Garrido and Gutiérrez *Lat Am Econ Rev* (2019) 28:1
https://doi.org/10.1186/s40503-019-0064-6

Latin American Economic Review

RESEARCH Open Access

Time goes by so slowly (for those who wait): a field experiment in health care

Sofía Garrido[1] and Emilio Gutiérrez[2*]

*Correspondence:
emilio.gutierrez@itam.mx
[2] Department of Economics, ITAM, Camino a Santa Teresa 930, Héroes de Padierna, CDMX CP 10700, Mexico
Full list of author information is available at the end of the article

Abstract

We exploit a unique field experiment to recover the willingness to pay (WTP) for shorter waiting times at a cataract detection clinic in Mexico City, and compare the results with those obtained through a hypothetical dichotomous choice questionnaire. The WTP to avoid a minute of wait obtained from the field experiment ranges from 0.59 to 0.82 Mexican pesos (1 USD = 12.5 Mexican pesos at the time of the survey), while that from the hypothetical choice experiment ranges from 0.33 to 0.48 Mexican pesos. WTP to avoid the wait is lower for lower income individuals, and it is larger the more accurately the announced expected waiting time matches the true values. Finally, we find evidence that the marginal disutility of waiting is not constant.

JEL Classification: I14, I18, I19

1 Background

Obtaining reliable measures of the cost implied by waiting times in the Mexican health-care sector seems particularly relevant given recent changes in the health-care system. For instance, after a law requiring a prescription for antibiotics was passed in 2010, pharmacy-adjacent doctors' offices (PADOs) expanded rapidly across the county (being almost inexistent in the mid-2000s). Their success[1] is attributed, among other things, to a very large difference in waiting times at doctors' offices between PADOs and public clinics (Pérez-Cuevas et al. 2012).[2]

Economists understand the willingness to pay (WTP) for a specific good or a good's attribute as a monetary measure of the value that consumers assign to it. Ideally, to recover such measures from data, econometricians can rely on variation in prices and goods' attributes readily available on the market and relate them to consumers' choices. Nonetheless, a consistent estimation of the WTP from observed choices requires sufficient variation in the goods' attributes and prices that is uncorrelated with other factors that could influence WTP. When this variation is unavailable from real-world data, WTP measures obtained from hypothetical choice experiments are very commonly used, and growingly so in the field of health economics. However, they have been widely criticized, as they are likely to deliver biased estimates for a variety of reasons.

[1] According to the industry's estimates, roughly the same number of patients visit PADOs every day in Mexico as the main social security system's outpatient clinics (about 300,000 daily visits).

[2] In the USA, some of the arguments regarding the potential cost of the health-care reform proposed and passed by the Obama administration revolved precisely around the potential increase in waiting times to receive health-care services.

This paper exploits a unique field experiment in which individuals were allowed to pay a randomly assigned price to avoid the waiting time to be seen by a physician at a cataract detection clinic in Mexico City to recover the WTP for shorter waiting times. In addition, it compares the estimates obtained from these real choices in the field with those obtained through a hypothetical dichotomous choice questionnaire administered to patients from the same clinic (throughout the text, we refer to this second as a contingent valuation (CV) exercise).

Our findings according to the field experiment indicate that the clinic's patients' WTP to avoid a minute of wait ranges from 0.59 to 0.82 Mexican pesos.[3] Participants in the hypothetical choice experiment are significantly less responsive to variations in price and waiting times, and the point estimates for the WTP to avoid a minute of wait in this case ranges from 0.33 to 0.48 Mexican pesos. While our experiment may suffer from lack of external validity, we present a series of heterogeneity tests to explore how informative our results may be for the Mexican context. For instance, while our sample of patients is drawn from the lower tail of the income distribution in the city, we find suggestive evidence that, for lower income individuals in our sample, the WTP to avoid the wait is lower. The WTP to avoid a minute of wait is larger the more accurately the announced expected waiting time matches the true parameters. Finally, we find evidence that the marginal disutility of waiting is not constant, casting doubt on the appropriateness of indirect measures of the cost of waiting times, such as forgone wages.

The rest of the paper is presented as follows. The next section motivates the need for consistent measures of WTP to shorten waiting times in health care. Section III discusses the existing methods to recover such measures, and the potential advantages and caveats associated with doing so in a dichotomous choice setting, both through hypothetical choice questionnaires and actual choices. Section IV describes the context for the field and hypothetical choice experiments conducted for this paper. Section V presents the empirical strategy and results. Section VI discusses concerns of external validity and performs heterogeneity analyses. The last section concludes the paper.

2 Motivation

Scholars have devoted their attention to understanding how different policies may have an impact on waiting times in health care (see for example Propper et al. 2002; Hurst and Siciliani 2003; Siciliani 2007; Siciliani et al. 2009; Brekke, et al. 2008). However, while measuring the direct impact of these policy interventions on waiting times may be useful per se, a full cost–benefit analysis requires quantifying welfare gains from changes in waiting times for the relevant population (Cullis et al. 2000).

Recovering consistent estimates of WTP is a challenging task. Ideally, to recover such measures from data, econometricians can rely on variation in prices and goods' attributes readily available on the market and relate them to consumers' choices. Nonetheless, a consistent estimation of the WTP from observed choices requires sufficient variation in the goods' attributes and prices that is uncorrelated with other factors that could influence WTP. When this variation is unavailable from real-world data, WTP measures

[3] 1 USD = 12.5 Mexican pesos at the time of the survey.

obtained from hypothetical choice experiments are very commonly used, and growingly so in the field of health economics. In light of this, an assessment of whether WTP measures recovered from hypothetical choices accurately reflect true valuations may contribute to the academic debate on the reliability of contingent valuation techniques.

