Reversing Reserves

Parag A. Pathak, Alex Rees-Jones, and Tayfun Sönmez

Abstract

Affirmative action policies are often implemented through reserve systems. We contend that the functioning of these systems is counterintuitive, and that the consequent misunderstanding leads individuals to support policies that ineffectively pursue their goals. We present 1,013 participants in the Understanding America Study with incentivized choices between reserve policies that vary in all decision-relevant parameters. Many subjects' choices are rationalized by a *nearly* correct decision rule, with errors driven solely by the incorrect belief that reversing the processing order has no effect. The prevalence of this belief helps to explain otherwise surprising decisions made in field applications of reserve systems.

Keywords: affirmative action, reserve systems, experimental economics, behavioral market design.

JEL Codes: C9, D9, D47.

Draft date: March 31, 2020. Pathak: Department of Economics, Massachusetts Institute of Technology, and NBER, email: ppathak@mit.edu, Rees-Jones: Department of Economics, Cornell University and NBER, email: arr34@cornell.edu. Sönmez: Department of Economics, Boston College, email: sonmezt@bc.edu. We are grateful to Ori Heffetz, Ted O'Donoghue, German Reyes, Ran Shorrer, Jason Somerville and seminar audiences at Cornell University and Pennsylvania State University for helpful conversations. We thank the National Science Foundation and the Wharton Behavioral Lab for financial support. The project described in this paper relies on data from surveys administered by the Understanding America Study, which is maintained by the Center for Economic and Social Research (CESR) at the University of Southern California. The content of this paper is solely the responsibility of the authors and does not necessarily represent the official views of USC or UAS.

Market organizers commonly seek to advantage some group in the course of an assignment procedure. Seats at a school may be granted according to a lottery, but with the desire to that do not appreciate this fact could deploy reserve systems in a manner that significantly blunts the degree of affirmative action achieved by a reserve of a fixed size.

To test this hypothesis, we deployed a preregistered online experiment to 1,013 members of a nationwide survey panel that is approximately representative on a broad range of demographic variables. In this experiment, subjects faced simple scenarios mirroring two high-profile applications of reserve systems: allocation of seats at a high school or allocation of work visas. In the scenarios, subjects are members of a group that will have positions reserved. Subjects face financial incentives to maximize the chance that their admission is attained in a simulation. They then choose how they would like the reserve system to be administered, selecting from pairs of policies that differ in the both the *number* of seats reserved and in the *order* that the reserve seats are processed.

Our experiment was designed to reveal the rate at which subjects adopt several competing decision rules. First, our design identifies the rate at which subjects choose between policies optimally, adopting a decision rule that correctly accounts for both the number of reserve seats and their processing order. Second, our design identifies the rate at which subjects adopt a decision rule that reflects more reserve seats being better but which treats processing order as irrelevant. Third, our design identifies the aggregate rate at which subjects adopt all other decision rules.

Our results illustrate that subjects often miss the importance of processing order. Our primary estimates suggest that 3% ($s.e. = 2pp$) of subjects apply the optimal decision rule; we are unable to reject the hypothesis that the optimal decision rule is never applied. In contrast, we estimate that 40% ($s.e. = 2pp$) of subjects adopt a decision rule that responds to reserve size but treats processing order as irrelevant. The widespread adoption of this decision rule helps explain the frequency of experimental decisions that are not payoff maximizing for subjects.

Beyond documenting the prevalent belief that processing order does not matter, we also document an important correlate of this belief: cognitive ability. Perhaps surprisingly, subjects with higher education, subjects with higher performance on cognitive ability tests external to our survey, and subjects with a higher performance on comprehension tests within our survey all show a greater likelihood of adopting our misguided decision rule of interest. This contrasts with a common finding in the behavioral market design literature that misreaction to matching-mechanisms' incentives is more prevalent among those of lower cognitive ability (see, e.g., Basteck and Mantovani, 2018; Rees-Jones, 2018; Rees-Jones and Skowronek, 2018; Shorrer and Sovago, 2018; Rees-Jones, Shorrer and Tergiman, 2020; Hassidim, Romm and Shorrer, 2020). In this instance, however, the finding may be rationalized by noting that adoption of this decision rule reflects a general understanding of incentives in this procedure. Our decision rule of interest is *almost* sophisticated, missing one subtle component of large ultimate importance.sophislv.u advice will clearly be insu cient to ensure that the nal policy adopted e ciently pursues

To fully specify the assignment procedure, the sole remaining requirement is to specify the processing order for reserved and open seats. Conceptually, any permutation is possible: one could process one reserved seat, followed by seven open seats, followed by two reserved seats, and so on. In practice, however, these systems are commonly administered in one of two configurations: processing all reserve seats either prior to all open seats or after all open seats. We will restrict attention to these two extremal policies.

1.2 Comparative Statics of Interest

In a system like that just specified, two key comparative statics govern the degree of advantage conferred to the reserve group.

> **Seat-number comparative static:** Hold fixed the priority order and the processing order. Increasing the number of reserved seats weakly increases the number of admitted reserve students.

The seat-number comparative static captures an obvious and intuitive determinant of assignments: saving more seats for a group helps the group. While some may harbor the intuition that this is the *only* relevant comparative static, a second more subtle comparative static follows from the work in Dur et al. (2018).

> **Processing-order comparative static:** Hold fixed the priority order and the number of reserved seats. Switching from processing the reserved seats first to processing the reserved seats last weakly increases the number of admitted reserve students.

Two forces contribute to the result in the processing-order comparative static. The first is a selection effect. When reserved seats are processed last, reserve applicants are admitted in the first-stage processing of open seats at a rate determined by their distribution of priorities relative to general-category applicants. Except for differences in priorities, competition for the open seats is effectively a level playing field between the two groups. In contrast, when reserved seats are processed first, the highest-priority members of the reserve group are removed from the applicant pool before the processing of the open seats. The competition for open seats is therefore between all members of the general category and the comparatively

families to select the school that best suits their needs was critically important. Such a policy would be particularly valuable to families living near a low-performing school, granting them a means of escaping a bad default assignment. An alternative viewpoint emphasized the importance of neighborhood schooling. Under this viewpoint, drawing the student population from the school's walk-zone bene ts the local community and the students themselves. Such a policy would be particularly valuable to families living near a high-performing school, allowing them to avoid intense competition for seats by restricting the admission of non-local students.

Consideration of these two opposing viewpoints led to the reservation of 50 percent of seats for walk-zone students. The remaining seats were open to all. Public accounts of this policy described it as an \uneasy compromise between neighborhood school advocates and those who want choice" (Daley, 1999). And indeed, the superintendent's memorandum presenting this policy explicitly described his desire to accommodate these two viewpoints, and his belief that the new policy \provides a fair balance" (BPS, 1999).

