NBER WORKING PAPER SERIES

USING PREFERENCE ESTIMATES TO CUSTOMIZE INCENTIVES: AN APPLICATION TO POLIO VACCINATION DRIVES IN PAKISTAN

James Andreoni
Michael Callen
Muhammad Yasir Khan
Karrar Jaffar
Charles Sprenger

Working Paper 22019
http://www.nber.org/papers/w22019

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
February 2016

We are grateful to Rashid Langrial (former Commissioner of Lahore), Muhammad Usman (District Coordinating Officer Lahore), and Zulfiqar Ali (Executive District Officer Health Lahore) for championing the project. We are grateful for financial support from the International Growth Center (IGC) and from the Department for International Development (DFID) Building Capacity for Research Evidence (BCURE) pilot program initiative. We are grateful to Evidence for Policy Design (EPoD), the Center for Economic Research Pakistan (CERP), Katy Doyle, and Sarah Oberst, for excellent logistical support. We are indebted to Michael Best, Eli Berman, Leonardo Bursztyn, Ernesto Dal Bó, Torun Dewan, Asim Fayaz, Frederico Finan, Ed Glaeser, Matthew Kahn, Supreet Kaur, Asim Khwaja, Aprajit Mahajan, Dilip Mookherjee, Andrew Newman, Gerard Padró-i-Miquel, Robert Powell, Matthew Rabin, Gautam Rao, Martin Rotemberg, Frank Schilbach, Erik Snowberg, Joshua Tasoff, Juan Vargas, Pedro Vicente, Nico Voigtlander, RomainWacziarg, Noam Yuchtman and seminar participants at UC San Diego, UC Berkeley, Claremont Graduate University, Harvard, the London School of Economics, Nova de Lisboa, Stanford, the Pacific Development Conference, the New England Universities Development Consortium, UCLA Anderson, Wharton, Princeton, and Boston University for extremely helpful comments. We are especially grateful to Rohini Pande for many conversations about this project. Hashim Rao and Danish Raza provided excellent research assistance. Research is approved by the Institutional Review Board at Harvard and at UC San Diego. The study is registered in the AEA RCT Registry as AEARCTR-0000417. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Using Preference Estimates to Customize Incentives: An Application to Polio Vaccination Drives in Pakistan
James Andreoni, Michael Callen, Muhammad Yasir Khan, Karrar Jaffar, and Charles Sprenger
NBER Working Paper No. 22019
February 2016, Revised August 2020
JEL No. D03,I1,O1

ABSTRACT

We use estimates of time preferences to customize incentives for polio vaccinators in Lahore, Pakistan. We measure time preferences using intertemporal allocations of effort, and derive the mapping between these estimates and individually optimized incentives. We evaluate the effect of matching contract terms to discounting parameters in a subsequent experiment with the same vaccinators. Our tailored policy is compared to alternatives that either rely on atheoretic reduced-form relationships for policy guidance or apply the same policy to all individuals. We find that contracts tailored to individual discounting outperform this range of policy alternatives.

James Andreoni
Department of Economics
University of California, San Diego
9500 Gilman Drive
La Jolla, CA 92093-0508
and NBER
andreoni@ucsd.edu

Michael Callen
London School of Economics
Houghton Street
London, WC2A 2AE
United Kingdom
and NBER
m.j.callen@lse.ac.uk

Muhammad Yasir Khan
University of California at Berkeley
2220 Piedmont Ave
Student Services Building 545
Berkeley, CA 94720
yasir.khan@berkeley.edu

Karrar Jaffar
USC Department of Economics
karrar.jaffar@gmail.com

Charles Sprenger
University of California, San Diego
Rady School of Business
9500 Gilman Drive
La Jolla, CA 93093
c.sprenger@gmail.com

A randomized controlled trials registry entry at https://www.socialscienceregistry.org/trials/417

1 Introduction

Intertemporal choice is a central topic in economics. The preference parameters governing such decisions affect a broad range of outcomes, justifying the considerable theoretical and empirical investments made to describe the level and shape of discounting.[1] An understanding of intertemporal preference parameters also provides valuable policy guidance. Indeed, a number of recent policies are motivated by empirical research on time preferences: commitment savings products, default retirement allocations, and the Save More Tomorrow retirement savings program are all partly motivated by the insight that time preferences may be 'present-biased' (for discussion and examples see, e.g., Laibson, 1997; Benartzi and Thaler, 2004; Beshears, Choi, Laibson and Madrian, 2009; Ashraf, Karlan and Yin, 2006; Blumenstock, Callen and Ghani, 2018).

Policy interventions such as those above may be further enhanced by using individualized, rather than broad, information on time preferences. Differences in experimental measures of time preferences correlate with differences in a number of policy-relevant behaviors such as take-up of commitment devices and credit card borrowing (examples include Chabris, Laibson, Morris, Schuldt and Taubinsky, 2008b; Meier and Sprenger, 2008, 2012, 2010; Ashraf et al., 2006; Dohmen, Falk, Huffman and Sunde, 2006; Castillo, Ferraro, Jordan and Petrie, 2011).[2] Such correlations suggest that interventions could leverage individual information on time preference to tailor unique policies for each person.

This paper studies the promise of theoretically-informed, individually-tailored policy interventions. We compare a tailored policy to alternatives that span the policy space on two

[1] Central examples of theoretical work include Samuelson (1937); Koopmans (1960); Laibson (1997) and O'Donoghue and Rabin (2001). Empirical exercises in field and laboratory settings focusing on parameter estimation include Hausman (1979); Lawrance (1991); Warner and Pleeter (2001); Cagetti (2003); Laibson, Repetto and Tobacman (2005); Mahajan and Tarozzi (2011); Harrison, Lau and Williams (2002); Andreoni and Sprenger (2012).

[2] It should be noted that none of these examples linking structural estimates of time preference to other behaviors provide an articulated model for what the precise correlation between the two values should be. Unlike our own efforts, such exercises could be conducted without appeal to structural estimation.

dimensions: broad vs. tailored and structural vs. atheoretic. A broad policy is one that is applied to all individuals, while a tailored policy is one individualized to each based on some characteristic. A structural policy is one that draws policy guidance from a theoretical model of preferences, while an atheoretic policy draws policy guidance from some reduced-form relationship or without prior information.[3]

Our project engages government health workers—termed Lady Health Workers (LHWs)—associated with polio eradication efforts for the Department of Health in Lahore, Pakistan.[4] The function of LHWs is to provide oral polio vaccine to children during monthly vaccination drives, which usually last two days. We introduce a monitoring and incentive system to measure intertemporal preferences using effort choices at work. Closely following the Convex Time Budget (CTB) design of Andreoni and Sprenger (2012); Augenblick, Niederle and Sprenger (2015), LHWs are asked to trade off work between the two days of a vaccination drive. Completion of their allocated work is tied to a bonus of 10 times their daily wage. For empirical realism, each LHW only makes a single CTB choice, which, under a set of assumptions, identifies their time preferences.[5] We then tailor policy based upon these measured preferences in a subsequent work decision. The tailored policy we examine attempts to equalize vaccinations over time by changing relative prices for each LHW. This structural, tailored policy is compared to the alternative polices noted above: broad policies which set relative prices to achieve the same objective based on aggregate structural estimates, or the reduced-form price-sensitivity of effort; and a tailored, but atheoretic policy, which sets relative prices based on a simple rule of giving higher relative prices to plausibly more patient individuals. These comparisons are facilitated by an overarching control group who receive a uniform random price, from which we

[3]Though far from an exhaustive labeling of potential policies, this 2-by-2 labeling helps to organize the comparisons we investigate.

[4]Polio is endemic in Pakistan. Of 350 new worldwide cases in 2014, 297 occurred in Pakistan, constituting a 'global public health emergency' according to the World Health Organization.Between 95 percent and 99 percent of individuals carrying polio are asymptomatic. One infection is therefore enough to indicate a substantial degree of ambient wild polio virus. The disease largely affects children under five.

[5]This design choice differs from most laboratory studies of time preferences where each subject makes multiple decisions.

draw subsets for comparison purposes.

In a sample of 337 LHWs, we document three principal findings. First, on aggregate, a present bias exists in allocations of vaccinations over time. LHWs choosing their work in advance of the first day of the vaccination drive allocate significantly fewer (more) vaccinations to the first (second) day than those allocating on the morning the drive actually commences. Corresponding estimates of aggregate present bias accord with those in prior laboratory studies. Second, there is substantial heterogeneity in time preferences across subjects. This sizable cross-sectional variation also resonates with prior experimental exercises. Third, and most importantly, our tailored policy works. Relative to a range of policy alternatives, our intervention generates behavior around 30% closer to the policy target of equal allocation. Interestingly, when focusing on conditions where subjects are asked to make allocations which take effect immediately – that is, when present bias is relevant – the tailored policy generates a roughly 50% improvement, eliminating the present bias in allocations.

This paper makes three contributions. First, our exercise uses field behavior about effort to examine time preferences, providing the first field implementation of the Augenblick et al. (2015) methods for measurement.[6] The consistency between our findings and prior lab experimental work indicates that the CTB design can be applied to real work decisions in the field. This finding helps to support the growing literature which identifies present bias from non-monetary choices in the field (Read and van Leeuwen, 1998; Sadoff, Samek and Sprenger, 2015; Read, Loewenstein and Kalyanaraman, 1999; Sayman and Onculer, 2009; Kaur, Kremer and Mullainathan, 2010, 2015; Carvalho et al., 2014).[7]

[6] Documenting dynamic inconsistency outside of the laboratory and outside of the standard experimental domain of time-dated monetary payments is particularly valuable given recent discussions on the elicitation of present-biased preferences using potentially fungible monetary payments (Cubitt and Read, 2007; Chabris, Laibson and Schuldt, 2008a; Andreoni and Sprenger, 2012; Augenblick et al., 2015; Carvalho, Meier and Wang, 2014).

[7]These studies include examination of present bias or dynamic inconsistency for food choices (Read and van Leeuwen, 1998; Sadoff et al., 2015); for highbrow and lowbrow movie choices (Read et al., 1999); for cafe reward choices (Sayman and Onculer, 2009); for completing survey items (Carvalho et al., 2014); and for fertilizer purchase decisions (Duflo, Kremer and Robinson, 2011). For discussion of this literature, see Sprenger (2015).

Second, we provide the first empirical evidence to date of value in structural, tailored policies for intertemporal choice. Given these results, there is clear opportunity to expand the scope of interventions beyond uniform strategies. The policy objective we consider attempts to implement a smooth allocation over time. Smoothing is a natural policy objective to consider for intertemporal decisions. For example, evidence suggests that consumption of Supplemental Nutritional Assistance Program (SNAP) benefits may be subject to present bias, leading to declining consumption during the benefit period (see, e.g., Shapiro, 2005; Hastings and Washington, 2010). Our results indicate that a tailored policy along the lines implemented here could help smooth benefit consumption. Such a policy could complement alternatives that have been discussed, including increasing the frequency of benefit payments (Shapiro, 2005).

Third, we provide techniques to deal with challenges that are likely to arise when trying to measure time preferences in the field using real work decisions. Unlike laboratory settings where sizable completion bonuses have been used to ensure near one-hundred percent completion rates, in our field setting even a completion bonus of 10 times the daily wage does not ensure uniform completion.[8] Roughly half of our subjects do not successfully complete their chosen allocations.[9] Such failure to complete has the potential to confound identification of time preferences as individuals may balance their true preferences against the failure probabilities induced by choice.[10] Recognizing this issue and the likelihood that other field elicitations will face a similar challenge, we develop and implement methodology to simultaneously esti-

[8]For college subjects Augenblick et al. (2015) employ bonuses \$100 in their six-week study and achieve 88% completion. In their follow-up work conditions they employ bonuses of \$60 for a three week study and achieve 95% completion. The lack of uniform completion was not a feature of the data we initially expected, but, in retrospect, is something we should have anticipated. Data from drives prior to our intervention, which were subject to almost no monitoring or scrutiny, showed that LHWs almost without exception hit their prescribed targets exactly. We believe these reports are at least partially driven by the fact that polio is a politicized issue in Pakistan, with a number of stakeholders and international donors being eager to demonstrate high numbers of vaccinations. Given our lack of foresight, neither the functional forms estimated for failure probabilities nor the implemented out-of-sample exercise predicting completion rates were in our study registration. As such, they should be viewed with appropriate caveats.

[9]Interestingly, our completion patterns also show hallmarks of present bias with immediate allocations being more frequently associated with non-completion than advance allocations.

[10]For example, in the extreme, if two full days worth of work are allocated to a single day, then the LHW should expect to fail.

mate parameters related to preferences and failure probabilities. Though in our setting, the required adjustments to measured time preferences due to completion issues are minimal, the methodology may be a valuable input to future field research.

The paper proceeds as follows: Section 2 presents our experimental design and corresponding theoretical considerations for estimating time preferences and tailoring contracts, Section 3 present results, Section 4 provides robustness tests, and Section 5 concludes.

2 Experimental Design

Our experiment has three components: implementing a high resolution smartphone monitoring system similar to that described in Callen, Gulzar, Hasanain, Khan and Rezaee (2019), eliciting individual discounting parameters using the Convex Time Budget (CTB) technique (Andreoni and Sprenger, 2012; Augenblick et al., 2015), and, after assigning tailored contracts to LHWs, testing whether these tailored contracts outperform comparison policies.

2.1 Vaccinations and Smartphone Monitoring

The Department of Health in Lahore, Pakistan, employs LHWs throughout the city to conduct polio vaccination drives. Every month there is a vaccination drive that is at least two days long. Prior to our study, the standard protocol for vaccination drives was to provide each LHW a fixed target for total vaccinations over the drive and a map of potential households (called a "micro-plan"). LHWs received no explicit benefits for reaching targets; they simply received a fixed daily wage of 100 rupees (around $1 at contemporary exchange rates). LHWs mapped their walk with pen and ink, knocking on each compound door, and vaccinating each child if their parents granted permission.[11] At the end of each day, LHWs in each neighborhood

[11] Vaccinating a child consists of administering a few drops of oral vaccine. As there is no medical risk of over-vaccination, LHWs are encouraged to vaccinate every child for whom permission is granted. For each attempted vaccination, LHWs were asked to mark information related to the attempt (number of children vaccinated, whether or not all children were available for vaccination, etc.) in chalk on the compound wall. Appendix

convened with their supervisor and self-reported their vaccination activity for the day.[12] In principle, a monitor could verify the claims.[13] In practice, however, there was virtually no monitoring, and reasons to suspect over-reporting.[14]

In collaboration with the Department of Health, we designed a smartphone-based monitoring system. Each LHW in our study was given a smartphone equipped with a vaccination monitoring application. The LHW was asked to record information related to each vaccination. Then, she was asked to take a picture of the home visited and her current vial of vaccine. An image of the main page of the application is provided as Figure 1, Panel A. Data from the smartphone system were aggregated in real-time on a dashboard available to senior health administrators.[15]

2.2 Drive 1: Intertemporal Bonus Contracts and Measurement of Preferences

We worked with the Department of Health to implement intertemporal bonus contracts in two-day drives in September, November and December of 2014. The contracts required workers to complete a present value total of $V = 300$ vaccination attempts in exchange for a fixed bonus of 1000 rupees. LHWs set daily targets, v_1 and v_2, corresponding to vaccinations on day 1 and day 2 of the drive, respectively. If either of the vaccination targets, v_1 or v_2, were not met, the

Figure A.1 provides an example of neighborhood micro-plan, Appendix Figure A.2 provides an example of a vaccination attempt, and Appendix Figure A.3 provides a picture of a chalk marking on a compound wall.

[12] Appendix Figure A.4 provides a picture of the form capturing the self-reports. The second column records the number of vaccinations for the day. The seventh column reports the number of vials of vaccine used in the process.

[13] This could potentially be done by walking the micro-plan and examining the chalk markings on each compound wall.

[14] We attempted to independently audit LHWs by following the trail of chalk markings, but our enumerators found the process too difficult to produce a reliable audit of houses visited. We do, however, know the targets associated with each micro-plan prior to our monitoring intervention and that LHWs almost always reported meeting their targets exactly. Even with a bonus incentive and smartphone monitoring in place, we find that LHWs on average achieve only 62 percent (s.d. = 58 percent) of the target given by their micro-plans. LHWs likely would achieve a smaller share of their target in the absence of both monitoring and financial incentives.

[15] This dashboard system is depicted in Appendix Figure A.5.

Panel A: Splash Page

Panel B: Slider Bar

Figure 1: Vaccination Monitoring Smartphone App

Notes: The picture is of two screenshots from the smartphone app used by LHWs. Panel A is depicted after partially scrolling down. The top bar in Panel A (white letters) translates to "polio survey." The next panel down (blue letters) translates to "Dashboard" (literally transliterated). The black letters under the top button translate to "new activity", the letters under the second button translate to "send activity" and the letters under the lowest button translate to "set target". The blue letters in panel B translate to "set target". The next line translates to "First day: 133; Second day: 133". The text next to the box translates to "finalize target" and the black letters on the bar translate to "set target."

1000 rupees would not be received, and the LHW would receive only her standard wage.

Each LHW was randomly assigned a relative price, R, translating vaccinations on day 1 to vaccinations on day 2. For each vaccination allocated to day 2, the number of vaccinations allocated to day 1 would be reduced by R. Hence, the targets v_1 and v_2 satisfy the intertemporal budget constraint

$$v_1 + R \cdot v_2 = V.$$

The bonus contract is identical to an experimental device termed a Convex Time Budget used to investigate time preferences (Andreoni and Sprenger, 2012; Augenblick et al., 2015).[16] The intertemporal allocation (v_1, v_2) potentially carries information on the time preferences of each LHW.

2.2.1 Experimental Variation and Measuring Preferences

Our design generates two sources of experimental variation. First, each LHW is randomly assigned a relative price, R, from the set $R \in \{0.9, 1, 1.1, 1.25\}$. These values were chosen following Augenblick et al. (2015). Operationally, experimental variation in R was implemented by providing each LHW with a slider bar on the introduction screen of the smartphone application. Figure 1, Panel B depicts the slider bar with R equal to 1.25. The LHW was asked to pull the slider bar to their desired allocation (v_1, v_2) and then submit. The allocation was required to be submitted before commencing vaccination.

Second, each LHW was randomly assigned to either submit their allocation in advance of day 1 of the drive or on the morning of day 1. We refer to the first of these as the 'Advance' treatment arm and the second as 'Immediate' treatment arm. The assignment to either the

[16]We also borrow an additional design element from such studies—minimum allocation requirements—from such studies. In order to avoid LHWs allocating all their vaccinations to a single day of the drive, we placed minimum work requirements of $v_1 \geq 12$ and $v_2 \geq 12$. The objective of minimum allocation requirements is to avoid confounds related to fixed costs. That is, by requiring LHWs to work on both days of the drive, we avoid confounding extreme patience or extreme impatience with LHWs simply not wishing to come to work on one of the two days.

Advance or the Immediate group was cross-randomized with the assignment of R, creating a 2 x 4 design. Section 2.4 describes the efforts taken to make everything else besides allocation timing equal between the Advance and Immediate conditions.

Random assignment to Advance or Immediate choice and random assignment of R are both critical design elements for identifying the discounting parameters of interest. We assume that individuals minimize the discounted costs of effort subject to the intertemporal budget constraint provided by their bonus contract. We make two further assumptions. First, we assume a stationary, power cost of effort function $c(v) = v^{\gamma}$, where v represents vaccinations performed on a given day and $\gamma > 1$ captures the convex costs of effort. Second, we assume that individuals discount the future quasi-hyperbolically (Laibson, 1997; O'Donoghue and Rabin, 1999). Hence, the worker's disutility of effort can be written as

$$v_1^{\gamma} + \beta^{\mathbf{1}_{d=1}} \delta \cdot v_2^{\gamma}.$$

The indicator $\mathbf{1}_{d=1}$ captures whether the decision is made in advance or immediately on day 1. The parameters β and δ summarize individual discounting with β capturing the degree of present bias, active for LHWs who make Immediate decisions, that is, $\mathbf{1}_{d=1} = 1$. If $\beta = 1$, the model nests exponential discounting with discount factor δ, while if $\beta < 1$ the decisionmaker exhibits a present bias, being less patient in Immediate relative to Advance decisions.