3 Existing literature recovering WTP

Many theorists have relied on measures of the cost of waiting times in health care using concrete measures of the opportunity cost that they entail, such as forgone wages. However, while useful from a theoretical perspective, such approximations fail to take into account that the specific conditions under which time is spent (or lost) may have an influence on the cost that waiting represents for patients. The disutility from sitting in a doctor's office waiting to be seen may differ substantially from the utility of time spent in another context. Moreover, the marginal disutility of waiting in a doctor's office may not be constant (an extra minute of wait after having waited for an hour may not cause the same disutility than that of the first minute in the waiting room). The cost of waiting may then be context specific and very different from forgone wages.[4]

Some researchers have exploited real-world data to recover estimates of the cost of waiting times. Deacon and Sonstelie (1985) exploit a mandated price decrease specifically for Chevron gas stations in California in the 1980s (which implied a large increase in the waiting times at these stations, and not at others where there was no price decrease) to estimate the value of time spent waiting for gasoline purchases. Apart from the fact that their study is performed in a very different context than a doctor's office, in their data the variation in waiting times and prices is only present across gas stations. Aguiar and Hurst (2007) exploit scanner data to structurally estimate the cost of time spent grocery shopping. Besley et al. (1999) show that longer waiting times in the British National Health Service (NHS) are associated with larger purchases of private health insurance, and calculate patients' willingness to pay for shorter waiting times. Their results assume that the characteristics of private health insurance do not vary with price, and they rely on the assumption that private and public care only differ in terms of the waiting times.

Since the WTP for some goods or goods' attributes (such as waiting times for health care) is hard to recover from real data, over the last decades and increasingly so in health economics, researchers have tried to recover such measures from hypothetical surveys, which use a wide array of techniques to ask individuals about their "reservation price" for a specific good or good's attributes. While widely used, an ongoing debate regarding the elicitation of such questions has been taking place for decades. Open-ended questions are widely believed to deliver biased estimates. However, since the National Oceanic and Atmospheric Administration (NOAA) panel (Arrow et al. 1993) performed a critical review of the existing methods to recover WTP through hypothetical surveys, dichotomous choice (DC) questionnaires are perhaps seen as the most reliable

[4] For instance, Aguiar and Hurst (2007) document that the least educated consume more leisure, which they find at odds with standard predictions from income and substitution effects, suggesting that the conditions under which leisure is spent may affect its marginal utility. Ramey and Francis (2009) also characterize the evolution of leisure in the USA during the past 100 years. Lee et al. (2012) show that, for the Japanese context, the marginal rate of substitution between work hours and different kinds of non-work activities (i.e., leisure and in-home production) differs.

alternative, both because of their simplicity and reduced incentives for strategic behavior (Hoehn and Randall 1987; Carson et al. 1999).

DC questionnaires have been widely used by health economists to recover measures of WTP for a variety of health-care attributes. For example, Propper (1990) estimates WTP for shorter waiting times in the NHS waiting list, and Bishai and Lang (2000) estimate differences in WTP for shorter waiting times for cataract surgery in Canada, Denmark and Spain. Johannesson et al. (1991) also estimate the WTP for shorter waiting times in the Swedish health-care system through an experiment of this kind.[5]

However, the literature that questions the validity of contingent valuation methods to recover WTP is also widespread (Portney 1994, Cummings et al. 1995; Klose 1999; Ryan et al. 2004; Donaldson and Shackley 2002; Smith 2005; Harrison and Rutström 2008; Hausman 2012).

Three main concerns arise when recovering WTP from hypothetical surveys: "hypothetical bias", which simply refers to the fact that individuals' responses to hypothetical choices may not fully correspond to their behavior in real life (for example, subjects may choose to please the interviewer or may infer that their answers could have a policy impact); the consistent difference in estimates obtained from "willingness to pay" and "willingness to accept" questionnaires, which can be generalized as evidence of subjects' sensitivity to the framing of the hypothetical choice questions; and the difficulty to correctly isolate in the questions' wording the good's attribute for which the WTP wants to be recovered (Hausman 2012).

While some theory-based techniques to assess the consistency of WTP estimates obtained through hypothetical surveys have been proposed (Diamond and Hausman 1994), to determine whether a hypothetical survey delivers biased estimates of WTP, one would ideally know the true values for respondents. The existing literature circumvents the problem by conducting laboratory experiments, in which respondents' true value is recovered through experiments where the choice is real, and then compared to those obtained through hypothetical surveys. But the debate on whether the results obtained from laboratory experiments can be extended to the real world is large and growing. In particular, while within a laboratory the researcher has full control over the environment under which choices are made, this is never the case in the real world. A bias in hypothetical surveys may arise in the field regardless of its absence in a laboratory. Smith and Mansfield (1998) do compare estimates from a hypothetical choice questionnaire and real choices, recovering the willingness to accept spending time answering to a phone interview, finding no significant differences between hypothetical and real choices.

This paper contributes then to this literature by effectively randomizing the price faced by individuals when making a decision in the field about whether to wait or not to be seen by a physician, and comparing the estimates obtained with those from a hypothetical choice experiment. Both exercises use patients of the same cataract detection clinic in Mexico City as subjects, and they both keep all characteristics of the service provided constant, except for the waiting time and the price for not having to wait.

[5] See Olsen and Smith (2001) and Diener et al. (1998) for a review of this literature.

4 The experiments

4.1 The field experiment

The field experiment was conducted during the last 3 weeks in October 2014, from Friday to Saturday, at a health-care facility in Mexico City, specializing in cataract detection and surgery. Patients of this clinic arrive at the reception desk and are announced the expected waiting time to be seen by a doctor. This waiting time is calculated by the clinic's personnel, based on the number of patients in the waiting room and the number of doctors at the clinic. Patients generally stay in a waiting room until their name is called. From the clinic's records, no patients chose to leave the clinic after this waiting time was announced, even before the option to pay to avoid the wait was offered.

In January 2013, this clinic introduced a new product, which consisted in the possibility of paying $300 Mexican pesos (approximately, 25 US dollars) to be seen by one of the doctors without having to wait. The field experiment consisted in randomizing the price at which this product was offered.

