control	Arm 1
	Yes	Arm 2	Arm 3

[7] Applied to PAP smears, Rivers et al. (2005) showed that when the exam was posed as a detection activity, loss-framed messages were most effective in encouraging uptake, and when posed as a prevention activity, gain-framed messages were most effective. This intervention, however, was in person (participants viewed informative videos in the context of non-PAP medical appointments), leaving a knowledge gap regarding the relative effectiveness of gain- or loss-framed messages.

- **Treatment Arm 2:** *Hi [Name]. Every three days a woman dies of cervical cancer. Not getting a Pap smear could cost you your life. Here is a link to schedule an appointment.*
- **Treatment Arm 3:** *Hi [Name]. The Pap smear can prevent cervical cancer. It is important to get screened. Schedule your test at your health center.*
- **Treatment Arm 4:** *Hi [Name]. Every three days a woman dies of cervical cancer. Not taking a Pap smear could cost you your life. Schedule your test at your health center.*

3.2. Eligibility

Our eligible sample comprised women who used ASSE, were between 30 to 70 years old, and lived in Montevideo. We considered women who used the public healthcare and were subsidized by the National Health Fund because of our partnership with ASSE. We focused on women aged 30 to 70 because the prevalence of cervical cancer increases at 30 and declines after 70 (Garau, Musetti, Alonso, et al., 2019).[8] The intervention was designed to take place in Montevideo because the city has the country's highest internet penetration rates (Rivoir and Landinelli, 2017) and internet connectivity was a condition for the intervention to work.

In May 2019, ASSE used its administrative records to prepare a dataset of the eligible women and shared it with us to conduct the randomization. At that time, ASSE had approximately 1.4 million members, 35 percent of whom were enrolled with the National Health Fund. Half (250,000) were women and approximately 76,000 resided in Montevideo. From those 76,000 women, 58,800 were between 30 and 70 years old and had at least one registered cell phone number[9] and thus were eligible for our study. The dataset included birth date and date of last Pap smear if it was done within five years. In accordance with Uruguayan data protection legislation, we were not provided phone numbers or any identifiable information. All of our analysis was done with anonymous identifiers generated by ASSE.

3.3. Randomization Procedures

As documented in our preregistered analysis plan,[10] we estimated we would need a sample of 5,000 women per treatment arm to achieve a minimal detectable effect size of 1 percentage point on attendance rates, at 5 percent statistical significance and 80 percent power.

---

[8] In addition, professional organizations (American Cancer Society, 2019; Saslow et al., 2012; Smith et al., 2016) recommend more frequent cervical cancer screening after age 30 and some also suggest that routine exams should be discontinued for older women (above 65), but academic articles have argued against a stopping age (e.g., White et al., 2017).

[9] To implement the text message treatments, we used the phone numbers gathered by ASSE. Numbers stay up to date because the information is checked every time a woman attends a health care facility.

[10] We registered our experiment in the American Economic Association Registry for Randomized Controlled Trials—#AEARCTR-0004716.

We ended up randomizing 5,700 women to each of our four treatment arms, totaling 22,800 women treated, and 24,800 women to a pure control group (Table 2, column 1).[11] We randomized the remaining eligible women (11,200) to a contingency sample to be drawn on in the event of a high rate of failed text message deliveries; we did not need to use any of this sample.

Table 2. Number of Eligible Women by Randomization Status

		Time Since Last PAP Smear	
Status	**(1) All**	**(2) On Time (< 5 years)**	**(3) Overdue (≥ 5 years)**
Treatment	22,800	11,400	11,400
Arm 1	5,700	2,850	2,850
Arm 2	5,700	2,850	2,850
Arm 3	5,700	2,850	2,850
Arm 4	5,700	2,850	2,850
Control	24,800	12,400	12,400
All Participating Women	47,600	23,800	23,800
Contingency	11,200	6,500	4,700
All Eligible Women	58,800	30,300	28,500

Notes: Table 2 shows the number of eligible women randomized to treatment (and each treatment arm), control and refreshment sample.

Following the recommended practice (e.g., Duflo et al., 2007; Athey and Imbens, 2017), we stratified our randomization using the available covariates in our administrative data, which were the time passed since last PAP (five years or more) and age (for which we used four brackets of 10 years each). This procedure ensures more precise inference compared to randomization without stratification.[12] In Table 2, columns 2 and 3 decompose column 1 by the time since the last Pap smear. From the 58,800 eligible women, 30,300 had received a PAP smear within the previous five years (the "on-time" group, Column 2) and 28,500 had not (the "overdue" group, Column 3).

WHO guidelines (WHO, 2013) state that the cervical cancer screenings should be conducted at least every five years, which is why we used the ASSE data to classify women who

[11] At the time of the randomization, we discussed with ASSE the possibility of having an additional treatment arm with 1,000 observations, which would have resulted in having 23,800 women in all treatment groups and 23,800 in the control group. Women in that treatment arm would have been offered the opportunity to conduct an HPV test at home. However, we could not implement this version due to logistic restrictions and operational costs. We decided to place the 1,000 women in the control group, which is why we ended up having 22,800 in treatment arms and 24,800 in the control group.

[12] As Duflo et al. (2007) explained, stratifying is more efficient than controlling for these variables ex post because it ensures an equal proportion of treated and untreated units within each strata, minimizing the variance of our estimates.

had a PAP smear less than five years ago as on time and those who had one five years ago or more as overdue.[13]

3.4. *Implementation and Timeline*

A text messaging firm sent our messages to the different treatment arms, recording which type of message was sent to whom, how many, and on which date. Messages were sent on a rolling basis each Tuesday afternoon,[14] starting on November 18, 2019, until February 11, 2020.

Messages were sent to women in the on time group during November and December 2019 and then to the overdue group during January and February 2020. Each woman could receive at most four messages over four weeks. If during that period the recipient scheduled an appointment, she stopped receiving messages.[15]

It is important to highlight that the intervention ended before the COVID-19 emergency started in the country, where the first cases were reported on March 13, 2020. Also, Uruguay was one of the countries with less confirmed cases of COVID-19 worldwide throughout 2020. According to the WHO Dashboard, Uruguay only started to record more than 500 confirmed cases by the end of 2020. The provision of public health services was not interrupted by the pandemic.

3.5. *Outcome Measures*

Main Outcome Measures

Our two main pre-specified outcomes measured intentions and actions. We measured whether women scheduled appointments for a Pap smear (intentions) and whether they attended those appointments (actions). These two variables were collected through ASSE's administrative records from the medical appointments system and measured between November 18, 2019 and March 3, 2020.[16]

Scheduling an appointment informs us about a woman's intention to follow the advice embedded in our treatment arms. If recipients already know that having a Pap smear is important and they are fully aware of the benefits of having (and the risks of not having) the exam, then we

[13] In Uruguay, the official guidelines of the Ministry of Health recommend women without specific risk factors between 21 and 69 years of age to have a PAP smear every three years after having received two consecutive negative annual PAPs. Nevertheless, we set the threshold at five years to separate women far surpassing the recommended time frame from those close to or within the recommended frequency, thus establishing greater clarity that the overdue group was in fact as such.

[14] Most of the messages were sent on Tuesday afternoons, except two dates in December (Mondays the 23th and 30th). Because of the high volume of SMS to be sent, the system sent messages in batches without recording a specific time of delivery for each individual message. However, we know that messages were sent between 1pm and 7pm during each delivery day.

[15] The online system also sent automatic reminders to all women who scheduled appointments. The reminder was sent 48 hours prior to the scheduled appointment time.

[16] We had information on whether each woman scheduled, attended, cancelled, or did not show up for her appointment.

should see no differences on this outcome between treatment arms 3 and 4 and the control group. In addition, if the transaction costs associated with making an appointment are sufficiently low, then we would not expect to find effects when comparing treatment arms 1 and 2 to the control group.

Because intentions do not always translate into actions, we complement the previous outcome with whether women attended their appointments. Finding effects associated to this outcome would tell us that some of the constraints on behavior were eased by our intervention. Results for the different treatment arms would tell which of the components of our intervention helped women overcome the obstacles to getting the Pap smear.

Secondary Outcomes

We also studied secondary outcomes to explore mechanisms through which treatment effects might operate. We surveyed a subgroup of participants on their knowledge about the Pap smear, their beliefs about the importance of testing, and costs and obstacles they face when making and attending medical appointments. Finding effects on these secondary outcomes compared with the control group, or between treatment arms, would suggest whether lack of information or transaction costs constrained behavior absent the intervention. We also asked interviewers to classify the respondents' patience and attitude during the survey, in the spirit of DellaVigna and Paserman (2005) and Cadena and Keys (2015).

4. Data

Our study used administrative records from ASSE supplemented with survey data. Table 3 shows descriptive statistics using administrative information for the full sample of women participating in the experiment. The variable "years since last PAP" was available for half of the sample—women who had had a PAP smear within the previous five years. The full sample of women studied comprised 47,600 women signed up with the public healthcare system. The average age was 49 years, with a range of 30 to 70 years. The WHO suggests that women older than 30 should have a Pap smear every five years. In our sample, 60 percent of the women had not had the exam in the previous three years. Consistent with our stratified randomization (see Section 3), half of the women had not had the test in the previous five years. For the on time group of women, on average their last PAP appointment 1.8 years previously. As detailed in the previous section, a total of 24,800 (52 percent) of the women studied were in the control group and 22,800 (48 percent) in the treatment group. Each treatment arm had 5,700 observations (12 percent of the total sample).

Table 3. Descriptive Statistics, Full Sample

Variable (% of sample unless otherwise noted)	**(1) Observations**	**(2) Mean**	**(3) Std. Dev.**	**(4) Min**	**(5) Max**
Age (years)	47,600	48.56	11.34	30.41	69.90
No PAP in previous 3 years	47,600	0.60	0.49	0.00	1.00
No PAP in previous 5 years	47,600	0.50	0.50	0.00	1.00
Years since last PAP appointment	23,800	1.79	1.33	0.01	5.00
Control Group	47,600	0.52	0.50	0.00	1.00
Treatment:					
Benef. Info. + Lower Trans. Cost	47,600	0.12	0.32	0.00	1.00
Risks Info. + Lower Trans. Cost	47,600	0.12	0.32	0.00	1.00
Benef. Info. + Status Quo	47,600	0.12	0.32	0.00	1.00
Risks Info. + Status Quo	47,600	0.12	0.32	0.00	1.00

Notes: Table 3 shows descriptive statistics using administrative information for the full sample of women participating in the experiment. The variable "years since last PAP" was available for half of the sample—women who had had a PAP smear within the previous five years.

We supplemented our administrative data with a survey of a subsample of just under 2,500 women, all of whom had had their most recent Pap smear more than five years before the study. Given a limited budget and their lower rate of PAP testing, we decided that it was policy relevant to learn more from this group. In addition, we expected our intervention to have less effect on this group and therefore believed we should collect additional data to get a deeper understanding of the obstacles to screening they faced.

To have the power to detect differences across treatment arms, we sampled the roughly 2,500 surveyed women to be equally distributed across treatment arms and the control group.[17]

Finally, we collected additional information to complement the limited number of covariates available from administrative sources. We asked the survey respondents about demographics such as household composition, income, education levels, and internet usage. Table 4 shows administrative and demographic statistics for the 2,462 women surveyed.

[17] We took five random subsamples of 2,500 observations each for the four treatment arms and the control group. ASSE merged this sample of 12,500 women with phone numbers and shared those with a specialized firm. The firm then implemented the survey by phone, with an overall response rate of 20 percent and no differential response by treatment arm or control group. The survey was implemented in two rounds, one in February 2020 and the other in April 2020.

Table 4. Descriptive Statistics, Survey Sample

Variable (% of sample unless otherwise noted)	**(1) Observations**	**(2) Mean**	**(3) Std. Dev.**	**(4) Min**	**(5) Max**
Administrative Information				-	
Age (years)	2,462	51.14	11.49	30.42	69.89
No PAP in previous 5 years	2,462	1.00	0.00	1.00	1.00
Control Group	2,462	0.20	0.40	0.00	1.00
Treatment					
Benef. Info. + Lower Trans. Cost	2,462	0.20	0.40	0.00	1.00
Risks Info. + Lower Trans. Cost	2,462	0.20	0.40	0.00	1.00
Benef. Info. + Status Quo	2,462	0.20	0.40	0.00	1.00
Risks Info. + Status Quo	2,462	0.20	0.40	0.00	1.00
Survey Demographics					
Years of Schooling	2,462	8.69	3.62	3.00	17.00
Husband/Partner Schooling (years)	980	8.37	3.43	3.00	17.00
Household Size (# of people)	2,462	3.24	1.80	1.00	17.00
At least one child (age 14 or less)	2,462	0.38	0.49	0.00	1.00
Household Income (URU$000)	2,462	18.4	11.7	5.0	55.0
Marital Status					
Married/Partner	2,462	0.40	0.49	0.00	1.00
Single	2,462	0.29	0.46	0.00	1.00
Divorced	2,462	0.22	0.42	0.00	1.00
Widow	2,462	0.09	0.28	0.00	1.00
Occupation					
Employed	2,462	0.43	0.50	0.00	1.00
Housewife	2,462	0.30	0.46	0.00	1.00
Retired	2,462	0.17	0.38	0.00	1.00
Unemployed	2,462	0.09	0.28	0.00	1.00
Contributing to BPS	996	0.48	0.50	0.00	1.00

Notes: Table 4 shows administrative and demographic statistics for the 2,462 women surveyed. No women in the sample took the PAP smear within the five last years. The husband/partner schooling is available for those who report to be married or have a partner, which is about 40 percent of the same (n=980). The variable “at least one child” reports the proportion of the women living in a household with at least one child aged 14 or younger. Household income is reported on a monthly basis and expressed in thousands of Uruguayan Pesos (URU$), with the URU$18,400 mean being equal to roughly US$433. The variables under “Marital Status” and “Occupation” each add to 100 percent. “Contributing to BPS” refers to those contributing to the state-owned social security institute BPS (*Banco de Prevision Social*).