Ultimately, this reserve system was abandoned in 2013. This abandonment was motivated in part by the discovery that this system only minimally advanced the admission of walk-zone applicants. Because a 50-50 reserve split was incorrectly (but widely) perceived to be an accommodation to both sides, the superintendent advocated for the usage of a new system that would be \honest and transparent" (Johnson, 2013).

The understanding that this system was misleading arose due to the intervention of market designers. In the course of studying this reserve system, Pathak and S•nmez discovered that software code used to determine the nal assignment processed all reserved seats before

U.S. Customs and Immigration Service (USCIS), initially chose to implement the policy as reserves-first. This decision is perhaps surprising: as is documented in Pathak, Rees-Jones and Sönmez (2020), this version of implementation results in the lowest degree of skill bias of all policies that comply with the legislation. This decision contrasts with the stated intents of the legislation itself, which was explicitly to introduce skill bias into this system.

Despite this initial plan, passage of the relevant act occurred at a time when application processing was already well underway. The reserves-first implementation was therefore considered impossible to administer in the first year of the new regime, and as a result the reserve seats were processed last. This version of implementation results in the highest degree of skill bias of all policies that comply with the legislation (matched only by a later policy adopted in FY2020). This policy was applied for one year only (FY2005), before the reserves-first version was adopted for a window of three years (FY2006-2008).

Over this initial window of the new regime, seats began filling earlier and earlier in the application season. This became a critical concern by FY2008, when all open seats were filled by applications that arrived on the first day that petitions would be considered. This motivated the regime adopted in FY2009 under which arrival time was replaced by lottery numbers as a means of determining priority. In contrast to the other settings considered thus far, a separate priority (i.e., lottery number) was generated for the reserve seats and the seats open to all. This adjustment eliminates the selection effect induced by processing order described in Section 1, but not the composition effect. As such, the USCIS's decision to continue processing advanced-degree applications first preserved a comparatively lower degree of skill bias in this system.

This regime persisted until its recent modification by the Trump administration. In the 2017 *Buy American and Hire American Executive Order*, the administration instructed the USCIS to switch to a reserves-last system for the explicit purpose of maximizing the degree of skill bias. Upon its implementation in FY2020, this restored the degree of skill bias in the reserve system to that achieved in its very first year — the theoretically maximal degree possible of all policies that comply with the legislation. Unlike prior reforms, discussion of this policy in the Federal Register included consideration of the effect of processing order on skill bias, as well as discussion of the policy's legality.

Across this period of 15 years, four different regime changes were put into effect, each influencing the level of skill bias. The reform proposed in 2017 was explicitly enacted for the intent of increasing the share of H-1Bs granted to highly educated applicants; estimates suggest that this reform granted approximately 5,000 more of the fixed 85,000 H-1Bs to advanced-degree applicants (an increase of 16% to the rate of advanced-degree awards granted). While this change is indeed substantial, we note that both of the preceding reforms — enacted without explicit intent to affect skill bias and seemingly motivated by logistical considerations — had even larger effects. The change applied between FY2005 and FY2006 is estimated to have resulted in a reduction of 14,000 annual awards granted to advanced-degree applicants. The change applied between FY2008 and FY2009 is estimated to have resulted in an increase of 9,000 annual awards granted to advanced-degree applicants. Unlike the 2020 reform, the effect of these reforms on skill bias was not contested despite being more pronounced.

Given that changes to immigration policy are often fiercely contested in U.S. politics, we view the lack of discussion and debate of these earlier reforms as suggesting that their importance was not widely understood.

For further details, this reserve system and its history are documented in Pathak, Rees-Jones and Sönmez (2020). The overview above draws on this work.

2.3 Summary

Across these field applications we observe motivated groups of stakeholders supporting or enacting versions of reserve policies that appear in contrast with their stated goals. In each case, we believe the history of these policies supports the idea that confusion regarding the functioning of reserve systems impacted the manner in which they were deployed. Furthermore, these two cases are not alone. There is similar potential for confusion in the deployment of reserve systems for school admissions in Chicago (see Dur, Pathak and Sönmez, 2019) and in New York City (NYCDOE, 2019). And as we will further discuss in Section 7, such worries are also present in the constitutionally mandated reserve systems for school choice and government employment used in India.

While we believe that misunderstanding is widespread in these environments, we note that

be most favorable to the individual if and only if its number of reserve seats exceeds the threshold. An individual who correctly analyzes the environment and chooses the policy in his best interest would therefore adopt the choice function

$$C(s^{RF}, s^{RL}) = \begin{cases} 1 & \text{if } s^{RF} > T(s^{RL}) \\ 0 & \text{if } s^{RF} \leq T(s^{RL}) \end{cases}.$$

Adopting this choice function would serve as strong evidence in support of a sophisticated understanding of the decision problem.[2]

Just as observation of the choice function would allow for the identification of sophistication, it is also useful for identification of the type of misunderstanding that we have posited. Consider next the choice function that would be observed among individuals who understand the seat-number comparative static but who are unaware of the processing-order comparative static. Such individuals adopt the choice function

$$C^n(s^{RF}, s^{RL}) = \begin{cases} 1 & \text{if } s^{RF} > s^{RL} \\ 0 & \text{if } s^{RF} \leq s^{RL} \end{cases}.$$

This choice function dictates choosing the policy that offers more seats, regardless of order. The superscript n denotes the fact this choice function reflects a degree of naïveté in his understanding of incentives.

Given these considerations, we formulate our approach to testing based on the aggregate choice function that would arise from a heterogeneous population of individuals making these decisions. Consider an individual's *average choice function*:

$$C_p(s^{RF}, s^{RL}) = $$

In this framework, we allow for the individual to probabilistically apply di erent choice functions at di erent times. The term p_i denotes the individual's probability of using the optimal choice function; p_i^n denotes the probability of using the na•ve choice function of interest; the p_i^k

$s^{RF} = T\ (s^{RL})$, these average rates of use may be isolated through the following relationships:

$$\lim_{\ \to 0} C(T\ (s^{RL}) + \ ; s^{RL}) \quad C(T\ (s^{RL}) \quad ; s^{RL}) = \mathrm{E}[p_i] \tag{1}$$

$$\lim_{\ \to 0} C(s^{RF} + \ ; s^{RL}) \quad C(s^{RF} \quad ; s^{RL}) = \mathrm{E}[p_i^n]: \tag{2}$$

(2):

> There are two groups of people, the Blue students and the Green students. Due to their historical underrepresentation, the school particularly values admitting Blue students.