Minimizing discounted costs subject to the intertemporal budget constraint of the experiment yields marginal condition:

$$\gamma v_1^{\gamma-1} - \frac{\beta^{\mathbf{1}_{d=1}} \delta}{R} \gamma v_2^{\gamma-1} = 0. \tag{1}$$

Interpreting this marginal condition as a moment requirement, time preferences can potentially be estimated with standard minimum distance estimation techniques (Hansen, 1982; Hansen and Singleton, 1982). Experimental manipulation of R and $\mathbf{1}_{d=1}$ provides identifying variation.[17]

Estimation based on equation (1) yields aggregate discounting parameters with each LHW's allocation contributing a single observation to the aggregate. Exercises exploring heterogeneity in time preferences document substantial differences across people, even from relatively homogeneous populations (see e.g., Harrison et al., 2002; Ashraf et al., 2006; Meier and Sprenger, 2015). Given only a single observation per LHW, estimation of all parameters at the individual level is infeasible. However, we can calculate each LHW's discount factor, which is either δ_i for those who make Advance decisions or $(\beta\delta)_i$ for those who make Immediate decisions. To make such a calculation, two further assumptions are required. First, we assume every LHW shares a common cost function, $\gamma = 2$, corresponding to quadratic cost. Second, we assume the relevant marginal condition (equation (1)) holds exactly. Let R_i be the value of R assigned to individual i, let $\mathbf{1}_{d=1,i}$ be their assignment to Advance or Immediate choice, and let $(v_{1,i}, v_{2,i})$

[17] A previous version of this paper expressed the Euler equation of (1) as

$$\left(\frac{v_1}{v_2}\right)^{\gamma-1} \frac{1}{\beta^{\mathbf{1}_{d=1}}\delta} = \frac{1}{R}.$$

Taking logs and rearranging yields

$$log\left(\frac{v_1}{v_2}\right) = \frac{log\delta}{\gamma-1} + \frac{log\beta}{\gamma-1}\mathbf{1}_{d=1} - \frac{1}{\gamma-1}logR.$$

If we assume that allocations satisfy the above equation subject to an additive error term, ϵ, we arrive at the linear regression equation

$$log\left(\frac{v_1}{v_2}\right) = \frac{log\delta}{\gamma-1} + \frac{log\beta}{\gamma-1}\mathbf{1}_{d=1} - \frac{1}{\gamma-1}logR + \epsilon,$$

which can also be estimated with standard techniques. This formulation provides intuition for the identification of structural parameters from LHW allocations, and make clear the purpose of our experimental variation in R and $\mathbf{1}_{d=1}$. Variation in the relative price, R, identifies the shape of the cost function, γ, while variation in $\mathbf{1}_{d=1}$ identifies β. Note that δ would be identified from the average level of v_1 relative to v_2 when decisions are made in advance (i.e., identified from the constant). An identical strategy for structurally estimating time preferences was introduced in controlled experiments by Andreoni and Sprenger (2012), and has precedents in a body of macroeconomic research identifying aggregate preferences from consumption data. See, for example, Shapiro (1984); Zeldes (1989); Lawrance (1991). Very similar results are obtained for our baseline estimates using this method and the minimum distance method now implemented.

be their allocation of vaccinations. Then,

$$\frac{R_i \cdot v_{1,i}}{v_{2,i}} = (\beta^{\mathbf{1}_{d=1,i}}\delta)_i, \tag{2}$$

the relative-price-adjusted ratio of allocated vaccinations identifies a discount factor for each individual, i.

The assumptions required for identification of aggregate and individual discount factors are potentially quite restrictive. Our research design, which involves tailoring contracts to individual discount factors, required commitment to the specific functional forms of equations (1) and (2). One critical assumption relates to the force of the implemented incentives. The contracts we implement feature a completion bonus of 1000 rupees paid the day after the drive if both targets, v_1 and v_2 are met. The choice of large bonuses (around 10 times daily wages) followed the design logic discussed in Augenblick et al. (2015). Not completing allocated vaccinations creates a sizable penalty at any given point in time. LHWs should forecast that they will indeed complete the required vaccinations and so allocate them according to their true preferences. If LHWs forecast not completing required vaccinations with some chance, the probability of completion has the potential to confound this approach to measuring preferences. In Appendix A.1 and sub-section 4.1 we analyze the potential effects of forecasted non-completion, and provide corresponding adjustments to our preference estimates. The theoretical exercise provided also generates predictions for the probability of completion, which are examined as well. Not being prespecified ex ante, these analyses should be viewed as exploratory. In sub-section 4.2, we also assess the validity of a set of required assumptions and present further exploratory analysis related to alternative functional forms.

2.3 Drive 2: Test of Structural, Tailored Contracts

In a second two-day drive, we investigate tailored contracts. All LHWs from the first drive were invited to participate in a second intertemporal bonus contract. LHWs were unaware that their previously measured behavior would be used to potentially inform their subsequent contracts. This sidesteps an important possibility that LHWs might alter their first drive behavior in order to receive a more desirable relative price in the second drive.

Half of LHWs were given an individually-tailored intertemporal bonus contract,

$$v_{1,i} + R_i^* \cdot v_{2,i} = V,$$

where R_i^* was either $(\beta\delta)_i$ or δ_i from equation (2) depending on whether they made Immediate or Advance decisions.[18] Setting the relative price, R_i^*, equal to the one period discount factor should lead LHWs to allocate an equal number of vaccinations to each day of the drive, $v_{1,i} = v_{2,i}$. Though LHWs in this group receive different relative prices, the contract is designed for each of them to achieve the same objective of smoothing vaccinations through time. Some LHWs' allocation behavior in the first drive implied extreme discount factors and hence extreme values of R_i^*. Our tailoring exercise focused only on a Tailoring Sample of LHWs with discount factors between 0.75 and 1.5.[19] LHWs outside of these bounds were given either the upper or lower bound accordingly.

The other half of LHWs were given a random intertemporal bonus contract,

$$v_{1,i} + \tilde{R}_i \cdot v_{2,i} = V,$$

where $\tilde{R}_i$ was drawn from a random uniform distribution $U[0.75, 1.5]$. The bounds on the

[18]Note that this tailoring exercise requires that LHWs remain in either the Immediate or Advance assignment across drives.

[19]Of our sample of 338 LHWs, 57 exhibit discount factors outside of this range. The Tailoring Sample consists of the remaining 281 LHWs.

distribution of $\tilde{R}_i$ were determined to match the bounds on R_i^*, while the choice of a random uniform control—rather than a single value of $\tilde{R}_i$ or some alternative distribution—was chosen to provide flexible scope for constructing a range of comparison policies by drawing subsets of LHWs assigned to the $\tilde{R}_i$ condition. Relevant subgroups that we draw from this group of LHWs are: 1) structural, broad: those with values of $\tilde{R}_i$ close to the value for achieving $v_1 = v_2$ implied by aggregate preferences; 2) atheoretic, broad: those with values of $\tilde{R}_i$ that are close to the optimal value for achieving $v_1 = v_2$ implied by a reduced form exercise; 3) atheoretic, tailored: those with values of $\tilde{R}_i$ that are generally increasing in patience but not required to be linear as in the structurally tailored policy, $R_i^* = (\beta^{\mathbf{1}_{d=1,i}}\delta)_i$. These comparisons span the policy space of being either atheoretic vs. structural and tailored vs. broad. Comparison is also provided for the full group of LHWs who received random bonus contracts.

Random assignment to structural tailoring in Drive 2 is stratified on the measure of absolute distance to equal provision $|\frac{v_1}{v_2} - 1|$, based on allocations from Drive 1.[20] This measure of distance to equal provision also serves as our eventual outcome measure when analyzing the effect of assignment to structural tailoring in Drive 2. Stratifying assignment on key outcomes of interest is standard practice in the field experimental literature (Bruhn and McKenzie, 2009), as it generally increases precision in estimating treatment effects.

2.4 Design Details

Our experiment is divided into two drives. The first drive took place November 10-11, 2014 with training on November 7. The second drive took place December 8-9, 2014 with training on December 5.

[20] Specifically, subjects are divided into terciles by this measure, with a roughly even number in each bin being assigned to the tailoring and to the control condition.

2.4.1 Training and Allocation Decisions

On November 7, all LHWs participating in the November 10-11 drive received two hours of training at one of three locations in central Lahore on using the monitoring features of the smartphone application and the process by which allocations were made. Both Advance and Immediate LHWs were given identical training.

At the end of the training, LHWs assigned to Advance decision were asked to select their allocations by using the page on their smartphone application. Assistance was available from training staff for those who required it. LHWs assigned to Immediate decision were told they would select their allocations using their smartphone application on Monday morning before beginning work. A hotline number was provided if assistance was required for those in the Immediate condition. The training activities on December 5, for the December 8-9 drive were identical. However, because LHWs had previously been trained on the smartphone application, this portion of the training was conducted as a refresher.

2.4.2 Experimental Timeline

Figure 2 summarizes our experimental timeline and the sample for each vaccination drive of our study.

Drive 0, Failed Drive, September 26-30, 2014: We had hoped to begin our study on Friday, September 26th, 2014 with a training session. 336 LHWs had been recruited, were randomized into treatments, and trained. Advance allocation decisions were collected from half of the subjects on Friday, September 26th. On Monday, September 29th, when we attempted to collect immediate allocation decisions, there was a disruption in the mobile network that prevented 82 of 168 Immediate decision LHWs from submitting their allocations. This caused us to abandon this drive for the purposes of measuring preferences for subsequent tailoring of contracts. The drive, however, was completed and intertemporal bonuses were paid. For the 82 individuals who did not make their allocations, we contacted them, allowed them to continue working, and paid bonuses for all. Figure 2 provides sample details.[21] For completeness, we present data from Drive 0 in Appendix Table A.4, but do not use Drive 0 for the purposes of tailoring contracts.

Drive 1, November 7-11, 2014: Of the original 336 LHWs in our failed drive, 57 did not participate in the next drive organized for November 7 - 11. We recruited replacements with the help of the Department of Health, identifying a total of 349 LHWs to participate in the intertemporal bonus program. The entire sample was re-randomized into R and allocation timing conditions. Training was conducted on November 7, and Advance allocation decisions were collected. The drive began on November 10, and Immediate allocation decisions were collected. 174 LHWs were assigned to the Advance Choice condition and 175 were assigned to the Immediate Choice condition. Bonuses were paid on November 12. While all 174 LHWs in the Advance Choice condition provided an allocation decision, only 164 of 175 in the Immediate Choice condition provided an allocation. Because 11 LHWs attrited from the Immediate Choice condition, we also provide bounds on the estimated effect of decision timing using the method of Lee (2009). In addition, for 232 LHWs, we have allocation decisions in both the failed drive, Drive 0, and Drive 1, forming a potentially valuable panel of response. Figure 2 provides sample details.

Drive 2, December 5-9, 2014: Of the 338 LHWs who participated in Drive 1 and provided an allocation, 337 again participated in Drive 2. These LHWs were randomly assigned to be structurally tailored or receive a random price in their Drive 2 bonus contracts. Importantly, LHWs retained their Advance or Immediate assignment, such that Drive 2 delivers a 2x2 design for structural tailoring and allocation timing. This allows us to investigate the effect of

[21] Appendix Table A.1 checks for balance by failure of the smartphone application in Drive 0. Only one of the eight comparison of means hypothesis tests reject equality at the 10 percent level.

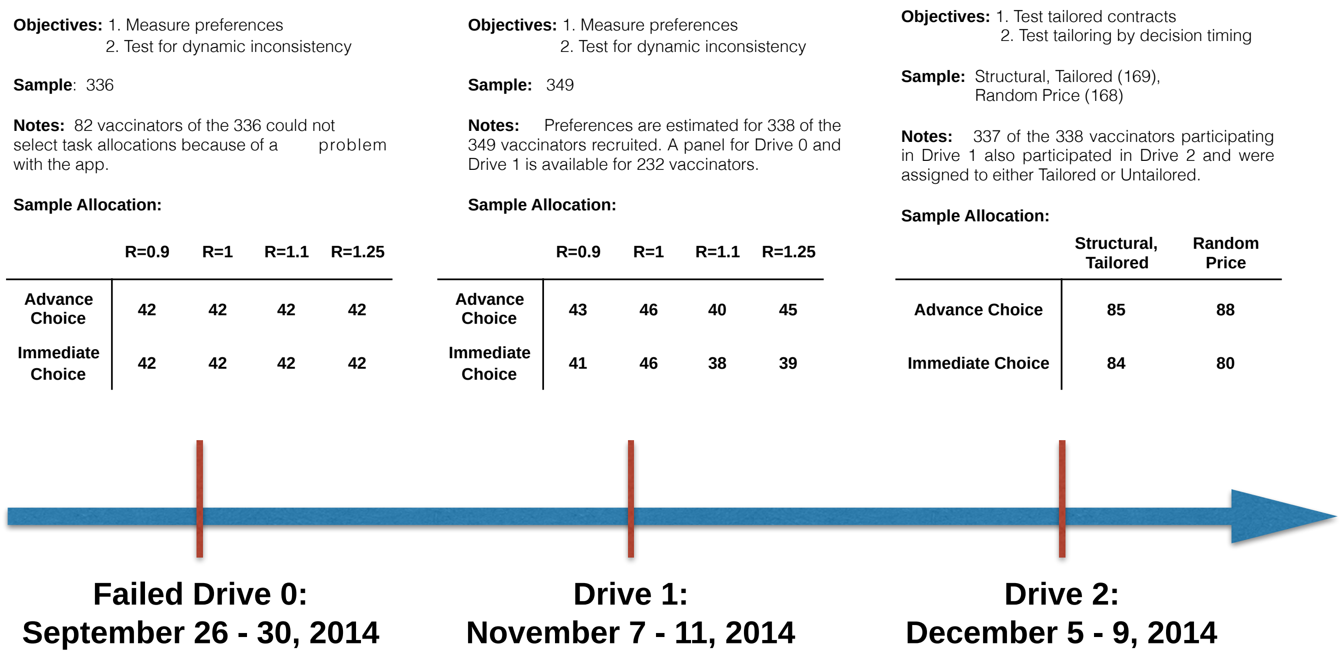


Figure 2: Experiment Overview

Notes: This figure provides an overview of the timing and sample breakdown of the experiment. Assignment to the advance choice and immediate choice condition in Drive 2 is inherited from vaccination Drive 1. Note that: (i) 57 LHWs participated only in Failed Drive 0; (ii) 6 LHWs participated in Drive 1 only; (iii) 1 LHW participated in Failed Drive 0 and Drive 1, but not in Drive 2 (iii) 67 LHWs participated in Drives 2 and 3 only; (iv) 271 LHWs participated in all three rounds.

structural tailoring in general, and if the effects depend on whether present bias is active.

2.4.3 Sample Details

Table 1 summarizes our sample of LHWs from Drive 1 and provides tests of experimental balance on observables. Column (1) presents the mean and standard deviation for each variable; columns (2) to (9) present the mean and standard error for each of our eight treatment arms, and column 10 presents a p-value corresponding to joint tests of equality. Our sample is almost exclusively female, more than 90 percent Punjabi in all treatment arms, and broadly without access to formal savings accounts. LHWs are generally highly experienced with an average of 10.5 years of health work experience and 10.4 years of polio work experience. Consistent with randomization, of the 8 tests performed, only the test performed on an indicator variable equal

Table 1: Summary Statistics and Covariates Balance

	Full Sample	Advance Decision				Immediate Decision				p-value
		R=0.9	R=1	R=1.1	R=1.25	R=0.9	R=1	R=1.1	R=1.25	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Demographics										
Gender (Female = 1)	0.985	1.000	1.000	1.000	0.978	0.975	0.978	0.947	1.000	0.284
	[0.121]	(0.000)	(0.000)	(0.000)	(0.022)	(0.025)	(0.022)	(0.037)	(0.000)	
Years of Education	10.415	10.767	10.652	10.650	10.279	9.850	10.565	10.184	10.282	0.500
	[2.291]	(0.416)	(0.273)	(0.462)	(0.330)	(0.298)	(0.368)	(0.238)	(0.395)	
Number of Children	3.424	3.419	3.422	3.538	3.286	3.605	3.391	3.421	3.333	0.997
	[1.826]	(0.279)	(0.301)	(0.309)	(0.296)	(0.286)	(0.274)	(0.243)	(0.294)	
Punjabi (=1)	0.952	0.930	0.932	1.000	0.955	0.950	0.978	0.917	0.947	0.022
	[0.215]	(0.039)	(0.038)	(0.000)	(0.032)	(0.035)	(0.022)	(0.047)	(0.037)	
Financial Background										
Has a Savings Account (=1)	0.269	0.310	0.250	0.275	0.302	0.350	0.283	0.189	0.179	0.630
	[0.444]	(0.072)	(0.066)	(0.071)	(0.071)	(0.076)	(0.067)	(0.065)	(0.062)	
Participated in a ROSCA (=1)	0.389	0.349	0.378	0.425	0.350	0.500	0.289	0.351	0.487	0.482
	[0.488]	(0.074)	(0.073)	(0.079)	(0.076)	(0.080)	(0.068)	(0.079)	(0.081)	
Health Work Experience										
Years in Health Department	10.520	10.605	10.578	10.211	11.549	9.050	10.678	10.395	11.026	0.456
	[4.961]	(0.777)	(0.695)	(0.685)	(0.792)	(0.695)	(0.846)	(0.867)	(0.808)	
Years as Polio Vaccinator	10.428	10.209	10.728	11.050	11.143	9.238	9.935	10.447	10.692	0.581
	[4.727]	(0.758)	(0.689)	(0.668)	(0.743)	(0.689)	(0.713)	(0.858)	(0.751)	
# LHWs	338	43	46	40	45	41	46	38	39	

Notes: This table checks balance across the eight treatment groups. Column 1 presents the mean for each variable based on our sample of 338 LHWs. These 338 LHWs comprise the estimation sample in Table 2, which reports tests of dynamic inconsistency. Standard deviations are in brackets. Columns 2 to 9 report the mean level of each variable, with standard errors in parentheses, for each treatment cell. For each variable, Column 10 reports the p-value of a joint test that the mean levels are the same for all treatment cells (Columns 2–9). The last row presents the number of observations in each treatment condition. A ROSCA is an informal Rotating Savings and Credit Association. Some calculations used a smaller sample size due to missing information. The proportion of subjects with missing information for each variable is never greater than 3.5 percent (8 LHWs did not report whether they had participated in a ROSCA).

to one for Punjabi subjects suggests baseline imbalance.

3 Results

Our project has two phases. The first phase measures intertemporal preferences. The second phase evaluates the effects of structural, tailored contracts.[22] The results are presented in two

[22]In addition, to test just the effect of providing the $10 bonus, we randomly assigned 85 LHWs in Drive 0 to carry a phone but not receive an incentive. 72 of these LHWs also participated in Drive 1, retaining the same ‘phone only’ treatment status. In Drive 1, LHWs in the ‘phone only’ group attempted 169.47 vaccinations (s.e.

sub-sections, corresponding to the two phases of the study.

3.1 Drive 1: Aggregate Behavior and Intertemporal Preferences

Figure 3 presents median behavior in Drive 1, graphing the allocation to the sooner work date, v_1, for each value of R.[23] Separate series are provided for Advance and Immediate choice. In Panel A we provide data for our Full Sample of 338 LHWs who provided allocations in Drive 1. In Panel B we focus only on our Tailoring Sample of 281 LHWs, trimming 57 LHWs with extreme allocation behavior that would imply individual discount factors from equation (2) outside of the range of [0.75, 1.5]. Two features of Figure 3 are notable. First, subjects appear to respond to the between-subject variation in relative prices. As the value R increases, vaccinations allocated to v_1 count relatively less towards reaching the two-day target of $V = 300$. LHWs respond to this changing incentive by reducing their allocation of v_1. Second, there is a tendency of present bias. LHWs appear to allocate fewer vaccinations to v_1 when making Immediate choice.

Table 2 presents corresponding median regression analysis for aggregate behavior in Drive 1.[24] In the Full and Tailoring Samples, LHWs assigned to Immediate choice allocate between 2 and 3 fewer vaccinations to v_1 than those assigned to Advance choice.[25]

Equation (1) links our experimental parameters, R and whether allocations are Immediate or Advance to the aggregate preference parameters governing choice. Estimating equation (1)

= 15.98) and LHWs in the phone plus incentives group attempted 205.82 vaccinations (se = 7.79) yielding an estimated increase of 36.35 attempts (s.e. = 18.42, p = 0.05). 49.3% of vaccination attempts were successful for the 'phone only' group while 49.1% of vaccinations were successful for the 'phone plus incentives' group. The difference in success rates between the two groups is small (0.2 percentage points) and statistically insignificant (p=0.69).

[23] We opt to provide medians as the average data are influenced by several extreme outliers in allocation behavior. Qualitatively similar patterns are, however, observed.

[24] Appendix Table A.4 presents identical analysis incorporating data from failed Drive 0, and identifies qualitatively similar effects.

[25] As discussed in Section 2.4.2 above, 11 LHWs attrited from the sample in the immediate choice condition in Drive 1. Bounding the effect of being assigned to the immediate choice condition on v_1 allocations using the method of Lee (2009) provides a lower bound of -3.78 tasks (s.e. = 2.06) and an upper bound of 0.205 (2.06) tasks.