On average, the surveyed women were 51 years old, which was consistent with the average age of the overdue group in the full sample. Overall, the characteristics of our survey sample were similar to those of the second quintile of national income according to official household surveys. In Uruguay, school attendance has been compulsory until nine years of education are completed (lower secondary education) since 1973 (Santiago, Ávalos, Burns, et al., 2016). Given their age, our respondents were educated under that regime and on average had about nine years of schooling. On average, their households comprised 3.2 persons compared to a national average of

2.8 and a mean of 3.2 for households in the 2nd income quintile (World Bank, 2018). Finally, 38 percent of the respondents lived in a household with at least one child (younger than 14 years old).

The average monthly household income about $18,400 Uruguayan Pesos (roughly US$420), which situates our sample in about the second quintile of the national income (mean $18,900 Uruguayan Pesos). Of the respondents, 40 percent indicated they were married or had a partner, 29 percent reported being single, 22 percent divorced, and 9 percent were widows. Only 43 percent reported being employed, while the remaining 57 percent were a housewife (30 percent), retired (17 percent), or unemployed (9 percent). Of those who were working, about 48 percent contributed to the state-owned social security institute (*Banco de Prevision Social*, BPS).

We used the information in Table 4 to show that our control and treatment groups were similar on a host of observable characteristics. We first conducted tests for joint orthogonality, regressing each treatment arm dummy on the set of covariates from Table 4, as shown in Table 5. Each column shows the estimation of a regression of the treatment arm dummy on the set of covariates from Table 4.

Table 5. Balance, Joint Test

	(1) Arm 1	(2) Arm 2	(3) Arm 3	(4) Arm 4
Age	−0.000	−0.001	0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.001)
Years of Schooling	−0.000	0.001	−0.001	−0.001
	(0.003)	(0.003)	(0.002)	(0.002)
Household Size	−0.002	−0.003	0.002	−0.002
	(0.005)	(0.006)	(0.006)	(0.006)
At Least One Child	0.013	0.003	0.000	−0.005
(age 14 or less)	(0.021)	(0.022)	(0.022)	(0.022)
Household Income	−0.001	0.001	0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.001)
Single	−0.009	−0.022	0.015	−0.004
	(0.174)	(0.179)	(0.161)	(0.181)
Married	−0.016	−0.008	−0.019	−0.003
	(0.174)	(0.178)	(0.181)	(0.181)
Divorced	−0.027	0.026	0.005	−0.018
	(0.174)	(0.179)	(0.181)	(0.181)
Partner	0.046	0.002	0.018	−0.011
	(0.029)	(0.028)	(0.027)	(0.028)
Widow	−0.016	0.074	0.001	−0.022
	(0.176)	(0.181)	(0.182)	(0.182)
Employed	−0.031	0.026	−0.066	−0.108
	(0.072)	(0.069)	(0.078)	(0.081)
Housewife	−0.007	0.017	−0.061	−0.112
	(0.073)	(0.071)	(0.080)	(0.082)
Retired	0.004	0.006	−0.067	−0.136
	(0.075)	(0.072)	(0.081)	(0.084)
Unemployed	−0.060	0.051	−0.127	−0.048
	(0.076)	(0.074)	(0.081)	(0.086)
Constant	0.269	0.208	0.250	0.304
	(0.200)	(0.201)	(0.205)	(0.206)
F-Test	0.669	0.578	0.787	0.782
Mean Dep. Variable	0.120	0.120	0.120	0.120
Observations	2,462	2,462	2,462	2,462
Sample	Survey	Survey	Survey	Survey

Notes: Table 5 shows the results of four separate regressions (in columns). Each column shows the estimation of a regression of the treatment arme dummy on the set of covariates from Table 4. For each estimation in the columns, we include a row with the p-values of the F-test for joint significance of all balance variables, all of which are 0.578 or higher. Robust standard errors are in parentheses. ***, **, and * indicate statistical significance at the 1, 5, and 10 percent levels, respectively.

Panels 1 and 2 of Table 6 show the results of individual tests of balance across the control group and the treatment arms. Each column shows the results of a separate regression of each variable on indicators for treatment arms and a constant. Columns 1 and 2 in Panel 1 show results for the full sample, using as dependent variables age and whether the women had had a PAP smear within the previous five years, respectively. Column 3 onwards do the same for the survey sample, for each variable in Table 4. Each regression also includes a formal test of equality of the treatment arm coefficients.

The results show no differences between the treatment and control groups that are statistically significant at the 1 percent level, with magnitudes that are small compared to the control mean. Two differences were significant, which is about what we would expect by chance. Since we compared differences in 15 variables in Table 5, we expected 1.5 to be significant at the 10 percent level due to chance. Women in treatment arm 3 were more likely to be unemployed (p=0.05) and those in arm 2 were more likely to be widows (p=0.10). As we show later, conditioning on these covariates does not change the coefficients of interest and serves mainly to improve the precision of our estimates, as is customary in well-implemented randomized controlled trials.

Table 6. Balance, Individual Tests

Panel 1

	(1) Age	(2) No PAP Last 3 Years	(3) Age	(4) Schooling	(5) Husband Schooling	(6) HH Size	(7) HH Children	(8) HH Income(b)
$\widehat{\beta_1}$: Benef. Info. + Lower Transaction Cost	-0.097	-0.001	-0.588	-0.206	-0.187	-0.011	0.028	-0.871
	(0.167)	(0.007)	(0.737)	(0.235)	(0.342)	(0.111)	(0.031)	(0.727)
$\widehat{\beta_2}$: Risks Info. + Lower Transaction Cost	0.069	-0.004	-0.449	0.042	-0.527	-0.079	0.010	0.217
	(0.167)	(0.007)	(0.720)	(0.239)	(0.358)	(0.116)	(0.031)	(0.772)
$\widehat{\beta_3}$: Benef. Info. + Status Quo	-0.113	0.004	-0.248	-0.092	0.194	-0.015	0.014	0.042
	(0.167)	(0.007)	(0.731)	(0.233)	(0.354)	(0.113)	(0.031)	(0.743)
$\widehat{\beta_4}$: Risks Info. + Status Quo	0.041	0.004	-0.477	-0.053	-0.553	-0.049	0.002	0.060
	(0.167)	(0.007)	(0.720)	(0.238)	(0.342)	(0.114)	(0.031)	(0.773)
Control Mean	48.576	0.603	51.494	8.752	8.589	3.275	0.373	18.534
Observations	47,600	47,600	2,462	2,462	980	2,462	2,462	2,462
Test: $\widehat{\beta_1} = \widehat{\beta_2} = \widehat{\beta_3} = \widehat{\beta_4}$ (p-value)	0.765	0.789	0.975	0.749	0.110	0.930	0.868	0.400
Sample	Full	Full	Survey	Survey	Survey	Survey	Survey	Survey

Panel 2

	(1) Married	(2) Single	(3) Divorced	(4) Widow	(5) Employed	(6) Housewife	(7) Unemployed	(8) Retired	(9) Contributing to BPS
$\widehat{\beta_1}$: Benef. Info. + Lower Transaction Cost	-0.001	-0.002	-0.017	0.020	-0.035	0.040	-0.018	0.001	-0.009
	(0.031)	(0.029)	(0.026)	(0.017)	(0.031)	(0.029)	(0.017)	(0.025)	(0.050)
$\widehat{\beta_2}$: Risks Info. + Lower Transaction Cost	-0.036	-0.030	0.022	0.044**	-0.003	-0.001	0.008	-0.013	0.021
	(0.031)	(0.028)	(0.027)	(0.018)	(0.032)	(0.029)	(0.018)	(0.024)	(0.049)
$\widehat{\beta_3}$: Benef. Info. + Status Quo	-0.040	0.014	0.006	0.020	-0.003	0.022	-0.028*	-0.007	0.066
	(0.031)	(0.029)	(0.027)	(0.017)	(0.032)	(0.029)	(0.017)	(0.024)	(0.049)
$\widehat{\beta_4}$: Risks Info. + Status Quo	-0.015	0.011	-0.006	0.010	-0.014	0.006	0.021	-0.032	0.028
	(0.031)	(0.029)	(0.026)	(0.017)	(0.032)	(0.029)	(0.019)	(0.024)	(0.049)
Control Mean	0.414	0.295	0.221	0.068	0.444	0.283	0.090	0.181	0.457
Observations	2,462	2,462	2,462	2,462	2,462	2,462	2,462	2,462	996
Test: $\widehat{\beta_1} = \widehat{\beta_2} = \widehat{\beta_3} = \widehat{\beta_4}$ (p-value)	0.564	0.412	0.514	0.345	0.702	0.500	0.021	0.561	0.535
Sample	Survey	Survey	Survey	Survey	Survey	Survey	Survey	Survey	Survey

Notes: Table 6 shows balance across the control and treatment arms. The columns show separate regressions of the covariates in Table 4 as dependent variables on treatment arm indicators and a constant. Robust standard errors are in parentheses. ***, ** and * indicate statistical significance at the 1, 5 and 10 percent level respectively.

5. Empirical Strategy

This section outlines our empirical strategy to compute intention-to-treat (ITT) and local average treatment effects (LATEs).

Intention-to-Treat Effects

Our main estimating equations were as follows:

$$(1)\ Y = \pi_1 + \boldsymbol{X}'\gamma_1 + \delta \sum_{j=1}^{4} Arm_j + \varepsilon_1$$

$$(2)\ Y = \pi_2 + \boldsymbol{X}'\gamma_2 + \sum_{j=1}^{4} Arm_j \cdot \alpha_1\,(j \leq 2) \cdot \alpha_2(j > 2) + \varepsilon_2$$

$$(3)\ Y = \pi_3 + \boldsymbol{X}'\gamma_3 + \sum_{j=1}^{4} \beta_j Arm_j + \varepsilon_3$$

where Y is an outcome variable, Arm_j is an indicator for random assignment to the treatment arm *j*, $\boldsymbol{X}$ is a vector of predetermined control variables, $(j \leq 2)$ and $(j > 2)$ are indicator variables, and ε_1, ε_2, and ε_3 are idiosyncratic error terms.

These equations allowed us to test different parameters of interest. Estimating the parameter δ in equation 1 provided us the overall ITT effect of our intervention. The overall effect is relevant because it tells us whether women randomized to receive our messages behaved differently than a scenario with no intervention. There should be no change in the behavior of informed women with sufficiently low transaction costs associated with appointments.

We were also particularly interested in the effects of promoting the online appointment system conditional on the information component in our messages. The estimates of parameters α_1 and α_2 in equation 2 represent the effects of being randomized to the group that received information plus a link to make appointments and those who received information and status quo encouragement, respectively, compared to the control group. We also implemented a formal test of equality of the treatment coefficients α_1 and α_2 . We expected α_1 to be higher because it estimates the effect of reducing transaction costs on top of providing information.

In equation 3 we augmented equation 1 by including separate indicators for each treatment arm to separately identify their effects. Each β_j gives us the average difference on outcome *Y* for women randomized to treatment arm j compared to those randomized to the control group. We also examined differences due to providing different types of information (benefits or risks) in our messages conditional on receiving the link ($H_0\!: \beta_1 = \beta_2$) and unconditionally ($H_0\!: \beta_3 = \beta_4$).

We estimated equations 1 to 3 for both the full sample and the survey sample. Following Bruhn and McKenzie (2009), we controlled for our randomization method in all of our estimations. For the full sample, the vector $\boldsymbol{X}$ includes stratum dummies based on age and time since last PAP. For our survey sample, $\boldsymbol{X}$ includes additional covariates, such as household size, children at home, household income, and marital and occupational status. Given the randomized nature of our treatment assignment, in our setup, inclusion of covariates does not change the estimated coefficients and rather serves to increase the precision of our estimates. We report Eicker-Huber-White standard errors for all of our estimates.[18]

Local Average Treatment Effects

The estimation of equations 1 to 3 provided us with ITT effects. We complemented these parameters with the estimation of two sets of policy relevant LATEs, taking into account successful delivery and receipt of the messages.