When patients arrived at the reception desk,[6] apart from being announced the expected waiting time to be seen by the physician, they were also informed that the clinic was offering a "no waiting time consultation" at promotional prices. This offer consisted in a lottery that assigned to each patient, randomly, a different price for this product. The prices offered were $200, $250 and $300 (the baseline price) Mexican pesos ($1 USD = 12.5 Mexican pesos). Because all patients could potentially interact in the waiting room, the receptionist informed them that the price offered was a "promotion aimed at improving their experience at the clinic" and that a different promotional price was offered, randomly, to each patient.[7] All patients were explicitly informed that the quality of the service, apart from the difference in waiting times, would be identical from the one offered to the rest of the patients.

All of the patients at the clinic, paying and non-paying, were asked by the physician to fill in a questionnaire that captures some basic socioeconomic characteristics. The sample obtained through the field experiment consists of 279 patients that arrived at the clinic individually. Table 1 presents the descriptive statistics.

Not surprisingly, patients were sensitive to the price of the non-waiting offer: 6, 12 and 24 percent of those offered this product for $300, $250 and $200 Mexican pesos, respectively, chose to pay for it. As expected, due to the random assignment of prices to patients, individuals did not seem to differ in any other observable characteristic, including the waiting time announced at the arrival to the clinic.

4.2 The contingent valuation exercise

Two weeks after the field experiment took place, a hypothetical contingent valuation (CV) questionnaire was administered to a comparable sample of 251 patients (all arriving to the clinic seeking an appointment with an ophthalmologist). The implementation of this survey also lasted three full weeks, from Friday to Saturday. Subjects participating

[6] Patients waited in line, outside the clinic, approximately for 10 min before being seen by the receptionist. All announced waiting times only considered the wait after patients registered at the reception desk.

[7] This feature of the experimental setting implies that our results should be interpreted with caution. Part of the patients' response to the price offered may be driven by its "promotional" nature.

Table 1 Descriptive statistics

	Price offered for the non-waiting time consult (Mexican pesos)					
	200		250		300	
	Mean	SD	Mean	SD	Mean	SD
Paid	0.24***		0.12		0.06***	
Waiting time (in minutes)	115.58	[6.40]	111.21	[5.33]	122.59	[6.98]
Age	60.68	[1.86]	61.43	[1.57]	61.78	[1.97]
Gender (male = 1)	0.39		0.34		0.36	
Household head	0.42		0.36		0.41	
Was accompanied	0.84		0.76		0.81	
Responsible for own health decisions	0.58		0.51		0.53	
Number of kids	3.16	[0.24]	3.27	[0.22]	3.19	[0.28]
Dirt floor	0.42		0.51		0.53	
Observations	95		99		85	

Standard errors of means in brackets

* Significant at 10%

** Significant at 5%

*** Significant at 1% for the test of that category against the rest

in the hypothetical choice survey did so before being informed that the non-waiting time consultation was available. Surveyors approached them before their arrival to the reception desk and stated that the clinic had an interest in improving the experience of future patients and asked if they were willing to answer a brief questionnaire. All subjects agreed[8] to participate in the short survey, which consisted in asking each patient the following:

> *"If the expected waiting time to be seen by the doctor today was T hours, would you pay P pesos to not have to wait and be seen by the doctor right away, or choose to wait T hours at no cost?"*

Prices (P) and waiting times (T) were randomly assigned to questionnaires in this hypothetical exercise. The point values for the price of the hypothetical non-waiting consultation were the same as those offered in the field experiment: 89, 84 and 82 subjects were assigned a 200, 250 and 300 pesos price, respectively. The waiting times randomly assigned to questionnaires were chosen to lie on the same range as those announced during the field experiment. In particular, 42, 45, 35, 30, 45, and 56 questionnaires stated a waiting time of 90, 120, 150, 180, 240, and 300 min, respectively. Because patients were not randomly assigned to the field experiment or the contingent valuation exercise, Table 2 shows the descriptive statistics of the socioeconomic variables listed in Table 1, this time comparing patients participating in each of the experiments.

The two samples differ significantly on the fraction of individuals that chose to pay to avoid waiting, although they also differ in the average waiting time announced, which

[8] The 100% compliance rate may seem surprising. Nonetheless, it is a result of the fact that participating individuals were waiting in line, outside the clinic, before being seen by the receptionist.

Table 2 Descriptive statistics

	Field experiment	CV exercise
Paid	0.14**	0.22**
Waiting time (in minutes)	116.16 [3.58]***	188.01 [4.85]***
Age	61.28 [1.03]	61.58 [1.03]
Gender (male = 1)	0.37	0.3
Household head	0.40 [0.03]**	0.50**
Was accompanied	0.8	0.75
Responsible for own health decisions	0.54	0.53
Number of kids	3.21 [0.14]	3.07 [0.14]
Dirt floor	0.48	0.42
Observations	279	251

Standard errors of means in brackets

* Significant at 10%

** Significant at 5%

*** Significant at 1% for the difference in means test

is higher for the CV sample[9] that chose to pay more frequently. Apart from that, small differences are observed in the rest of the variables, and only the fraction of interviewed individuals that declared to be the head of their household differs significantly between samples.

5 Empirical strategy and results

5.1 Random utility model (RUM) framework

In the context analyzed, both for the field and hypothetical choice experiments, the basic formulation of consumers' utility that can allow for recovering their WTP to avoid waiting at the doctor's office can be embodied in a random utility model (RUM) of the following form:

$$U_{ia}(t_{ia}, p_{ia}, X_i) = V_{ia}(t_{ia}, p_{ia}, X_i) + \varepsilon_{ia},$$

where U_{ia} is the utility that individual i derives from choosing alternative a, at a price p_{ia}, with a waiting time of t_{ia} minutes, X_i is a vector of observable characteristics, and ε_{ia} is an error term, which captures unobserved heterogeneity in individuals' preferences.