As is illustrated by this text, \Blue" and \Green" labeling dictated group membership. We chose to avoid the usage of more standard racial, gender-based, or income-based group de nitions to avoid inviting the subject to rely on beliefs about the desirability of a rmative action for these groups. While the two groups are always labeled Blue and Green, we randomly assign which of these groups is chosen to be favored.

This introduction was followed by an initial presentation of possible reserve policies:

> In order to meet its goal of admitting Blue students, the school is considering two policies. In this example, both policies will involve reserving 30 seats for the Blue students. When applying either policy, students will be admitted one at a time.
>
> Admissions will happen in two stages.
>
> In one stage, seats are available to both Blue and Green students. When each seat is assigned, it will be given to the student with the highest lottery number who has not yet been admitted. Color will not be considered.
>
> In the other stage, seats are reserved for Blue students only. When each seat is assigned, it will be given to the Blue student with the highest lottery number who has not yet been admitted.
>
> The policies that the school is considering di er in the order of these stages.
>
> **Policy 1: Save the last 30 seats for the Blue students.**
>
> > Stage 1: The rst 70 seats will be assigned to the 70 students who have the highest lottery numbers, regardless of color.
> >
> > Stage 2: The remaining 30 seats will be assigned to the 30 Blue students who have the highest lottery numbers of all Blue students not yet admitted.
>
> **Policy 2: Save the rst 30 seats for the Blue students.**

> Stage 1: The rst 30 seats will be assigned to the 30 Blue students who have the highest lottery numbers.
>
> Stage 2: The remaining 70 seats will be assigned to the 70 students who have the highest lottery numbers of all students not yet admitted, regardless of color.

The assignment of the RF and RL policies to policy 1 and policy 2 was randomized at the subject level. After the initial randomization, these number assignments remained constant

Simulation Details:

All six of the choices you face will have the same basic set-up.

Consider a setting where 200 students are applying to the school. 100 students are Blue and 100 students are Green. You are one of the Blue students.

As before, only 100 students can be admitted. Admissions decisions are still made based on lottery numbers and on diversity considerations. Lottery numbers will be simulated by assigning each student a random number between 1 and 100. All students' numbers, regardless of color, are randomly drawn from the same uniform distribution, so there are no di erences across groups in lottery numbers. If two students have the same lottery number, ties will be broken randomly.

Compensation Details:

One of the six choices you make will be randomly selected to be the choice that \counts." After you answer all six questions, we will reveal the question that \counts" and simulate the admissions decision in the scenario you chose. If you are admitted based on this simulation, an additional $5 will be added to your bonus.

Since you do not know which of the six choices will be chosen to \count," it is in your best interest to answer all six carefully.

Following these screens, subjects faced six screens presenting choices as described above. Each screen took the following format:

Consider the following two ways in which the school could implement its admissions policy.

Policy 1: Save the last (s^{RL}) seats for the Blue students.

Stage 1: The rst (100-s^{RL}) seats will be assigned to the (100-s^{RL}) students who have the highest lottery numbers, regardless of color.

Stage 2: The remaining (s^{RL}) seats will be assigned to the (s^{RL}) Blue students who have the highest lottery numbers of all Blue students not yet admitted.

Policy 2: Save the rst (s^{RF}) seats for the Blue students.

Stage 1: The rst (s^{RF}) seats will be assigned to the (s^{RF}) Blue students who have the highest lottery numbers.

Stage 2: The remaining (100-s^{RF}) seats will be assigned to the (100-s^{RF}) students who have the highest lottery numbers of all students not yet admitted, regardless of color.

As a Blue student, which policy would you prefer?

As described in the prior section, our empirical strategy relies on observing choices between RF and RL policies for a range of ($s^{RF}; s^{RL}$) tuples. To that end, these values were randomly generated for each choice the subject faced. The six decisions presented six values of s^{RL}, assigned deterministically but in random order: 40, 44, 48, 52, 56, and 60 seats. For each of these scenarios, the required number of seats needed for the RF policy to be optimal was 70, 72, 74, 76, 78, and 80, respectively. For each s^{RL} value, s^{RF} was uniformly sampled from 13 potential values: -5, -3, -1, +1, +3, or +5 seats relative to both the optimal and na•ve thresholds, as well as an additional point approximately between the two thresholds. By sampling values in the R(g)-31-311(R(g)-31-311(R(g)-31-311(R(g)-31-311(R(g)-31-315lottery)Td [(Ba(

commit to our sample size and exclusion restrictions. While we will also present some exploratory analyses that were not preregistered, we do not deviate from this preregistration in our presentation of primary results.

5 Experimental Deployment and Sample

5.1 The Understanding America Study

We deployed our experiment in the Understanding America Study (UAS).[5] The UAS is an online panel of American Households recruited for their demographic diversity. The advantage of this panel is its established infrastructure for reaching a broad group of respondents and its substantial e orts to achieve representative sampling. Additionally, by using this panel we can merge data from many other surveys into our analyses, which enables our analysis of the demographic predictors of the behaviors we study.

The UAS panel is recruited through address-based sampling. Respondents are targeted for recruitment based on a random draw from postal records. Once targeted for recruitment, substantial e orts to integrate the individual into the panel are pursued. After an initial attempt to recruit a targeted respondent to the panel, follow-up continues over an approximately six-month period. This follow-up involves attempts to resolve common barriers to survey participation. For example, targeted respondents who do not have internet access are provided with a tablet and broadband internet access so they may participate. Additionally, all UAS materials are available in Spanish to allow for the recruitment of solely Spanish-speaking targeted respondents.

In principle, such a sampling approach can approximate census-level quality in representative sample construction. In practice, however, recruitment of this variety is challenging, and the ultimate panel-entry rate among targeted respondents typically ranges from 10% to 15%. This does introduce the possibility of selection in the sample. However, the UAS's quarterly collection of a very broad set of demographics permits testing for selection on observables, and the construction of sample weights that correct for it. Selection on unob-

[5] For a detailed description of the UAS, see Alattar, Messel and Rogofsky (2018).

servables remains possible. Despite this concern, we note that the procedures described here minimize this worry relative to other commonly-used experimental platforms. Furthermore, we will reconstruct our primary analyses making use of sampling weights aimed to correct for these issues in Section 6.3.1.