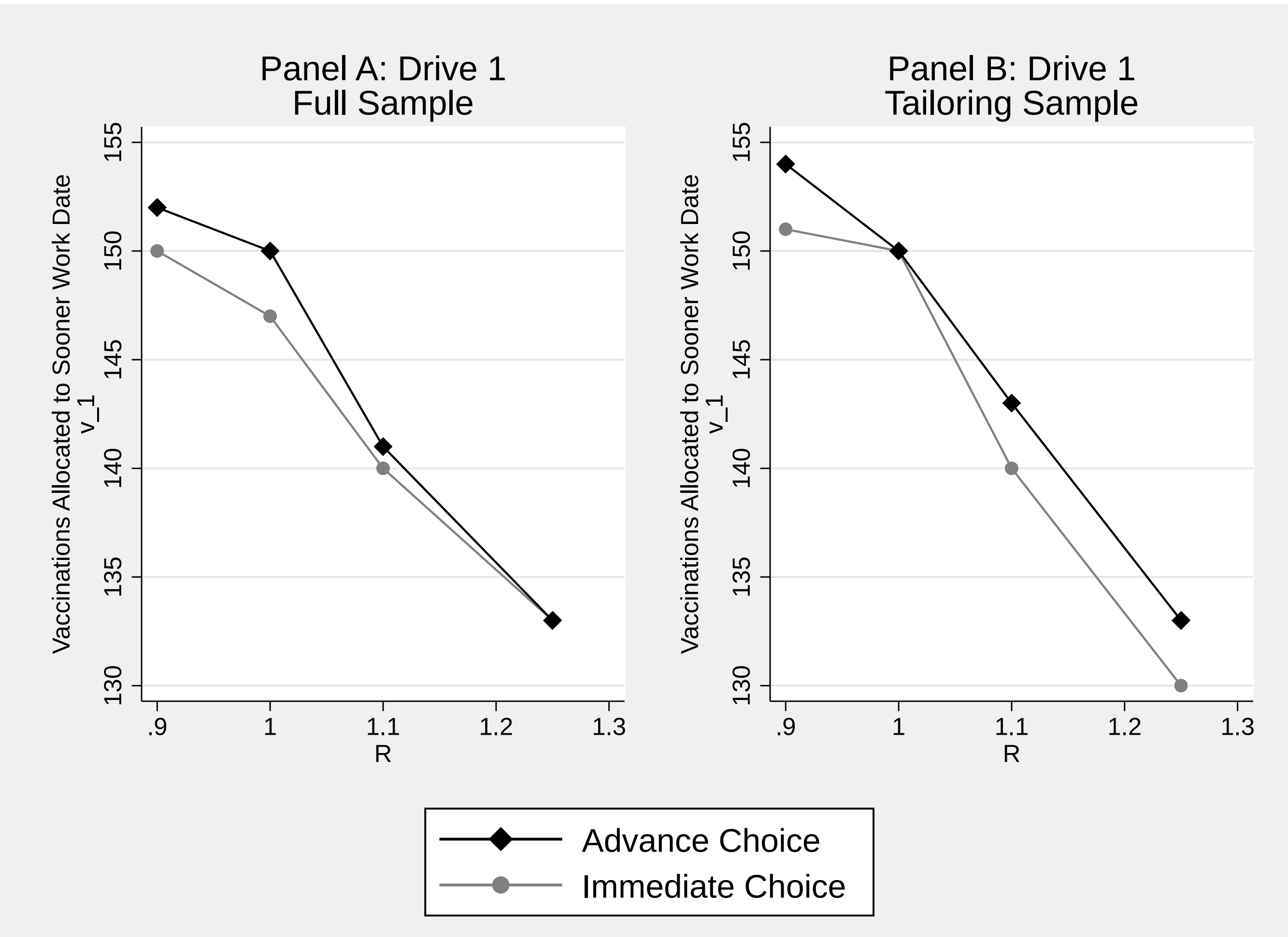


Figure 3: Aggregate Experimental Response

Notes: This figure examines whether tasks assigned to the sooner work date and completion respond to the experimental variation in the relative price, R, and in decision timing. Allocation data represent medians for each of the eight treatment groups and completion data represent group averages. Panel A depicts the Full Sample and Panel B depicts the tailoring sample (LHWs with $R^* < 0.75$ or $R^* > 1.5$). Black series are advance choice groups and gray series are immediate choice groups.

Table 2: Aggregate Drive 1 Behavior

	(1) Full Sample	(2) Tailoring Sample
Dependent Variable: v_1		
Immediate Decision (=1)	-2.00*	-3.00***
	(1.13)	(0.91)
Relative Price (R)	-54.29***	-66.67***
	(4.38)	(3.66)
Constant	201.86***	216.33***
	(4.72)	(3.93)
Median Advance Choice	146.5	148
# Observations	338	281

Notes: This table reports on the effects of decision timing and relative price variation on vaccinations allocated to the first day of the drive. Median regression. *Levels of Significance:* $*p < 0.1$, $**p < 0.05$, $***p < 0.01$.

via minimum distance on our Full Sample, with the restriction that the cost parameter $\gamma = 2$ (for comparison with subsequent individual analysis) we estimate $\delta = 0.985$ (s.e. $= 0.017$) and $\beta = 0.935$ (s.e. $= 0.030$), and reject the null hypothesis of no present bias ($\chi^2(1) = 4.841$, ($p = 0.028$). In the Tailoring Sample, we estimate $\delta = 1.017$ (s.e. $= 0.013$) and $\beta = 0.969$ (s.e. $=$ 0.018), and marginally reject the null hypothesis of no present bias ($\chi^2(1) = 2.88$, ($p = 0.090$). Appendix Table A.2 provides these estimates along with a variety of additional estimations with alternate values of the assumed cost parameter, γ, reaching broadly similar conclusions.[26]

Following equation (2), we calculate individual discount factors for each LHW assuming quadratic costs. For those LHWs assigned to Advance choice, this discount factor corresponds

[26] In principle, variation in the relative price R, should provide an opportunity to identify γ without restriction. Unfortunately, our minimum distance estimators did not reliably converge without restrictions. This highlights a potentially important issue with respect to the estimates of Table A.2: the estimated parameters predict more sensitivity to R than truly exists in the data. Appendix Figure A.8 reproduces Figure 3, with in-sample predictions from Table A.2, column (2). Though the estimates do match the responsiveness of behavior from $R = 1$ to $R = 1.1$, they do not generate the lack of sensitivity for other changes in R. This mis-specification presents a clear challenge for using individual preference parameters for tailored contracts. Having committed to a possibly mis-specified functional form ex-ante, any success in tailoring contracts should likely be viewed as a lower bound on the potential benefits of such initiatives.

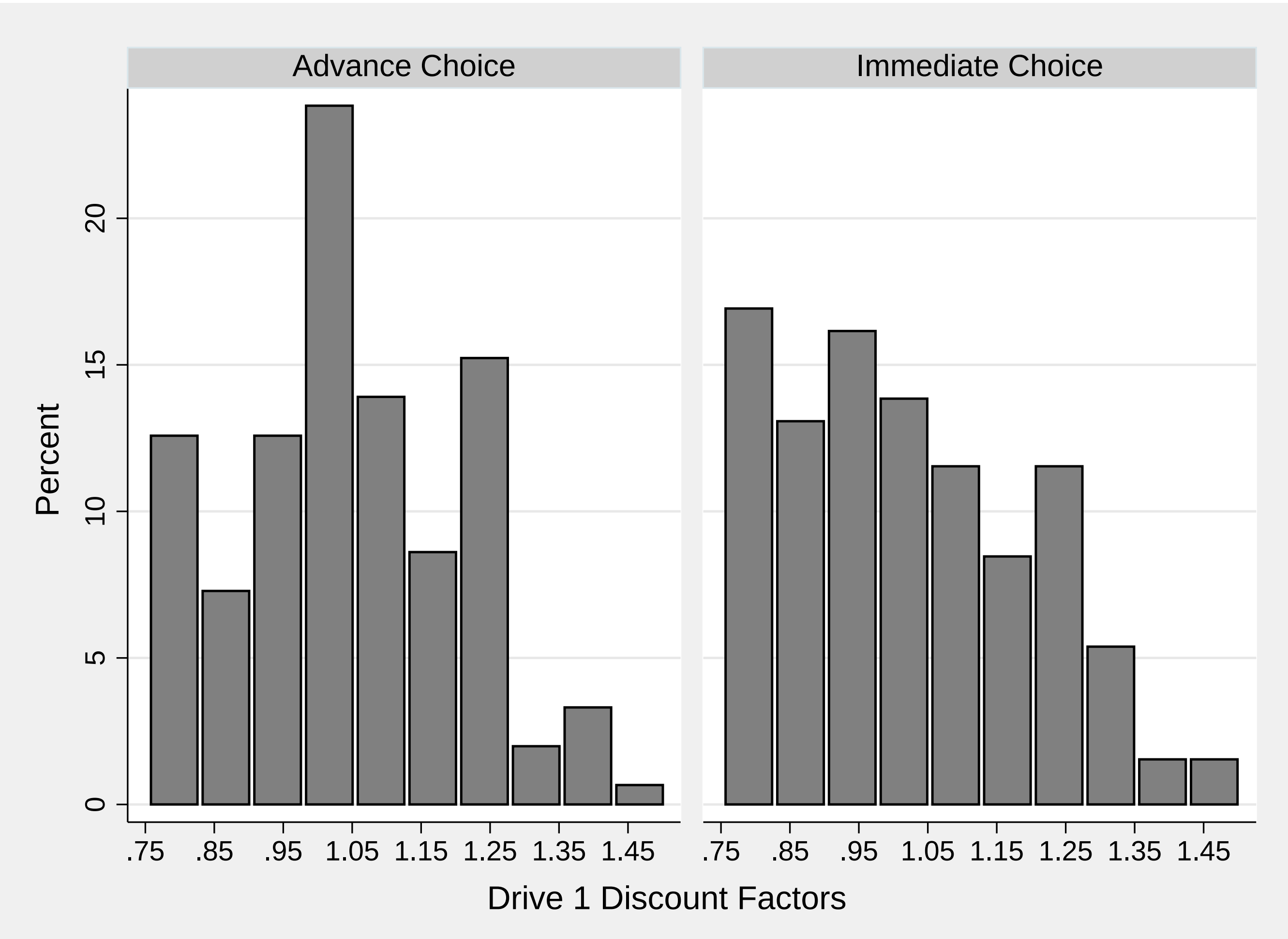


Figure 4: Individual Discount Factors in the Tailoring Sample

Notes: This figure provides histograms of one period discount factors calculated from equation (2) separately for subjects in the Advance Choice condition (left panel) and the Immediate Choice condition (right panel). The sample is restricted to LHWs in the Tailoring Sample (LHWs with $R_i^* \geq 0.75$ or $R_i^* \leq 1.5$).

to δ_i, while for those assigned to Immediate choice it corresponds to $(\beta\delta)_i$. The median [25th, 75th percentile] discount factor in Advance choice is 1.015 [0.88, 1.18], while the median discount factor in Immediate choice is 1 [0.84, 1.21]. As noted above, fifty-seven of 338 subjects in Drive 1 have implied discount factors either above 1.5 or below 0.75.[27] We term such LHWs the 'Boundary Sample.' As our tailoring exercise focuses on individuals with discount factors

[27]Such extreme behavior is slightly more pronounced in Immediate choice (34 LHWs) relative to Advance choice (23 LHWs), ($t = 1.84,\ p = 0.07$).

between 0.75 and 1.5, we restrict our individual analysis to the 281 LHWs in the Tailoring Sample, and discuss the Boundary Sample in robustness tests (see section 4.3). Figure 4 presents histograms of implied discount factors for the Tailoring Sample in Advance and Immediate decisions. In addition to substantial heterogeneity in discounting, Figure 4 highlights broad evidence of present bias. The one period discount factors are skewed below 1 in Immediate relative to Advance choice. A Kolmogorov-Smirnov test for equality of distributions sits at the cusp of statistical significance, $D_{KS} = 0.15,\ (p = 0.10)$.

The observed present bias in work allocations at the aggregate and individual level echoes recent laboratory work eliciting time preferences over effort (Augenblick et al., 2015; Augenblick and Rabin, 2015). Additionally, heterogeneity in discount factors across LHWs resonates with prior exercises demonstrating heterogeneity of preferences even with relatively homogeneous samples (see e.g., Harrison et al., 2002; Ashraf et al., 2006; Meier and Sprenger, 2015).

3.2 Drive 2: Evaluating Structural, Tailored Contracts

Individual discount factors from Drive 1 in hand, we evaluate contracts tailored to individual discounting parameters. Of the 281 LHWs in the Tailoring Sample, 280 participated in Drive 2.[28] Of these, 142 LHWs were assigned a value of R equal to their discount factor. That is, tailored LHWs were assigned $R_i^* = (\beta^{\mathbf{1}_{d=1,i}}\delta)_i$, from equation (2), which should induce equal allocation of effort through time, $v_{1,i} = v_{2,i}$. The remaining 138 LHWs provide the basis for the different comparison policies that we consider, and were assigned a uniform random relative price $\tilde{R}_i \in U[0.75, 1.5]$.[29]

We examine differences in the distance from the 45-degree line using the metric $|\frac{v_{1,i}}{v_{2,i}} - 1|$,

[28] LHWs from the boundary sample were allowed to participate in Drive 2 and were either assigned $\tilde{R}_i \in U[0.75, 1.5]$ if they were in the untailored control group (31 subjects) or assigned $R_i = 0.75$ or $R_i = 1.5$ if they were in the tailored group and had $R_i^* < 0.75$ (15 subjects) or $R_i^* > 1.5$ (11 subjects). See section 4.3 for analysis of the boundary sample.

[29] As noted in section 2.3, assignment to the tailored or the untailored group was conducted via stratified randomization with strata based upon the tercile of differences from equal provision of effort in Drive 1.

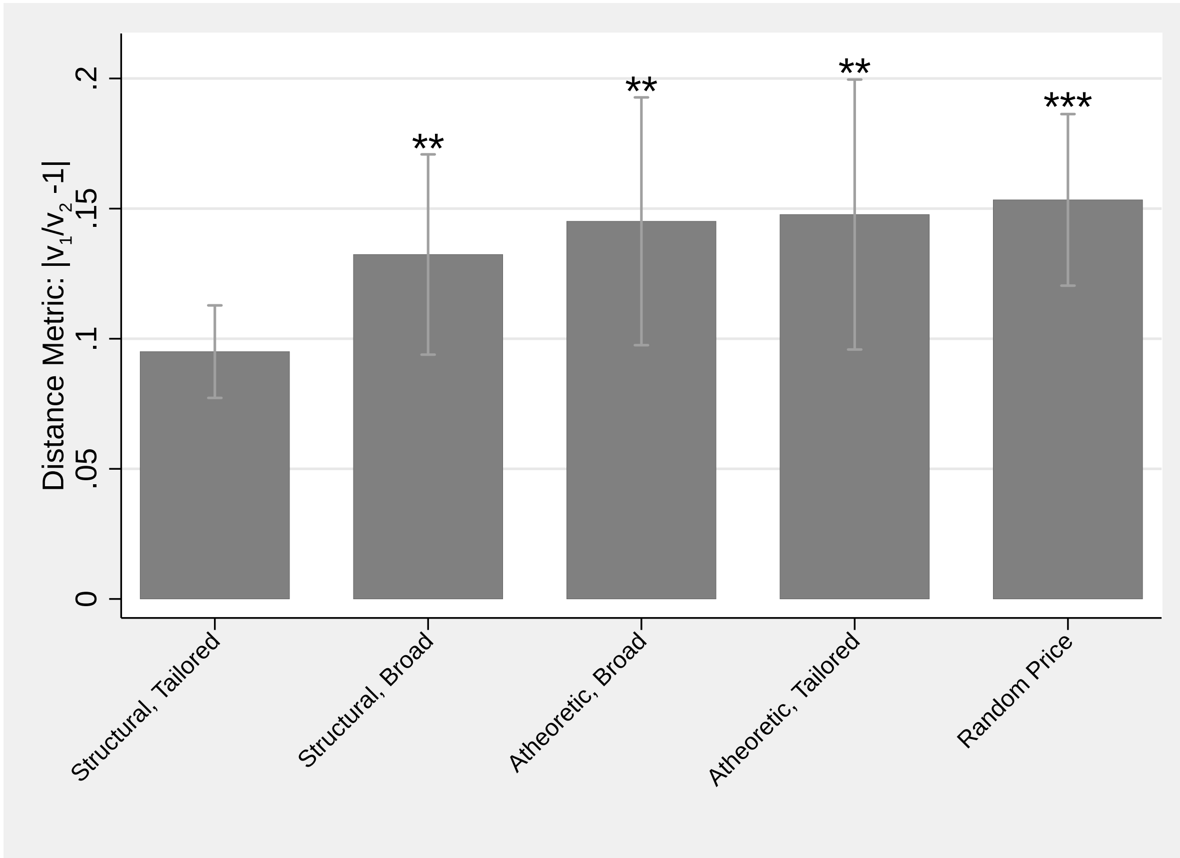


Figure 5: Policy Comparison

Notes: This figure reports the mean absolute distance from the 45-degree line using the metric $|\frac{v_{1,i}}{v_{2,i}} - 1|$ for five comparison groups in Drive 2: Structural, Tailored; Structural, Broad; Atheoretic, Broad; Atheoretic, Tailored; and Random Price. Definitions of comparisons provided in section 3.2.

the absolute percentage difference between v_1 and v_2. The mean distance in Drive 2 for the Tailoring Sample is 0.37 (s.d. = 2.57). The large average distance is driven by several extreme outliers. Trimming the top and bottom 1% of the sample of Drive 2 allocations, the mean distance is 0.12 (s.d. = 0.16). We focus our analysis on this trimmed sample, but provide results corresponding to the complete Tailoring Sample in Appendix Table A.3.

Figure 5 provides mean distance measures in Drive 2 across a range of subgroups. The mean distance for the subsample that was tailored to receive $R_i^* = (\beta^{\mathbf{1}_{d=1,i}}\delta)_i$ is 0.10 (s.d. = 0.11), an average 10% deviation from the policy benchmark of smooth provision. In the subsample that received $\tilde{R}_i \in U[0.75, 1.5]$, the mean distance is 0.15 (s.d. $= 0.19$). Individuals who receive structural, tailored contract terms are significantly closer to the policy target than those receiving a random relative price, $t_{265} = 3.07$, ($p < 0.01$).

The random price treatment can be disaggregated into further subsamples corresponding to important policy comparisons. Three further policy comparisons are analyzed in Figure 5.

1. *Structural, Broad Policy:* The aggregate estimation of preferences based on equation (1) for the Tailoring Sample yielded $\delta = 1.017$ (s.e. $= 0.013$) and $\beta = 0.969$ (s.e. $= 0.018$). The value of R required to to achieve $v_1 = v_2$ at these aggregate values is thus $R = 1.017$ in Advance Choice and $R = 1.017 * 0.969 = 0.985$ in Immediate Choice. To approximate this broad, structural policy we select the 54 individuals from the condition who received $\tilde{R}_i \in U[0.75, 1.5]$ within one standard deviation of these prices. Appendix Figure A.9, Panel A indicates the exact assignments for this subgroup. The relative prices implied by the aggregate model are structural, informed by an estimation of preferences in Drive 1, but not tailored to each individual. The mean distance for this group is 0.13 (s.d. $= 0.15$), which is significantly further from the target of smooth provision than the structural, tailored policy, $t_{192} = 1.98$, ($p < 0.05$).

2. *Atheoretic, Broad Policy:* The reduced-form evidence in Table 2 provides regression coefficients indicating the sensitivity of the allocation v_1 to R and whether the decision is immediate. The reduced-form relationship estimated is $v_1 = 216.33 - 3.00 \times \mathbf{1}_{d=1} - 66.67 \times R$. In order to equate $v_1 = v_2$ under the constraint $v_1 + Rv_2 = 300$, one requires $(1+R)v_1 - 300 = 0$. Substituting in for the reduced-form relationship for v_1, one obtains $f(R) = (1+R)(216.33 - 3.00 \times \mathbf{1}_{d=1} - 66.67 \times R) - 300 = 0$. Note that $f(R)$ is quadratic in R. In Advance Choice, it obtains the value of zero at $R = 1.05$ and $R = 1.19$. In Immediate Choice, $f(R)$ does not achieve the value zero, but has a maximum value of $f(R) = -6.01$ at $R = 1.10$. To approximate this broad atheoretic policy, we select the 54 individuals from the subsample who received $\tilde{R}_i \in U[0.75, 1.5]$ within one standard

deviation of $R = 1.1$. Appendix Figure A.9, Panel B indicates the exact assignments for this subgroup. The relative prices implied in this case are atheoretic, informed only by the estimated sensitivities of Table 2, and not tailored to each individual. The mean distance for this group is 0.15 (s.d. = 0.19), which is significantly further from the target of smooth provision than the tailored policy, $t_{192} = 2.37$, ($p < 0.05$).

3. *Atheoretic, Tailored Policy:* From the subsample who received $\tilde{R}_i \in U[0.75, 1.5]$, random assignment generates a match between the random price received, $\tilde{R}_i$, and Drive 1 allocation behavior. Even without structural guidance on the correct value of R to achieve equal allocations, random assignment will at times assign higher values of $\tilde{R}_i$ to individuals with higher values of $(\beta^{\mathbf{1}_{d=1,i}}\delta)_i$. These assignments give higher prices to more patient LHWs, but do not require that the relationship between prices and patience be linear as in the structural, tailored policy, which gives $R_i^* = (\beta^{\mathbf{1}_{d=1,i}}\delta)_i$. For each LHW who received a random price in Drive 2, we count the percent of LHWs who were more patient in Drive 1 but received a lower value of $\tilde{R}_i$. From the subsample who received $\tilde{R}_i \in U[0.75, 1.5]$, we select the 49 LHWs for whom this number is less than or equal to 10%, being effectively in order with at least 90% of the sample. Appendix Figure A.9, Panel C indicates the exact assignments for this subgroup. The relative prices implied in this case are atheoretic — loosely related to patience, but not designed to achieve a specific objective beyond giving more patient LHWs higher prices — and tailored to each individual. The mean distance for this group is 0.15 (s.d. = 0.19), which is significantly further from the target of smooth provision than the tailored policy, $t_{182} = 2.41$, ($p < 0.5$).

Figure 5 highlights significant differences in distance to the policy target of smooth allocation between our tailored policy and three natural alternatives spanning the policy space of atheoretic vs. structural and broad vs. tailored. In Table 3, we provide corresponding least squares regression analysis. Following best practice for such analysis (Bruhn and McKenzie, 2009), we control for fixed effects for each stratum in the stratified randomization. We additionally control for the value of R_i^* or $\tilde{R}_i$ assigned in Drive 2. The regressions identify whether s tailoring generates more equal provision for a given value of R, and hence controls for any differences in R across groups. In the odd columns, we provide comparisons between our structurally tailored group and the alternatives noted in Figure 5. In each case the tailored group is around one-third closer to the policy target compared to the relevant alternative.