For empirical purposes, it is common practice to assume that time and money are linearly separable in the individuals' utility function, and that the marginal utility of both is constant across alternatives and individuals. For the specific context analyzed, we can incorporate these assumptions by describing the utility from each alternative as:

$$U_{i1}(t_{i1}, p_{i1}, X_i) = \alpha_1 + \beta_1 t_{i1} + \beta_2 p_{i1} + F_1(X_i) + \varepsilon_{i1}$$

and

$$U_{i2}(t_{i2}, p_{i2}, X_i) = \alpha_2 + \beta_1 t_{i2} + \beta_2 p_{i2} + F_2(X_i) + \varepsilon_{i2}.$$

Given this setup, β_1 and β_2 are the marginal utilities of waiting times and price, respectively, and the WTP to avoid a unit of wait can be easily computed as:

[9] This is the result of the fact that higher and lower waiting times were assigned to the CV questionnaires with similar probabilities, while, in the field experiment, longer waiting times are relatively less common.

$$\text{WTP} = -\text{MRS}_{t,p} = \frac{\beta_1}{\beta_2},$$

which is interpreted in consumer choice theory as the units of currency that the individual is willing to pay to avoid a unit of wait.

In our setting, an individual will choose to pay to avoid waiting (alternative 1) over not paying and waiting to be seen by the physician (alternative 2) when:

$$U_{i1}(t_{i1}, p_{i1}, X_i) \geq U_{i2}(t_{i2}, p_{i2}, X_i)$$
$$\text{or}$$
$$\varepsilon_{i2} - \varepsilon_{i1} \leq V_{i1}(t_{i1}, p_{i1}, X_i) - V_{i2}(t_{i2}, p_{i2}, X_i).$$

Given the parametrization of the individuals' utility function above, and given that t_{i1}, the expected waiting time for the non-waiting time consultation, and p_{i2}, the price individuals pay for the consultation if they choose to wait, are both equal to zero; individuals will choose to pay for the non-waiting time consultation if:

$$\varepsilon_{i2} - \varepsilon_{i1} \leq (\alpha_1 - \alpha_2) - \beta_1 t_{i2} + \beta_2 p_{i1} + F_1(X_i) - F_2(X_i).$$

Assuming a specific distribution for $(\varepsilon_{i2} - \varepsilon_{i1})$ and a functional form for $(F_1(X_i) - F_2(X_i))$, it is then possible to estimate the marginal disutility of time spent waiting (β_1) and the marginal disutility of the price paid for the consultation (β_2), and thus compute the WTP to avoid waiting: $\left(\frac{\beta_1}{\beta_2}\right)$.

It is common practice to assume (as we do in our empirical analysis) that the ε_i are independently and identically extreme-value distributed, so that the difference $(\varepsilon_{i2} - \varepsilon_{i1})$ is distributed logistically. Then, defining P_1 as the probability that an individual will choose alternative 1 (the non-waiting time consultation, in this case) implies that:

$$P_1 = \frac{1}{1 + \mathrm{e}^{-((\alpha_1 - \alpha_2) - \beta_1 t_{i2} + \beta_2 p_{i1} + F_1(X_i) - F_2(X_i))}}$$

and all the relevant parameters can be estimated through a logit regression.[10]

5.2 Empirical specification

In this particular case, to allow preferences to vary with respect to individual characteristics, and to directly test for potential differences in WTP in the field and in the hypothetical surveys, we assume that the difference in utility between the non-waiting and waiting consultation alternatives can be parametrized as:

$$\Delta U_i = \alpha_1 + \alpha_2 \text{CV}_i + \beta_1 \text{Time}_i + \beta_2 \text{Price}_i + \beta_3 \text{Time}_i * \text{CV}_i + \beta_4 \text{Price}_i * \text{CV}_i + \sum_n \delta_n \text{Control}_{ni} + e_i,$$

where Time measures the announced waiting time (in minutes) for individuals participating in the field experiment, and the hypothetical waiting time listed in the hypothetical choice setting for the participants in the CV exercise. Price indicates the price (in

[10] For a thorough discussion of random utility models and the use of logit to estimate them, see Train (2009).

Table 3 Logit regression results

Dependent variable:	Dummy = 1 if paid		
	1	2	3
Price	− 0.014638 [0.003328]***	− 0.014665 [0.003369]***	− 0.016325 [0.005320]***
Waiting time	0.009147 [0.002274]***	0.008618 [0.002270]***	0.013968 [0.004578]***
Price*contingent valuation	0.003635 [0.002286]	0.004291 [0.002305]*	0.008134 [0.006710]
Waiting time*contingent valuation	− 0.004812 [0.003025]	− 0.005167 [0.003025]*	− 0.010074 [0.005038]**
Observations	530	530	530
WTP-field experiment	0.62	0.59	0.86
Chi squared for test WTP = 0)	12.39***	10.98***	5.03**
WTP-CV exercise	0.39	0.33	0.48
Chi squared for test WTP = 0)	4.48**	3*	2.03
Difference in WTP between field and CV	0.23	0.25	0.38
Chi squared for test of difference in WTP = 0	1.97	2	0.57
Socioeconomic controls	No	Yes	Yes
Date fixed effects	No	No	Yes

Robust standard errors in brackets

For participants in the field experiment, the dependent variables are dummy, taking value of one if individuals paid for the non-waiting consult. For participants in the CV exercise, the dependent variables are dummy taking value of one if individuals declared they would have paid for the non-waiting consult, given the hypothetical price and waiting time announced

* Significant at 10%

** Significant at 5%

*** Significant at 1%

Mexican pesos) randomly assigned to each patient for the non-waiting time consultation in the field experiment, and the hypothetical price of the non-waiting time consultation for those participating in the CV exercise. CV is a dummy variable taking a value of one if the individual corresponds to the CV sample. $Control_n$ are the control variables listed in Table 1, and e_i is an error term.

As outlined above, this then implies that the probability of choosing the non-waiting time consultation can be written as:

$$P_1 = \Pr(\Delta U_i < 0) = \frac{1}{1 + e^{-(\alpha_1 + \alpha_2 CV_i + \beta_1 Time_i + \beta_2 Price_i + \beta_3 Time_i * CV_i + \beta_4 Price_i * CV_i + \sum_n \delta_n Control_{ni})}},$$

where the functional form follows from the common assumption on the distribution of the difference in the error terms. We can then estimate this equation via a logit regression, using the dummy variable indicating if individuals chose the non-waiting time consultation as our dependent variable.