[18] As noted in Section 3, we randomized our intervention at the individual level. If we had randomized a cluster of individuals (like health centers) to treatment instead, then we would need to cluster at that unit level (e.g., the health center level). That is not the case in the present design (see Abadie, Athey, Imbens, et al., 2017).

The messaging system produced data on whether messages were successfully sent to each of the phone numbers randomized to treatment. These data were available for the full sample. We instrument the successful delivery of messages using random assignment to estimate a first set of LATEs.

We estimated a second set of LATEs for our survey sample. According to the messaging system, messages were successfully delivered to all women in this sample; however, delivery does not guarantee that messages were opened by recipients. They could save the message to read it later (and forget to do so), delete it before reading, etc. Thus, we went further and asked survey respondents to acknowledge the receipt of our messages. We instrument the reporting of receiving each type of message using the random assignment.

We conservatively interpreted our LATE estimates as upper-bound effects of our intervention because by construction they scale up our ITT estimates. For both sets of LATEs, we used two-stage least squares to estimate:

$$(4)\ Y = \eta_1 + \boldsymbol{X}'\theta_1 + \delta_K \sum_{j=1}^{4} RM_j + u_1$$

$$(5)\ Y = \eta_2 + \boldsymbol{X}'\theta_2 + \sum_{j=1}^{4} RM_j \cdot \alpha_{K1}\,(j \leq 2) \cdot \alpha_{K2}(j > 2) + u_2$$

$$(6)\ Y = \eta_3 + \boldsymbol{X}'\theta_3 + \sum_{j=1}^{4} \beta_{Kj} RM_j + u_3$$

where RM_j is defined as whether the message for treatment arm *j* was sent (i.e., the phone number was not rejected by the messaging system) or whether the survey respondents reported having received the type *j* message, and the rest of variables are defined as described above. The parameters of interest were the estimates of δ_K, α_{K1}, α_{K2}, and β_{Kj}, where K=D,R for delivery and receipt of messages, respectively. For equations 4 to 6, we reported the corresponding Cragg-Donald (Cragg and Donald, 1993) weak instrument test statistics, which jointly test the rank of the instruments. With one instrument (as in equation 4), the CD statistic is equivalent to the F-statistic. The CD values are well beyond critical values (e.g., Stock and Yogo, 2005), rejecting our instruments being weak.

6. Results

Women in the treatment groups significantly increased their probability of both scheduling and attending appointments compared with the control group. Our most conservative estimates showed the magnitude of the overall effects to be sizable, more than doubling the control group's rates.

One of our key findings was that most of the overall effect was driven by the lower transaction cost component, independent of the framing of messages. For women encouraged to make appointments online, the attendance and scheduling rates were about 3.0 times higher than the control group. For those encouraged to schedule appointments as usual, rates were 1.5 times higher than the control group.

In this section, first we document treatment effects on our main outcomes for the full sample of participating women. Then, we provide the results for the women in our survey sample. For them, we also estimate effects on secondary outcomes, exploring their roles as mechanisms driving effects.

6.1. Full Sample

Average Effects

Tables 7 and 8 show the ITT and LATE estimates for the full sample of 47,600 women randomized to the control group or treatment arms. Both tables present the estimates from equations 1, 2, and 3 on our two main outcomes of interest: scheduling and attending Pap smear appointments. The difference between effects on these two outcomes is an estimation of declaring intentions versus acting.

Table 7. ITT Estimates for the Full Sample

	Scheduling			Attendance		
	(1)	(2)	(3)	(4)	(5)	(6)
$\widehat{\delta}$: Treatment	0.033***			0.023***		
	(0.002)			(0.002)		
$\widehat{\alpha_1}$: Info. + Lower Trans. Cost		0.052***			0.032***	
		(0.003)			(0.002)	
$\widehat{\alpha_2}$: Info. + Status Quo		0.013***			0.014***	
		(0.002)			(0.002)	
$\widehat{\beta_1}$: Benef. Info. + Lower Trans. Cost			0.059***			0.033***
			(0.004)			(0.003)
$\widehat{\beta_2}$: Risks Info. + Lower Trans. Cost			0.046***			0.031***
			(0.004)			(0.003)
$\widehat{\beta_3}$: Benef. Info. + Status Quo			0.014***			0.015***
			(0.003)			(0.003)
$\widehat{\beta_4}$: Risks Info. + Status Quo			0.013***			0.013***
			(0.003)			(0.002)
$H_0 : \widehat{\alpha_1} = \widehat{\alpha_2}$		0.000			0.000	
$H_0 : \widehat{\beta_1} = \widehat{\beta_2}$			0.009			0.615
$H_0 : \widehat{\beta_3} = \widehat{\beta_4}$			0.902			0.538
Observations	47,600	47,600	47,600	47,600	47,600	47,600
Control Group Mean	0.027	0.027	0.027	0.019	0.019	0.019
Sample	Full	Full	Full	Full	Full	Full
Strata	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Table 7 shows the intention-to-treat results from estimating equations (1), (2) and (3) for the full sample of women participating in our study. In columns 1 to 3, the dependent variable is an indicator for whether women made a PAP appointment. In columns 4 to 6, the dependent variable is an indicator for whether women attended the PAP appointment. In columns 1 and 4, we include a single treatment indicator equal to one if women were randomized to any treatment arm. In columns 2 and 4, we include a separate indicator for whether women were randomized to received information plus online encouragement, and those who received information and status-quo encouragement, respectively. In columns 3 and 6 we included a separate indicator for each of the four treatment arms. The lower panel displays the p-values of F-tests for the differences between treatment parameters. All regressions control for our randomization method, by including stratum dummies based on age and time since the last pap. Robust standard errors are in parentheses. ***, ** and * indicate statistical significance at the 1, 5 and 10 percent level respectively.

The estimates in columns 1 and 4 of Table 7 show the overall ITT effect ($\widehat{\delta}$) on each outcome. On average, women in the treatment groups were 3.3 percentage points (pp) more likely to schedule appointments (over a control mean of 2.7 percent) and 2.3 pp more likely to attend (over a control mean of 1.9 pp?). These ITT effects represent 122 percent and 121 percent increases over the control group, respectively.

Columns 2 and 5 display the estimates of equation 2, showing that effects are larger for women randomized to receive the lower transaction cost component ($\widehat{\alpha_1}$) compared to the status quo encouragement ($\widehat{\alpha_2}$), both conditional on receiving information. Women in the status quo group were 1.3 pp and 1.4 pp more likely than the control group to schedule and attend appointments. When the lower transaction cost is added, the coefficients jump to 5.2 pp and 3.2

pp, respectively. The lower panel shows that the differences between the respective treatment parameters in columns 2 and 5 are statistically significant (p-value<0.01).

Our design allowed us to further decompose effects using the framing of the information sent to women. We report the estimates of equation 3 in columns 3 and 6. The estimates show that when encouraging women to use the online tool, including a benefits message has a higher effect on scheduling than a risks message (5.9 pp vs. 4.6 pp, p-value=0.01). This differential effect vanishes when we estimate effects on attending the appointments (3.3 pp vs. 3.1 pp, p-value=0.61). Benefits-focused messages appear to pique interest but, at the end of the day, are no better than risk messages at getting women to act and attend their appointments.

Table 8 shows LATE results from estimating equations 4, 5, and 6 for the full sample. Our estimated LATEs are about 11 percent larger than the ITT estimates because the first-stage coefficient is about 0.9 and therefore LATE=ITT/0.9=1.11 ITT. This result indicates that the effects of our intervention would have been 11 percent higher had the messaging system delivered all messages successfully. We thus think of our ITT effects as providing a lower bound of the effect of our intervention.

Table 8. LATE Estimates for the Full Sample

	Scheduling			Attendance		
	(1)	(2)	(3)	(4)	(5)	(6)
$\widehat{\delta_D}$: Treatment	0.037***			0.026***		
	(0.002)			(0.002)		
$\widehat{\alpha_{D1}}$: Info. + Lower Trans. Cost		0.059***			0.036***	
		(0.003)			(0.002)	
$\widehat{\alpha_{D2}}$: Info. + Status Quo		0.015***			0.016***	
		(0.002)			(0.002)	
$\widehat{\beta_{D1}}$: Benef. Info. + Lower Trans. Cost			0.066***			0.037***
			(0.004)			(0.003)
$\widehat{\beta_{D2}}$: Risks Info. + Lower Trans. Cost			0.051***			0.035***
			(0.004)			(0.003)
$\widehat{\beta_{D3}}$: Benef. Info. + Status Quo			0.015***			0.017***
			(0.003)			(0.003)
$\widehat{\beta_{D4}}$: Risks Info. + Status Quo			0.015***			0.015***
			(0.003)			(0.003)
$H_0 : \widehat{\alpha_{D1}} = \widehat{\alpha_{D2}}$		0.000			0.000	
$H_0 : \widehat{\beta_{D1}} = \widehat{\beta_{D2}}$			0.010			0.632
$H_0 : \widehat{\beta_{D3}} = \widehat{\beta_{D4}}$			0.892			0.529
Observations	47,600	47,600	47,600	47,600	47,600	47,600
Control Group Mean	0.027	0.027	0.027	0.019	0.019	0.019
Sample	Full	Full	Full	Full	Full	Full
Strata	Yes	Yes	Yes	Yes	Yes	Yes
1st Stage Cragg-Donald Statistic	212,044	106,009	52,987	212,044	106,009	52,987

Notes: Table 8 shows LATE results, from estimating equations (4), (5) and (6) for the full sample of women participating in our study. We instrumented the successful delivery of messages with random assignment. About 90% of all messages within each treatment arm was successfully delivered. No messages were delivered to women in the control group. In columns 1 to 3 the dependent variable is an indicator for whether women made a PAP appointment. In columns 1 and 4, we include a single treatment indicator equal to one if women were randomized to any treatment arm. In columns 2 and 4, we include a separate indicator for whether women were randomized to received information plus online encouragement, and those who received information and status-quo encouragement, respectively. In columns 3 and 6 we included a separate indicator for each of the four treatment arms. The lower panel displays the p-values of F-tests for the differences between treatment parameters. All regressions control for our randomization method, by including stratum dummies based on age and time since the last pap. The bottom row reports the Cragg-Donald (1993) weak instrument test statistics, which jointly tests the rank of the instruments. With one instrument, the Cragg-Donald statistic is equivalent to the F-statistic. The values are well beyond critical values (see Stock and Yogo, 2005) so we reject that our instruments are weak. Robust standard errors are in parentheses. ***, ** and * indicate statistical significance at the 1, 5 and 10 percent level respectively.

Heterogeneity by Time since Last Pap Smear and Age Group

We documented whether treatment effects varied by the time passed since the previous Pap smear and age group. The main takeaway is that effects were larger for women in the on-time group compared to those in the overdue group. In other words, women who had had the exam at some point in the previous five years appeared to be more receptive to our intervention than those who had not had a Pap smear in that timeframe. In addition, our estimated effects were smaller for the oldest women in the overdue group.

Effects by On-Time versus Overdue Groups

Table 9 shows the ITT results of estimating equations 1 to 3 for women in the on-time and overdue groups. Columns 1 and 2 show the coefficients from separate regressions on appointment scheduling, and columns 4 and 5 do the same for attendance rates. Columns 3 and 6 present the coefficients from fully interacted regressions, indicating the difference between coefficients for the respective groups.

Table 9. ITT Estimates for the On-Time and Overdue Groups, Full Sample

	Scheduling			Attendance		
	(1)	(2)	(3)	(4)	(5)	(6)
Coefficient	On-Time	Overdue	Difference (1)-(2)	On-Time	Overdue	Difference (4)-(5)
$\widehat{\delta}$: Treatment	0.038***	0.028***	0.010***	0.027***	0.019***	0.008**
	(0.003)	(0.002)	(0.004)	(0.003)	(0.002)	(0.003)
$\widehat{\alpha_1}$: Info. + Lower Transaction Cost	0.065***	0.039***	0.026***	0.039***	0.025***	0.015***
	(0.004)	(0.003)	(0.005)	(0.004)	(0.003)	(0.004)
$\widehat{\alpha_2}$: Info. + Status Quo	0.011***	0.016***	-0.005	0.015***	0.014***	0.002
	(0.003)	(0.003)	(0.004)	(0.003)	(0.002)	(0.004)
$\widehat{\beta_1}$: Benef. Info. + Lower Transaction Cost	0.077***	0.041***	0.036***	0.040***	0.026***	0.013**
	(0.006)	(0.004)	(0.008)	(0.005)	(0.004)	(0.006)
$\widehat{\beta_2}$: Risks Info. + Lower Transaction Cost	0.054***	0.038***	0.016**	0.039***	0.023***	0.016***
	(0.006)	(0.004)	(0.007)	(0.005)	(0.003)	(0.006)
$\widehat{\beta_3}$: Benef. Info. + Status Quo	0.011**	0.016***	-0.005	0.016***	0.015***	0.001
	(0.004)	(0.003)	(0.006)	(0.004)	(0.003)	(0.005)
$\widehat{\beta_4}$: Risks Info. + Status Quo	0.011**	0.016***	-0.005	0.014***	0.012***	0.002
	(0.004)	(0.003)	(0.006)	(0.004)	(0.003)	(0.005)
$H_0 : \widehat{\alpha_1} = \widehat{\alpha_2}$:	0.000	0.000		0.000	0.000	
$H_0 : \widehat{\beta_1} = \widehat{\beta_2}$:	0.004	0.643		0.926	0.472	
$H_0 : \widehat{\beta_3} = \widehat{\beta_4}$:	0.975	0.879		0.758	0.545	
Observations	23,800	23,800		23,800	23,800	
Control Group Mean	0.038	0.016		0.027	0.010	

Notes: Table 9 shows the intention-to-treat results, from estimating equations (1), (2) and (3) for women in the on-time group (columns 1 and 4) and the overdue group (columns 2 and 5), respectively. Columns 3 and 6 display the coefficients from fully interacted regressions, indicating the difference between coefficients by subgroups. Outcomes are the same as described in Table 7. The lower panel displays the p-values of F-tests for the differences between treatment parameters. All regressions control for age stratum dummies. Robust standard errors are in parentheses. ***, ** and * indicate statistical significance at the 1, 5 and 10 percent level respectively.