5.2 Deployment

Our survey was deployed to the UAS population in December 2019 and January 2020. With the help of UAS personnel, our study was integrated into their online platform and translated into Spanish for the relevant respondents. To achieve our targeted sample size of 1,000 responses, the UAS drew a random subsample of 1,500 respondents from their full panel. These 1,500 respondents received invitations both through the UAS online platform and by mail to take our study, with periodic reminders provided. The survey was closed shortly after the target sample size was attained, ultimately resulting in 1,013 complete observations and a 67% response rate.

5.3 Demographic Properties of Sample

Table 1 summarizes basic demographics of our respondents. As is seen across panels of this table, our sample is demographically diverse. However, due to the selection that occurs in the process of recruitment to online panels, our sample differs from the general U.S. population in several ways. Compared to the general adult population of the U.S., members of our sample are somewhat more likely to be female, married, and U.S. citizens. Our sample also skews to be somewhat older and somewhat more likely to be white.

While there is some evidence of selection on observables influencing the general UAS population, we find little evidence that such effects influence which UAS participants respond to our survey. In the final column of this table, we present formal tests for differences in the demographic variable across respondents who did and did not participate. Only two of the nine tests conducted reach significance at traditional levels. First, participants are slightly less likely to be employed (59.2% vs 66.1%; $p = 0.01$), consistent with the possibility that those not working have more time to complete online studies. Second, participants who completed our study have a notably di erent age distribution. On average, those who completed our survey are 3.79 years older than those who did not ($s.e. = 0.90$; $p = 0.00$).

We additionally examine the geographic distribution of respondents. Figure 1 presents the number of observations obtained for respondents residing in each U.S. state. As is observed in the gure, our survey reached a broad populace: the only U.S. state with no representation in our sample is Delaware. Furthermore, we see no evidence of selection by geography: a chi-squared test for di erences in state of residency by completion status yields a p-value of 0.24. A similar lack of selection is observed based on place of birth (by country: $p = 0.42$; by state: $p = 0.28$).

6 Experimental Results

6.1 Primary Test of Misguided Policy Choices

In this subsection, we present the preregistered tests of our primary hypothesis: that a substantial fraction of respondents mistakenly believe that processing order does not matter in these assignment mechanisms.

To test this hypothesis, we estimate models of the form

$$Y_{ij} = \quad + \quad {}^{n} N_{ij} + \quad O_{ij} + f(s_{ij}^{RF}; s_{ij}^{RL}) + \quad {}_{ij}. \tag{3}$$

Subscripts i and j index the respondent and choice number, respectively. In this model, the dependent variable Y is an indicator for whether the RF policy was chosen in a given binary choice. Variables N_{ij} and O_{ij} provide the value of Y dictated by the na•ve or optimal choice function. Formally, $N_{ij} = I(s_{ij}^{RF} > s_{ij}^{RL})$ and $O_{ij} = I(s_{ij}^{RF} > T(s_{ij}^{RL}))$, where $I()$ denotes the indicator function taking the value of 1 when the statement in parentheses is true. $(s_{ij}^{RF}; s_{ij}^{RL})$ denote the number of seats assigned to each policy, as before, and $f(s_{ij}^{RF}; s_{ij}^{RL})$ denotes a function meant to control for the number of each type of seats assigned. Across speci cations, we will consider a variety of approaches to handling this control, including modeling f as a local polynomial, a cubic spline, or a fth-order polynomial.

Interpreted in lightther4a ^{n}

serves as an estimate of $E[p_i]$. Despite this interpretation, the model above does not constrain the sign of ^{n} or to be positive. In principle, this means that these estimates could yield invalid probabilities. We would interpret the detection of a (statistically significant) negative value for these parameters as a rejection of our framework for type estimation.

Table 2 presents our estimates of this model. In columns 1 and 2, we report estimates of this model with the data restricted to s_{ij}^{RF} values that are within 5 seats of the two thresholds. This amounts to a simple difference in means of the rate of choosing the RF policy when s_{ij}^{RF} is immediately above versus immediately below each threshold. Formally, no term controlling for $f(s_{ij}^{RF}; s_{ij}^{RL})$ is included in the regression; instead, the influence of this term assumed to be nearly constant for a sufficiently narrow region of s_{ij}^{RF} values, and the estimation sample is correspondingly restricted to a narrow region near the threshold.

Interpreting the results from column 1, we see that on average, the RF policy is 40 percentage points ($s.e. = 2pp$

this figure, each dot illustrates the average rate of choosing the RF policy for the the number RF seats illustrated on the x-axis, with the six dots above each point summarizing choices under the six RL seat amounts. The solid line presents a fitted spline analogous to that in column 5 of Table 2. This figure illustrates a stark change in the rate of choosing RF at the naïve threshold of interest. In contrast, there is no apparent discontinuity at the threshold where it should occur among optimizing agents.

In principle, our estimates of the rate of choice-function adoption could differ across the school-choice and visa-allocation versions of our scenarios. In practice, however, the estimated differences are small in magnitude. Appendix Table A1 reproduces Table 2, restricting the data to each of these scenarios in turn. The estimates in these tables typically are within 3 percentage points of the estimates of Table 2,[6] and the difference never exceeds 6 percentage points. Furthermore, in our primary specifications, we find no statistically significant interaction between the estimated discontinuities and the scenario version ($p = 0.18$ and $p = 0.63$ for the column 1 and 2 analysis, respectively). In short, we find no evidence of differences in choice-rule adoption based on the framing of the scenario.

6.1.1 Summary of Primary Findings

We estimate that a large fraction of respondents (40% in our primary regression) adopt a choice function that reflects an understanding of the seat-number comparative static while reflecting ignorance of the processing-order comparative static. These respondents understand that more seats are better, but do not see the benefits of the reserves-last design.

6.2 Predictors of Optimal and Naïve Choices

In this subsection we explore cross-group differences in policy choices. In contrast to the previous section, which presents pregistered analyses, most analysis here is exploratory.

To help assess the predictors of the choice functions of interest, we reconduct the primary analysis of Table 2 while allowing the estimated parameters to vary by group. Interpreted in light of our empirical model, this allows us to infer the rate of use of the two focal choice

[6] More specifically, they are no larger than 3 percentage points for 17 of the 24 estimates.

functions within each group.

Formally, we estimate regressions of the following form.

$$Y_{ij} = \quad + \; {}^{n}N_{ij} + \; G_i + \; G_i \quad N_{ij} + \; {}_{ij} \tag{4}$$

$$Y_{ij} = \quad + \; {}^{n}O_{ij} + \; G_i + \; G_i \quad O_{ij} + \; {}_{ij} \tag{5}$$

In these regressions, the term G_i is an indicator variable indicating membership in the relevant group. In groups where classification is not binary, we will split the group into two approximately equal-sized bins. For example, in one regression the group variable will take the value of 1 for male respondents; in another, it will take the value of 1 for respondents of age 50 or greater. The terms $G_i \quad N_{ij}$ and $G_i \quad O_{ij}$ capture the interaction between this indicator variable and the choice function of interest (which itself is an indicator variable taking the value of 1 when the relevant threshold is surpassed). Except for the terms involving G_i, these regressions are the same as columns 1 and 2 of Table 2. Importantly, we maintain the same sample restriction, estimating the regression only from observations in which the number of RF seats is no more than 5 away from the relevant threshold.