Under the assumptions listed so far, we can recover estimates of the individuals' utility function arguments: the marginal disutility of time and money in the field experiment ($-\beta_1$ and β_2, respectively), the marginal disutility of hypothetical time and money in the CV exercise ($-\beta_1 - \beta_3$ and $\beta_2 + \beta_4$, respectively), and the implied WTP to avoid a minute of wait implied by the field experiment $\left(\frac{-\beta_1}{\beta_2}\right)$ and by the CV exercise $\left(\frac{-\beta_1 - \beta_3}{\beta_2 + \beta_4}\right)$.

5.3 Main results

Table 3 shows the results of the logit regression.[11] Column 1 includes no controls. Column 2 includes all control variables listed in Table 1, and column 3 additionally includes date fixed effects. As can be seen, throughout specifications, the coefficients on price and announced waiting time have the expected signs. Participants are less likely to pay for the non-waiting time consult when its (randomly assigned) price is higher, and more likely to do so when the announced expected waiting time is higher. The coefficients for the interactions between price and waiting time with the dummy variable indicating if the individuals' responses correspond to those in the CV exercise roughly suggest that individuals are less responsive to both price and waiting time in the hypothetical scenario. Throughout specifications, the coefficient associated with the assigned price and the interaction with the CV indicator is positive, and the coefficient associated with the interaction between the CV dummy and the announced waiting time is negative. When including socioeconomic variables as controls, both of these coefficients are significantly different from zero at a 10% confidence level. When we additionally include date fixed effects, although the coefficient for the interaction between price and the CV dummy loses significance, its sign and magnitude remain relatively stable.

As shown in Table 3, the implied WTP to avoid a minute of wait ranges between 0.59 and 0.86 Mexican pesos, while for the CV sample, it ranges from 0.33 to 0.48 Mexican pesos. The WTP recovered from the hypothetical survey is considerably lower than the one recovered from true choices. While we cannot reject the hypothesis that the WTP recovered from the field differs significantly from that recovered from the CV exercise, taking the results from column 2, we can reject the hypotheses that individuals respond similarly to price and announced waiting time in a hypothetical setting and when facing true choices. The results then cast doubt on the validity of WTP measures recovered from hypothetical choices.

To better put into perspective the WTP estimates recovered from the field experiment, a simple back-of-the-envelope calculation may be useful. Approximately, ten million patients are seen monthly only by physicians in The Mexican Social Security Institute[12] (IMSS, for its acronym in Spanish), which provides health care to 39.2 percent of the Mexican population.[13] According to the 2016 Mexican Health and Nutrition Survey (ENSANUT), the average wait to be seen by a physician in the IMSS system is approximately 70 min. Given our estimate of the cost per minute of wait (0.59 Mexican pesos), the monthly cost associated to waiting times in the IMSS system could roughly account to 413 million pesos. Given that in Mexico there were approximately 225,000 practicing physicians in 2014, and that the average monthly salary of a doctor in Mexico is 12,722 pesos,[14] the cost associated with waiting times only in the IMSS system is then approximately equivalent to 15 percent of the sum of all physicians' salaries in the country.

[11] The analogous results estimating a linear probability model through OLS are presented in Appendix Table 6.

[12] http://www.imss.gob.mx/conoce-al-imss/memoria-estadistica-2017.

[13] http://www.beta.inegi.org.mx/temas/derechohabiencia/.

[14] http://www.beta.inegi.org.mx/contenidos/saladeprensa/aproposito/2014/medico0.pdf.

6 External validity and heterogeneity analysis

In this section, we further explore the data recovered from the field experiment to investigate to what extent our results may be informative to public policy. All the estimates presented thus far are based on the assumption that the marginal disutility of time and money is constant for individuals with different observable characteristics. Testing for this assumption is particularly relevant in the context analyzed, as the experiment exploited in this paper was performed in a very particular setting: a private clinic where typical patients lack any kind of health insurance. The clinic's promotion strategies around the time we conducted the experiment were to visit remote, low income neighborhoods in the city to inform its citizens of the availability of the no-cost consultations. It is then perhaps not surprising that, according to Table 1, socioeconomic characteristics of participating subjects differ considerably from those of Mexico City's population. For example, more than 45 percent of participating subjects' dwellings have a dirt floor, while this number for Mexico City's population is less than 2 percent, according to the 2010 census. The extent to which the WTP estimates recovered through the experiment are informative for the Mexican health-care system crucially depends on how the estimated WTP differs by subjects' characteristics.

The model's assumptions can be easily relaxed to allow for heterogeneity in the WTP based on observables. In particular, we can then allow the marginal utility of time and money to vary with respect to observable characteristics by assuming that the differences in utility from alternative 1 versus alternative 2 can be expressed as:

$$\Delta U_i = \alpha_1 + \beta_1 \text{Time}_i + \beta_2 \text{Price}_i + \beta_3 \text{Time}_i * X_i + \beta_4 \text{Price}_i * X_i + \sum_n \delta_n \text{Control}_{ni} + e_i,$$

where all variables are defined as above and X is any observable characteristic of interest. The coefficients associated with the interaction between this variable and Time or Price (β_3 and β_4, respectively) will indicate if the marginal utility of time or money differs for individuals for which X takes different values. We explore if this is the case for men vs. women, and for individuals whose dwelling has a dirt floor. The latter is particularly relevant to explore if the WTP varies with the closest proxy for income available in our dataset. In addition, we explore if the WTP varies with respect to the uncertainty about whether the announced waiting time at the time of arrival to the clinic reflects the true waiting times that patients face. In particular, we compute the standard deviation in the announced waiting times within each of the days the experiment was conducted. Our assumption is that the variation we observe in the announced waiting times within a day can proxy for how much waiting times varied within each day and, thus, how close to the true waiting time the announced wait was. For ease of interpretation, we compute a dummy variable taking value of one if the standard deviation of waiting times within each day is above or below the sample median.