Though the relative importance of treatment effects within each group is similar to the pattern for the full sample, we also found differences. The results in the first row indicate that the overall effects on both scheduling and attendance rates are approximately 1 pp higher for the on-time group. The second row shows that this difference is completely driven by the lower transaction cost component.

Effects are differential when lower transaction costs are part of the treatment. Women in on-time group were 2.6 pp more likely to schedule an appointment than women in the overdue group and 1.5 pp more likely to attend the appointment. For women who received status quo encouragement, the effects remain the same independent of the framing of the information sent. Effects appear to be higher for scheduling appointments when combined with benefits information, but that difference vanishes for attendance rates (as shown by the coefficient on $ß_1$ for scheduling vs attendance)..

Effects by Age Group within the On-Time and Overdue Groups

Table 10 shows ITT estimates by age for the on-time (panel 1) and overdue (panel 2) groups. Columns 1 to 4 show the coefficients from separate regressions on appointment scheduling for four age groups: 30–39, 40–49, 50–59, and 60–70. Columns 6 to 9 do the same using attendance rates as the outcome. Columns 5 and 10 show the difference between the younger and older groups for both outcomes, from fully interacted regressions.

Table 10. ITT Estimates by Age and On-Time and Overdue Groups, Full Sample

Panel 1: On-Time Group

Coefficient	Scheduling (1) Age [30,40)	(2) Age [40,50)	(3) Age [50,60)	(4) Age [60,70]	(5). Dif. 1-4	Attendance (6) Age [30,40)	(7) Age [40,50)	(8) Age [50,60)	(9) Age [60,70]	(10) Dif. 6-10
$\hat{\delta}$: Treatment	0.032***	0.039***	0.041***	0.043***	-0.010	0.020***	0.030***	0.030***	0.032***	-0.011
	(0.006)	(0.006)	(0.006)	(0.006)	(0.009)	(0.005)	(0.005)	(0.005)	(0.006)	(0.007)
$\widehat{\alpha_1}$: Info. + Lower Trans. Cost	0.060***	0.069***	0.065***	0.068***	-0.008	0.031***	0.044***	0.041***	0.045***	-0.014
	(0.008)	(0.008)	(0.009)	(0.010)	(0.013)	(0.007)	(0.007)	(0.007)	(0.008)	(0.011)
$\widehat{\alpha_2}$: Info. + Status Quo	0.005***	0.008***	0.018***	0.018***	-0.014	0.010***	0.017***	0.018***	0.019***	-0.009
	(0.006)	(0.006)	(0.007)	(0.007)	(0.009)	(0.006)	(0.006)	(0.006)	(0.007)	(0.009)
$\widehat{\beta_1}$: Benef. Info. + Lower Trans. Cost	0.071***	0.082***	0.078***	0.078***	-0.007	0.022***	0.047***	0.047***	0.050***	-0.028**
	(0.011)	(0.012)	(0.013)	(0.014)	(0.015)	(0.008)	(0.010)	(0.010)	(0.012)	(0.013)
$\widehat{\beta_2}$: Risks Info. + Lower Trans. Cost	0.049***	0.057***	0.052***	0.058***	-0.009	0.040***	0.040***	0.036***	0.040***	0.000
	(0.011)	(0.011)	(0.011)	(0.013)	(0.015)	(0.009)	(0.009)	(0.010)	(0.011)	(0.013)
$\widehat{\beta_3}$: Benef. Info. + Status Quo	-0.001***	0.008***	0.020***	0.026***	-0.027*	0.005***	0.017***	0.024***	0.025***	-0.020
	(0.008)	(0.008)	(0.009)	(0.010)	(0.015)	(0.007)	(0.008)	(0.009)	(0.010)	(0.013)
$\widehat{\beta_4}$: Risks Info. + Status Quo	0.010***	0.008***	0.015***	0.011***	-0.001	0.015***	0.016***	0.013***	0.013***	0.002
	(0.009)	(0.008)	(0.009)	(0.008)	(0.014)	(0.008)	(0.008)	(0.008)	(0.008)	(0.012)
Test: $\widehat{\alpha_1} = \widehat{\alpha_2}$: p-Value	0.000	0.000	0.000	0.000		0.006	0.001	0.010	0.008	
Test: $\widehat{\beta_1} = \widehat{\beta_2}$: p-Value	0.129	0.110	0.117	0.290		0.122	0.567	0.420	0.514	
Test: $\widehat{\beta_3} = \widehat{\beta_4}$: p-Value	0.292	0.963	0.639	0.238		0.290	0.938	0.304	0.299	
Observations	7,378	7,055	5,556	3,811		7,378	7,055	5,556	3,811	
Control Group Mean	0.049	0.042	0.032	0.019		0.035	0.029	0.025	0.015	

Panel 2: Overdue Group

Coefficient	Scheduling (1) Age [30,40)	(2) Age [40,50)	(3) Age [50,60)	(4) Age [60,70]	(5). Dif. 1-4	Attendance (6) Age [30,40)	(7) Age [40,50)	(8) Age [50,60)	(9) Age [60,70]	(10) Dif. 6-10
$\hat{\delta}$: Treatment	0.034***	0.031***	0.026***	0.020***	0.013**	0.023***	0.018***	0.020***	0.015***	0.009*
	(0.005)	(0.005)	(0.005)	(0.003)	(0.006)	(0.004)	(0.004)	(0.004)	(0.003)	(0.005)
$\widehat{\alpha_1}$: Info. + Lower Trans. Cost	0.044***	0.048***	0.041***	0.026***	0.018**	0.027***	0.027***	0.028***	0.018***	0.009
	(0.007)	(0.007)	(0.007)	(0.005)	(0.008)	(0.005)	(0.006)	(0.006)	(0.004)	(0.007)
$\widehat{\alpha_2}$: Info. + Status Quo	0.024***	0.014***	0.011***	0.014***	0.009	0.020***	0.010***	0.013***	0.011***	0.009
	(0.006)	(0.005)	(0.005)	(0.004)	(0.007)	(0.005)	(0.004)	(0.005)	(0.003)	(0.006)
$\widehat{\beta_1}$: Benef. Info. + Lower Trans. Cost	0.028***	0.060***	0.043***	0.034***	-0.007	0.016***	0.035***	0.029***	0.026***	-0.010
	(0.008)	(0.011)	(0.009)	(0.007)	(0.010)	(0.006)	(0.008)	(0.008)	(0.006)	(0.008)
$\widehat{\beta_2}$: Risks Info. + Lower Trans. Cost	0.060***	0.037***	0.038***	0.017***	0.043***	0.037***	0.018***	0.026***	0.010***	0.027***
	(0.010)	(0.009)	(0.009)	(0.006)	(0.010)	(0.008)	(0.007)	(0.008)	(0.005)	(0.008)
$\widehat{\beta_3}$: Benef. Info. + Status Quo	0.029***	0.018***	0.013***	0.005***	0.024**	0.026***	0.015***	0.014***	0.003***	0.023***
	(0.008)	(0.008)	(0.007)	(0.004)	(0.010)	(0.007)	(0.007)	(0.006)	(0.003)	(0.008)
$\widehat{\beta_4}$: Risks Info. + Status Quo	0.018***	0.009***	0.010***	0.024***	-0.006	0.013***	0.005***	0.011***	0.019***	-0.006
	(0.008)	(0.007)	(0.007)	(0.007)	(0.010)	(0.006)	(0.005)	(0.006)	(0.006)	(0.008)
Test: $\widehat{\alpha_1} = \widehat{\alpha_2}$: p-Value	0.016	0.000	0.000	0.056		0.328	0.011	0.027	0.186	
Test: $\widehat{\beta_1} = \widehat{\beta_2}$: p-Value	0.012	0.084	0.716	0.058		0.034	0.096	0.773	0.037	
Test: $\widehat{\beta_3} = \widehat{\beta_4}$: p-Value	0.321	0.358	0.736	0.012		0.150	0.204	0.721	0.013	
Observations	6,279	5,768	5,520	6,233		6,279	5,768	5,520	6,233	
Control Group Mean	0.022	0.018	0.016	0.007		0.013	0.012	0.011	0.005	

Notes: Table 10 shows the intention-to-treat results, from estimating equations (1), (2) and (3) by age groups for women in the on-time group (panel 1) and the overdue group (panel 2), respectively. In each panel, columns 1 to 4 show the coefficients from separate regressions on appointments scheduling for four age groups: 30-39, 40-49, 50-59 and 60-70. Columns 6 to 9 do the same using attendance rates as the outcome. Column 5 and 10 show the difference between the youngest and oldest groups for both outcomes, from fully interacted regressions. Robust standard errors are in parentheses. ***, ** and * indicate statistical significance at the 1, 5 and 10 percent level respectively.

The magnitude of the coefficients decreases for both scheduling and attendance for the oldest women in the overdue group. When comparing the effects against the respective control group, the effects represent an increase in scheduling ranging from 154 percent to 285 percent, and an increase in attendance ranging from of 150 percent to 300 percent. Note also that the relative magnitude of the effects increases with age because the control group mean is smaller for older groups.

6.2. Survey Sample

As noted above, all women in this sample belonged to the overdue group, having not had a Pap smear within the previous five years.

Average Effects

Table 11 shows the ITT estimates from equations 1, 2, and 3 for our survey sample. Columns 1 and 4 include a single treatment indicator equal to one if women were randomized to any treatment arm. Columns 2 and 5 include a separate indicator for whether women received information with online encouragement or with status quo encouragement. Columns 3 and 6 include a separate indicator for each of the four treatment arms. The lower panel displays the p-values of F-tests for the differences between treatment parameters.

Table 11. ITT Estimates for the Survey Sample

	Scheduling			Attendance		
	(1)	(2)	(3)	(4)	(5)	(6)
$\widehat{\delta}$: Treatment	0.064***			0.047***		
	(0.008)			(0.007)		
$\widehat{\alpha_1}$: Info. + Lower Trans. Cost		0.087***			0.059***	
		(0.011)			(0.009)	
$\widehat{\alpha_2}$: Info. + Status Quo		0.040***			0.035***	
		(0.009)			(0.008)	
$\widehat{\beta_1}$: Benef. Info. + Lower Trans. Cost			0.089***			0.059***
			(0.015)			(0.013)
$\widehat{\beta_2}$: Risks Info. + Lower Trans. Cost			0.085***			0.059***
			(0.015)			(0.012)
$\widehat{\beta_3}$: Benef. Info. + Status Quo			0.043***			0.039***
			(0.012)			(0.011)
$\widehat{\beta_4}$: Risks Info. + Status Quo			0.037***			0.030***
			(0.012)			(0.010)
$H_0 : \widehat{\alpha_1} = \widehat{\alpha_2}$		0.000			0.021	
$H_0 : \widehat{\beta_1} = \widehat{\beta_2}$			0.827			0.992
$H_0 : \widehat{\beta_3} = \widehat{\beta_4}$			0.659			0.521
Observations	2,462	2,462	2,462	2,462	2,462	2,462
Control Group Mean	0.016	0.016	0.016	0.010	0.010	0.010
Sample	Survey	Survey	Survey	Survey	Survey	Survey
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Table 11 shows the intention-to-treat results, from estimating equations (1), (2) and (3) for our survey sample. In columns 1 to 3, the dependent variable is an indicator for whether women made a PAP appointment. In columns 4 to 6, the dependent variable is an indicator for whether women attended the PAP appointment. In columns 1 and 4, we include a single treatment indicator equal to one if women were randomized to any treatment arm. In columns 2 and 4, we include a separate indicator for whether women were randomized to received information plus online encouragement, and those who received information and status-quo encouragement, respectively. In columns 3 and 6 we included a separate indicator for each of the four treatment arms. The lower panel displays the p-values of F-tests for the differences between treatment parameters. All regressions control for age and the set of demographics displayed in Table 4. Robust standard errors are in parentheses. ***, ** and * indicate statistical significance at the 1, 5 and 10 percent level respectively.