Table 3: The Effect of Tailoring Intertemporal Incentives

Dependent variable:	$\lvert\frac{v_{1,i}}{v_{2,i}} - 1\rvert$							
Policy Comparison Group	Random Price		Structural, Broad		Atheoretic, Broad		Atheoretic, Tailored	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Structural Tailored (=1)	-0.049***	-0.014	-0.039*	-0.018	-0.051**	-0.010	-0.054*	-0.014
	(0.018)	(0.019)	(0.021)	(0.024)	(0.025)	(0.022)	(0.029)	(0.023)
Immediate Choice		0.117***		0.092**		0.137***		0.169***
		(0.035)		(0.041)		(0.051)		(0.062)
Structural Tailored x Immediate		-0.084**		-0.054		-0.102*		-0.132**
		(0.040)		(0.045)		(0.055)		(0.064)
Stratum FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Exclude 99th and 1st Percentiles	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Drive 2 R_i^* or $\tilde{R}_i$	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-Squared	0.082	0.154	0.081	0.135	0.114	0.194	0.085	0.193
Mean in Comparison Contract	0.153	0.153	0.132	0.132	0.145	0.145	0.148	0.148
Mean in Comparison Advance		0.098		0.087		0.091		0.100
Mean in Comparison Immediate		0.222		0.223		0.271		0.180
# LHWs	267	267	194	194	194	194	184	184
# Comparison LHWs	132	132	59	59	59	59	49	49

Notes: This table reports the effects of tailoring on the equality of effort provision over time. The measure $\lvert\frac{v_{1,i}}{v_{2,i}} - 1\rvert$ (the percentage difference between tasks allocated to day 1 and day 2 of the drive) reflects the distance of the task allocation (v_1,v_2) from equality ($v_1 = v_2$). Section 3.2 provides definitions of comparison groups. Ordinary least squares regressions. Heteroskedasticity robust White standard errors reported in parentheses. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

LHWs assigned to Advance choice in Drive 1 remain in Advance choice in Drive 2, while those assigned to Immediate choice remain in Immediate choice. In the even columns of Table 3, we examine differential effects across these two groups. Due to present bias, one might expect larger distance measures in Immediate conditions (and hence greater benefits to structural tailoring). This is precisely what is observed. Immediate choice is associated with significantly larger distance measures and structural tailoring in Immediate choice significantly reduces these distances. Across comparisons we find that structural tailoring in Immediate choice reduces distance from equal provision by around one-half. Note that the effect size is similar to the effect of moving a comparison LHW from advance to immediate choice. That is, structural tailoring effectively eliminates present bias in allocations.

Table 3 and Figure 5 indicate benefits to structural, tailored interventions over a set of policy alternatives. Naturally, the comparisons carry less weight if the assignments to different comparison groups are exactly overlapping. Appendix Figure A.9 provides the exact assignment of relative prices for each of the alternatives considered. The counterfactuals differ markedly between the atheoretic, tailored policy and the two broad policies. The correlation in inclusion probability between the atheoretic, tailored group and the atheoretic, broad and structural, broad groups is 0.10 and 0.19, respectively.[30] The policy comparisons are distinct and in each case clearly outperformed by our structural, tailored intervention.

The analysis to this point indicates three key findings. First, there appears to be a present bias in LHW allocation behavior. Those individuals making Immediate choice allocate fewer vaccinations to v_1 than those making Advance choice. Second, along with the general tendency towards less patience in Immediate choice, substantial heterogeneity in discounting is observed. Both of these effects resonate with prior experimental findings and highlight the potential for policy interventions tailored to individual preferences. Third, structural, tailored contracts work. Those individuals given a tailored price equal to their previously measured discount factor provide smoother service than a set of alternatives that span the policy space of structural vs. atheoretic and broad vs. tailored. In the following section, we explore robustness and provide a set of additional examinations.

[30]The correlation in inclusion probability for the two broad policies is 0.8, reflecting the fact that the relative prices implied in these two exercises are in the neighborhood of $R = 1.1$ and $R = 1$, respectively.

4 Robustness Tests and Additional Exercises

4.1 Probabilistic Completion

One critical assumption in our theoretical development revolves around the force of the implemented incentives. The contracts we implement feature a completion bonus of 1000 rupees paid the day after the drive if both targets, v_1 and v_2 are met. The choice of large bonuses (around 10 times daily wages) followed the design logic discussed in Augenblick et al. (2015). Not completing allocated vaccinations creates a sizable penalty at any given point in time. LHWs should forecast that they will indeed complete the required vaccinations and so allocate them according to their true preferences. If LHWs forecast not completing required vaccinations with some chance, the probability of completion has the potential to confound this approach to measuring preferences.

In Appendix Section A.1, we analyze the potential effects of forecasted non-completion. Specifically, we assume the LHW believes that they they will be able to complete v allocations on a given date with probability $p(v) = \frac{1}{1+\alpha v}$. This probabilistic completion introduces a wedge in the LHW's marginal condition. Discounted relative marginal costs are no longer equated to a relative price, but rather to a relative price adjusted for marginal completion probabilities. We propose methodology to simultaneously estimate the key belief parameter, α, along with discounting parameters, β and δ, at the aggregate level. This methodology assumes that agents know the correct empirical completion probability for each allocation level and uses this to quantify the wedge induced in the marginal condition. Using the aggregate estimate of α, this methodology can also be used to re-calculate individual discounting parameters.

The development of Appendix Section A.1 is critical because— unlike laboratory settings where sizable completion bonuses have been used to ensure near one-hundred percent completion rates— in our field setting even a bonus of 10 times daily wages does not ensure uniform completion. We determine completion by examining the records obtained from each LHW's cell phone application. Of 338 LHWs in Drive 1, 288 registered activity in their cell phone application during the drive, while 50 generated no data. The cellular network in Lahore is known to have some coverage gaps. As such, we consider a subject to have successfully completed their work if they completed an average of 90% or more of their required tasks.[31] One-hundred seventy-four (51.5%) subjects successfully completed by this measure. Appendix Figure A.6 presents the histogram of average completion percentages across subjects, showing a bimodal distribution of success and failure. In Appendix Table A.6, we examine the determinants of completion in Drive 1 with linear probability models and an indicator for completion, $Complete(=1)$, as dependent variable. We find no discernible relationship between R and completion. However, individuals assigned to Immediate choice are between 9 and 13 percentage points less likely to satisfactorily complete their allocated vaccinations. This evidence is clearly supportive of a present-biased interpretation. Subjects in the Immediate choice condition postpone more work, which they are subsequently unable to satisfactorily complete.

Implementing the methodology of Appendix Section A.1 on our Full Sample with the restriction that the cost parameter $\gamma = 2$, we estimate $\delta = 0.992$ (s.e. $= 0.017$) and $\beta = 0.952$ (s.e. $= 0.029$). In the Tailoring Sample, we estimate $\delta = 1.018$ (s.e. $= 0.013$) and $\beta = 0.970$ (s.e. $= 0.018$). Accounting for probabilistic completion does little to alter the aggregate discounting estimates previously discussed (see Appendix Table A.2 for further detail). At the individual level, there is also broad consistency between the parameters recovered with and without accounting for completion. Without accounting for completion In Drive 1, the median [25th, 75th percentile] discount factor in Advance choice is 1.015 [0.88, 1.18], while the median discount factor in Immediate choice is 1 [0.84, 1.21]. Accounting for completion, the median [25th,75th percentile] discount factor in Advance choice is again 1.015 [0.88, 1.18], while the median discount factor in Immediate choice is again 1 [0.84, 1.21]. The correlation in discount factors with and without accounting for completion is effectively 1, indicating probabilistic

[31]Average completion rates are calculated as $1/2(min(Completed_1/v_1, 1) + min(Completed_2/v_2, 1))$.

completion does not dramatically confound any individual calculations. Indeed, the difference between the implied discount factor with and without accounting for completion has a median [25th-75th %-ile] value of 0.00004 [-0.0002, 0.0001].[32]

Our tailored policy links individual discount factors to assigned prices. Impatient subjects 'pay' for their impatience with lower values of R_i^*. Given the intertemporal constraint $v_1 + R_i^* \cdot v_2 = 300$, this means less patient individuals will ultimately have to do weakly more work. Under our assumed functional form for completion probabilities, this generates a relationship between completion probabilities and patience. In the structural, tailored group, less patient individuals should be substantially less likely to successfully complete their Drive 2 targets, while in the random control group no such relationship should exist. Table 4 provides corresponding logit regressions showing a significant positive relationship between discounting and completion only for the tailored group of subjects. Table 4 also provides actual and predicted completion rates separately for the structural, tailored and random price groups. The two groups are predicted to differ in their completion rates due to different contract terms[33], and they indeed do differ in the predicted direction. Overall, predicted and actual completion rates are significantly correlated, as those individuals who actually do fail were predicted to do so with higher probability, 0.521 (s.e. = 0.002) vs 0.509 (0.002), $t_{335} = 3.80,\ (p < 0.01)$.[34]

4.2 Structural Assumptions

As in any structural exercise, a set of assumptions are required to infer discounting parameters from LHW allocation behavior. Five assumptions are relevant for the present discussion, which we discuss below.

[32]Note that this close correspondence implies that the relative price required to generate smooth provision should also not be dramatically altered to account for probabilistic completion.

[33]Due to the random uniform assignment the random price group has relatively more relative prices above 1.

[34]For the Tailoring Sample alone, these values are 0.520 (s.e. = 0.003) vs 0.512 (0.002), $t_{278} = 2.30,\ (p < 0.05)$

Table 4: Tailoring, Discount Factors, and Completion

Dependent variable:	Drive 2 Completed (=1)			
	Random Price		Structural, Tailored	
	(1)	(2)	(3)	(4)
Drive 1 Discount Factor	0.693	0.152	2.506**	2.312**
	(1.049)	(1.086)	(1.043)	(1.069)
Constant	-0.533	0.118	-2.780**	-2.516**
	(1.082)	(1.131)	(1.120)	(1.150)
# LHWs	138	132	142	135
Log-Likeliihood	-94.908	-90.254	-95.022	-91.078
Exclude 99th and 1st Percentiles	No	Yes	No	Yes
Actual Completion Rate	0.543	0.568	0.458	0.474
Predicted Completion Rate	0.524	0.523	0.508	0.508

Notes: This table reports logit regressions for successful completion in Drive 2 on Drive 1 discount factor for Structural Tailored and Random Price subjects. Individual discount factor calculated from equation (2) based on Drive 1 allocation. Predicted completion rate calculated as $p(\hat{v}_{1,i}, \hat{v}_{2,i})$ at predicted Drive 2 allocation $(\hat{v}_{1,i}, \hat{v}_{2,i})$. Heteroskedasticity robust White standard errors reported in parentheses. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

Assumption 1: Stationarity of the Cost Function: We assume the cost function is the same for day 1 and day 2 of each drive. If sooner costs are forecasted to be more severe than later costs, LHWs may appear disproportionately impatient, while if later costs are forecasted to be more severe, they may appear disproportionately patient. Further, if perceived costliness of vaccinations changes from Advance to Immediate choice, present bias measured by β is conflated with non-stationarity.

Importantly, our monitoring technology provides time-stamps and geo-stamps for vaccination activity. Time stamps are recorded for every vaccination attempt, while geo-stamps are collected approximately every 10 vaccination attempts. This may provide independent means for assessing the costliness of tasks from time use. For each LHW, we identify the median time lapse between vaccination attempts and the median distance covered per 30 minute window each day.[35] Of our 338 LHWs, measures for median time lapse between vaccination attempts are available for 277 on either Day 1 or Day 2 and for 228 LHWs on both days of Drive 1.[36] Of our 338 LHWs, measures for median distance traveled every 15 minutes are available for 274 on either Day 1 or Day 2 and for 226 LHWs on both days of Drive 1.[37]

[35]We focus only on the distance traveled and time taken for vaccinations between 8 am and 6pm each day. The distribution of time taken and distance traveled carried some extreme outliers for some subjects. As such, we felt the median was an appropriate summary statistic. Though we had expected to receive geo-stamp data approximately every 10 vaccination attempts, when the monitoring data arrived we noted substantial variance in the number of vaccinations with common geo-stamps and sequences of geo-stamps which 'bounced' back and forth between geographic coordinates. In order to not overstate subject movements, we opted to take average coordinates within a 15 minute window and calculate direct-line distance between window-average coordinates as our measures of distance.

[36]265 LHWs have Day 1 lapse data while 240 have Day 2 lapse data. Of the 73 LHWs with missing Day 1 data, 68 completed either zero or one vaccination on Day 1 such that time lapse between vaccination attempts is not calculable. The remaining 5 conducted vaccinations but did not have phones that interacted with the server to report time use. Of the 98 LHWs with missing Day 2 data, 92 of them completed either zero or one vaccination on Day 2 and the remaining 6 did not have phones that interacted with the server to report time use. Those LHWs who completed vaccinations but did not have interaction with the server had their vaccination records pulled manually from their phones after the drive.

[37]257 LHWs have Day 1 distance data while 240 have Day 2 distance data. Of the 81 LHWs with missing Day 1 data, 75 completed four or fewer vaccination attempts on Day 1 such that distance traveled between 15 minute windows is not calculable. The remaining 6 conducted vaccinations but either did not have phones that interacted with the server to report location or had faulty Global Position Systems (GPS) in their phones. Of the 98 LHWs with missing Day 2 data, 96 of them completed four or fewer vaccination attempts on Day 2 and the remaining 2 did not have phones that interacted with the server to report location or had faulty GPS.

LHWs take around 3.4 minutes between vaccination attempts and walk around 0.06 miles per 15 minutes on Day 1. Focusing on individuals with measures on both days of the drive, we find that time taken and distance traveled are uncorrelated both with Advance choice and with discount factors within condition. Time and distance are also uncorrelated with Advance choice and discount factors on Day 2 of the drive. Further, differences in time taken or distance walked are statistically indistinguishable from zero, uncorrelated with allocation timing, and uncorrelated with discount factors within condition. These data indicate stability in required average effort per vaccination which is unrelated to assignment to Advance or Immediate choice, and that changes in efficacy are unrelated to measured preferences. This suggests that perceived changes in costs likely do not drive our measures of patience or our finding of present bias.[38] These results are all presented in Appendix Table A.7.

Assumption 2: No Idiosyncratic Costs: We assume that vaccinations are the only argument of costs when identifying time preferences. However, there may be idiosyncratic costs across time or individuals that could influence measured patience. For example, a LHW with an appointment lasting 2 hours on Day 1 and no appointments on Day 2 may find it extremely costly to allocate vaccinations to Day 1. This may appear to the researcher as impatience, but only reflects the LHW's idiosyncratic costs across days. Further, if such idiosyncratic events are easier to re-organize when making Advance choice, present bias may be conflated with ease of scheduling.

Here, again, the additional data on LHW time use available from the monitoring application is potentially valuable. We can investigate whether extended periods of non-vaccination exist and if they are correlated with measured preferences and allocation timing. As in the example above, a LHW with an extended period of non-vaccination may well be experiencing forecasted

[38]Ultimately, such stationarity is likely to be expected given that LHWs are already well-versed in vaccination procedures, have an average of 10.5 years of experience as LHWs, and received a half day's training on the vaccination monitoring application.

idiosyncratic costs unrelated to vaccinations. Appendix Table A.8 repeats the analysis from Appendix Table A.7, with dependent variables of the maximum daily time lapse between vaccination attempts and whether the longest daily break is in excess of two hours. Longest daily breaks are, on average, around 59 minutes on Day 1 with around 13% of LHWs taking longest breaks in excess of 2 hours. Focusing on individuals with measures on both days of the drive, we find that the length of longest breaks and the probability of 2 hour breaks are uncorrelated with Advance choice and uncorrelated with discount factors within condition. Almost identical patterns are observed on Day 2 of the drive. Differences in break behavior across days are statistically indistinguishable from zero, uncorrelated with allocation timing, and uncorrelated with discount factors within condition. These data suggest that idiosyncratic costs identified from taking extended breaks do not explain the extent of impatience in the sample, and that potential difficulties in rescheduling do not explain observed present bias.

Assumption 3: Identical Cost Functions: Our aggregate exercise assumes identical costs across subjects, and our individual elicitation assumes identical quadratic costs. Though these assumptions allow for straightforward estimation and calculation of time preferences, any violation would lead us to confound differences in patience across individuals or across allocation timing with differences in costs. One natural view would be to assume that individuals do not discount at all, $\delta = 1$ and $\beta = 1$, such that allocations identify only the shape of the cost function. In this case, when $R = 1$, all LHWs, regardless of allocation timing, should exhibit $v_1 = v_2 = 150$ for all values of γ.[39] Examining the Drive 0 and Drive 1 data, we find that for 163 LHWs who were assigned $R = 1$, the mean allocation is $v_1 = 140.84$ ($s.d. = 24.76$).[40] Though the median allocation is indeed 150, responses range widely with 5th-95th percentiles of response being 103 to 160. If heterogeneity in costs were driving response and discounting was not a key feature of the data, one would not expect to see this extent of variation in

[39]This is because the Euler equation reduces to $(\frac{v_1}{v_2})^\gamma = R = 1$, which implies $\frac{v_1}{v_2} = 1$.

[40]42 of 163 LHWs allocated exactly $v_1 = v_2 = 150$.

response when $R = 1$. Further, given random assignment to allocation timing, heterogeneity in costs does not easily rationalize the observed present bias in the data.

Assumption 4: Only Failure, No Shirking: Our structural exercise assumes individuals know their likelihood to succeed and work only some minimal amount (e.g., that associated with the outside option) in the case where their target is not attainable. Appendix Figure A.6 demonstrates the plausibility of this assumption with a bimodal pattern of almost complete success and almost complete failure. Another possibility is that subjects find an alternate way to renege on their contracts by shirking and still receiving pay. Not all vaccination attempts are equally challenging. In Appendix Figure A.7 we plot for each half-hour of Drive 1 the total number of attempted vaccinations along with the probability of successful vaccination and the probability that no child was reported as present. Reporting that no child was present is likely to be less time consuming than a successful vaccination and easier to falsify. The vast majority of vaccination activity occurs before 3:00pm, there exists no sharp uptick in activity as days end, and we find evidence that LHWs' proportion of successful or failed vaccination attempts remains largely steady throughout the workday. This suggests that allocated vaccination attempts are conducted with due diligence.

Assumption 5: No Biases in Choice: Our study assumes that the allocation environment itself induces no biases in choice such that LHW allocations are directly informative of preferences. A substantial literature in experimental economics suggests that aspects of the decision environment may deeply influence measures of preferences (for recent examples, see Harrison, Lau, Rutstrom and Sullivan, 2005; Beauchamp, Benjamin, Chabris and Laibson, 2015). One common view is that subjects are biased towards the middle of a choice set. In our environment, this could involve subjects opting for either equal allocations of $v_1 = v_2$, or choosing an allocation in the middle of their budget constraint, $v_1 = Rv_2$. Only 31 of 338 LHWs (9%)

exhibit $v_1 = v_2$. Taking a less conservative measure of $v_2 - 2.5 \leq v_1 \leq v_2 + 2.5$, we find that still only 58 of 338 LHWs (17%) are within 5 vaccinations of $v_1 = v_2$.[41] Only 35 of 338 LHWs (10.3%) exhibit $v_1 = Rv_2$. Taking a less conservative measure of $Rv_2 - 2.5 \leq v_1 \leq Rv_2 + 2.5$, we find that 83 of 338 LHWs (25%) are within 5 vaccinations of $v_1 = Rv_2$.[42] Taken together, this suggests that biases towards the middle of the budget constraint or towards equal allocation are unlikely to be driving substantial portions of allocation behavior.

4.3 Tailoring Robustness Tests

Our Drive 2 data show that LHWs who are given bonus contracts with a value of R equal to their estimated discount factors provide significantly smoother service than a number of policy comparison groups. Here we examine robustness of this result to alternative measures for smoothness in service provision and alternative measures for tailoring. We also provide an analysis of tailoring by completion and a discussion of alternate policy preferences.

4.3.1 Alternative Measures for Smooth Provision

Our analysis measures the distance to equal provision using the metric $|\frac{v_{1,i}}{v_{2,i}} - 1|$. In Table A.11, we reconduct the analysis of Table 3, using five alternate measures for smoothness. Panel A presents the Euclidean distance to the 45 degree line, $\frac{|v_{1,i} - v_{2,i}|}{\sqrt{2}}$. Panel B presents the Euclidean distance normalized by the total number of vaccinations allocated, $\frac{|v_{1,i} - v_{2,i}|}{\sqrt{2}(v_{1,i} + v_{2,i})}$. Panel C presents the number of sooner vaccinations that would need to be reallocated to reach the 45 degree line, $|v_{1,i} - \frac{300}{1+R}|$. Panel D presents probit regressions for needing to reallocate more than 25 vaccinations, $|v_{1,i} - \frac{300}{1+R}| > 25$. And finally, Panel E presents the value, $min[v_{1,i}, v_{2,i}]$. Across all specifications, the main conclusions are reproduced. However, the results with respect to additional structural tailoring benefits in Immediate choice fall, at times, outside the range of

[41] As an even less conservative measure, 145 of 338 (43%) satisfy $v_2 - 10 \leq v_1 \leq v_2 + 10$.

[42] As an even less conservative measure, 137 of 338 (40.5%) satsify $Rv_2 - 10 \leq v_1 \leq Rv_2 + 10$.

statistical significance. These alternative measures of smooth provision indicate that our results are not an artifact of how one measures the outcome of interest.