The results of the logistic regression restricting the sample to participants in the field experiment are presented in Table 4.[15] Column 1 does not have interaction of the price and time variables with any other. Column 2 presents the results by gender (interacting with an indicator for males). Column 3 shows estimates by type of floor in the

[15] The analogous results estimating a linear probability model through OLS are presented in Appendix Table 7.

Table 4 Logit regression results

Dependent variable:		Dummy = 1 if paid		
X		Male = 1	Dirt floor = 1	SD (time) > median
Price	− 0.018096 [0.005143]***	− 0.024269 [0.007189]***	− 0.014244 [0.007879]*	− 0.020965 [0.006027]***
Waiting time	0.008913 [0.002704]***	0.008095 [0.003407]**	0.013381 [0.004359]***	0.02406 [0.009977]**
Price*$X=1$		0.013637 [0.010504]	− 0.006929 [0.010473]	0.004882 [0.005208]
Waiting time*$X=1$		0.001624 [0.005709]	− 0.007391 [0.005604]	− 0.015095 [0.010421]
Observations	279	279	279	279
WTP − $X=0$	0.49	0.33	0.94	1.15
Chi squared for test WTP = 0	6.84***	4.17**	2.9*	7.1***
WTP − $X=1$		0.91	0.28	0.56
Chi squared for test WTP = 0		1.79	2.44	5.63**
Difference in WTP between $X=0$ and $X=1$		− 0.58	0.66	0.59
Chi squared for test of difference in WTP = 0		0.68	1.27	3.57*
Socioeconomic controls	Yes	Yes	Yes	Yes
Date fixed effects	Yes	Yes	Yes	Yes

Robust standard errors in brackets

The dependent variables are dummy taking value of one if individuals paid for the non-waiting consult

Column 2 presents the results by gender (interacting with an indicator for males). Column 3 shows estimates by type of floor in the individual's dwelling and column 4 shows interactions of the price and time variables with an indicator of whether the standard deviation in the announced waiting time during the day of the patient's visit was higher than the median in our sample

* Significant at 10%

** Significant at 5%

*** Significant at 1%

individual's dwelling, and column 4 shows interaction of the price and time variables with an indicator of whether the standard deviation in the announced waiting time during the day of the patient's visit was higher than the median in our sample. Results in column 2 suggest that women have a lower WTP for shorter waiting times than men, although the difference between the estimates is not significantly different from zero. Column 3 shows that lower income individuals are more responsive to the price, and less responsive to waiting times (although, again, the coefficients associated with these variables and the interaction between the dummy variable for dirt floor are not significantly different from zero). These estimates also imply that their willingness to pay for shorter waiting times is considerably lower than for the rest of the sample. Finally, column 4 shows that the WTP for shorter waiting times is significantly higher when the standard deviation in waiting times announced during the patients' day of visit is larger than the median. We interpret this last result as evidence that, if the uncertainty in the actual waiting times faced given the announcement can bias our estimates of WTP, this bias is likely to be reflecting a lower bound of the WTP in the context analyzed.

Table 5 Logit regression results

Dependent variable:	Dummy = 1 if paid		
	1	2	3
Price	−0.018514 [0.005022]***	−0.019084 [0.005248]***	−0.017525 [0.005634]***
Waiting time	0.041575 [0.012955]***	0.043092 [0.013647]***	0.044853 [0.016113]***
Waiting time squared	−0.00009 [0.000035]***	−0.000094 [0.000037]**	−0.000085 [0.000043]**
Observations	279	279	279
WTP-field experiment (at $T = 150$)	0.79	0.78	1.10
Chi squared for test WTP = 0	8.64***	8.44***	5.98**
Socioeconomic controls	No	Yes	Yes
Date fixed effects	No	No	Yes

Robust standard errors in brackets

WTP is calculated as the ratio of marginal utilities between time and money given the coefficients presented. The marginal utility of time is computed at $T = 150$

* Significant at 10%

** Significant at 5%

*** Significant at 1%

In addition, it is perhaps worth testing the assumption that the marginal disutility of waiting is constant, as rejecting this hypothesis may be informative of the appropriateness of using indirect measures for the cost of waiting, such as forgone wages. For this purpose, we rewrite the difference in utility from choosing alternative 1 vs 2 as:

$$\Delta U_i = \alpha_1 + \beta_1 \text{Time}_i + \beta_2 \text{Price}_i + \beta_3 \text{Time}_i^2 + \sum_n \delta_n \text{Control}_{ni} + \varepsilon_i.$$

We test the null hypothesis that $\beta_3 = 0$. In this particular context, rejection of the null indicates that the marginal disutility of waiting may not be constant.

Results are presented in Table 5.[16] Column 1 includes no controls, column 2 controls for all observable characteristics listed in Table 1, and column 3 additionally includes date fixed effects. Across specifications, we reject the hypothesis that the marginal disutility of waiting is constant, as the coefficient associated with the square of the waiting time is always negative and statistically different from zero. We interpret this last result as strong evidence against using indirect measures for the cost of time, such as forgone wages, for the welfare analysis of policies that may have an impact on waiting times.

7 Conclusions

In this paper, we recover estimates of the WTP for shorter waiting times at a cataract detection clinic in Mexico City through a field experiment and a contingent valuation exercise. Results from the field experiment indicate that the clinic's patients' WTP to avoid a minute of wait ranges from 0.59 to 0.82 Mexican pesos, and further heterogeneity analysis suggests that this estimate is likely a lower bound of the WTP of Mexico City's population in similar contexts. A simple back-of-the-envelope calculation suggests

[16] The analogous results estimating a linear probability model through OLS are presented in Appendix Table 8.

that reducing the average waiting time to be seen by a doctor at IMSS (Mexico's Social Security System) by 70 min, on average, would imply welfare gains for patients equivalent to 15 percent of the sum of all physicians' salaries in the country.