6.2.1 Predictors of Adopting the Naïve Choice Function

We begin by examining estimates of equation (4), capturing differences in the rate of application of the naïve choice function. When interpreting the results of this estimating equation, note that term measures the difference in the discontinuity seen at the naïve threshold, and thus estimates the difference in the rate of adoption of the naïve choice function between those in and out of this group. Furthermore, note that in the immediate vicinity of the naïve threshold, the optimal decision is to choose the RL policy. Since a negative value of indicates a higher propensity to choose the RL policy, this should be interpreted as indicating on average "better" decisions by this group, holding fixed their rate of adoption of the naïve choice function.

Estimates of these equations are presented in Table 3. In panel A, we split the sample by the demographic groups previously considered in Table 1. We omit only the variables related to race or citizenship status: these classifications yield small subgroups in which our analysis is substantially less powered. Examining the estimates of the term , we find some evidence of cross-group differences in the rate of adopting the naïve choice function. Focusing attention on estimates reaching significance at the 5% -level, we find that married respondents are 10 percentage points more likely at adopt this choice function ($s.e. = 4pp$); working respondents are 9 percentage points more likely ($s.e. = 4pp$); respondents with an Associate's degree or above are 20 percentage points more likely ($s.e. = 4pp$); and respondents with annual household income of at least \$50,000 are 21 percentage points more likely ($s.e. = 4pp$). No good are you at working with fractions?").[7] These measures come from independent modules deployed to the UAS sample with broad coverage. Each measure is available for at least 92% of our sample.[8] In addition to these measures, we analyze one measure internal to our study that is plausibly related to cognitive ability: passing the first-stage comprehension check described in Section 4.2.

Panel B of Table 3 reports analysis of these variables. Across these measures, a consistent picture emerges: higher cognitive performance is associated with a higher rate of adoption of the naïve choice function. These results are statistically significant for all cognitive measures except that measuring the breadth of vocabulary — the measure we believe to be the least related to general logical ability. Furthermore, these differences are large in magnitude: higher ability respondents are estimated to be 17 to 31 percentage points more likely to adopt the naïve decision rule across measures, excluding the measure of breadth of vocabulary. Individuals with high cognitive performance appear to face a pitfall when attempting to choose optimal policies. Note, however, that if this pitfall is avoided, those of high cognitive performance choose comparatively well in this region: estimates of reveal that, among those not responding to the threshold, the rate of incorrectly choosing the RF policy is lower.[9]

6.2.2 Predictors of Adopting the Optimal Choice Function

We next examine estimates of equation (5), which measure differences in the rate of application of the optimal choice function. This analysis and its interpretation closely follow that just presented above.

Table 4 shows relatively small differences in the rate of optimal choice function adoption across groups. Interaction effects that are significant at the 5% -level are only detected by marital status (married respondents are 8 percentage points less likely to adopt the optimal

[7] For complete documentation of these measures, see Moldoff and Becker (2019). We apply the aggregate Wave-12 measures discussed under topics N, V, and A: n12nsa_score, a12vea_score, and v12pva_score. Additionally, the subjective numeracy measure discussed below is documented under Topic C: c12avgsnsscore.

[8] Whenever these data are used, we conduct our analyses on all observations for which these measures are available, consistent with an assumption that these measures are missing at random.

[9] For comparability to the panel A results, panel B presents analyses of a discrete above/below median indicator for the cognitive performance measures. Note that the same qualitative results arise from examination of the underlying continuous measures (see Appendix Tables A2 and A3).

choice function; $s.e. = 3pp$) and by education (respondents with an Associate's degree or higher are 7 percentage points more likely to adopt the optimal choice function; $s.e. = 3pp$). Despite this di erence by education, insigni cant and quantitatively small di erences are seen for all cognitive measures examined in panel B | i.e., these results do not suggest that more cognitively able respondents are more likely to adopt the optimal choice function. Overall, while some cross-group di erences are observed in the baseline rate of choosing the RF policy (as measured by parameter), these analyses generally support a much smaller degree of heterogeneity in adoption of the optimal choice rule as compared to the na•ve choice rule. This lower degree of heterogeneity is perhaps expected given the lower overall adoption of the optimal choice rule.

6.2.3 Implications for Payo Maximization

Our results on cross-group di erences in choice-function adoption motivate a practical question: how do these di erences in inferred perceptions [(F47(optimal)-318bp)-27(hao)27(imo)1(l)-318maps
[(0(o)1optimal)-400(c)27(hoic?:)-74(Sience)-400(tho)-369(optimal)-400(c)28(hoice)-400(functioo)-369(is)-4
wer to This queptionipndis the pefi of the na•ve choice functioo

Table5e we(exlmore)-347(This)-325(queption)-325(with)-347mpactiularnn tho di erences inco(es)]TJ -7

hgree chulenise weredene resulthsholbtesinfg either remains stable or declines in magnitude.

Overall, these results illustrate a consequence of con icting ndings from the prior sections. On the one hand, cognitive performance predicts adoption of the na•ve choice function | a behavior that pushes respondents to make suboptimal choices in some circumstances. On the other hand, conditional on not responding to the threshold associated with the na•ve choice function, cognitive performance predicts better choices in the vicinity of the na•ve threshold. The results presented here show the the bene ts of wisdom inherent in this latter nding are mostly o set by the costs of the na•vete in the former. Adopting a choice function that is *nearly* optimal | failing to attend only to the processing-order comparative static | o sets the comparatively high rate of payo -maximizing choices that would be realized in the absence of this pitfall.

Finally, column 7 of this table presents results using only our demographic variables to predict choices. Again, cross-group heterogeneity is shown to be quite modest.

6.2.4 Summary

Taken together, these ndings demonstrate that misunderstanding of the importance of processing order in reserve systems is a prevalent, cross-group phenomenon. Across a wide range of demographic variables available, some variation in decision rules exists; however, adoption of the na•ve choice function remains common among all groups studied. Indeed, the subjects who traditionally would be expected to be the most likely to avoid this pitfall | the highly educated, the comparatively rich, the cognitively able, and those who pass our internal comprehension checks | are those that are most susceptible to it in our data.