4.3.2 Alternative Sample Restrictions and Treatment Measures

Our exercise focuses on LHWs with discount factors between 0.75 and 1.5. Of 337 LHWs in Drive 2, 280 satisfied this requirement. Those LHWs whose discount factors fell outside of this range were given either $R = 0.75$ or 1.5 depending on which bound they were closest to. For such individuals, structural tailoring is not a binary treatment, but rather a continuous difference between their discount factor and the exogenously given price. Indeed, for all LHWs in the untailored group, treatment is also a continuous measure. In Table A.12, Panel A, we reconduct the analysis of Table 3, columns (1) and (2) using as the measure of treatment the absolute difference between each LHW's discount factor and their assigned price, which we label *Structural Tailoring Intensity*. The main results are reproduced; the closer discount factors are to assigned prices, the smoother is provision.[43] In Panel B, we include those individuals in the Boundary Sample with discount factors that lie outside of the bounds of the price assignment. Including these observations does not alter the conclusions; however, it should be noted that treatment is no longer orthogonal to individual preferences as extremely patient and impatient LHWs will receive larger intensity measures on average.[44]

4.3.3 Tailoring and Completion

Our analysis to of probabilistic completion in sub-section 4.1 evaluates completion through the lens of a model and attempts to assess the trade-off between marginal completion probabilities and discounted marginal costs. Though this analysis seems both tractable and yields valuable

[43] Restricting attention only to the untailored group reveals directionally similar, though insignificant, results across all specifications.

[44] Using the indicator for tailoring would not be an appropriate solution to this problem as tailored LHWs with extreme patience or impatience may actually receive relative prices that are further from their policy-optimal values than those in the untailored condition.

predictive insights, an alternative interpretation for non-completion exists. If the *outcome* of failure rather than its probability is perfectly forecasted by the LHW, there is no incentive to respond truthfully. As such, the targets set in Drive 1, and our corresponding inference on time preferences, would be systematically inaccurate for individuals expecting to fail. In effect, successful LHWs are allocating according to equation (2), while unsuccessful LHWs are providing only noisy response. Under this assumption, we should be dramatically less able to predict allocation behavior for LHWs who fail in Drive 1.

Table A.13 repeats the analysis of Table 3, columns (1) and (2) separately for LHWs who completed and failed to complete their Drive 1 targets. Similar magnitude effects are observed for both sets of LHWs, with structural tailoring serving to reduce distance from the equal provision by around one third. Focusing only on the completing subjects, we would reach effectively the same conclusion as our initial analysis. Furthermore, the fact that predictive accuracy remains for subjects who fail to complete demonstrates that there is content to the allocations subjects make regardless of ex-post completion.

4.4 Repeated Measurement and Within-Subject Variation

In Drive 1, when relying on between subjects tests, statistical tests of present bias fall at the cusp of significance. Given the wide heterogeneity in observed patience regardless of decision timing, one may fail to statistically identify present bias even if it exists on average. Indeed, most studies of present bias and dynamic inconsistency are conducted as within-subject exercises with more choices, potentially because of such wide heterogeneity.

Fortunately, our failed Drive 0 and the corresponding re-randomization in Drive 1 allows us to identify present bias using both more data and within-subject variation for LHWs who changed from Advance to Immediate choice (or vice versa) across drives. Appendix Table A.5, provides aggregate estimates following Appendix Table A.2, columns (1) and (2) using this augmented data set. First, we analyze all potential observations, drawing from 622 choices made by 390 LHWs in either Drive 0 or Drive 1. There, we estimate $\beta = 0.895$ (clustered s.e. $= 0.030$) and reject the null hypothesis of no present bias at all conventional levels, $\chi^2(1) = 12.375,\ (p = 0.001)$. The estimated degree of present bias corresponds closely with other recent estimates of working over time from laboratory studies Augenblick et al. (2015); Augenblick and Rabin (2015). Examining only our panel of 232 individuals who participated in both Drive 0 and Drive 1, a similar estimate is obtained, $\beta = 0.930$ (0.032), $\chi^2(1) = 4.624,\ (p = 0.032)$. This significant degree of present bias is driven by within-subject variation. For those individuals who transition from Advance to Immediate Choice or vice-versa across drives we find $\beta = 0.906$ (0.043), and reject the null hypothesis of no present bias $\chi^2(1) = 4.895,\ (p = 0.027)$.

5 Conclusion

This paper examines the potential for policy interventions to be tailored to individual time preferences. We couch this question in an effort to customize contracts for 337 vaccination workers who spend two days each month attempting to deliver polio vaccines in the neighborhoods of Lahore, Pakistan.

We monitor workers' efforts using a smartphone application developed especially for our project, and elicit preferences using a Convex Time Budget design (Andreoni and Sprenger, 2012; Augenblick et al., 2015). Workers in an Advance condition allocate vaccinations over a two day drive prior to the beginning of the drive, while workers in an Immediate condition state their allocations at the beginning fo the first day. Each worker also faces a randomized relative price for converting vaccinations across days. Worker behavior in this drive is used to identify individual time preferences. In a subsequent drive, we tailor contract terms to individual time preferences for half of the workers. This is done by choosing a relative price designed to generate equal provision of effort over the two days of the drive. The other half of workers is given a random uniform price. We contrast our structural, tailored policy with three alternatives drawn from the random uniform condition that span the policy space in two dimensions: atheoretic vs. structural and broad vs. tailored.

Our findings are encouraging. Those workers who receive structural, tailored contract terms are substantially closer to the policy objective than the alternate policies considered. Using individual discounting parameter estimates to form a new incentive contract does indeed have the predicted effect on allocation behavior. To date, little research makes use of such predictive value of discounting estimates. Our results show not only that estimates are predictive, but also that useful parameter estimates are identifiable from a very limited number of experimental choices. This suggests that the substantial effort of articulating and estimating models in this domain has been well-invested. Policymakers should be encouraged by these findings to consider such tailored interventions. In the domain of intertemporal choice, the specific intervention we consider may be of interest for policymakers wishing to achieve smoothness in allocation behavior or consumption over time.

This paper also speaks to a recent discussion on the external validity of randomized control trials. Developing structural models through which to interpret experimental treatment effects potentially provides a means for generalizing results to other settings (Acemoglu, 2010; Banerjee, Chassang and Snowberg, 2016).[45] In our setting, translating from our reduced form experimental treatment effects to a structural model of choice requires a set of potentially strong (and implausible) assumptions.[46] Nonetheless, the findings of predictive validity in this case suggests there is indeed potential for using structure as a means of increasing the external validity of results obtained from a single sample.

Separately, our results link to the growing literature on the personnel economics of the state (Ashraf, Bandiera and Lee, 2015; Bertrand, Burgess, Chawla and Xu, 2016; Finan, Olken and Pande, Forthcoming; Dal Bó, Finan and Rossi, 2013; Deseranno, 2016; Callen et al., 2019).

[45] Attanasio and Meghir (2012), Duflo, Hanna and Ryan (2012), and Duflo, Greenstone, Pande and Ryan (2016) provide examples in development of using experiments to estimate key policy parameters.

[46] Banerjee et al. (2016) discuss how the plausibility of such identifying assumptions might limit external validity.

Within this literature, there is interest in understanding whether heterogeneity in competencies and in motivation of state actors is linked to meaningful differences in state performance or service provsion (Ashraf et al., 2015; Dal Bó et al., 2013; Deseranno, 2016; Callen et al., 2019). We take the additional step of asking not only whether this heterogeneity matters for outcomes, but also whether it can be acknowledged and reflected in the design of individual incentives.

There are a number of clear limitations to our study which should be addressed by future research. First, our study sidesteps the critical issue of incentive compatibility by not informing subjects of the possibility that their initial behavior would potentially be subsequently used to inform their own contract terms. The mechanism design problem of eliciting preferences and tailoring on said preferences with complete information will be critical if one wishes to implement such contracts repeatedly in the field. Second, future research should seek to gain more precise estimates of preferences. Our exercise requires restrictive assumptions that could be relaxed in the presence of more data. If our results point to a lower bound in the promise of structural, tailored contracts, it is important to know how much more can be achieved. Third, alternative policy objectives and contract types should be investigated to ensure robustness of the identified predictive validity. Our findings have natural extensions to piece rate contracts, multi-period settings, and alternative policy targets that are worthy of study. Notable contributions in this vein include the recent work of Bai, Handel, Miguel and Rao (2019) and Aggarwal, Dizon-Ross and Zucker (2019).

References

Acemoglu, Daron, "Theory, General Equilibrium, and Political Economy in Development Economics," *Journal of Economic Perspectives*, 2010, *24* (3), 17–32.

Aggarwal, Shilpa, Rebecca Dizon-Ross, and Ariel Zucker, "Incentivizing Behavioral Change: The Role of Time Preferences," 2019.

Andreoni, James and Charles Sprenger, "Estimating Time Preferences with Convex Budgets," *American Economic Review*, 2012, *102* (7), 3333–3356.

Ashraf, Nava, Dean Karlan, and Wesley Yin, "Tying Odysseus to the Mast: Evidence from a Commitment Savings Product in the Philippines," *Quarterly Journal of Economics*, 2006, *121* (1), 635–672.

_ , Oriana Bandiera, and Scott S. Lee, "Do-Gooders and Go-Getters: Career Incentives, Selection, and Performance in Public Sector Service Delivery," 2015.

Attanasio, Orazio and Costas Meghir, "Education Choices in Mexico: Using a Structural Model and a Randomized Experiment to Evaluate PROGRESA," *Review of Economic Studies*, 2012, *79* (1), 37–66.

Augenblick, Ned and Matthew Rabin, "An Experiment on Time Preference and Misprediction in Unpleasant Tasks," 2015.

_ , Muriel Niederle, and Charles Sprenger, "Working Over Time: Dynamic Inconsistency in Real Effort Tasks," *Quarterly Journal of Economics*, 2015, *130* (3), 1067–1115.

Bai, Liang, Benjamin Handel, Edward Miguel, and Gautam Rao, "Self-Control and Demand for Preventiv Health: Evidence from Hypertension in India," 2019.

Banerjee, Abhijit, Sylvain Chassang, and Erik Snowberg, "Decision Theoretic Approaches to Experiment Design and External Validity," 2016.

Beauchamp, Jonathan P., Daniel J. Benjamin, Christopher F. Chabris, and David I. Laibson, "Controlling for Compromise Effects Debiases Estimates of Preference Parameters," *Working Paper*, 2015.

Benartzi, Shlomo and Richard H. Thaler, "Save More Tomorrow: Using Behavioral Economics to Increase Employee Saving," *Journal of Political Economy*, 2004, *112* (1), 164–87.

Bertrand, Marianne, Robin Burgess, Arunish Chawla, and Guo Xu, "Determinants and Consequences of Bureaucratic Effectiveness: Evidence from the Indian Administrative Service," 2016.

Beshears, John, James J. Choi, David Laibson, and Brigitee C. Madrian, "The Importance of Default Options for Retirement Savings Outcomes: Evidence from the United States," *National Bureau of Economic Research Working Paper 12009*, 2009.

Blumenstock, Joshua, Michael Callen, and Tarek Ghani, "Why Do Defaults Affect Behavior? Experimental Evidence from Afghanistan," *American Economic Review*, 2018, *108* (10), 2868–2901.

Bruhn, Miriam and David McKenzie, "In Pursuit of Balance: Randomization in Practice in Development Field Experiments," *American Economic Journal: Applied Economics*, 2009, *1* (4), 200–232.

Cagetti, Marco, "Wealth Accumulation Over the Life Cycle and Precautionary Savings," *Journal of Business and Economic Statistics*, 2003, *21* (3), 339–353.

Callen, Michael, Saad Gulzar, Ali Hasanain, Yasir Khan, and Arman Rezaee, "Data and Policy Decisions: Experimental Evidence from Pakistan," 2019.

Carvalho, Leandro S., Stephan Meier, and Stephanie W. Wang, "Poverty and Economic Decision-Making: Evidence from Changes in Financial Resources at Payday," *Working Paper*, 2014.

Castillo, Marco, Paul Ferraro, Jeffrey Jordan, and Ragan Petrie, "The Today and Tomorrow of Kids: Time Preferences and Educations Outcomes of Children," *Journal of Public Economics*, 2011, *95* (11-12), 1377–1385.

Chabris, Christopher F., David Laibson, and Jonathon P. Schuldt, "Intertemporal Choice," in Steven N. Durlauf and Larry Blume, eds., *The New Palgrave Dictionary of Economics*, London: Palgrave Macmillan, 2008.

_ , _ , Carrie Morris, Jonathon Schuldt, and Dmitry Taubinsky, "Individual Laboratory-Measured Discount Rates Predict Field Behavior," *Journal of Risk and Uncertainty*, 2008, *37* (2-3), 237–269.

Cubitt, Robin P. and Daniel Read, "Can Intertemporal Choice Experiments Elicit Preferences for Consumption?," *Experimental Economics*, 2007, *10* (4), 369–389.

Dal Bó, Ernesto, Frederico Finan, and Martín A. Rossi, "Strengthening State Capabilities: The Role of Financial Incentives in the Call to Public Service," *Quarterly Journal of Economics*, 2013, *forthcoming.*

DellaVigna, Stefano and Ulrike Malmendier, "Paying Not to Go to the Gym," *American Economic Review*, 2006, *96* (3), 694–719.

Deseranno, Erika, "Financial Incentives as Signals: Experimental Evidence From the Recruitment of Health Promoters," 2016.

Dohmen, Thomas, Armin Falk, David Huffman, and Uwe Sunde, "Dynamic Inconsistency Predicts Self-Control Problems in Humans," *Working Paper*, 2006.

Duflo, Esther, Michael Greenstone, Rohini Pande, and Nicholas Ryan, "The Value of Regulatory Discretion: Estimates from Environmental Inspections in India," 2016.

_ , Michael Kremer, and Jonathan Robinson, "Nudging Farmers to Use Fertilizer: Theory and Experimental Evidence from Kenya," *American Economic Review*, 2011, *101* (6), 2350–2390.

_ , Rema Hanna, and Stephen P. Ryan, "Incentives Work: Getting Teachers to Come to School," *The American Economic Review*, 2012, *102* (4), 1241–1278.

Finan, Frederico, Benjamin A. Olken, and Rohini Pande, "The Personnel Economics of the State," *Annual Review of Economics*, Forthcoming.

Hansen, Lars Peter, "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, 1982, *50* (4), 1029–1054.

_ and Kenneth J. Singleton, "Generalize Instrumental Variables Estimation of Nonlinear Rational Expectations Models," *Econometrica*, 1982, *50* (5), 1269–1286.

Harrison, Glenn W., Morten I. Lau, and Melonie B. Williams, "Estimating Individual Discount Rates in Denmark: A Field Experiment," *American Economic Review*, 2002, *92* (5), 1606–1617.

_ , _ , Elisabet E. Rutstrom, and Melonie B. Sullivan, "Eliciting Risk and Time Preferences Using Field Experiments: Some Methodological Issues," in Jeffrey Carpenter, Glenn W. Harrison, and John A. List, eds., *Field experiments in economics*, Vol. Vol. 10 (Research in Experimental Economics), Greenwich and London: JAI Press, 2005.

Hastings, Justine and Ebonya Washington, "The First of the Month Effect: Consumer Behavior and Store Responses," *American Economic Journal - Economic Policy*, 2010, *2* (2), 142–162.

Hausman, Jerry A., "Individual Discount Rates and the Purchase and Utilization of Energy-Using Durables," *The Bell Journal of Economics*, 1979, *10* (1).

Kaur, Supreet, Michael Kremer, and Sendhil Mullainathan, "Self-Control and the Development of Work Arrangements," *American Economic Review, Papers and Proceedings*, 2010, *100* (2), 624–628.

_ , _ , and _ , "Self-Control at Work," *Journal of Political Economy*, 2015, *123* (6), 1227–1277.

Koopmans, Tjalling C., "Stationary Ordinal Utility and Impatience," *Econometrica*, 1960, *28* (2), 287–309.

Laibson, David, "Golden Eggs and Hyperbolic Discounting," *Quarterly Journal of Economics*, 1997, *112* (2), 443–477.

_ , Andrea Repetto, and Jeremy Tobacman, "Estimating Discount Functions with Consumption Choices Over the Lifecycle," *Working Paper*, 2005.

Lawrance, Emily C., "Poverty and the Rate of Time Preference: Evidence from Panel Data," *Journal of Political Economy*, 1991, *99* (1), 54–77.

Lee, David S., "Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects," *Review of Economic Studies*, 2009, *76* (3), 1071–1102.

Mahajan, Aprajit and Alessandro Tarozzi, "Time Inconsistency, Ecpectations and Technology Adoption: The Case of Insecticide Treated Nets," *Working Paper*, 2011.

Meier, Stephan and Charles Sprenger, "Discounting Financial Literacy: Time Preferences and Participation in Financial Education Programs," *IZA Discussion Paper Number 3507*, 2008.

_ **and** _ , "Present-Biased Preferences and Credit Card Borrowing," *American Economic Journal - Applied Economics*, 2010, *2* (1), 193–210.

_ **and** _ , "Time Discounting Predicts Creditworthiness," *Psychological Science*, 2012, *23*, 56–58.

_ **and** _ , "Temporal Stability of Time Preferences," *Review of Economics and Statistics*, 2015, *97* (2), 273–286.

O'Donoghue, Ted and Matthew Rabin, "Doing It Now or Later," *American Economic Review*, 1999, *89* (1), 103–124.

_ **and** _ , "Choice and Procrastination," *The Quarterly Journal of Economics*, 2001, *116* (1), 121–160.

Parente, Paulo and João Santos Silva, "Quantile Regression with Clustered Data," *Journal of Econometric Methods*, 2016, *5* (1), 1–15.

Read, Daniel and Barbara van Leeuwen, "Predicting Hunger: The Effects of Appetite and Delay on Choice," *Organizational Behavior and Human Decision Processes*, 1998, *76* (2), 189–205.

_ **, George Loewenstein, and Shobana Kalyanaraman**, "Combining the Immediacy Effect and the Diversification Heuristic," *Journal of Behavioral Decision Making*, 1999, *12*, 257–273.

Sadoff, Sally, Anya Samek, and Charles Sprenger, "Dynamic Inconsistency in Food Choice: Experimental Evidence from a Food Desert," *Working Paper*, 2015.

Samuelson, Paul A., "A Note on Measurement of Utility," *The Review of Economic Studies*, 1937, *4* (2), 155–161.

Sayman, Serdar and Ayse Onculer, "An Investigation of Time Inconsistency," *Management Science*, 2009, *55* (3), 470–482.

Shapiro, Jesse, "Is There a Daily Discount Rate? Evidence From the Food Stamp Nutrition Cycle," *Journal of Public Economics*, 2005, *89* (2-3), 303–325.

Shapiro, Matthew D., "The Permanent Income Hypothesis and the Real Interest Rate: Some Evidence from Panel Data," *Economics Letters*, 1984, *14* (1), 93–100.

Sprenger, Charles, "Judging Experimental Evidence on Dynamic Inconsistency," *American Economic Review, Papers and Proceedings*, 2015.

Warner, John and Saul Pleeter, "The Personal Discount Rate: Evidence from Military Downsizing Programs," *American Economic Review*, 2001, *91* (1), 33–53.

Zeldes, Stephen B., "Consumption and Liquidity Constraints: An Empirical Investigation," *The Journal of Political Economy*, 1989, *97* (2), 305–346.

A Appendix

A.1 Probabilistic Completion

One critical assumption in our theoretical development above revolves around the force of the implemented incentives. The contracts we implement feature a completion bonus of 1000 rupees paid the day after the drive if both targets, v_1 and v_2 are met. The choice of large bonuses (around 10 times daily wages) followed the design logic discussed in Augenblick et al. (2015). Not completing allocated vaccinations creates a sizable penalty at any given point in time. LHWs should forecast that they will indeed complete the required vaccinations and so allocate them according to their true preferences. If LHWs forecast not completing required vaccinations with some chance, the probability of completion has the potential to confound this approach to measuring preferences.

Consider a LHW with probability $p(v_1, v_2)$ of successfully completing her allocated targets. Hence, the expected disutility of effort is

$$p(v_1, v_2)[v_1^\gamma + \beta^{\mathbf{1}_{d=1}}\delta \cdot v_2^\gamma] + (1 - p(v_1, v_2))[v_{1,n}^\gamma + \beta^{\mathbf{1}_{d=1}}\delta \cdot v_{2,n}^\gamma],$$

where $(v_{1,n}, v_{2,n})$ are expected work to be completed on days one and two when not able to complete the contract (e.g., perhaps the standard work-load). Similarly, the expected bonus utility is

$$p(v_1, v_2)\delta^2 u(1000) + (1 - p(v_1, v_2))\delta^2 u(0).$$

For simplicity, we normalize the net utility under non-completion $\delta^2 u(0) - v_{1,n}^\gamma - \beta^{\mathbf{1}_{d=1}}\delta \cdot v_{2,n}^\gamma$ to be zero. Under this assumption, allocations are delivered by the constrained optimization problem

$$max_{v_1,v_2} p(v_1, v_2)[\delta^2 u(1000) - v_1^\gamma - \beta^{\mathbf{1}_{d=1}}\delta \cdot v_2^\gamma]$$
$$s.t.\ v_1 + Rv_2 = V.$$

The corresponding marginal condition,

$$\gamma v_1^{\gamma-1} - \frac{\beta^{\mathbf{1}_{d=1}}\delta}{R}\gamma v_2^{\gamma-1} = \left(\frac{\frac{\partial p(v_1,v_2)}{\partial v_1} - \frac{1}{R}\frac{\partial p(v_1,v_2)}{\partial v_2}}{p(v_1,v_2)}\right)[\delta^2 u(1000) - v_1^\gamma - \beta^{\mathbf{1}_{d=1}}\delta \cdot v_2^\gamma],$$

highlights a central tradeoff between discounted marginal costs and marginal completion probabilities. Of course, if the probability of success is independent of choice, $\frac{\partial p(v_1,v_2)}{\partial v_1}, \frac{\partial p(v_1,v_2)}{\partial v_2} = 0$, the formulation provided in equation (1) is maintained. Otherwise, probabilistic completion can create a wedge, influencing choice and biasing resulting inference on time preference if equation (1) is assumed.