The estimates obtained cast doubt on the appropriateness of indirect measures of the cost of time, such as forgone wages. For instance, the average daily salary of full-time formal workers in Mexico City in 2014 was 366 pesos,[17] or 0.76 pesos per minute. Given that our sample is drawn from the lower tail of Mexico City's income distribution and that WTP is lower for relatively lower income participants in our experiment, indirect estimates of the cost of time obtained from forgone wages may heavily underestimate them. Moreover, our results suggest that the marginal disutility of waiting is not constant.

In addition, estimates recovered from the field experiment differ considerably from those obtained through the contingent valuation exercise. Participants in the hypothetical choice experiment are significantly less responsive to variations in price and waiting times, and the WTP to avoid a minute of wait according to the hypothetical choice experiment ranges from 0.33 to 0.48 Mexican pesos.

While our results then also cast doubt on the appropriateness of contingent valuation techniques to consistently recover WTP for goods or goods' attributes unavailable on the market, they should be taken with caution. Field experiments to recover WTP as the one exploited in this paper are generally very hard to implement, and even if implementable, they can be prohibitively costly. Rather than completely dismissing the appropriateness of CV questionnaires to recover WTP, we hope that in cases where the cost of large-scale field experiments is too costly, small-scale experiments as the one analyzed in this paper can be exploited to identify the best correction techniques to approximate WTP from data generated from CV questionnaires (Harrison 2006).

Authors' contributions
Both authors participated in the experimental design. SG supervised the field work and collected the data. Both authors worked in the data analysis. EG wrote the final version of the paper. Both authors read and approved the final manuscript.

Author details
[1] General Atlantic, Mexico City, Mexico. [2] Department of Economics, ITAM, Camino a Santa Teresa 930, Héroes de Padierna, CDMX CP 10700, Mexico.

Acknowledgements
We are grateful to Javier Okhuysen and Carlos Orellana for the opportunity to conduct the field experiment. Sofia Cerrilla provided outstanding assistance throughout the experiment's implementation. All errors are ours.
The title of this paper makes a reference to the lyrics of Madonna's song "Hung Up".

Competing interests
SG was employed at the clinic in which the experiment described in this paper was carried out. She received no compensation for conducting it. EG declares no competing interests.

Availability of data and materials
All data generated or analyzed during this study are included in this article.

Funding
Emilio Gutierrez is grateful for support from the Asociacion Mexicana de Cultura.

Appendix

See Tables 6, 7 and 8.

[17] http://archivo.eluniversal.com.mx/finanzas-cartera/2014/impreso/cinco-estados-concentran-los-salarios-mas-elevados-107715.html.

Table 6 OLS regression results

Dependent variable:	Dummy = 1 if paid		
	1	2	3
Price	− 0.001848 [0.000414]***	− 0.001804 [0.000402]***	− 0.001777 [0.000568]***
Waiting time	0.001293 [0.000385]***	0.001189 [0.000377]***	0.001461 [0.000450]***
Price*contingent valuation	0.000333 [0.000283]	0.000417 [0.000285]	0.000521 [0.000804]
Waiting time*contingent valuation	− 0.000543 [0.000500]	− 0.000588 [0.000493]	− 0.000827 [0.000541]
Observations	530	530	530
WTP-field experiment	0.70	0.66	0.82
F-statistic for test WTP = 0	8.98***	8.06***	5.56**
WTP-CV exercise	0.50	0.43	0.50
F-statistic for test WTP = 0	5.04**	3.57*	2.45
Difference in WTP between field and CV	0.20	0.23	0.32
F-statistic for test of difference in WTP = 0	0.339	0.326	0.5042
Socioeconomic controls	No	Yes	Yes
Date fixed effects	No	No	Yes

Robust standard errors in brackets

For participants in the field experiment, the dependent variables are dummy taking value of one if individuals paid for the non-waiting consult. For participants in the CV exercise, the dependent variables are dummy taking value of one if individuals declared they would have p aid for the non-waiting consult, given the hypothetical price and waiting time announced

Table 3 in the main text is the analog of this table. However, in this case we assume a linear probability model and estimate the regression through ordinary least squares (tables in the main text are estimated through a logit regression)

* Significant at 10%

** Significant at 5%

*** Significant at 1%

Table 7 OLS regression results

Dependent variable:		**Dummy = 1 if paid**		
X		**Male = 1**	**Dirt floor = 1**	**SD (time) > median**
Price	−0.001932 [0.000503]***	−0.002326 [0.000633]***	−0.001215 [0.000703]*	−0.002098 [0.000540]***
Waiting time	0.001209 [0.000336]***	0.001038 [0.000419]**	0.001607 [0.000500]***	0.002405 [0.000945]**
Price*$X = 1$		0.000987 [0.001036]	−0.001535 [0.001018]	0.000303 [0.000435]
Waiting time*$X = 1$		0.000438 [0.000719]	−0.000734 [0.000672]	−0.001218 [0.001005]
Observations	279	279	279	279
WTP − X	0.63	0.45	1.32	1.15
F-statistic for test WTP = 0	7.21***	4.2**	2.47	6.46**
WTP − $X = 1$		1.10	0.32	0.66
F-statistic for test WTP = 0		2.16	3.02*	6.51**
Difference in WTP between $X = 0$ and $X = 1$		− 0.66	1.01	0.49
F-statistic for test of difference in WTP = 0		0.71	1.36	2.18
Socioeconomic controls	Yes	Yes	Yes	Yes
Date fixed effects	Yes	Yes	Yes	Yes

Robust standard errors in brackets

The dependent variables are dummy taking value of one if individuals paid for the non-waiting consult Column 2 presents the results by gender (interacting with an indicator for males). Column 3 shows estimates by type of floor in the individual's dwelling and column 4 shows interactions of the price and time variables with an indicator of whether the standard deviation in the announced waiting time during the day of the patient's visit was higher than the median in our sample