6.3 Robustness Considerations

6.3.1 Sample Weights

As emphasized in Section 5, the UAS follows a variety of good practices to target representative sampling, but some selection into the survey panel remains. To help assess the importance of this issue to our primary estimates, we reproduce all main analyses with the

the UAS, account for both the adaptive sampling procedure used in recruitment as well as any di erences in attrition seen across measured demographics (for complete details, see

adopt a *nearly* sophisticated choice function, demonstrating general understanding of the decision environment but a lack of understanding of the importance of processing order. In contrast to many other environments, we do not find that this misguided behavior is tempered by education or cognitive ability, but rather that it is primarily driven by the educated and cognitively able.

Given the rapid proliferation of reserve systems — used to enact affirmative action policies in a wide variety of settings — the tendency for misunderstanding that we document is unfortunate. We believe this misunderstanding serves as a primary explanation for several surprising elements of the history of the school-choice reserve system in Boston and the H-1B visa allocation system in the United States. Furthermore, we believe the potential for this misunderstanding to influence policy (or policy's reception by the public) extends well beyond these two examples. Indeed, when reserve systems are deployed, this type of misunderstanding may be the rule and not the exception.

When facing such a situation, a well-meaning market organizer may benefit from taking steps to help his constituency clearly assess the consequences of potential implementations of a reserve policy. One potential solution that we view as promising is to have stakeholders vote on policies with transparent forecasts of their degree of affirmative action provided. Even the mere requirement to provide such a forecast imposes discipline on the process: a forecast cannot be made without specifying processing order, eliminating the possibility that this component will be left undefined. Furthermore, when the degree of affirmative action is made transparent, we speculate that failures to pursue one's own best interests would be reduced. As a concrete illustration, we believe that proponents of neighborhood schooling in Boston would have been substantially less supportive of the original 50-50 reserve system if it had been presented alongside forecasts showing its lack of advancement of walk-zone students.

Of course, information interventions like these are only possible when market organizers actively and intelligently attempt to improve their constituents' understanding. If market organizers themselves do not understand reserve systems, these steps will not be taken. As was illustrated in the case of H-1B policy, it's not obvious that administrators are always aware of these issues. However, as this literature continues to evolve and as market designers continue their interactions with market organizers, we believe that the probability that market organizers are informed will be higher. We hope that papers like this one will help.

Even in cases where market organizers are informed, however, the assumption that they will be motivated to debias the populace is a strong one. When policy makers benefit from misunderstanding, we believe there is relatively little to stop them from using it to their advantage. This may mark one of the most potentially costly implications of the behavior we have documented.

While our interpretation is only speculative, we believe that a version of this story played out in recent reassessments of reserve policy in India.[11] In India, a reserve for members of historically disadvantaged castes is applied in some school-assignment and government-job allocation procedures.[12] The implementation of these reserves was considered in the landmark Supreme Court case *Indra Sawhney and others v. Union of India (1992)*. In this case, the court interpreted constitutional support for the "the reservation of appointments or posts in favor of any backward class of citizens"[13] to specify that a reserves-last policy should apply, providing these groups with the most effective policy for achieving affirmative action. It also specified that other reserves promoting equality of opportunity[14] should be implemented as reserves-first, granting them a lower degree of affirmative action for the same number of seats. We view this court case as a rare demonstration of clear understanding of the use of reserve order as a policy lever.

In the lead-up to the 2019 election, this reserve system became the topic of public debate and criticism. Many economically disadvantaged Indians do not come from a historically disadvantaged caste. Based on their economic disadvantage, it seemed unreasonable to many that their admission was deprioritized relative to more affluent members of historically disadvantaged castes. In response to these concerns, incumbent President Modi widely publicized his pursuit of a 10% reserve for the "economically weaker sections" (EWS). Partially motivated by a desire to pass this policy before the spring election, the One Hundred and Third Amendment of the Constitution of India went from its first presentation in the lower house

[11] For extensive market-design analysis of these systems see Sönmez and Yenmez (2019*a*,*b*).

[12] Formally, the primary groups considered are the "scheduled castes," "scheduled tribes," and "other backwards castes." Each label is precisely defined in law.

[13] See Article 16(4) in the Constitution of India (1949).

[14] As specified in Article 16(1) in the Constitution of India (1949).

References

Alattar, Laith, Matt Messel, and David Rogofsky. 2018. \An Introduction to the Understanding America Study Internet Panel." *Social Security Bulletin*, 78(2).

Angrisani, Marco, Arie Kapteyn, Erik Meijer, and Htay-Wah Saw. 2019. \Sampling and Weighting the Understanding America Study." Working Paper.

Artemov, Georgy, Yeon-Koo Che, and Yinghua He. 2020. \Strategic `Mistakes': Implications for Market Design Research." Working Paper.

Ashraf, Ajaz. 2019. \\The Supreme Court has Erred": Former Madras HC Judge K Chandru on EWS and Maratha Reservations." *The Caravan*. Available at: https://caravanmagazine.in/law/k-chandru-madras-hc-judge-ews-maratha-reservations-supreme-court, last accessed: March 24, 2020.

Basteck, Christian, and Marco Mantovani. 2018. \Cognitive Ability and Games of School Choice."

modi-governments-quota-move-may-not-yield-the-results-it-wants/174197/, last accessed: March 24, 2020.

Ding, Tingting, and Andrew Schotter. 2017. \Matching and Chatting: An Experimental Study of the Impact of Network Communication on School-Matching Mechanisms." *Games and Economic Behavior*

Johnson, Carol R. 2013. \Speech to Boston School Committee, March 13." Available at: http://bostonschoolchoice. les.wordpress.com/2013/03/3-13-13-superintendent-sc-memo-on-assignment.pdf, last accessed: April 14, 2013.

Khemka, Shreenath A. 2019. \The State's Social Policy on Economic Disadvantage is Simply Inconsistent." *The Indian Express*. Available at: https://indianexpress.com/article/%20opinion/columns/124th-constitutional-amendment-bill-reservation-5534333/, last accessed: March 24, 2020.

Mathew, Pramod. 2019. \Ahead of Polls, India Clears Job and Educational Quotas for the Economically Weak." *Quartz India*. Available at: https://qz.com/india/1519912/modi-government-clears-10-quota-for-general-category-in-india/, last accessed: March 24, 2020.

Mishra, Satish. 2019. \Who Will Gain from 10% Reservation for EWS?" *Observer Research Foundation*. Available at: https://www.orfonline.org/expert-speak/who-will-gain-from-10-reservation-for-ews-47312/, last accessed: March 24, 2020.