The challenge created by probabilistic completion in settings like ours can be overcome with additional assumptions of functional form and internal consistency. Provided a functional form for $p(v_1, v_2)$, we assume LHWs know the correct mapping,

$$p(v_1, v_2) = p^*(v_1, v_2),$$

where $p^*(v_1, v_2)$ is the true completion probability induced by a given allocation (v_1, v_2). The researcher observes either success or failure as draws from the distribution $p^*(v_1, v_2)$.[47] To provide a functional form for $p(v_1, v_2)$, we assume that the probability of completing a target

[47]Hence, the function $p(v_1, v_2)$, known to the LHW, can be recovered from choice and observed success. It is as if $p(v_1, v_2)$ represents the physical possibility of achieving a given allocation. Given that we assume all LHWs know this mapping, we assume away failures of rational expectations such as believing one can achieve with higher probability than the truth. Intuitively, as in DellaVigna and Malmendier (2006) such misguided beliefs about efficacy would carry quite similar predictions to those of present-biased preferences.

of v on day 1 or 2 is

$$p_1(v) = p_2(v) = \frac{1}{1+\alpha v}.$$

Provided $\alpha > 0$, this completion function assumes that success is assured at $v = 0$ and diminishes as v increases. As such $p(v_1, v_2) = \frac{1}{1+\alpha v_1}\frac{1}{1+\alpha v_2}$.

Under such probabilistic completion and internal consistency, two moment conditions obtain:

$$\gamma v_1^{\gamma-1} - \frac{\beta^{\mathbf{1}_{d=1}}\delta}{R}\gamma v_2^{\gamma-1} - \left(\frac{-\alpha}{(1+\alpha v_1)} - \frac{1}{R}\frac{-\alpha}{(1+\alpha v_2)}\right)[\delta^2 u(1000) - v_1^\gamma - \beta^{\mathbf{1}_{d=1}}\delta \cdot v_2^\gamma] = 0, \tag{3}$$

$$\frac{1}{1+\alpha v_1}\frac{1}{1+\alpha v_2} - p^*(v_1, v_2) = 0. \tag{4}$$

Standard minimum distance methods can be applied to simultaneously estimate the parameters of $p(v_1, v_2)$ and the discounting parameters of interest.[48] In effect, imposing internal consistency on completion rates allows the researcher to quantify the wedge induced by considering marginal completion probabilities. It is important to note that without quality data on actual completion, the exercise would be effectively impossible; highlighting the value of our implemented monitoring technology.[49]

An additional issue generated by probabilistic completion is the presence of monetary utility, $u(1000)$. This value partially pins down the magnitude of the wedge created by marginal completion probabilities. Indeed the net utility of completion, $[\delta^2 u(1000) - v_1^\gamma - \beta^{\mathbf{1}_{d=1}}\delta \cdot v_2^\gamma]$, can be set to any number with suitable definition of $u(1000)$. Of course, for allocations to carry any information, an obvious participation constraint, $[\delta^2 u(1000) - v_1^\gamma - \beta^{\mathbf{1}_{d=1}}\delta \cdot v_2^\gamma] \geq \delta^2 u(0) - v_{1,n}^\gamma - \beta^{\mathbf{1}_{d=1}}\delta \cdot v_{2,n}^\gamma = 0$, needs to be satisfied.[50] To understand how slack this constraint

[48]Considering completion alone, equation (3) could be estimated with non-linear least squares in a similar way to linear probability models with ordinary least squares. Though, in principle, one might predict completion probabilities outside of the bounds [0,1], in practice this does not occur.

[49]Naturally, the predictions may be sensitive to the imposed functional form of $p(v_1, v_2)$. As such, below we discuss several alternative forms for $p(v_1, v_2)$.

[50]Otherwise the individual would want to set v_1, v_2 to increase the probability of non-completion.

was, we asked our LHWs survey questions attempting to identify the minimum bonus they would require to participate in the program again. Of 330 respondents, 329 said they would participate again for the same 1000 rupees bonus while only 42 said they would participate again if the bonus were 900 rupees. Of course, such responses can be difficult to interpret given a lack of incentives, but one view is that the value $[\delta^2 u(1000) - v_1^\gamma - \beta^{\mathbf{1}_{d=1}}\delta \cdot v_2^\gamma]$ may be only slightly higher than the normalized non-participation value of zero. When assessing probabilistic completion, we set $[\delta^2 u(1000) - v_1^\gamma - \beta^{\mathbf{1}_{d=1}}\delta \cdot v_2^\gamma] = 100$.

To estimate aggregate preferences under probabilistic completion we use minimum distance methods to estimate equations (3) and (4) simultaneously. Corresponding estimates are reported in subsection 4.1 and in Appendix Table A.2. To calculate individual preferences, we assume equation (3) is satisfied with equality at the aggregate estimate of α. As such, individual discounting parameters are calculated as

$$\frac{R_i \cdot \left(v_{1,i} - \left(\frac{-\alpha}{(1+\alpha v_{1,i})} - \frac{1}{R}\frac{-\alpha}{(1+\alpha v_{2,i})}\right)[100]\right)}{v_{2,i}} = (\beta^{\mathbf{1}_{d=1,i}}\delta)_i. \tag{5}$$

These values are reported in subsection 4.1.

Our exercise assumes that individuals know the mapping from vaccinations to completion probabilities and trade off discounted marginal costs and marginal failure probabilities. Two important functional form assumptions inform our development. First, we assume the failure probability (known to the LHW) is given by $p(v_1, v_2) = \frac{1}{\alpha v_1}\frac{1}{\alpha v_2}$. In Appendix Table A.9, we reconduct the analysis of Table A.2, Panel A with two alternate functional forms for $p(v_1, v_2)$. First, we assume $p(v_1, v_2) = \frac{1}{1+\alpha'(v_1^2+v_2^2)}$. Second, we assume $p(v_1, v_2) = \frac{1}{1+\alpha''(v_1^3+v_2^3)}$. Both functional forms carry the property that failure probabilities are declining with the volume of work as long as $\alpha', \alpha'' > 0$. They differ only in the marginal tradeoffs they entail. Very limited differences are observed in aggregate estimates across these functional forms and the one used in the main text. Additionally, when assuming $\gamma = 2$, the pairwise correlations between individual discount factor measures using these three functional forms all exceed 0.99.

Our exercise additionally restricts the net utility of completion, $[\delta^2 u(1000) - v_1^\gamma - \beta^{\mathbf{1}_{d=1}} \delta \cdot v_2^\gamma]$, to be equal to 100. In Appendix Table A.10, we reconduct the analysis of Table A.2, Panel A assuming this net utility equal to 1000 or to 10000. Only small changes in the aggregate estimates are observed. Furthermore, at the individual level when assuming $\gamma = 2$, the pairwise correlations between individual discount factor measures using these three net completion utility values all exceed 0.99.

A.2 Appendix Figures

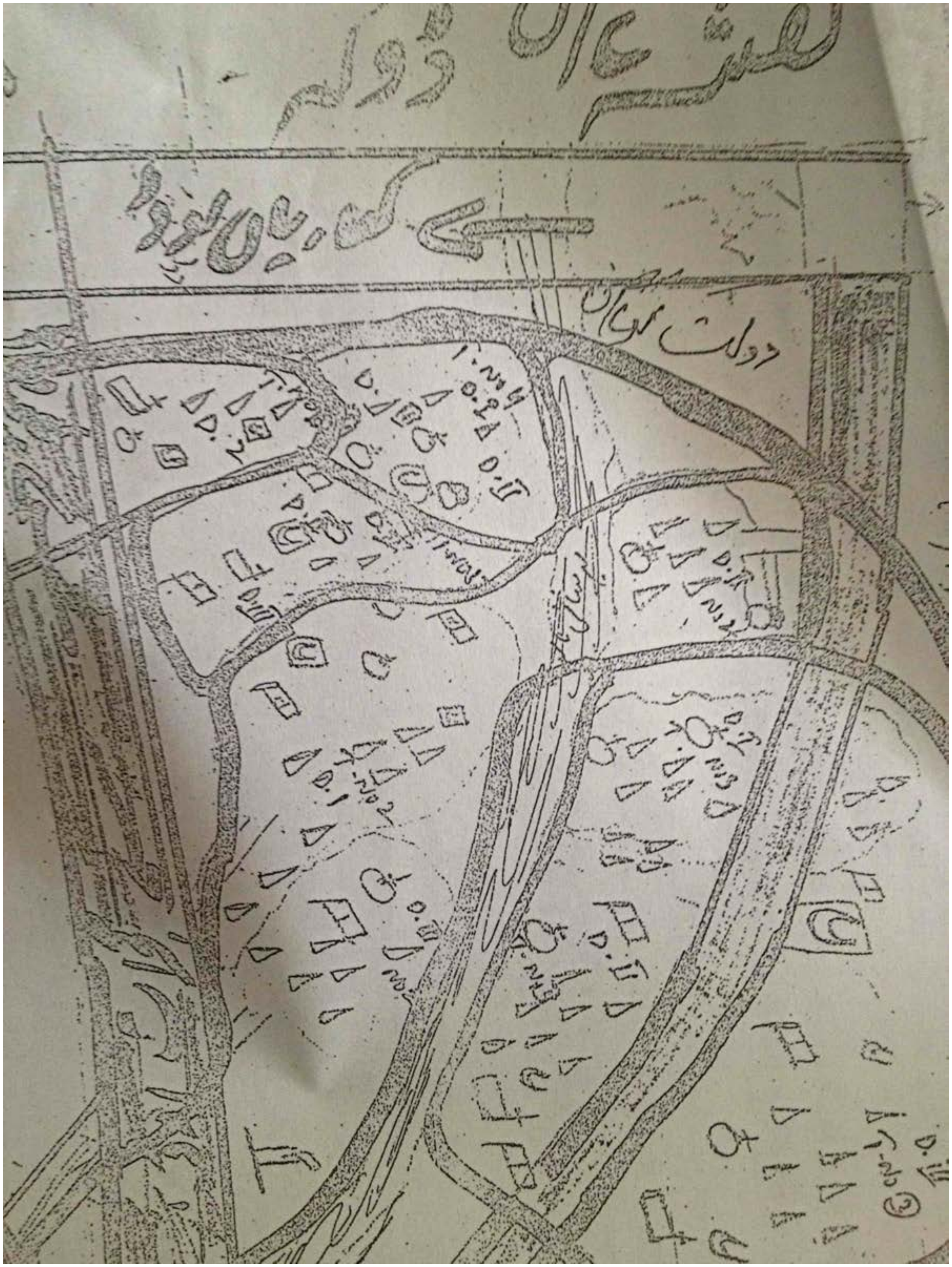


Figure A.1: Map Given to LHWs to Plan Route

Figure A.2: Picture of a Door-to-Door Vaccination During a Drive

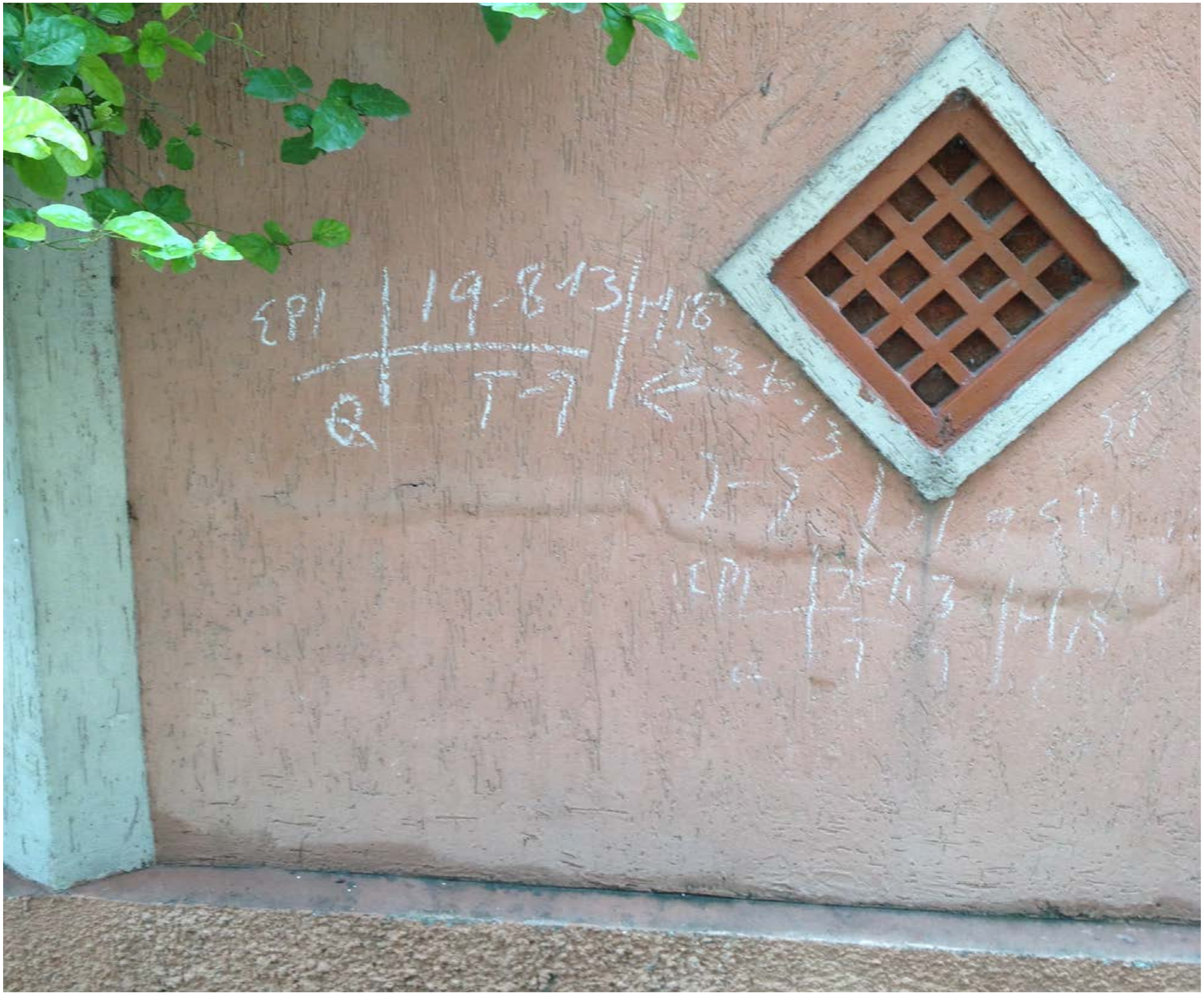
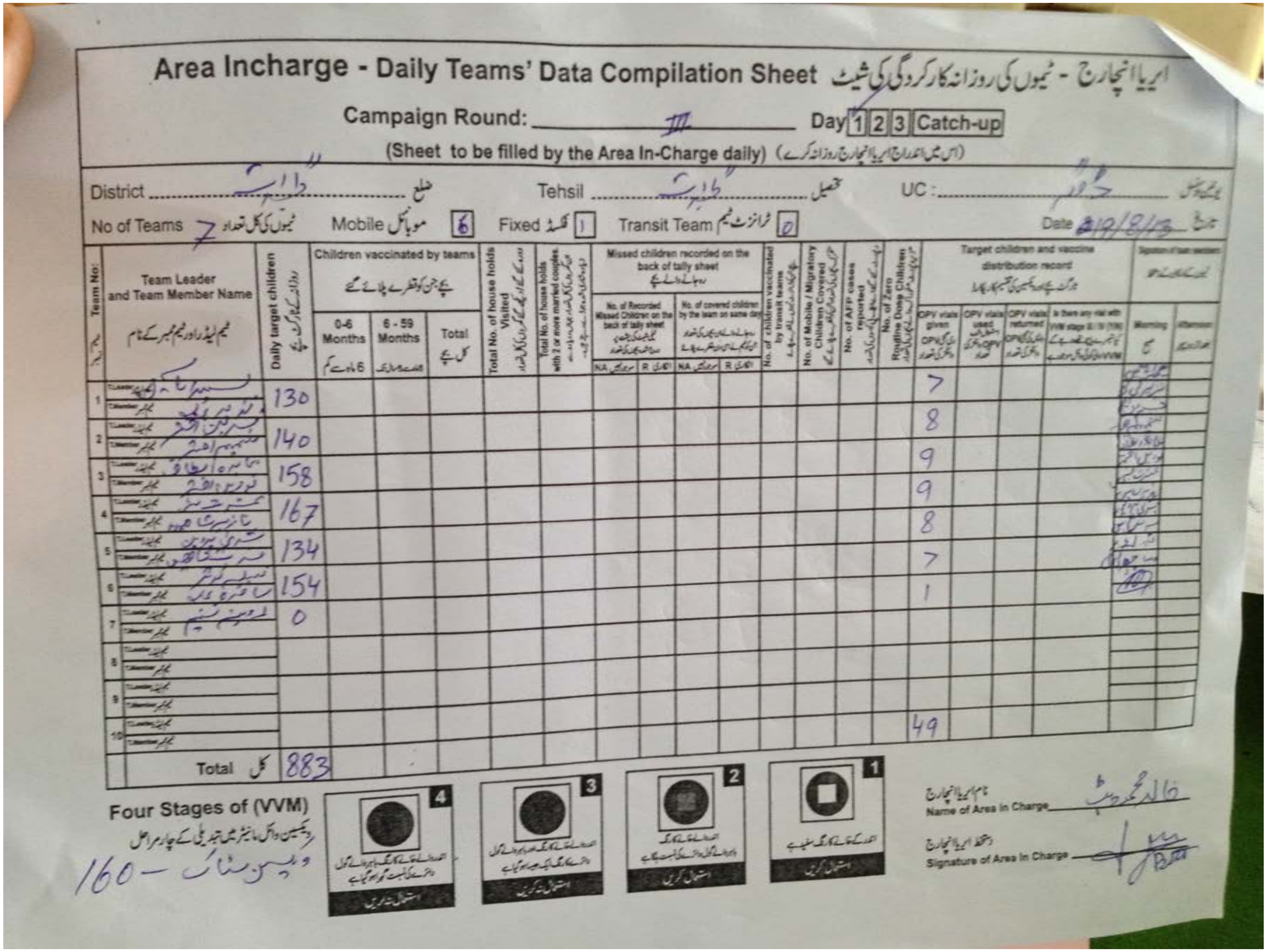


Figure A.3: Chalk Marking to Record Visit by Vaccination Team



Area Incharge - Daily Teams' Data Compilation Sheet اریاانچارج - ٹیموں کی روزانہ کارکردگی کی شیٹ

Campaign Round: III Day 1 2 3 Catch-up

(Sheet to be filled by the Area In-Charge daily)

District: .... Tehsil: .... UC: ....