The ITT effects are higher than those for the women in the overdue group for the full sample. In particular, the overall effects on appointment scheduling and attendance were 6.4 and 4.7 pp higher than the control group compared to 2.8 and 1.9 pp for the full overdue group (Table 9), respectively. The data suggests that the control groups for both the survey and overdue groups are similar since they display the same mean for appointment scheduling (1.6 percent) and attendance (1.0 percent). As before, effects are significantly larger (about twice) on both outcomes for women encouraged to use the online system compared to those encouraged to make appointments at their clinics.

Table 12 shows LATE results from estimating equations 4, 5, and 6 for the survey sample. Columns 1 and 4 include a single treatment indicator equal to one if women were randomized to any treatment arm. Columns 2 and 5 include a separate indicator for whether women received information with online encouragement or with status quo encouragement. Columns 3 and 6 include a separate indicator for each of the four treatment arms. The lower panel displays the p-values of F-tests for the differences between treatment parameters.

We instrumented the receipt of messages (as reported in the survey) with random assignment. On average about 60 percent of the women in the treatment arms and 1.6 percent of those in the control group reported having received messages.

Table 12. LATE Estimates for the Survey Sample

	Scheduling			Attendance		
	(1)	(2)	(3)	(4)	(5)	(6)
$\widehat{\delta_R}$: Treatment	0.108***			0.080***		
	(0.014)			(0.012)		
$\widehat{\alpha_{R1}}$: Info. + Lower Trans. Cost		0.146***			0.099***	
		(0.019)			(0.016)	
$\widehat{\alpha_{R2}}$: Info. + Status Quo		0.070***			0.060***	
		(0.016)			(0.013)	
$\widehat{\beta_{R1}}$: Benef. Info. + Lower Trans. Cost			0.134***			0.088***
			(0.023)			(0.019)
$\widehat{\beta_{R2}}$: Risks Info. + Lower Trans. Cost			0.155***			0.107***
			(0.026)			(0.022)
$\widehat{\beta_{R3}}$: Benef. Info. + Status Quo			0.068***			0.061***
			(0.019)			(0.016)
$\widehat{\beta_{R4}}$: Risks Info. + Status Quo			0.066***			0.054***
			(0.021)			(0.018)
$H_0 : \widehat{\alpha_{R1}} = \widehat{\alpha_{R2}}$		0.000			0.023	
$H_0 : \widehat{\beta_{R1}} = \widehat{\beta_{R2}}$			0.522			0.494
$H_0 : \widehat{\beta_{R3}} = \widehat{\beta_{R4}}$			0.940			0.769
Observations	2,462	2,462	2,462	2,462	2,462	2,462
Control Group Mean	0.027	0.027	0.027	0.019	0.019	0.019
Sample	Full	Full	Full	Full	Full	Full
Strata	Yes	Yes	Yes	Yes	Yes	Yes
1st Stage Cragg-Donald Statistic	721	361	188	721	361	188

Notes: Table 12 shows LATE results, from estimating equations (4), (5) and (6) for our survey sample. We instrumented the receipt of messages (as reported in the survey) with random assignment. On average about 60% of women in the treatment arms and a 1.6% of those in the control group report to have received messages. In columns 1 to 3 the dependent variable is an indicator for whether women made a PAP appointment. In columns 1 and 4, we include a single treatment indicator equal to one if women were randomized to any treatment arm. In columns 2 and 4, we include a separate indicator for whether women were randomized to received information plus online encouragement, and those who received information and status-quo encouragement, respectively. In columns 3 and 6 we included a separate indicator for each of the four treatment arms. The lower panel displays the p-values of F-tests for the differences between treatment parameters. All regressions control for age and the set of demographics displayed in Table 4. The bottom row reports the Cragg-Donald (1993) weak instrument test statistics, which jointly tests the rank of the instruments. With one instrument, the Cragg-Donald statistic is equivalent to the F-statistic. The values are well beyond critical values (see Stock and Yogo, 2005) so we reject that our instruments are weak. Robust standard errors are in parentheses. ***, ** and * indicate statistical significance at the 1, 5 and 10 percent level respectively.

The overall LATE estimates are 10.8 pp for scheduling and 8.0 pp for attending appointments. For both ITT and LATE estimates, the pattern of treatment effects for surveyed women was similar to the results for the full sample. Effects were larger for women targeted with the lower transaction cost component than women targeted with the status quo encouragement. As for the full sample, the framing of the information did not have a differential effect on the main outcomes for women in the survey sample.

Mechanisms

To explore mechanisms, we included questions in the survey related to the transaction costs of scheduling appointments, knowledge about the Pap smear, and perceptions about public services. Our results suggest that our intervention translated into changes in behavior by making the PAP smear more salient to women and, mostly, making it easier to schedule appointments.

Transaction Costs. One component of our intervention targeted transaction costs, as such, easing the process of scheduling appointments. Table 13 shows ITT estimates for secondary outcomes related to perceptions about appointment scheduling. We asked respondents how they would usually schedule their appointments and used their answers as dependent variables in columns 1 to 5. Column 6 assesses the women's perceptions of how easy it was to schedule appointments (on a 1–5 scale), and column 7 indicates whether women though it was very easy to do so (5 on the 1–5 scale).

Table 13. ITT Effects on Perceptions about Appointment Scheduling

	How would you schedule an appointment?:					How easy is scheduling?:	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Via Internet	In Person	By Phone	Other Way	Do not know	On a 1-5 Scale	Very Easy (Value 5)
$\widehat{\delta}$: Treatment	0.037***	-0.041*	-0.014	0.007	-0.009	0.099*	0.069***
	(0.006)	(0.022)	(0.016)	(0.008)	(0.015)	(0.055)	(0.025)
$\widehat{\alpha_1}$: Info. + Lower Trans. Cost	0.050***	-0.063**	-0.007	0.012	-0.014	0.120**	0.080***
	(0.008)	(0.025)	(0.018)	(0.009)	(0.017)	(0.059)	(0.027)
$\widehat{\alpha_2}$: Info. + Status Quo	0.023***	-0.019	-0.021	0.002	-0.004	0.077	0.058**
	(0.007)	(0.024)	(0.018)	(0.009)	(0.017)	(0.060)	(0.027)
$\widehat{\beta_1}$: Benef. Info. + Lower Trans. Cost	0.053***	-0.075***	-0.005	0.010	-0.010	0.139**	0.085***
	(0.011)	(0.029)	(0.021)	(0.011)	(0.019)	(0.068)	(0.032)
$\widehat{\beta_2}$: Risks Info. + Lower Trans. Cost	0.047***	-0.051*	-0.008	0.014	-0.018	0.102	0.074**
	(0.011)	(0.029)	(0.021)	(0.011)	(0.019)	(0.069)	(0.032)
$\widehat{\beta_3}$: Benef. Info. + Status Quo	0.026***	-0.045	-0.015	0.006	0.011	0.147**	0.080**
	(0.009)	(0.029)	(0.020)	(0.010)	(0.020)	(0.068)	(0.032)
$\widehat{\beta_4}$: Risks Info. + Status Quo	0.021**	0.008	-0.027	-0.001	-0.018	0.006	0.036
	(0.009)	(0.028)	(0.020)	(0.010)	(0.019)	(0.071)	(0.032)
$H_0 : \widehat{\alpha_1} = \widehat{\alpha_2}$	0.004	0.033	0.318	0.208	0.434	0.366	0.343
$H_0 : \widehat{\beta_1} = \widehat{\beta_2}$	0.695	0.405	0.892	0.756	0.634	0.577	0.743
$H_0 : \widehat{\beta_3} = \widehat{\beta_4}$	0.622	0.066	0.543	0.493	0.149	0.044	0.172
Observations	2,462	2,462	2,462	2,462	2,462	2,462	2,462
Control Group Mean	0.01	0.73	0.13	0.02	0.11	4.11	0.47
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Table 13 shows intention-to-treat estimates on perceptions about appointment scheduling, for our survey sample. The panels in the table show estimates of equations (1), (2) and (3). In columns 1 to 5, the dependent variables are indicators to how women report they would make an appointment. In columns 6 and 7, the dependent variable is their perception of how easy is to schedule appointments (on a 1-5 scale), and an indicator for whether women think it is very easy to do so (value 5 on the 1-5 scale). The lower panel displays the p-values of F-tests for the differences between treatment parameters. All regressions control for age and the set of demographics displayed in Table 4. Robust standard errors are in parentheses. ***, ** and * indicate statistical significance at the 1, 5 and 10 percent level respectively.

Those in the treatment groups were more likely to report that they would usually schedule appointments using the internet (3.7 pp over less than 1 percent in the control group). Consistent with our design, these effects were two times as high for women who were encouraged to use the online system (5.0 pp) compared to those in the status quo group (2.3 pp). We hypothesize that the message motivated some of the women to go online and discover the online tool, which ASSE made available for everyone to sign into.[19]

Another consistent result is that women in the treatment groups were less likely to say they would make an appointment in person. In a similar pattern, those who received the link to

[19] Another piece of evidence that supports this hypothesis is that no women among those randomized to treatment in the first round that did not receive the link made appointments through the web system. Note that the online system was not available to everyone during the first round.

make appointments online were 6.3 pp less likely to say they would make appointments in person, which is about a 10 percent decrease compared to the control mean of 73 percent. This result suggests that the online system could help to decongest health services, reducing the number of people who would otherwise visit the clinic centers to schedule appointments.

The magnitude of the effects for other ways of making appointments, such as phone calls and not knowing how to make appointments, is small and the effects are measured with less precision. With that caveat in mind, the sign of the coefficients is consistent with our intervention, as women randomized to treatment reported being less likely to make appointments by phone and less likely to not know how to schedule.

Columns 6 and 7 show that treated women reported that it was easier to schedule appointments compared to those in the control group. The coefficient in column 6 (0.099) is equivalent to 0.9 standard deviations, as the mean rating in the control group was of 4.11 with a standard deviation of 1.09. Similar to the results for the other secondary outcomes, the magnitude of this effect appears to be higher for those who received the link versus those who did not (0.12 versus 0.077), though we lack statistical power to differentiate both coefficients at conventional levels.

Column 7 shows that a higher proportion of women reported that it was very easy to make appointments (6.9 pp over 47 percent). These effects were higher for women who received the link to make an appointment online (8.0 pp) versus those who did not (5.8 pp). Overall, these results suggest that the larger impacts are a result of reduced transactions costs (i.e., making it easier to schedule appointments), which translates into behavior later.

Information. The information component of our intervention either emphasized the importance of having a Pap smear to prevent cervical cancer or was explicit about the mortality risk of not having the test. Table 14 shows ITT estimates on Pap smear knowledge for the survey sample. Column 1 assesses women's perceptions of the importance of PAP smears (on a 1-5 scale), and column 2 indicates whether women think it is very important (5 on the 1-5 scale). Column 3 looks at women’s perceptions of whether PAP smears can save lives, and column 4 indicates whether women think they definitely do (5 on the 1-5 scale). Columns 5 to 8 are estimates for additional variables related to PAP knowledge: whether women report knowing about PAP smears, whether they report knowing that the test is free at health centers, whether they are aware that they can skip one day of work to take the test, and whether they believe that they should have the PAP smear every year.