Moldo , Michael, and Andrew Becker. 2019. \UAS Comprehensive File Data Description." USC Dornsife Center for Economic and Social Research Working Paper.

NYCDOE. 2019. \Diversity in our Schools." https://www.schools.nyc.gov/about-us/vision-and-mission/diversity-in-our-schools.

Pathak, Parag A., Alex Rees-Jones, and Tayfun Sönmez. 2020. \Immigration Lottery Design: Engineered and Coincidental Consequences of H-1B Reforms." NBER Working Paper 26767.

Pathak, Parag A., and Tayfun Sönmez. 2013. \Recommendation on Algorithm Processing, Public Testimony to EAC." Available at: https://economics.mit.edu/ les/16966.

Rees-Jones, Alex. 2018. “Suboptimal Behavior in Strategy-Proof Mechanisms: Evidence from the Residency Match.” *Games and Economic Behavior*, 108: 317–330.

Rees-Jones, Alex, and Samuel Skowronek. 2018. “An Experimental Investigation of Preference Misrepresentation in the Residency Match.” *Proceedings of the National Academy of Sciences*, 115(45): 11471–11476.

Rees-Jones, Alex, Ran I. Shorrer, and Chloe Tergiman. 2020. “Correlation Neglect in Student-to-School Matching.” NBER Working Paper 26734.

Shorrer, Ran I., and Sandor Sovago. 2018. “Obvious Mistakes in a Strategically Simple College Admissions Environment: Causes and Consequences.” SSRN Working Paper 2993538.

Sönmez, Tayfun, and Bumin Yenmez. 2019*a*. “Affirmative Action in India via Vertical and Horizontal Reservations.” Working paper.

Sönmez, Tayfun, and Bumin Yenmez. 2019*b*. “Constitutional Implementation of Vertical and Horizontal Reservations in India: A Unified Mechanism for Civil Service Allocation and College Admissions.” Working paper.

Table 1: Demographic Information and Sample Selection

	(1)	(2)	(3)	(4)
	Survey Completion Status			Test for Di erence
	Complete	Incomplete	All Recruits	
Basic Demographics				
Female	56.2	57.2	56.5	$p = 0.71$
Married	58.4	59.1	58.7	$p = 0.80$
Working	59.2	66.1	61.4	$p = 0.01$
U.S. Citizen	97.9	97.9	97.9	$p = 0.98$
Hispanic or Latino	10.8	13.6	11.7	$p = 0.12$
Race				
White Only	82.2	77.1	80.5	
Black Only	9.0	10.1	9.4	
Am. Indian or Alaska Native Only	1.3	2.3	1.6	$p = 0.11$
Asian Only	2.8	2.7	2.7	
Hawaiian/Paci c Islander Only	0.6	0.5	0.5	

Table 2: Estimates of Choice Functions Governing Policy Preferences

Table 3: Cross-Group Di erences in Na•ve-Choice-Function Adoption

Panel A: Demographic Groups

Group Indicates:	(1) Male	(2) Married	(3) Working	(4) High Education	(5) High Income	(6) High Age
: Constant	0.30	0.32	0.33	0.37	0.37	0.28
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
^{n}: N_{ij}	0.39	0.35	0.35	0.29	0.28	0.40
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
: Group	-0.04	-0.06	-0.08	-0.15	-0.14	0.02
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
: Interaction	0.03	0.10	0.09	0.20	0.21	0.00
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
Respondents	990	990	990	990	988	989
N	2865	2865	2865	2865	2859	2863
R^2	0.164	0.165	0.166	0.176	0.174	0.163

Panel B: Cognitive Performance Measures

Cog. Measure:	(1) Number Sequence	(2) Analogies	(3) Picture Vocab.	(4) Subjective Numeracy	(5) Comp. Check
: Constant	0.37	0.36	0.31	0.34	0.39
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
^{n}: N_{ij}	0.30	0.32	0.37	0.33	0.24
	(0.03)	(0.03)	(0.02)	(0.03)	(0.03)
: High Cog. Perf.	-0.17	-0.16	-0.06	-0.11	-0.20
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
: Interaction	0.23	0.20	0.07	0.17	0.31
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
Respondents	968	943	956	914	990
N	2811	2724	2772	2640	2865
R^2	0.178	0.176	0.161	0.170	0.190

Notes: This table reports regressions analogous to that in column 1 of Table 2, but additionally including a control for group a liation and an interaction with the estimated discontinuity. High education indicates that the respondent completed an Associate's degree or higher. High income indicates that the respondent's household income is $50,000 per year or more. High age indicates that the respondent is 50 years old or higher. In panel B, we present similar analyses based on splitting the sample by tests of cognitive performance. Standard errors, clustered by respondent, are reported in parentheses.

Table 4: Cross-Group Di erences in Optimal-Choice-Function Adoption

Panel A: Demographic Groups

Group Indicates:	(1) Male	(2) Married	(3) Working	(4) High Education	(5) High Income	(6) High Age
: Constant	0.79	0.75	0.79	0.76	0.74	0.77
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
: O_{ij}	0.03	0.08	0.04	-0.01	0.01	0.04
	(0.02)	(0.03)	(0.03)	(0.02)	(0.03)	(0.02)
: Group	0.00	0.07	0.01	0.06	0.09	0.04
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
: Interaction	0.00	-0.08	-0.01	0.07	0.05	-0.02
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Respondents	991	991	991	991	989	990
N	2709	2709	2709	2709	2703	2705
R^2	0.002	0.006	0.002	0.019	0.021	0.003

Panel B: Cognitive Performance Measures

Cog. Measure:	(1) Number Sequence	(2) Analogies	(3) Picture Vocab.	(4) Subjective Numeracy		(5) Comp. Check
: Constant	0.75	0.75	0.78	0.75		0.70
	(0.02)	(0.02)	(0.02)	(0.02)		(0.02)
: O_{ij}	0.02	0.03	0.03	0.03		0.04
	(0.02)	(0.02)	(0.02)	(0.02)		(0.02)
: High Cog. Perf.	0.10	0.10	0.04	0.10		0.17
	(0.03)	(0.03)	(0.03)	(0.03)		(0.03)
: Interaction	0.02	0.00	0.03	0.00		-0.00
	(0.03)	(0.03)	(0.03)	(0.03)		(0.03)
Respondents	969	943	957	916		991
N	2642	2579	2614	2510		2709
R^2	0.023	0.017	0.007	0.018		0.051