Table 4 in the main text is the analog of this table. However, in this case we assume a linear probability model and estimate the regression through ordinary least squares (tables in the main text are estimated through a logit regression)

* Significant at 10%

** Significant at 5%

*** Significant at 1%

Table 8 OLS regression results

Dependent variable:	**Dummy = 1 if paid**		
	1	**2**	**3**
Price	−0.001967 [0.000498]***	−0.001954 [0.000499]***	−0.001801 [0.000518]***
Waiting time	0.003999 [0.001282]***	0.004076 [0.001282]***	0.003718 [0.001377]***
Waiting time squared	−0.000008 [0.000004]**	−0.000009 [0.000004]**	−0.000007 [0.000004]*
Observations	279	279	279
WTP-field experiment (at $T = 150$)	0.81	0.70	0.90
F-statistic for test WTP = 0	9.04***	8.65***	7.62***
Socioeconomic controls	No	Yes	Yes
Date fixed effects	No	No	Yes

Robust standard errors in brackets

WTP is calculated as the ratio of marginal utilities between time and money given the coefficients presented

The marginal utility of time is computed at $T = 150$

Table 5 in the main text is the analog of this table. However, in this case we assume a linear probability model and estimate the regression through ordinary least squares (tables in the main text are estimated through a logit regression)

* Significant at 10%

** Significant at 5%

*** Significant at 1%

Publisher's Note

Received: 27 March 2017 Accepted: 28 January 2019
Published online: 19 February 2019

References

Aguiar M, Hurst E (2007) Life-cycle prices and production. Am Econ Rev 97(5):1533–1559
Arrow K, Solow R, Portney PR, Leamer EE, Radner R, Schuman H (1993) Report of the NOAA panel on contingent valuation. Federal Register 58(10):4601–4614
Besley T, Hall J, Preston I (1999) The demand for private health insurance: do waiting lists matter? J Public Econ 72(2):155–181
Bishai DM, Lang HC (2000) The willingness to pay for wait reduction: the disutility of queues for cataract surgery in Canada, Denmark, and Spain. J Health Econ 19(2):219–230
Brekke KR, Siciliani L, Straume OR (2008) Competition and waiting times in hospital markets. J Public Econ 92(7):1607–1628
Carson RT, Flores NE, Mitchell RC (1999) The theory and measurement of passive-use value. Oxford University Press, Oxford, pp 97–130
Cullis JG, Jones PR, Propper C (2000) Waiting lists and medical treatment: analysis and policies. Handb Health Econ 1:1201–1249
Cummings RG, Harrison GW, Rutström EE (1995) Homegrown values and hypothetical surveys: is the dichotomous choice approach incentive-compatible? Am Econ Rev 85(1):260–266
Deacon RT, Sonstelie J (1985) Rationing by waiting and the value of time: results from a natural experiment. J Polit Econ 93(4):627–647
Diamond PA, Hausman JA (1994) Contingent valuation: Is some number better than no number? J Econ Perspect 8(4):45–64
Diener A, O'Brien B, Gafni A (1998) Health care contingent valuation studies: a review and classification of the literature. Health Econ 7(4):313–326
Donaldson C, Shackley P (2002) Should we use willingness to pay to elicit community preferences for health care?: New evidence from using a 'marginal' approach. J Health Econ 21(6):971–991
Harrison GW (2006) Experimental evidence on alternative environmental valuation methods. Environ Resour Econ 34(1):125–162
Harrison GW, Rutström EE (2008) Experimental evidence on the existence of hypothetical bias in value elicitation methods. Handb Exp Econ Results 1:752–767
Hausman J (2012) Contingent valuation: from dubious to hopeless. J Econ Perspect 26(4):43–56
Hoehn JP, Randall A (1987) A satisfactory benefit cost indicator from contingent valuation. J Environ Econ Manag 14(3):226–247
Hurst J, Siciliani L (2003) Explaining waiting times variations for elective surgery across OECD countries OECD economic studies, OECD Publishing, 2004(1), 95–123
Johannesson M, Jönsson B, Borgquist L (1991) Willingness to pay for antihypertensive therapy—results of a Swedish pilot study. J Health Econ 10(4):461–473
Klose T (1999) The contingent valuation method in health care. Health Policy 47(2):97–123
Lee J, Kawaguchi D, Hamermesh DS (2012) Aggregate impacts of a gift of time. Am Econ Rev 102(3):612–616
Olsen JA, Smith RD (2001) Theory versus practice: a review of 'willingness-to-pay'in health and health care. Health Econ 10(1):39–52
Pérez-Cuevas R, Doubova SV, Wirtz VJ, Dreser A, Mori EES, Ávila MH (2012) Consultorios médicos en farmacias privadas: efectos inesperados en el uso de servicios de salud y el acceso a medicamentos. INSP, Cuernavaca
Portney PR (1994) The contingent valuation debate: why economists should care. J Econ Perspect 8(4):3–17
Propper C (1990) Contingent valuation of time spent on NHS waiting lists. Econ J 100(400):193–199
Propper C, Croxson B, Shearer A (2002) Waiting times for hospital admissions: the impact of GP fundholding. J Health Econ 21(2):227–252
Ramey VA, Francis N (2009) A century of work and leisure. Am Econ J Macroecon 1(2):189–224
Ryan M, Scott DA, Donaldson C (2004) Valuing health care using willingness to pay: a comparison of the payment card and dichotomous choice methods. J Health Econ 23(2):237–258
Siciliani L (2007) Optimal contracts for health services in the presence of waiting times and asymmetric information. BE J Econ Anal Policy 7(1):1–27
Siciliani L, Stanciole A, Jacobs R (2009) Do waiting times reduce hospital costs? J Health Econ 28(4):771–780
Smith RD (2005) Sensitivity to scale in contingent valuation: the importance of the budget constraint. J Health Econ 24(3):515–529
Smith VK, Mansfield C (1998) Buying time: real and hypothetical offers. J Environ Econ Manag 36(3):209–224
Train KE (2009) Discrete choice methods with simulation. Cambridge University Press, Cambridge