Table 14. ITT Effects on Pap Smear Knowledge

	PAP Importance: (1) On a 1-5 Scale	(2) Very High (Value 5)	PAP Saves Lives: (3) On a 1-5 Scale	(4) Definitely Yes (Value 5)	(5) Knows the PAP	(6) PAP is Free	(7) Skip Work Day	(8) PAP Every Year
$\widehat{\delta}$: Treatment	-0.018	-0.009	0.055	0.029	-0.004	0.023	0.004	-0.009
	(0.032)	(0.021)	(0.045)	(0.022)	(0.007)	(0.022)	(0.014)	(0.021)
$\widehat{\alpha_1}$: Info. + Lower Trans. Cost	-0.035	-0.020	0.054	0.034	-0.008	0.009	0.009	-0.011
	(0.036)	(0.023)	(0.050)	(0.024)	(0.008)	(0.024)	(0.015)	(0.023)
$\widehat{\alpha_2}$: Info. + Status Quo	-0.001	0.002	0.056	0.024	0.000	0.038	-0.001	-0.007
	(0.035)	(0.023)	(0.050)	(0.024)	(0.008)	(0.024)	(0.015)	(0.023)
$\widehat{\beta_1}$: Benef. Info. + Lower Trans. Cost	-0.028	-0.020	0.026	0.024	-0.011	-0.003	-0.006	-0.006
	(0.041)	(0.027)	(0.058)	(0.028)	(0.010)	(0.028)	(0.018)	(0.026)
$\widehat{\beta_2}$: Risks Info. + Lower Trans. Cost	-0.043	-0.020	0.081	0.044	-0.004	0.020	0.025	-0.016
	(0.043)	(0.026)	(0.057)	(0.027)	(0.009)	(0.028)	(0.016)	(0.026)
$\widehat{\beta_3}$: Benef. Info. + Status Quo	-0.022	-0.002	0.036	0.028	-0.001	0.022	-0.012	-0.006
	(0.043)	(0.026)	(0.059)	(0.028)	(0.009)	(0.028)	(0.018)	(0.026)
$\widehat{\beta_4}$: Risks Info. + Status Quo	0.019	0.006	0.076	0.021	0.001	0.054	0.010	-0.008
	(0.039)	(0.026)	(0.055)	(0.028)	(0.009)	(0.027)	(0.017)	(0.026)
$H_0: \widehat{\alpha_1} = \widehat{\alpha_2}$	0.261	0.241	0.952	0.618	0.241	0.133	0.377	0.842
$H_0: \widehat{\beta_1} = \widehat{\beta_2}$	0.729	0.990	0.333	0.444	0.541	0.411	0.059	0.709
$H_0: \widehat{\beta_3} = \widehat{\beta_4}$	0.330	0.767	0.478	0.809	0.820	0.230	0.205	0.925
Observations	2,462	2,462	2,462	2,462	2,462	2,462	2,462	2,462
Control Group Mean	4.73	0.78	4.51	0.73	0.98	0.73	0.92	0.79

Notes: Table 14 shows intention-to-treat estimates on PAP smear knowledge, for our survey sample. The panels in the table show estimates of equations (1), (2) and (3). In columns 1 and 2 the dependent variable is women's perception about the PAP smear importance (on a 1-5 scale), and an indicator for whether women think it is very important (value 5 on the 1-5 scale). In columns 3 and 4 the dependent variable is their perception on whether the PAP smear can save lives (on a 1-5 scale), and an indicator for whether women think it definitely does (value 5 on the 1-5 scale). The columns 5 to 8 include estimates on additional variables related to PAP knowledge: whether women report to know the PAP, whether they report to know that the PAP is free at health centers, whether they are aware they can skip one day of work to take the test and whether they believe that they should take the PAP smear every year. The lower panel displays the p-values of F-tests for the differences between treatment parameters. All regressions control for age and the set of demographics displayed in Table 4. Robust standard errors are in parentheses. ***, ** and * indicate statistical significance at the 1, 5 and 10 percent level respectively.

The results in columns 1 to 4 indicate that, independent of treatment status, women perceived the Pap smear to be very important and that most believed that having the test can save lives. None of the coefficients are significant at conventional levels, and their magnitude is very small relative to the control mean. For example, women in the control group rated the importance of the Pap 4.7 out of 5, and 78 percent reported the Pap test is very important. Similarly, when asked to rate whether the Pap could save lives, women in the control group reported a mean of 4.5, and 73 percent reported that they believed the Pap smear would definitely save lives.

The results for columns 5 to 8, which include estimates for additional variables related to PAP knowledge, also show zero effects. Essentially women in the control group reported knowing about PAP tests (98 percent), with 73 percent indicating they knew that the PAP is free, 92 percent being aware that they can take one day off work to take the test, and 79 percent believing that they should have a Pap smear every year.

These results suggest that the information component in our messages did not provide women additional knowledge. This finding implies that information did not play a main role in our treatment effects. We hypothesize that rather than providing new information, our messages brought the PAP exam to their minds, encouraging women to have the test.

Perceptions about Public Services. We also explored whether our intervention had effects beyond the content of the messages (i.e., beyond information and transaction costs). The hypothesis is that by contacting users, our intervention may have improved women's perceptions of public health services and the government, which in turn could mediate some of our main treatment effects. Table 15 shows small, non-significant effects on women's perceptions of the quality or safety of public services and on the confidence users have in ASSE services or whether they believe that the government cares about them.

Table 15. ITT Effects on Perceptions about e-Government and ASSE Services

	ASSE Quality:		ASSE Confidence:		eGovt Safety:		Govt Cares:	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	On a 1-5 Scale	High+ (4, 5 Value)	On a 1-5 Scale	High+ (4, 5 Value)	On a 1-5 Scale	High+ (4, 5 Value)	On a 1-5 Scale	High+ (4, 5 Value)
$\widehat{\delta}$: Treatment	0.052	0.018	0.014	0.001	-0.005	-0.026	0.112	0.034
	(0.054)	(0.023)	(0.052)	(0.024)	(0.070)	(0.025)	(0.072)	(0.025)
$\widehat{\alpha_1}$: Info. + Lower Trans. Cost	0.071	0.011	0.056	0.020	0.027	-0.017	0.114	0.028
	(0.059)	(0.025)	(0.056)	(0.026)	(0.077)	(0.027)	(0.079)	(0.027)
$\widehat{\alpha_2}$: Info. + Status Quo	0.034	0.026	-0.028	-0.019	-0.037	-0.035	0.110	0.041
	(0.059)	(0.025)	(0.057)	(0.026)	(0.077)	(0.027)	(0.079)	(0.027)
$\widehat{\beta_1}$: Benef. Info. + Lower Trans. Cost	0.077	0.019	0.068	0.025	0.117	0.019	0.109	0.030
	(0.068)	(0.029)	(0.064)	(0.030)	(0.089)	(0.031)	(0.092)	(0.031)
$\widehat{\beta_2}$: Risks Info. + Lower Trans. Cost	0.064	0.003	0.044	0.015	-0.063	-0.052*	0.118	0.025
	(0.067)	(0.029)	(0.065)	(0.030)	(0.088)	(0.031)	(0.090)	(0.031)
$\widehat{\beta_3}$: Benef. Info. + Status Quo	0.053	0.025	0.004	-0.004	-0.048	-0.039	0.121	0.055*
	(0.068)	(0.029)	(0.066)	(0.030)	(0.089)	(0.031)	(0.092)	(0.031)
$\widehat{\beta_4}$: Risks Info. + Status Quo	0.014	0.026	-0.061	-0.034	-0.027	-0.031	0.099	0.027
	(0.069)	(0.029)	(0.067)	(0.030)	(0.088)	(0.031)	(0.090)	(0.031)
$H_0 : \widehat{\alpha_1} = \widehat{\alpha_2}$	0.443	0.470	0.070	0.066	0.299	0.414	0.956	0.554
$H_0 : \widehat{\beta_1} = \widehat{\beta_2}$	0.844	0.571	0.713	0.739	0.038	0.024	0.919	0.871
$H_0 : \widehat{\beta_3} = \widehat{\beta_4}$	0.574	0.994	0.343	0.338	0.810	0.819	0.812	0.381
Observations	2,462	2,462	2,462	2,462	2,462	2,462	2,462	2,462
Control Group Mean	3.86	0.69	3.93	0.65	3.05	0.54	2.98	0.40
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Table 15 shows intention-to-treat estimates on PAP smear knowledge, for our survey sample. The panels in the table show estimates of equations (1), (2) and (3). In columns 1 and 2 the dependent variable is women's perception about the PAP smear importance (on a 1-5 scale), and an indicator for whether women think it is very important (value 5 on the 1-5 scale). In columns 3 and 4 the dependent variable is their perception on whether the PAP smear can save lives (on a 1-5 scale), and an indicator for whether women think it definitely does (value 5 on the 1-5 scale). The columns 5 to 8 include estimates on additional variables related to PAP knowledge: whether women report to know the PAP, whether they report to know that the PAP is free at health centers, whether they are aware they can skip one day of work to take the test and whether they believe that they should take the PAP smear every year. The lower panel displays the p-values of F-tests for the differences between treatment parameters. All regressions control for age and the set of demographics displayed in Table 4. Robust standard errors are in parentheses. ***, ** and * indicate statistical significance at the 1, 5 and 10 percent level respectively.

Women rated the quality of ASSE services 3.9 out of 5, and almost 70 percent indicated that the quality was high or very high. In addition, women reported having high or very high confidence in ASSE (65 percent). Results in columns 1 to 4 show that there are non-detectable differences between the control group and any of the treatment groups on these secondary outcomes.

We also asked women to rate how safe they perceived electronic public services to be. The average rating was 3 out of 5, and 54 percent believed that those services are safe or very safe to use. Columns 5 and 6 show that the estimates for perceived safety of e-government services are indistinguishable from zero at conventional significance levels. Finally, we asked respondents whether they felt that, more generally, the government cared about them. The average rating was 3 out of 5; however, only 40 percent reported believing that the government really cares (the high or very high categories). Again there was with no distinguishable between the control group and the treatment groups.

These estimates suggest the treatment effects on our main outcomes are not driven by enhanced perceptions about ASSE or the government. Overall, the results from exploring mechanisms suggest that reduced transaction costs for appointments is the major driver behind our treatment effects on women's behavior.

Heterogeneity with Survey Covariates

Finally we studied heterogeneous treatment effects by women's characteristics that are invariant to our intervention. Tables 16 and 17 show ITT estimates from separate regressions by subgroup for scheduling (columns 1 and 2) and attending (columns 4 and 5) PAP appointments. We divided the women into subgroups using the corresponding sample median. For example, women were considered older if their age was equal to or higher than the sample median and younger if their age was lower than the median. All the coefficients reported in columns 1, 2, 4, and 5 are individually statistically significant. Columns 3 and 6 present the difference between these subgroup effects, computed from fully interacted regressions. Table 16 presents results by age, education, income, and household size and composition (proxied by whether children live at home), while Table 17 provides estimates for frequency of internet use, use of cell phone to connect, and commuting time from home to the local clinic center.

Table 16. Heterogeneous ITT Estimates

Panel 1: Age

	Appointments			Take-up		
	(1)	(2)	(3)	(4)	(5)	(6)
Coefficient	Younger	Older	Dif. (1)-(2)	Younger	Older	Dif. (4)-(5)
$\widehat{\delta}$: Treatment	0.095	0.044	0.051***	0.066	0.035	0.031**
	(0.013)	(0.011)	(0.017)	(0.011)	(0.009)	(0.014)
$\widehat{\alpha_1}$: Information + Lower Transaction Cost	0.098	0.081	0.018	0.058	0.060	-0.002
	(0.017)	(0.015)	(0.023)	(0.013)	(0.013)	(0.018)
$\widehat{\alpha_2}$: Information Only	0.091	0.006	0.085***	0.073	0.009	0.064***
	(0.017)	(0.010)	(0.020)	(0.015)	(0.009)	(0.017)

Panel 2: Education

	Appointments			Take-up		
	(1)	(2)	(3)	(4)	(5)	(6)
Coefficient	Higher	Lower	Dif. (1)-(2)	Higher	Lower	Dif. (4)-(5)
$\widehat{\delta}$: Treatment	0.076	0.044	0.031*	0.053	0.036	0.017
	(0.009)	(0.017)	(0.019)	(0.008)	(0.013)	(0.015)
$\widehat{\alpha_1}$: Information + Lower Transaction Cost	0.099	0.068	0.031	0.069	0.043	0.026
	(0.013)	(0.021)	(0.024)	(0.011)	(0.016)	(0.019)
$\widehat{\alpha_2}$: Information Only	0.052	0.019	0.033	0.038	0.030	0.008
	(0.010)	(0.018)	(0.021)	(0.009)	(0.015)	(0.017)

Panel 3: Household Income

	Appointments			Take-up		
	(1)	(2)	(3)	(4)	(5)	(6)
Coefficient	Higher	Lower	Dif. (1)-(2)	Higher	Lower	Dif. (4)-(5)
$\widehat{\delta}$: Treatment	0.068	0.051	0.017	0.047	0.047	0.000
	(0.009)	(0.018)	(0.021)	(0.008)	(0.013)	(0.015)
$\widehat{\alpha_1}$: Information + Lower Transaction Cost	0.100	0.053	0.047*	0.064	0.045	0.020
	(0.013)	(0.022)	(0.026)	(0.011)	(0.016)	(0.020)
$\widehat{\alpha_2}$: Information Only	0.037	0.050	-0.012	0.030	0.049	-0.019
	(0.010)	(0.022)	(0.024)	(0.009)	(0.017)	(0.019)

Panel 4: Household Size

	Appointments			Take-up		
	(1)	(2)	(3)	(4)	(5)	(6)
Coefficient	Larger	Smaller	Dif. (1)-(2)	Larger	Smaller	Dif. (4)-(5)
$\widehat{\delta}$: Treatment	0.081	0.038	0.043**	0.054	0.037	0.017
	(0.010)	(0.014)	(0.017)	(0.008)	(0.012)	(0.014)
$\widehat{\alpha_1}$: Information + Lower Transaction Cost	0.107	0.058	0.049**	0.067	0.046	0.021
	(0.014)	(0.018)	(0.023)	(0.012)	(0.015)	(0.019)
$\widehat{\alpha_2}$: Information Only	0.054	0.020	0.034*	0.040	0.028	0.012
	(0.012)	(0.015)	(0.019)	(0.010)	(0.013)	(0.017)

Panel 5: Children at Home

	Appointments			Take-up		
	(1)	(2)	(3)	(4)	(5)	(6)
Coefficient	Yes	No	Dif. (1)-(2)	Yes	No	Dif. (4)-(5)
$\widehat{\delta}$: Treatment	0.082	0.053	0.029*	0.053	0.044	0.009
	(0.013)	(0.011)	(0.017)	(0.012)	(0.008)	(0.014)
$\widehat{\alpha_1}$: Information + Lower Transaction Cost	0.096	0.083	0.013	0.057	0.061	-0.004
	(0.017)	(0.015)	(0.023)	(0.015)	(0.012)	(0.019)
$\widehat{\alpha_2}$: Information Only	0.067	0.024	0.043**	0.048	0.027	0.022
	(0.016)	(0.011)	(0.019)	(0.014)	(0.009)	(0.017)

Notes: Table 16 show ITT estimates from separate regressions by subgroup on PAP appointments (columns 1 and 2) and take up rates (columns 4 and 5) in Panels 1 to 5 . We divide women in subgroups using the corresponding sample median. For example, women are in the older group if their age is equal or higher than the sample median and in the younger group if not. All the coefficients reported in these columns are individually statistically significant. In columns 3 and 6, we report the difference between these subgroup effects, computed from fully interacted regressions. Robust standard errors are in parentheses. For columns 3 and 6, ***, ** and * indicate statistical significance at the 1, 5 and 10 percent level respectively.