Table 5: Cross-Group Di erences in Rate of Payo -Maximizing Choice

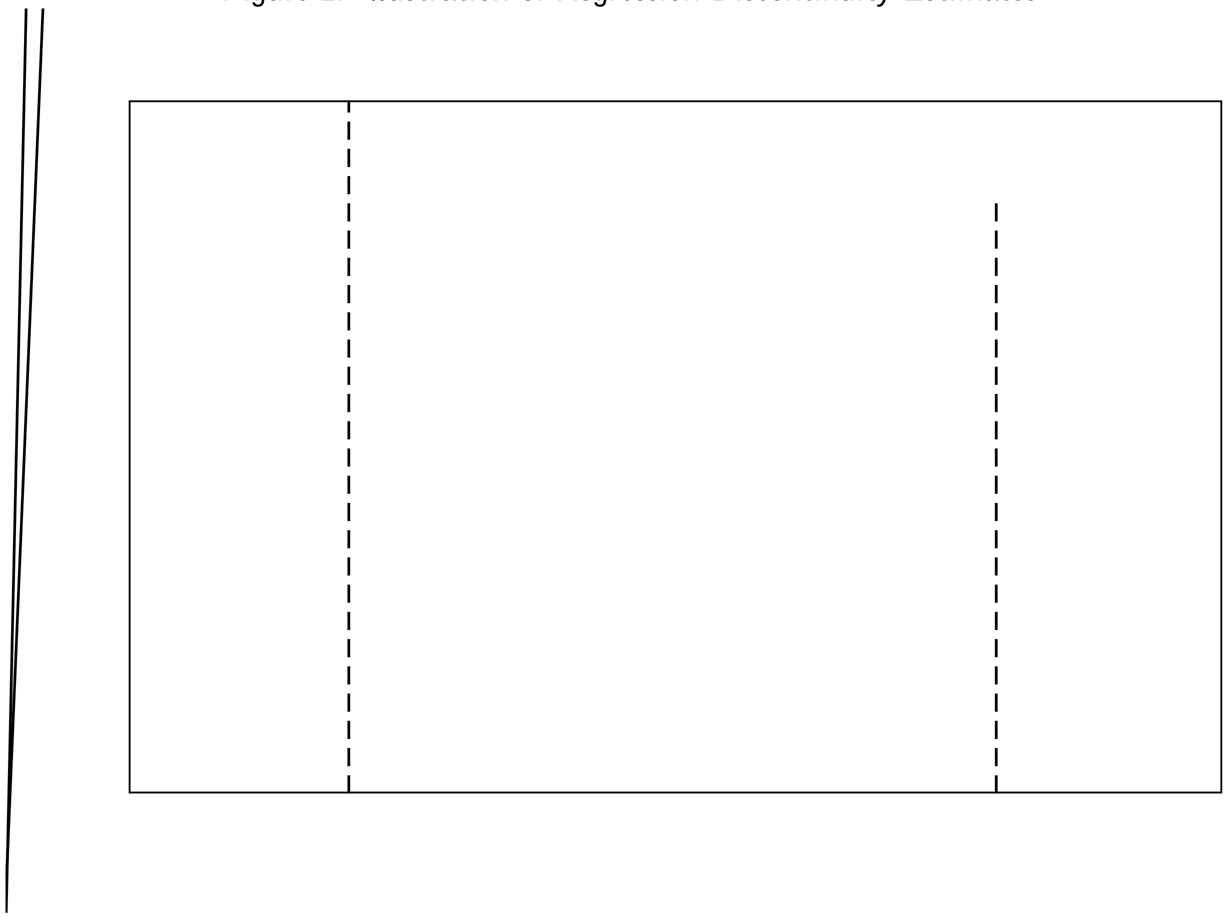
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
High Performance:	0.03			0.03	0.03	0.03	
Number Sequences	(0.02)			(0.02)	(0.02)	(0.02)	
High Performance:		0.03		0.02	0.03	0.02	
Analogies		(0.02)		(0.02)	(0.02)	(0.02)	
High Performance:			0.01	0.00	-0.00	-0.00	
Picture Vocab.			(0.02)	(0.02)	(0.02)	(0.02)	
High Performance:					-0.01	-0.02	
Subjective Numeracy					(0.02)	(0.02)	
Passed Comp. Check					-0.01	-0.01	
					(0.02)	(0.02)	
Male						0.02	0.02
						(0.02)	(0.02)
Married						-0.03	-0.03
						(0.02)	(0.02)
Working						0.01	0.00
						(0.02)	(0.02)
High Education						0.03	0.04
						(0.02)	(0.02)
High Income						0.02	0.01
						(0.02)	(0.02)
High Age						-0.01	-0.02
						(0.02)	(0.02)
Respondents	991	964	979	964	921	917	1009
N	5946	5784	5874	5784	5526	5502	6054

Notes: This table reports average marginal e ects of logit regressions predicting the choice of the payo maximizing policy with cognitive performance and demographic measures. The \high performance" measures are indicator variables indicating above-median performance on the cognitive measure of interest. High education indicates that the respondent completed an Associate's degree or higher. High income indicates that the respondent's household income is $50,000 per year or more. High age indicates that the respondent is 50 years old or higher. All other variables are indicators of their respective title. Standard errors are reported in parentheses, and are calculated by applying the delta-method to the clustered (by respondent) standard errors of the logit coe cient estimates.

Figure 1: Geographic Distribution of Survey Respondents

Notes: This gure presents the number of respondents in our sample who reside in each state. Four respondents are omitted: 1 from Alaska, 1 from Hawaii, and 2 with unknown states of residence. The ranges of values indicated in the legend are split to form quartiles.

Figure 2: Illustration of Regression Discontinuity Estimates



Notes: This gure illustrates the discontinuities in choice probabilities that occur at the thresholds of interest. In our experiment, subjects faced six scenarios containing choices between reserves- rst and reserves-last policies. The scenarios always contained the same six reserves-last policies. In each scenario, the number of seats in the reserves- rst policy was randomly drawn from 13 values spanning the the x-axis, de ned by their position relative to two thresholds. Vertical dashed lines demarcate these thresholds: the point where the number of reserves- rst seats comes to exceed the number of reserves-last seats (the na•ve threshold), and the point where the number of reserves- rst seats comes to exceed the amount needed to make choosing the reserves- rst policy optimal (the optimal threshold). The six dots above each point on the x-axis illustrate the average rate of choosing the reserves- rst policy across the six scenarios. As seen in this gure, subjects' average propensity to choose the reserves- rst policy increases substantially when the na•ve threshold is exceeded, but does not change substantially when the optimal threshold is exceeded. The plotted line is a tted cubic spline over these points, with its associated 95% con dence interval.