No of Teams 7 Mobile 6 Fixed 1 Transit Team 0 Date 2/12/8/13

Team No.	Daily target children	Children vaccinated by teams: 0-6 Months	6-59 Months	Total	OPV vials given
1	130				7
2	140				8
3	158				9
4	167				9
5	134				8
6	154				7
7	0				1
8					
9					
10					
Total	883				49

Four Stages of (VVM)

160 —

Name of Area In Charge

Signature of Area In Charge

Figure A.4: End-of-Day Compilation of Self-Reports by Vaccination Teams

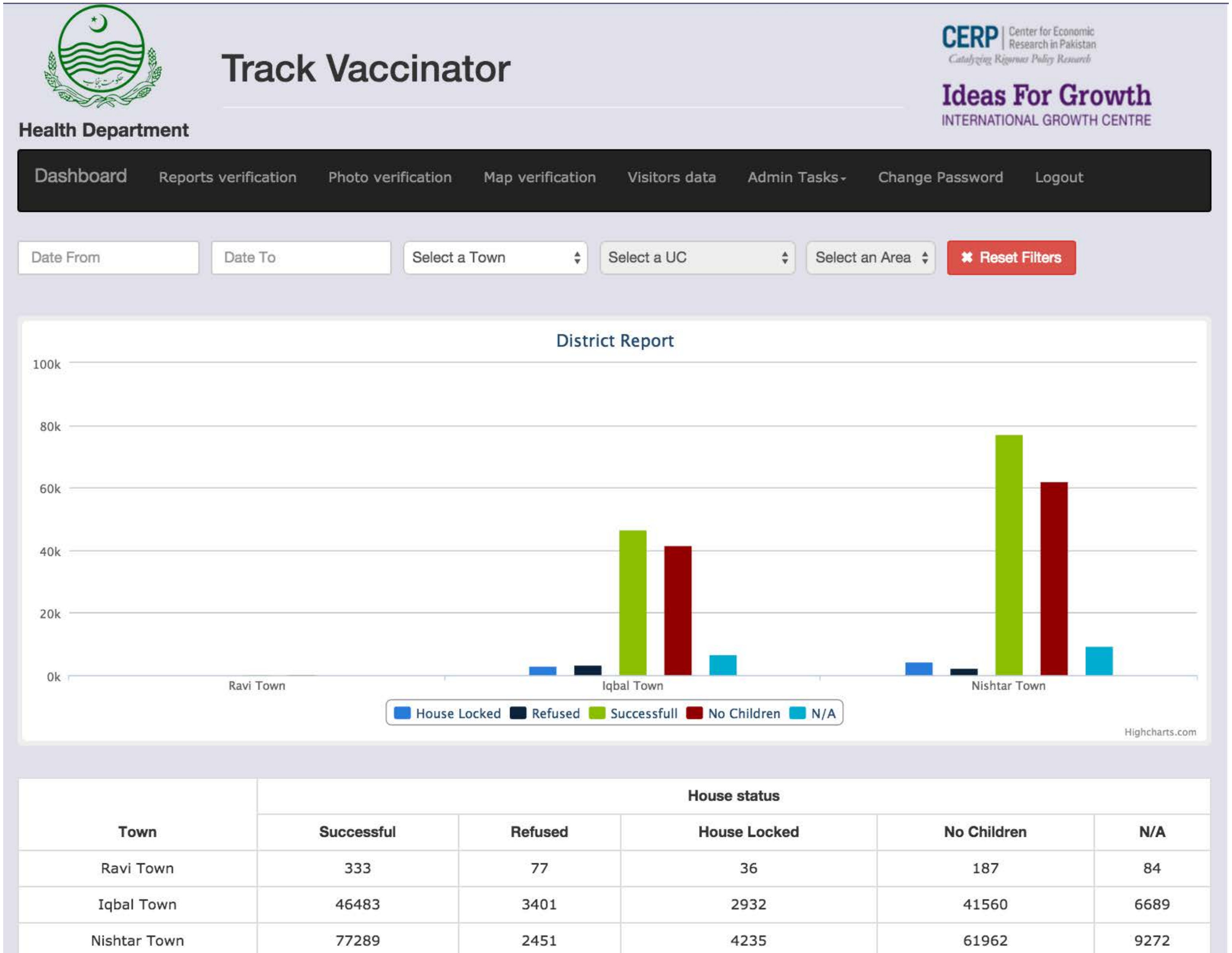


Figure A.5: Screenshot of the Track LHW Dashboard

A.3 Appendix Tables

Table A.1: No Allocation Provided in Drive 0

	Allocation Provided (1)	No Allocation Provided (2)	p-value (3)
Gender (Female = 1)	0.965	1.000	0.082
	(0.020)	(0.000)	
Years of Education	10.294	10.146	0.608
	(0.220)	(0.185)	
Number of Children	3.268	3.388	0.695
	(0.239)	(0.188)	
Punjabi (=1)	0.952	0.975	0.440
	(0.023)	(0.018)	
Has a Savings Account (=1)	0.317	0.305	0.867
	(0.052)	(0.051)	
Participated in a Rosca (=1)	0.446	0.378	0.380
	(0.055)	(0.054)	
Years in Health Department	10.135	10.886	0.337
	(0.554)	(0.547)	
Years as Polio Vaccinator	9.994	10.531	0.467
	(0.538)	(0.502)	
# LHWs	86	82	

Notes: This table tests whether the failure of the smartphone app during Drive 0 was systematic. Standard errors reported in parentheses. Column 3 reports a p-value corresponding to the null that the mean in the Did Not Fail group is equal to the Failed group.

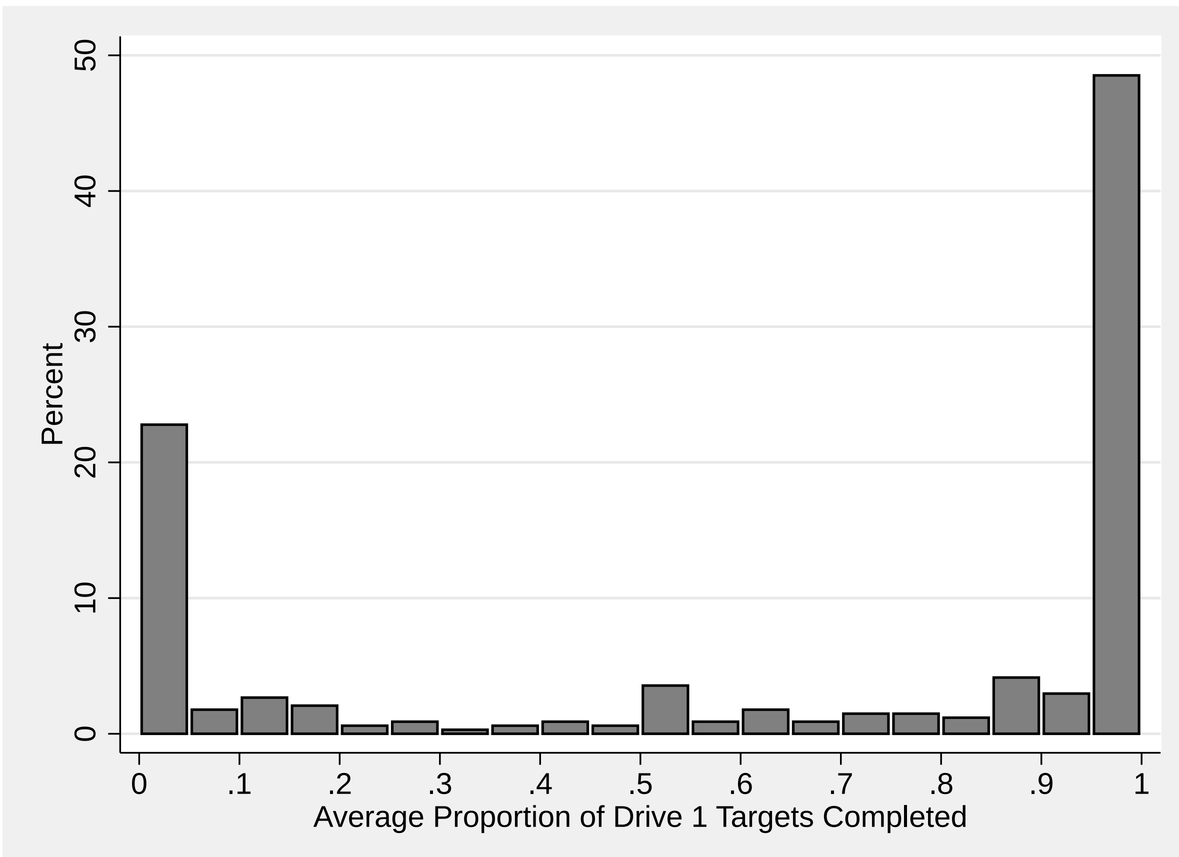


Figure A.6: Individual Completion Rates

Notes: This figure reports individual average completion rates in Drive 1. The individual average completion rate is calculated as $1/2(min(Completed_1/v_1, 1) + min(Completed_2/v_2, 1))$.

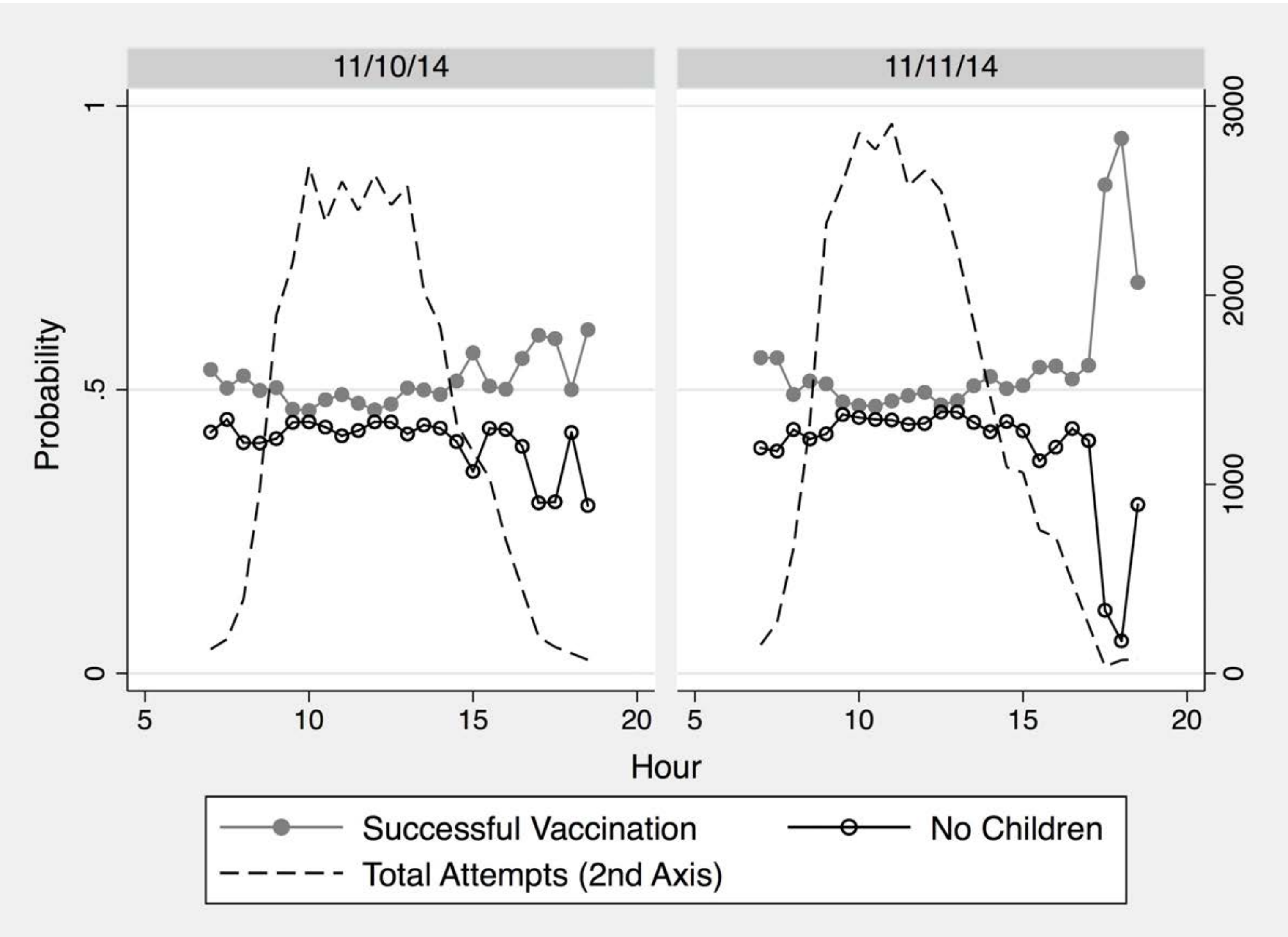


Figure A.7: Drive 1 Vaccination Activity

Notes: The solid light grey circles are the share of all vaccination attempts that reflect a successful vaccination during the indicated hour. The hollow dark black circles are the share of all vaccination attempts that report no children being available during the attempt. These quantities are compared against the left axis. The dotted line indicates the total number of vaccination attempts for all LHWs in the sample. This quantity is compared against the right axis.

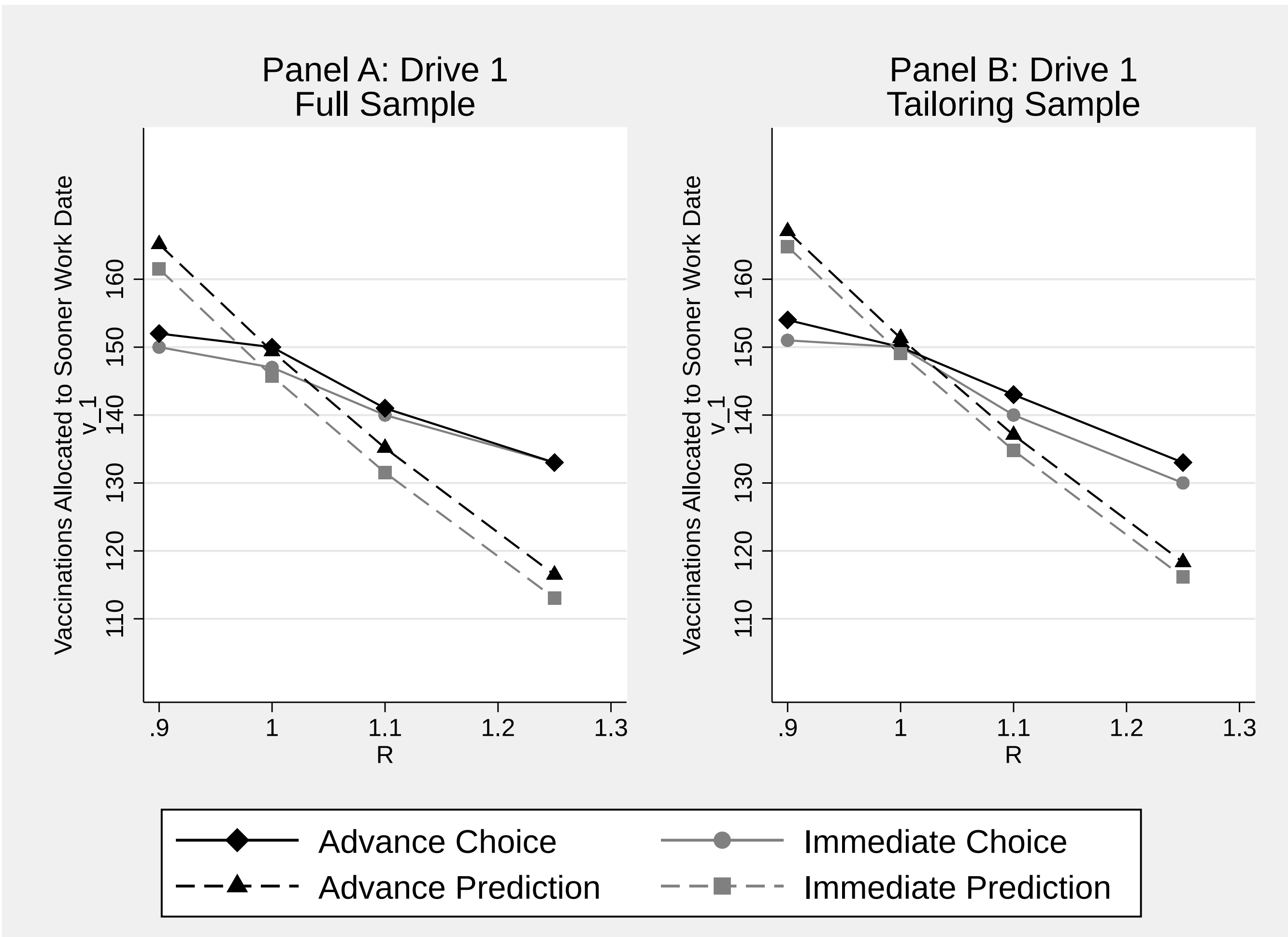


Figure A.8: Predicted and Actual Experimental Response

Notes: Points in the plots are medians for each of the eight treatment groups respectively. Panel A depicts the Full Sample and Panel B depicts the tailoring sample (LHWs with $R^* < 0.75$ or $R^* > 1.5$). Black are advance choice groups and gray are immediate choice groups. The series for predictions correspond to Table A.2, column (2).

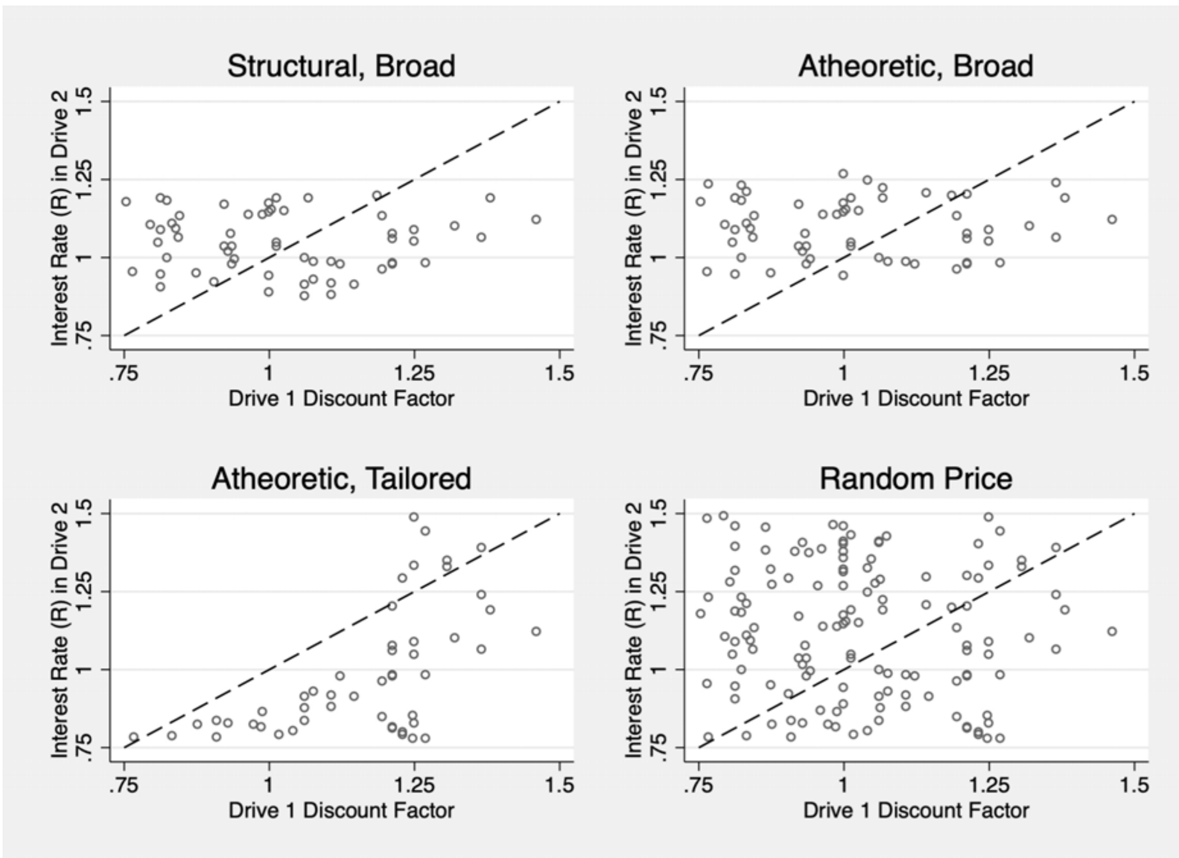


Figure A.9: Exact Assignment of Comparison Policies

Notes: This figure presents the exact assignments of LHWs to four policy comparison groups based on their Drive 1 discount factor: Structural, Broad; Atheoretic, Broad; Atheoretic, Tailored; and Random Price.

Table A.2: Aggregate Parameter Estimates, Drive 1

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	$\gamma = 2$		$\gamma = 2.5$		$\gamma = 3$		$\gamma = 3.25$	
Panel A: Full Sample								
β	0.935	0.952	0.922	0.946	0.906	0.938	0.896	0.934
	(0.030)	(0.029)	(0.040)	(0.039)	(0.050)	(0.050)	(0.055)	(0.055)
δ	0.985	0.992	0.958	0.967	0.932	0.942	0.919	0.931
	(0.017)	(0.017)	(0.022)	(0.022)	(0.027)	(0.027)	(0.029)	(0.030)
α		0.003		0.003		0.003		0.003
		(0.000)		(0.000)		(0.000)		(0.000)
$H_0 : \beta = 1.$ $(\chi^2(1))$	4.841	2.637	3.889	1.856	3.600	1.531	3.609	1.449
[p-value]	[0.028]	[0.104]	[0.049]	[0.173]	[0.058]	[0.216]	[0.057]	[0.229]
Criterion Value	0.278	0.310	0.217	0.249	0.180	0.212	0.166	0.199
# LHWs	338	338	338	338	338	338	338	338
Panel B: Tailoring Sample								
β	0.969	0.970	0.962	0.963	0.954	0.955	0.949	0.951
	(0.018)	(0.018)	(0.023)	(0.023)	(0.028)	(0.028)	(0.031)	(0.031)
δ	1.017	1.018	1.003	1.004	0.990	0.991	0.984	0.985
	(0.013)	(0.013)	(0.016)	(0.016)	(0.019)	(0.019)	(0.020)	(0.020)
α		0.002		0.003		0.003		0.003
		(0.000)		(0.000)		(0.000)		(0.000)
$H_0 : \beta = 1.$ $(\chi^2(1))$	2.880	2.713	2.716	2.575	2.729	2.592	2.754	2.618
[p-value]	[0.090]	[0.100]	[0.099]	[0.109]	[0.099]	[0.107]	[0.097]	[0.106]
Criterion Value	0.370	0.384	0.268	0.284	0.196	0.213	0.170	0.186
# LHWs	281	281	281	281	281	281	281	281

Notes: This reports structural estimates of β, δ, and α obtained using minimum distance estimation of equations (1) in even columns or (3) and (4) in odd columns. Standard errors are reported in parentheses. Test statistic for $\beta = 1$ with p-value in brackets. Panel A provides estimates for the Full Sample, Panel B provides estimates for the Tailoring Sample.