Table 17. Heterogeneous ITT Estimates

Panel 1: Internet Use

	Appointments			Take-up		
	(1)	(2)	(3)	(4)	(5)	(6)
Coefficient	1	0	Dif. (1)-(2)	1	0	Dif. (4)-(5)
$\widehat{\delta}$: Treatment	0.070	0.020	0.051**	0.051	0.020	0.031
	(0.009)	(0.024)	(0.025)	(0.007)	(0.024)	(0.025)
$\widehat{\alpha_1}$: Information + Lower Transaction Cost	0.095	0.038	0.056*	0.062	0.038	0.023
	(0.012)	(0.030)	(0.032)	(0.010)	(0.030)	(0.031)
$\widehat{\alpha_2}$: Information Only	0.046	0.003	0.043*	0.040	0.003	0.037
	(0.010)	(0.024)	(0.026)	(0.009)	(0.024)	(0.026)

Panel 2: Cell Phone Use

	Appointments			Take-up		
	(1)	(2)	(3)	(4)	(5)	(6)
Coefficient	1	0	Dif. (1)-(2)	1	0	Dif. (4)-(5)
$\widehat{\delta}$: Treatment	0.071	0.025	0.046**	0.051	0.025	0.026
	(0.009)	(0.018)	(0.020)	(0.007)	(0.018)	(0.019)
$\widehat{\alpha_1}$: Information + Lower Transaction Cost	0.095	0.043	0.053*	0.062	0.043	0.019
	(0.012)	(0.024)	(0.027)	(0.010)	(0.024)	(0.026)
$\widehat{\alpha_2}$: Information Only	0.046	0.007	0.039*	0.040	0.007	0.033
	(0.010)	(0.018)	(0.021)	(0.009)	(0.018)	(0.020)

Panel 3: Distance to Health Center (less 30 min or more)

	Appointments			Take-up		
	(1)	(2)	(3)	(4)	(5)	(6)
Coefficient	Close	Far	Dif. (1)-(2)	Close	Far	Dif. (4)-(5)
$\widehat{\delta}$: Treatment	0.070	0.015	0.056	0.049	0.032	0.017
	(0.009)	(0.033)	(0.034)	(0.007)	(0.026)	(0.026)
$\widehat{\alpha_1}$: Information + Lower Transaction Cost	0.093	0.041	0.053	0.058	0.062	-0.004
	(0.012)	(0.039)	(0.041)	(0.010)	(0.033)	(0.034)
$\widehat{\alpha_2}$: Information Only	0.047	-0.011	0.057	0.039	0.003	0.037
	(0.010)	(0.034)	(0.035)	(0.009)	(0.025)	(0.027)

Notes: Table 17 show ITT estimates from separate regressions by subgroup on PAP appointments (columns 1 and 2) and take up rates (columns 4 and 5) in Panels 1 to 5 . We divide women in subgroups using the corresponding sample median. For example, women are in the older group if their age is equal or higher than the sample median and in the younger group if not. All the coefficients reported in these columns are individually statistically significant. In columns 3 and 6, we report the difference between these subgroup effects, computed from fully interacted regressions. Robust standard errors are in parentheses. For columns 3 and 6, ***, ** and * indicate statistical significance at the 1, 5 and 10 percent level respectively.

The main finding was heterogeneous effects on scheduling appointments that shrink and are not statistically distinguishable from zero when examining appointment attendance rates. The ITT effects on scheduling appointments tended to be higher for younger women, for those who are more educated or have higher income levels, and for women living in larger households or with children. Effects are also higher for women who reported using the internet with higher frequency, those who use their cell phone to connect, and for women living closer to the health center.

6.3. Estimated Costs and Cost Effectiveness

Table 18 provides estimates of intervention costs and compares them with the estimated results produced by the intervention. Costs fall into four categories: (i) development of the online appointment tool, (ii) sending and monitoring the text messages, (iii) incremental service delivery costs for ASSE, and (iv) the private time costs for women of scheduling and attending appointments.

Table 18. Estimated Costs and Cost Effectiveness

Estimated Costs	
Number of treatment arm participants	22,800
Total cost of delivering intervention (online tool + text messages + incremental service)	US$74,396
Online Appointment Tool	
Firm contract	$50,000.00
ASSE and AGESIC staff time	$8,498.88
Total online tool costs	$58,498.88
Online tool cost per participant	$2.57
Text Messages	
Text message platform and message credits	$2,227.30
ASSE staff time	$5,269.68
Total text message delivery costs	$7,496.98
Text message cost per participant	$0.33
Incremental Service Delivery	
Additional ASSE capacity	$8,400.00
Incremental service delivery costs per participant	$0.37
Private Costs (per participant)	
Schedule appointment	$0.07
Schedule and attend appointment	$0.34
Total Cost per Participant	
All costs	$3.60
Excluding cost of online appointment tool	$1.04
Cost Effectiveness	
Impact on participants attending appointments (ITT)	0.023
Cost per appointment attended because of intervention (ITT) (all costs)	$156.67
Cost per appointment attended because of intervention (ITT) (without online tool)	$45.12
Impact on participants attending appointments (LATE)	0.08
Cost per appointment attended because of intervention (LATE) (all inclusive)	$45.04
Cost per appointment attended because of intervention (LATE) (without online tool)	$12.97

Design and implementation of the online appointment tool was outsourced to an external firm, supervised by staff from ASSE and AGESIC. The US$58,499 cost for the online tool included the direct costs of hiring the external firm as well as the opportunity costs of staff involved in supervising the firm. We report cost-effectiveness results with and without this cost because

ASSE continues to use the online appointment tool for purposes beyond this experiment, thus diluting the extent to which its total costs should be imputed to this project alone. Our preferred cost-effectiveness estimate excludes this cost because, in the long run, use of the tool for this specific experiment will likely be a small fraction of its use by ASSE affiliates at large for the full array of appointment types.

The US$7,497 cost for text messages was similarly broken down into the direct cost of the text messaging platform and messages, procured on a per-message basis from an external firm, and the opportunity cost of ASSE staff involved in preparing the messages for delivery and consolidating basic data on each delivery.

The incremental service delivery costs were associated with the short-term need to boost ASSE's capacity to analyze the additional PAP smears generated by the intervention. The cost of the contracts for four professionals hired for this purpose was US$8,400. All other actions associated with the intervention were managed within ASSE's existing capacity and therefore there were no additional costs incurred with the intervention.

The total cost of delivering the intervention of US$74,396 (the total of the three components above) divided by the number of participants in all treatment arms (N=22,800) provides the estimated per participant cost of delivering the intervention of US$3.26. As discussed above, eliminating the cost of designing and implementing the online appointment tool reduces the cost to an estimated US$0.70 per participant.

To provide a fuller understanding of the costs associated with the intervention, we also estimated the private costs incurred by participants to schedule and attend appointments. Generating these estimates required assumptions about how much time it took to make an appointment online (15 minutes) or in person (60 minutes, including transportation to and from), how long an appointment took (2 hours, including transportation to and from), and an appropriate average opportunity cost (US$3 per hour), which we based on responses provided in the survey. Though we used a uniform opportunity cost for all women, this cost was assumed by employers for those who were working and by the women themselves for those not working. We estimated the average private cost per scheduled appointment to be US$0.07 and per attended appointment to be US$0.34. Adding the total private costs to the costs discussed above provides an overall estimated cost of US$3.60 per participant or US$1.04 excluding the cost of the online appointment tool. Note that, to maintain the uniformity of the analysis, this exercise distributes the private costs of those who made and attended appointments over all treatment arm participants.

Dividing the estimated costs per participant (US$3.60 and US$1.04, with and without the cost of the appointment tool, respectively) by the overall ITT effect on appointments attended

(0.023) produces an estimated cost per PAP appointment attended as a result of our intervention. This cost-effectiveness ratio is US$156.67 including all costs or US$45.12 excluding the cost of the online appointment tool. For our survey sample, the ITT estimate on attendance is 4.7 pp; therefore, the ratios are US$76.67 and US$22.12, respectively. These estimates are conservative because we used the ITT estimate in the denominator. If we used the LATEs estimate, the ratios could be about as half as small. Overall, the strategy appears to be highly cost effective compared to the related literature (e.g., Scoggins et al., 2010).

7. Conclusions

A critical issue in public policy and development economics is how to increase the take up of public services. This experiment addressed this problem with an intervention designed to increase demand for an important preventive health care service: cervical cancer screening.

We partnered with the Government of Uruguay to design and implement an online appointment system. We then conducted a field experiment at scale with 47,600 participating women, randomized to a pure control group and four treatment arms. We encouraged women to either schedule an appointment for a Pap smear using the online system or to make an appointment as usual. The messages also included either benefit or risk information regarding PAP smears.

Our empirical results show that women randomized to the online appointment system were three times more likely to attend an appointment they scheduled for a Pap smear than the pure control group and twice as likely as those reminded to schedule an appointment as usual. We also found a precise zero difference when comparing benefit and risk messages.

The fact that the women in the treatment groups scheduled and attended appointments at a sizably higher rate than those in the pure control group is relevant because it suggests that women were facing barriers that were at least partially diminished by our treatment arms.

Exploring mechanisms, we found evidence that suggests a priori that women were well informed about the Pap smear and its importance. Moreover, women in the control group and the treatment arms were also aware that taking a Pap smear is cost-free and of the mortality rates associated with cervical cancer. Therefore, the information component of the intervention does not appear to have played a defining role.

Our results suggest that the encouragement element of our messages played the main role in driving our treatment effects. For those who received the status quo encouragement, we hypothesize that our messages brought the PAP exam to their minds, encouraging women to have the test. We interpret their change in behavior as a result of a salience effect, which has been shown to be important in family health behaviors (Fadlon and Nielsen, 2019) and also in other contexts, such as parent-school communications (Bettinger, Cunha, Lichand, et al., 2020), energy consumption (Allcott and Taubinsky, 2015), and taxation (Farhi and Gabaix, 2020).

For the women who were encouraged to make appointments using the online tool, the salience effects were supplemented with the benefits of diminishing transaction costs. We see this piece of evidence as an important contribution to the literature because it highlights the potential of investing in digital government to improve the uptake of public services. Our research complements recent studies assessing the use of technology to improve state capacity in developing countries (Callen et al., 2020; Muralidharan et al., 2016, 2020).

Our findings have direct policy implications. In a context of rapid advances in technology, governments are developing the capacity to store and use abundant data. The results suggest that investing in digital government tools that provide more than just information, reminders, or planning interventions might generate the most meaningful impacts on public service delivery.

References

Abadie, A, S Athey, GW Imbens, and J Wooldridge. 2017. When Should You Adjust Standard Errors for Clustering? NBER Working Papers 24003. Cambridge, MA: National Bureau of Economic Research (NBER).

AGESIC. 2019. Nuevos datos: resultados de la 5ª Encuesta de Conocimientos, Actitudes y Prácticas de Ciudadanía Digital (New data: results of the 5th Survey of Knowledge, Attitudes and Practices of Digital Citizenship). Montevideo, Uruguay: Agencia de Gobierno Electrónico y Sociedad de la Información y del Conocimiento (Agency of Electronic Government and Society of Information and Knowledge). https://www.gub.uy/agencia-gobierno-electronico-sociedad-informacion-conocimiento/comunicacion/noticias/nuevos-datos-resultados-5a-encuesta-conocimientos-actitudes-practicas

Allcott, H, and DTaubinsky. 2015. Evaluating Behaviorally Motivated Policy: Experimental Evidence from the Lightbulb Market. *American Economic Review* 105(8):2501–38.

Altmann, S, and C Traxler. 2014. Nudges at the Dentist. *European Economic Review* 72:19–38.

American Cancer Society. 2019. Cancer Prevention & Early Detection Facts & Figures 2019-2020. Atlanta, GA: American Cancer Society.

Arbulo, V, G Castelao, I Oreggioni, and JP Pagano. 2015. Uruguay: Building Up the National Integrated Health System. Health Systems Governance & Financing, WHO.

Artagaveytia, C, and A Toledo. 2018. Informe Cobertura Poblacional Del SNIS Según Prestador (SNIS Population Coverage Report According to the Provider). Montevideo: Ministerio de Salud de Uruguay.

Athey, S, and G Imbens. 2017. Chapter 3: The Econometrics of Randomized Experiments. In: E Duflo and A Banerjee (eds.), *Handbook of Economic Field Experiments*, Volume 1, pp.73–140.

Barrios, E, and M Garau. 2017. Cáncer: magnitud del problema en el mundo y en Uruguay, aspectos epidemiológicos (Cancer: Magnitude of the Problem in the World and in Uruguay, Epidemiological Aspects). Annals of the Faculty of Medicine, University of the Republic of Uruguay 4(1):9–46. https://doi.org/10.25184/anfamed2017.4.1

Benia, W, and G Tellechea. 2000. Estudio de los factores de riesgo para cáncer de mama y cuello uterino en mujeres usuarias de tres policlínicas barriales de Montevideo (Study of Risk Factors for Breast and Cervical Cancer in Female Users of Three Neighborhood Polyclinics of Montevideo). *Medical Journal of Uruguay* 16:103–13.

Berterretche, R, and A Sollazzo. 2012. El abordaje de la Atención Primaria de Salud, modelos organizativos y prácticas: caso de un Centro de Salud público urbano de Montevideo, Uruguay 2011 (The Approach to Primary Health Care, Organizational Models and Practices: Case of an Urban Public Health Center in Montevideo, Uruguay 2011). *Saúde em Debate* 36(94):461–72. https://dx.doi.org/10.1590/S0103-11042012000300017

Beshears, J, KL Milkman, and J Schwartzstein. 2016. Beyond Beta-Delta: The Emerging Economics of Personal Plans. *American Economic Review* 106(5):430–34.

Bettinger, E, N Cunha, G Lichand, and R Madeira. 2021. Are the Effects of Informational Interventions Driven by Salience? University of Zurich, Department of Economics, Working Paper No. 350. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3633821

Beuermann, DW, P Garcia, JP Lu, R Anta, A Maffioli, and MF Rodrigo. 2020. Information and Communication Technologies, Prenatal Care Services, and Neonatal Health. *Journal of Economics, Race, and Policy* 3(1):41–59.

Blanco, M, and J Vargas. 2014. Can SMS Technology Improve Low Take-up of Social Benefits? *Peace Economics, Peace Science and Public Policy* 20(1):61–81. https://doi.org/10.1515/peps-2013-0060

Bossuroy, T, C Delavallade, and V Pons. 2019. Biometric Tracking, Healthcare Provision, and Data Quality: Experimental Evidence from Tuberculosis Control. NBER Working Paper No. 26388. Cambridge, MA: NBER.

Buchmueller, TC, and L Goldzahl. 2018. The Effect of Organized Breast Cancer Screening on Mammography Use: Evidence from France. *Health Economics* 27(12):1963–80. https://doi.org/10.1002/hec.3813

Busso, M, J Cristia, and S Humpage. 2015. Did You Get Your Shots? Experimental Evidence on the Role of Reminders. *Journal of Health Economics* 44:226–37. doi:10.1016/j.jhealeco.2015.08.005. PMID: 26519909.

Bronchetti, E, D Huffman, and E Magenheim. 2015. Attention, Intentions, and Follow-Through in Preventive Health Behavior: Field Experimental Evidence on Flu Vaccination. *Journal of Economic Behavior & Organization* 116(C):270–91.

Bruhn, M, and D McKenzie. 2009. In Pursuit of Balance: Randomization in Practice in Development Field Experiments. *American Economic Journal: Applied Economics* 1(4):200–32. Available at: http://www.aeaweb.org/articles.php?doi=10.1257/app.1.4.200

Cadena, B, and B Keys. 2015. Human Capital and the Lifetime Costs of Impatience. *American Economic Journal: Economic Policy* 7(3):126–53. Available at: https://www.aeaweb.org/articles?id=10.1257/pol.20130081

Callen, M, S Gulzar, A Hasanain, MY Khan, and A Rezaee. 2020. Data and Policy Decisions: Experimental Evidence from Pakistan. *Journal of Development Economics* 146. Available at: https://doi.org/10.1016/j.jdeveco.2020.102523

Comisión honoraria de lucha contra el cancer. 2016. Registro Nacional de Cáncer. Incidencia del Cáncer en Uruguay. Periodo 2012–2016. Available at: http://www2.comisioncancer.org.uy/uc_515_1.html

Cragg, JG, and SG Donald. 1993. Testing Identifiability and Specification in Instrumental Variables Models. *Econometric Theory* 9(2):222–40.

Currie, J. 2004.The Take Up of Social Benefits. NBER Working Paper 10488. Cambridge, MA: NBER. Available at: http://www.nber.org/papers/w10488

DellaVigna, S, and MD Paserman. 2005. Job Search and Impatience. *Journal of Labor Economics* 23(3):527–88.

Duflo, E, R Glennerster, and M Kremer. 2007. Using Randomization in Development Economics Research: A Toolkit. *Handbook of Development Economics* 4:3895–962.

Dupas, P. 2011. Health Behavior in Developing Countries. *Annual Review of Economics* 3:425–49.

Fadlon, I, and TH Nielsen. 2019. Family Health Behaviors. *American Economic Review* 109(9):3162–91.

Farhi, E, and X Gabaix. 2020. Optimal Taxation with Behavioral Agents. *American Economic Review* 110(1):298–336.

Ferlay, J, E Steliarova-Foucher, J Lortet-Tieulent, S Rosso, JW Coebergh, H Comber, et al. 2012. Cancer Incidence and Mortality Patterns in Europe: Estimates for 40 Countries in

2012. *European Journal of Cancer*49(6):1374–403. doi:10.1016/j.ejca.2012.12.027. PMID: 23485231

Figlio, D, K Karbownik, and K Salvanes. 2017. The Promise of Administrative Data in Education Research. *Education Finance and Policy* 12(2):129–36.

Finkelstein, A, and MJ Notowidigdo. 2019. Take-Up and Targeting: Experimental Evidence from SNAP. *The Quarterly Journal of Economics* 134(3):1505–56. https://doi.org/10.1093/qje/qjz013

Garau, M, C Musetti, R Alonso, and E Barrios. 2019. Trends in Cancer Incidence in Uruguay: 2002–2015. *Colombia Medica* 50(4):224–38. Available at: https://doi.org/10.25100/cm.v50i4.4212

Goldfarb, A, and C Tucker. 2019. Digital Economics. *Journal of Economic Literature* 57(1):3–43. DOI:10.1257/jel.20171452

Huf, S, RS Kerrison, D King, T Chadborn, A Richmond, D Cunningham, et al. 2020. Behavioral Economics Informed Message Content in Text Message Reminders to Improve Cervical Screening Participation: Two Pragmatic Randomized Controlled Trials. *Preventive Medicine* 139. Available at: https://doi.org/10.1016/j.ypmed.2020.106170

Kremer, M, G Rao, and F Schilbach. 2019. Chapter 5: Behavioral Development Economics. In: BD Bernheim, S DellaVigna, and D Laibson (eds.), *Handbook of Behavioral Economics: Applications and Foundations*, Volume 2, pp.345-458. Available at: https://doi.org/10.1016/bs.hesbe.2018.12.002

Milkman, KL, J Beshears, JJ Choi, D Laibson, and BC Madrian. 2011. Using Implementation Intentions Prompts to Enhance Influenza Vaccination Rates. *Proceedings of the National Academy of Sciences* 108(26):10415–20.

Milkman, KL, J Beshears, JJ Choi, D Laibson, and BC Madrian. 2012. Following Through on Good Intentions: The Power of Planning Prompts. Cambridge, MA: NBER.

Muralidharan, K, P Niehaus, and S Sukhtankar. 2016. Building State Capacity: Evidence from Biometric Smartcards in India. *American Economic Review* 106(10):2895–929. DOI:10.1257/aer.20141346

Muralidharan, K, and P Niehaus. 2017. Experimentation at Scale. *Journal of Economic Perspectives* 11(31):103–24.

Muralidharan, K, P Niehaus, S Sukhtankar, and J Weaver. 2020. Improving Last-Mile Service Delivery Using Phone-Based Monitoring. *American Economic Journal: Applied Economics* (Forthcoming).

Muralidharan, K, M Romero, and K Wüthrich. 2020. Factorial Designs, Model Selection, and (Incorrect) Inference in Randomized Experiments. Working Paper 26562. Cambridge, MA: National Bureau of Economic Research (NBER). Available at: https://www.nber.org/papers/w26562.

OECD/The World Bank. 2020. *Health at a Glance: Latin America and the Caribbean 2020*. Paris: OECD Publishing. Available at: https://doi.org/10.1787/6089164f-en

Rivers, SE, P Salovey, DA Pizarro, J Pizarro, and TR Schneider. 2005. Message Framing and Pap Test Utilization among Women Attending a Community Health Clinic. *Journal of Health Psychology* 10(1):65–77. https://doi.org/10.1177/1359105305048556

Rivoir, A, and J Landinelli. 2017. ICT-Mediated Citizen Engagement, Case Study: Open Government National Action Plan in Uruguay. ACM International Conference Proceeding Series, v.: Part F1280, 214–17. New Delhi, India.

Rodríguez, G, C Caviglia, R Alonso, A Sica, S Segredo, I León, and I Musé. 2015. Conocimientos, actitudes y prácticas sobre el test de Papanicolaou y estadificación del cáncer de cuello uterino. *Revista Médica del Uruguay* 31(4):231–40.

Roseth, B, A Reyes, and C Santiso. 2018. *Wait No More: Citizens, Red Tape and Digital Government*. Washington, DC: Inter-American Development Bank. Available at: https://publications.iadb.org/en/wait-no-more-citizens-red-tape-and-digital-government

Santiago, P, B Ávalos, T Burns, A Morduchowicz, and T. Radinger. 2016. Chapter 1: School Education in Uruguay. In: *OECD Reviews of School Resources: Uruguay*. Paris: OECD Publishing.

Saslow, D, D Solomon, H Lawson, et al. 2012. American Cancer Society, American Society for Colposcopy and Cervical Pathology, and American Society for Clinical Pathology Screening Guidelines for the Prevention and Early Detection of Cervical Cancer. *A Cancer Journal for Clinicians* 62:147–72.

Scoggins, JF, SD Ramsey, JC Jackson, and VM Taylor. 2010. Cost Effectiveness of a Program to Promote Screening for Cervical Cancer in the Vietnamese-American Population. *Asian Pacific Journal of Cancer Prevention* 11(3):717–22.

Smith, RA, K Andrews, D Brooks, CE DeSantis, SA Fedewa, J Lortet-Tieulent, et al. 2016. Cancer Screening in the United States, 2016: A Review of Current American Cancer Society Guidelines and Current Issues in Cancer Screening. *A Cancer Journal for Clinicians* 66(2):96–114. DOI: 10.3322/caac.21336

Stock, JH, and M Yogo. 2005. Testing for Weak Instruments in Linear IV Regression. In: JH Stock and DW Andrews (eds.), *Identification and Inerence for Econometric Models: Essays in Honor of Thomas J. Rothenberg*, Chapter 5. Cambridge University Press.

United Nations. 2018. United Nations e-Government Survey 2018: E-government in Support of Sustainable Development. New York: United Nations.

White, MC, ML Shoemaker, and VB Benard. 2017. Cervical Cancer Screening and Incidence by Age: Unmet Needs Near and After the Stopping Age for Screening. *External American Journal of Preventive Medicine* 53(3):392–5.

WHO. 2013. *WHO Guidelines for Screening and Treatment of Precancerous Lesions for Cervical Cancer Prevention*. Geneva: World Health Organization.

WHO. 2014. *Comprehensive Cervical Cancer Control: A Guide to Essential Practice, Second edition*. Geneva: WHO.

WHO. 2020. Global Strategy to Accelerate the Elimination of Cervical Cancer as a Public Health Problem. Geneva: WHO. Available at: https://www.who.int/publications/i/item/9789240014107

World Bank. 2003. World Development Report 2004: Making Services Work for Poor People. Washington, DC: World Bank.

World Bank. 2019. World Development Indicators Online database. Washington, DC: World Bank. Available at: https://datacatalog.worldbank.org/dataset/world-development-indicators

World Bank. 2018. Socio-Economic Database for Latin America and the Caribbean (CEDLAS and The World Bank). Washington, DC: World Bank. Accessed in July 2020 from: http://www.cedlas.econo.unlp.edu.ar/wp/en/estadisticas/sedlac/estadisticas