Table A.3: Tailoring Intertemporal Incentives, Complete Tailoring Sample

Dependent variable:	$\lvert\frac{v_{1,i}}{v_{2,i}} - 1\rvert$							
Policy Comparison Group	Random Price		Structural, Broad		Atheoretic, Broad		Atheoretic, Tailored	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Structural Tailored (=1)	-0.346	-0.002	-0.409	-0.051	-0.402	-0.031	-0.019	-0.025
	(0.234)	(0.086)	(0.389)	(0.052)	(0.378)	(0.042)	(0.033)	(0.025)
Immediate Choice		0.866*		0.958		0.989		0.165***
		(0.496)		(0.814)		(0.802)		(0.063)
Structural Tailored x Immediate		-0.782		-0.843		-0.873		-0.032
		(0.532)		(0.834)		(0.819)		(0.075)
Constant	-0.244	-0.416	0.476	0.326	0.740	0.558	0.112	0.062
	(0.992)	(1.009)	(0.356)	(0.271)	(0.608)	(0.473)	(0.102)	(0.100)
Stratum FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Exclude 99th and 1st Percentiles	No	No	No	No	No	No	No	No
Drive 2 R_i^* or $\tilde{R}_i$	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-Squared	0.047	0.061	0.032	0.053	0.031	0.053	0.030	0.132
Mean in Comparison Contract	0.612	0.612	0.563	0.563	0.575	0.575	0.148	0.148
Mean in Comparison Advanced		0.098		0.100		0.088		0.091
Mean in Comparison Immediate		1.190		1.163		1.167		0.265
# LHWs	280	280	204	204	204	204	191	191
# Comparison LHWs	138	138	62	62	62	62	49	49

Notes: This table reports the effects of tailoring on the equality of effort provision over time. The measure $\lvert\frac{v_{1,i}}{v_{2,i}} - 1\rvert$ (the percentage difference between tasks allocated to day 1 and day 2 of the drive) reflects the distance of the task allocation (v_1,v_2) from equality ($v_1 = v_2$). Section 3.2 provides definitions of comparison groups. Ordinary least squares regressions. Heteroskedasticity robust White standard errors reported in parentheses. $*p < 0.1$, $**p < 0.05$, $***p < 0.01$.

Table A.4: Aggregate Drive 0 and 1 Behavior

Dependent variable:	Tasks Allocated to the First Day of the Drive (v_1)	
	Full Sample	Tailoring Sample
	(1)	(2)
	Median	Median
Immediate Decision (=1)	-2.00**	-3.00***
	(0.95)	(0.88)
Relative Price (R)	-40.00***	-60.00***
	(6.04)	(4.12)
Constant	188.00***	210.00***
	(6.06)	(4.34)
Median Advance Choice	150	150
# Observations	622	475

Notes: This table reports on the effects of making Immediate allocation decisions and R on Drive 0 and Drive 1 vaccinations allocated to the first day of the drive. Median regression coefficients with standard errors reported in parentheses. Standard errors are clustered at the LHW level. Clustered standard errors for quantile regressions are calculated using the approach in Parente and Santos Silva (2016). Immediate Decision is an indicator equal to one for LHWs selecting their allocations on the morning of the vaccination drive. The relative price R takes the values $R \in \{0.9, 1, 1.1, 1.25\}$. *Levels of Significance:* $*p < 0.1$, $**p < 0.05$, $***p < 0.01$.

Table A.5: Within Subject Parameter Estimates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	All Observations		Panel Only		Panel No Change		Panel Change	
β	0.895	0.903	0.930	0.931	0.946	0.944	0.906	0.912
	(0.030)	(0.030)	(0.032)	(0.032)	(0.048)	(0.048)	(0.043)	(0.042)
δ	0.979	0.979	0.988	0.988	0.967	0.967	1.012	1.010
	(0.017)	(0.017)	(0.016)	(0.016)	(0.023)	(0.023)	(0.023)	(0.023)
α		0.003		0.002		0.003		0.002
		(0.000)		(0.000)		(0.000)		(0.000)
$H_0 : \beta = 1.\ (\chi^2(1))$	12.375	10.571	4.624	4.497	1.283	1.399	4.895	4.344
[p-value]	[0.000]	[0.001]	[0.032]	[0.034]	[0.257]	[0.237]	[0.027]	[0.037]
Criterion Value	0.182	0.191	0.198	0.200	0.175	0.177	0.217	0.221
# Observations	622	622	464	464	212	212	252	252
# LHWs	390	390	232	232	106	106	126	126

Notes: This reports structural estimates of β, δ, and α obtained using minimum distance estimation of equations (1) in odd columns or (3) and (4) in even columns. Estimates provided for either all possible observations or panel of individuals participating in both Failed Drive 0 and Drive 1. Only Drive 1 data used for estimation of completion parameter, α. Standard errors clustered on individual level are reported in parentheses. Test statistic for $\beta = 1$ with p-value in brackets.

Table A.6: Drive 1 Completion

	(1) Full Sample	(2) Tailoring Sample
Dependent Variable: v_1		
Immediate Decision (=1)	-0.087	-0.128**
	(0.054)	(0.060)
Relative Price (R)	0.143	0.095
	(0.210)	(0.241)
Constant	0.405*	0.483*
	(0.227)	(0.258)
Advance Completion Probability	0.557	0.582
# Observations	338	281

Notes: This table reports on the effects of decision timing and relative price variation on completion in Drive 1. Linear probability models with robust standard errors. *Levels of Significance:* $*p < 0.1$, $**p < 0.05$, $***p < 0.01$.

Table A.7: Testing Stationarity of Costs Across Days

Panel A: Time Lapse Between Vaccinations (in minutes)								
Dependent variable:	Day 1 Med. Time Lapse			Day 2 Med. Time Lapse			Day 1 - Day 2 Med. Time Lapse	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Advance Choice (=1)	0.519	1.134	1.011	-0.910	-1.161	-0.829	2.295	1.840
	(2.492)	(1.163)	(1.045)	(3.164)	(3.324)	(3.182)	(3.527)	(3.343)
Discount Factor			-3.697			10.004		-13.701
			(3.504)			(8.247)		(9.000)
Constant	3.370*	1.422***	5.337	4.447*	4.540*	-6.053	-3.118	11.390
	(1.851)	(0.084)	(3.708)	(2.372)	(2.501)	(6.558)	(2.501)	(7.581)
R-Squared	0.000	0.004	0.016	0.000	0.001	0.013	0.002	0.022
# Observations	265	228	228	240	228	228	228	228
Panel B: Distance Walked Between Vaccinations (in Kilometers)								
Dependent variable:	Day 1 Med. Distance			Day 2 Med. Distance			Day 1 - Day 2 Med. Distance	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Advance Choice (=1)	0.112	0.146	0.132	-0.148	-0.171	-0.154	0.317	0.286
	(0.144)	(0.154)	(0.139)	(0.152)	(0.161)	(0.144)	(0.223)	(0.199)
Discount Factor			-0.444			0.509		-0.953
			(0.466)			(0.516)		(0.697)
Constant	0.059**	0.038***	0.507	0.201	0.201	-0.337	-0.164	0.844
	(0.026)	(0.010)	(0.492)	(0.151)	(0.161)	(0.388)	(0.162)	(0.629)
R-Squared	0.002	0.004	0.014	0.004	0.005	0.020	0.009	0.033
# Observations	257	226	226	240	226	226	226	226

Notes: This table reports on the relationship between decision timing and the one period discount factor with two proxies of the cost of performing a vaccination (the amount of time that lapses between vaccinations and the distance traveled between vaccinations). Heteroskedasticity robust White standard errors reported in parentheses. $*p < 0.1$, $**p < 0.05$, $***p < 0.01$.

Table A.8: Testing for Idiosyncratic Shocks

Panel A: Maximum Daily Time Lapse Between Vaccinations (in Minutes)								
Dependent variable:	Max Day 1 Time Lapse			Max Day 2 Time Lapse			Day 1 - Day 2 Max Time Lapse	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Advance Choice (=1)	-0.702	0.578	0.739	8.067	5.274	5.574	-4.695	-4.836
	(9.783)	(9.104)	(9.026)	(8.885)	(9.114)	(9.021)	(12.319)	(12.211)
Discount Factor			4.831			9.054		-4.223
			(14.139)			(15.570)		(19.824)
Constant	59.258***	54.880***	49.764***	53.437***	54.362***	44.774***	0.518	4.990
	(7.920)	(7.254)	(15.266)	(5.178)	(5.404)	(15.351)	(8.724)	(20.662)
R-Squared	0.000	0.000	0.000	0.003	0.001	0.003	0.001	0.001
# Observations	265	228	228	240	228	228	228	228
Panel B: Maximum Time Lapse > 2 hours								
Dependent variable:	Max Day 1 Lapse > 2hr.			Max Day 2 Time Lapse > 2hr.			Day 1 > 2hr. - Day 2 > 2hr.	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Advance Choice (=1)	0.051	0.042	0.044	0.026	0.001	0.002	0.041	0.042
	(0.045)	(0.047)	(0.047)	(0.041)	(0.042)	(0.041)	(0.060)	(0.060)
Discount Factor			0.053			0.032		0.021
			(0.077)			(0.071)		(0.100)
Constant	0.133***	0.127***	0.071	0.103***	0.109***	0.075	0.018	-0.004
	(0.029)	(0.032)	(0.085)	(0.028)	(0.030)	(0.075)	(0.043)	(0.111)
R-Squared	0.005	0.004	0.005	0.002	0.000	0.001	0.002	0.002
# Observations	265	228	228	240	228	228	228	228

Notes: This table reports on the relationship between decision timing and the one period discount factor with two proxies for experiencing a shock during the drive (the maximum time lapse between vaccinations and whether a lapse of more than 2 hours occurred). Heteroskedasticity robust White standard errors reported in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A.9: Aggregate Parameter Estimates, Alternate Probabilistic Completion

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	$\gamma = 2$		$\gamma = 2.5$		$\gamma = 3$		$\gamma = 3.25$	
$p(v_1, v_2) =$	$\frac{1}{1+\alpha'(v_1^2+v_2^2)}$	$\frac{1}{1+\alpha''(v_1^3+v_2^3)}$	$\frac{1}{1+\alpha'(v_1^2+v_2^2)}$	$\frac{1}{1+\alpha''(v_1^3+v_2^3)}$	$\frac{1}{1+\alpha'(v_1^2+v_2^2)}$	$\frac{1}{1+\alpha''(v_1^3+v_2^3)}$	$\frac{1}{1+\alpha'(v_1^2+v_2^2)}$	$\frac{1}{1+\alpha''(v_1^3+v_2^3)}$
β	0.944	0.936	0.935	0.924	0.924	0.910	0.917	0.883
	(0.030)	(0.030)	(0.040)	(0.039)	(0.050)	(0.050)	(0.055)	(0.053)
δ	0.992	0.992	0.967	0.966	0.943	0.942	0.931	0.926
	(0.017)	(0.017)	(0.022)	(0.022)	(0.027)	(0.027)	(0.030)	(0.028)
α'	0.000022		0.000022		0.000022		0.000022	
	(0.000002)		(0.000002)		(0.000002)		(0.000002)	
α''		$1.48e^{-7}$		$1.50e^{-7}$		$1.50e^{-7}$		6.58
		$(1.63e^{-8})$		$(1.64e^{-8})$		$(1.65e^{-8})$		(.)
$H_0 : \beta = 1.\ (\chi^2(1))$	3.583	4.650	2.700	3.681	2.341	3.299	2.261	4.904
[p-value]	[0.058]	[0.031]	[0.100]	[0.055]	[0.126]	[0.069]	[0.133]	[0.027]
Criterion Value	0.308	0.306	0.248	0.247	0.211	0.210	0.198	0.551
# LHWs	338	338	338	338	338	338	338	338

Notes: This reports structural estimates of β, δ, and α obtained using minimum distance estimation of equations (3) and (4). Standard errors are reported in parentheses. Test statistic for $\beta = 1$ with p-value in brackets.

Table A.10: Aggregate Parameter Estimates, Alternate Net Completion Utility

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	$\gamma = 2$		$\gamma = 2.5$		$\gamma = 3$		$\gamma = 3.25$	
$[\delta^2 u(1000) - v_1^\gamma - \beta^{\mathbf{1}_{d=1}}\delta \cdot v_2^\gamma] =$	1000	10000	1000	10000	1000	10000	1000	10000
β	0.952	0.952	0.946	0.946	0.938	0.940	0.934	0.906
	(0.029)	(0.029)	(0.039)	(0.039)	(0.050)	(0.050)	(0.055)	(0.053)
δ	0.992	0.996	0.967	0.967	0.942	0.943	0.931	0.915
	(0.017)	(0.017)	(0.022)	(0.022)	(0.027)	(0.027)	(0.030)	(0.029)
α	0.003	0.003	0.003	0.003	0.003	0.003	0.003	-0.016
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$H_0 : \beta = 1.\ (\chi^2(1))$	2.644	2.710	1.856	1.855	1.525	1.466	1.425	3.070
[p-value]	[0.104]	[0.100]	[0.173]	[0.173]	[0.217]	[0.226]	[0.233]	[0.080]
Criterion Value	0.311	0.326	0.249	0.250	0.212	0.212	0.199	0.185
# LHWs	338	338	338	338	338	338	338	338

Notes: This reports structural estimates of β, δ, and α obtained using minimum distance estimation of equations (3) and (4). Standard errors are reported in parentheses. Test statistic for $\beta = 1$ with p-value in brackets.

Table A.11: Robustness Tests for Tailoring Intertemporal Incentives

Policy Comparison Group	Random Price		Structural, Broad		Atheoretic, Broad		Atheoretic, Tailored	
Panel A: Dependent variable $\frac{\lvert v_{1,i}-v_{2,i}\rvert}{\sqrt{2}}$								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Structural Tailored (=1)	-4.481**	-1.868	-4.854*	-3.040	-4.852*	-1.734	-5.313	-1.511
	(2.068)	(2.229)	(2.718)	(3.227)	(2.469)	(2.571)	(3.223)	(3.058)
Immediate Choice		10.597***		9.488*		12.325**		17.503**
		(3.449)		(5.147)		(4.868)		(7.720)
Structural Tailored x Immediate		-6.220		-4.803		-7.933		-12.911
		(4.136)		(5.645)		(5.375)		(8.140)
Constant	16.412**	14.128**	32.313***	30.893***	26.219***	24.051***	21.422***	16.129**
	(6.857)	(6.671)	(9.981)	(10.234)	(7.690)	(7.545)	(8.096)	(7.874)
Panel B: Dependent variable $\frac{\lvert v_{1,i}-v_{2,i}\rvert}{\sqrt{2}(v_{1,i}+v_{2,i})}$								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Structural Tailored (=1)	-0.016**	-0.007	-0.017*	-0.010	-0.018**	-0.006	-0.019*	-0.005
	(0.007)	(0.008)	(0.009)	(0.011)	(0.009)	(0.009)	(0.011)	(0.010)
Immediate Choice		0.037***		0.033*		0.044**		0.059**
		(0.012)		(0.018)		(0.017)		(0.024)
Structural Tailored x Immediate		-0.023*		-0.019		-0.030		-0.045*
		(0.014)		(0.019)		(0.019)		(0.026)
Constant	0.033	0.025	0.088***	0.083**	0.070***	0.062**	0.052**	0.034
	(0.022)	(0.022)	(0.031)	(0.032)	(0.025)	(0.025)	(0.026)	(0.025)
Panel C: Dependent variable $\lvert v_{1,i} - \frac{300}{1+R}\rvert$								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Structural Tailored (=1)	-3.445**	-1.405	-3.661*	-2.267	-3.856**	-1.452	-3.942*	-1.121
	(1.459)	(1.591)	(1.936)	(2.289)	(1.825)	(1.891)	(2.294)	(2.135)
Immediate Choice		7.844***		6.874*		9.095**		12.664**
		(2.509)		(3.743)		(3.636)		(5.287)
Structural Tailored x Immediate		-4.850		-3.667		-6.092		-9.512*
		(2.974)		(4.072)		(3.970)		(5.582)
Constant	7.571	5.871	19.194***	18.146***	15.291***	13.666***	11.378**	7.539
	(4.735)	(4.622)	(6.750)	(6.917)	(5.332)	(5.208)	(5.542)	(5.409)
Panel D: Dependent variable $\lvert v_{1,i} - \frac{300}{1+R}\rvert > 25$								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Structural Tailored (=1)	0.206	0.146	0.135	0.124	0.245	0.175	0.137	0.021
	(0.168)	(0.234)	(0.215)	(0.298)	(0.212)	(0.302)	(0.242)	(0.313)
Immediate Choice		-0.573**		-0.517		-0.640*		-0.866**
		(0.238)		(0.378)		(0.362)		(0.415)
Structural Tailored x Immediate		0.144		0.064		0.188		0.409
		(0.340)		(0.445)		(0.438)		(0.487)
Constant	1.530***	1.705***	0.164	0.250	0.617	0.720	0.950	1.253**
	(0.513)	(0.524)	(0.779)	(0.797)	(0.796)	(0.814)	(0.607)	(0.635)
Panel E: Dependent variable $min[v_{1,i}, v_{2,i}]$								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Structural Tailored (=1)	2.540*	0.843	4.329**	2.986	4.404**	2.179	2.711	0.263
	(1.416)	(1.567)	(1.922)	(2.279)	(1.766)	(1.885)	(2.187)	(2.178)
Immediate Choice		-6.815***		-6.452*		-8.383**		-11.173**
		(2.332)		(3.746)		(3.513)		(5.149)
Structural Tailored x Immediate		4.037		3.521		5.639		8.292
		(2.806)		(4.070)		(3.844)		(5.451)
Constant	208.758***	210.228***	200.999***	201.990***	204.405***	205.905***	208.831***	212.213***
	(4.541)	(4.433)	(6.724)	(6.882)	(5.357)	(5.228)	(5.476)	(5.306)
Stratum FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Exclude 99th and 1st Percentiles	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Drive 2 R_i^* or $\tilde{R}_i$	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
# LHWs	267	267	194	194	194	194	184	184
# Comparison LHWs	132	132	59	59	59	59	49	49

Notes: This table reports the effects of tailoring on the equality of effort provision over time. The measure $\lvert \frac{v_{1,i}}{v_{2,i}} - 1\rvert$ (the percentage difference between tasks allocated to day 1 and day 2 of the drive) reflects the distance of the task allocation (v_1,v_2) from equality ($v_1 = v_2$). Section 3.2 provides definitions of comparison groups. Ordinary least squares regressions. Heteroskedasticity robust White standard errors reported in parentheses. $*p < 0.1$, $**p < 0.05$, $***p < 0.01$.

Table A.12: Structural Tailoring Intensity

Dependent variable:	$\lvert\frac{v_{1,i}}{v_{2,i}} - 1\rvert$			
	(1)	(2)	(3)	(4)
Structural Tailoring Intensity	0.110*	0.089	0.124*	0.025
	(0.063)	(0.076)	(0.065)	(0.054)
Immediate Choice		0.068***		0.064**
		(0.022)		(0.025)
Structural Tailoring Intensity x Immediate		0.057		0.154
		(0.131)		(0.114)
Constant	-0.009	-0.018	0.044	0.016
	(0.058)	(0.058)	(0.065)	(0.063)
# LHWs	267	267	320	320
Include Boundary Sample	No	No	Yes	Yes
Stratum FEs	Yes	Yes	Yes	Yes
Exclude 99th and 1st Percentiles	Yes	Yes	Yes	Yes
Drive 2 R	Yes	Yes	Yes	Yes

Notes: This table reports the effects of tailoring on the equality of effort provision over time. The measure $\lvert\frac{v_{1,i}}{v_{2,i}} - 1\rvert$ (the percentage difference between tasks allocated to day 1 and day 2 of the drive) reflects the distance of the task allocation (v_1,v_2) from equality ($v_1 = v_2$). Structural Tailoring Intensity measured as the absolute difference between each LHW's discount factor and their assigned price in Drive 2. Ordinary least squares regressions. Heteroskedasticity robust White standard errors reported in parentheses. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

Table A.13: Structural Tailoring and Completion

Dependent variable:	$\lvert\frac{v_{1,i}}{v_{2,i}} - 1\rvert$							
Policy Comparison Group	Random Price		Structural, Broad		Atheoretic, Broad		Atheoretic, Tailored	
Panel A: Completed Drive 1								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Structural Tailored (=1)	-0.042*	-0.004	-0.034	-0.003	-0.012	0.011	-0.107**	-0.051
	(0.024)	(0.025)	(0.029)	(0.031)	(0.024)	(0.023)	(0.049)	(0.037)
Immediate Choice		0.119***		0.129**		0.111**		0.160*
		(0.045)		(0.057)		(0.050)		(0.089)
Structural Tailored x Immediate		-0.096*		-0.092		-0.076		-0.135
		(0.051)		(0.062)		(0.054)		(0.094)
Constant	-0.051	-0.077	0.221*	0.199	0.110	0.091	-0.043	-0.088
	(0.088)	(0.090)	(0.118)	(0.121)	(0.084)	(0.076)	(0.097)	(0.111)
# LHWs	142	142	99	99	101	101	93	93
Panel B: Failed Drive 1								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Structural Tailored (=1)	-0.065**	-0.034	-0.046	-0.038	-0.100**	-0.045	-0.023	0.008
	(0.029)	(0.031)	(0.031)	(0.039)	(0.046)	(0.047)	(0.035)	(0.032)
Immediate Choice		0.096*		0.048		0.135		0.154*
		(0.056)		(0.061)		(0.088)		(0.081)
Structural Tailored x Immediate		-0.066		-0.019		-0.104		-0.124
		(0.063)		(0.066)		(0.091)		(0.086)
Constant	0.032	0.018	0.156*	0.151*	0.216**	0.193*	0.143*	0.107
	(0.079)	(0.077)	(0.086)	(0.087)	(0.102)	(0.097)	(0.076)	(0.066)
# LHWs	125	125	95	95	93	93	91	91
Stratum FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Exclude 99th and 1st Percentiles	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Drive 2 R_i^* or $\tilde{R}_i$	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports the effects of tailoring on the equality of effort provision over time. The measure $\lvert\frac{v_{1,i}}{v_{2,i}} - 1\rvert$ (the percentage difference between tasks allocated to day 1 and day 2 of the drive) reflects the distance of the task allocation (v_1,v_2) from equality ($v_1 = v_2$). Ordinary least squares regressions. Section 3.2 provides definitions of comparison groups. Heteroskedasticity robust White standard errors reported in parentheses. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.