Econ 219B
Psychology and Economics: Applications
(Lecture 14 and last)

Stefano DellaVigna

April 29, 2020
Section 1

Behavioral Health Economics
Health economics is a thriving area of application of behavioral economics

Present slides introducing some themes from overview chapter by Chandra, Handel, and Schwartzstein (Handbook 2019)
Behavioral Health Economics

There is now substantial work incorporating insights from behavioral economics across a wide range of key topics in health economics

- Consumer choice of insurance
- Consumer use of insurance
- Consumer health care treatment choices w/o provider
- Consumer health care treatment choices w/ provider
- Consumer adherence to medication / treatments
- Diet
- Exercise
- Addiction
- End-of-life Care
- Provider treatment choices / interaction with patients
- Provider responses to financial incentives, quality targets, info on peers
- Health care organization design/use of technology (patients + providers)
Key Themes

- **Key theme**: Focus on consumer behavior
- **Key theme**: Focus on mistakes that consumers make across different health care contexts, and reasons / mechanisms for those mistakes.
  - This rules out, e.g., a comprehensive discussion of non-standard prefs
- **Key theme**: How do consumers’ mistakes interact with market failures due to asymmetric information, e.g., adverse selection & moral hazard
- **Key theme**: How do consumers’ mistakes influence how we measure adverse selection and moral hazard?
- **Key theme**: How the health economics implications of mistakes driven by systematic errors often differ dramatically from those of mistakes driven by imperfect information
- **Key theme**: What do supply-side responses imply about consumers’ mistakes?
- **Key theme**: What does consumer behavior imply about market regulation / care structures currently in place or being actively considered?
Outline

1. Insurance Choice (Point of entry into health care)
   -- Inertial mistakes, “active” decision mistakes, attention to key choice dimensions, insurance literacy

2. Treatment Selection
   -- Consumer responses to non-linear insurance contracts (myopia, insurance literacy, information about risk)
   -- Decision to go to doctor (procrastination/present bias, behavioral hazard, false beliefs)
   -- Decisions across options with providers (framing of info, false beliefs)

3. Adherence to Treatments / Drugs
   -- Behavioral hazard, myopia, symptom salience, false beliefs

4. Pre-system health behaviors (e.g. diet and exercise)
   -- Present bias, projection bias, interactions of errors with habit formation
# Insurance Choice

**TABLE 1. Literature on Health Insurance Choices**

## Studies of Consumer Choice in Health Insurance Markets

<table>
<thead>
<tr>
<th>Study</th>
<th>Market</th>
<th>Key Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Handel (2013)</td>
<td>Large-employer</td>
<td>Investigates inertia in health plan choice, and shows that in the large employer setting studied, consumers leave approximately $2,000 on the table due to inertia, on average. Many consumers remain in dominated plans, where they lose a substantial amount of money for sure in inertial choice environments. Handel finds that, if consumer inertia is reduced, adverse selection would likely increase in a marketplace with no insurer risk-adjustment transfers.</td>
</tr>
<tr>
<td>Bhargava, Loewenstein, and Sydnor (2015)</td>
<td>Large-employer</td>
<td>Studies employees who actively choose from 48 plans with a lot of flexibility to build their own plan on financial dimensions. Employees frequently choose dominated options, resulting in an average excess spending of 42 percent of annual premium. Choices do not improve over time. Lab intervention ties results to fundamental lack of understanding of insurance products.</td>
</tr>
<tr>
<td>Handel and Kolstad (2015)</td>
<td>Large-employer</td>
<td>Studies role of limited information in plan choice by investigating consumer choice between plans at large-employer using claims, choice, and survey data related to consumer information about plan options. Consumers lack information about plan provider networks, financial characteristics, and hassle costs that can cause them to leave thousands of dollars on table in choice.</td>
</tr>
<tr>
<td>Stromborn, Buchmueller, and Feldstein (2002)</td>
<td>Large-employer</td>
<td>Documents evidence of inertia in a large-employer setting, related to (i) whether choices are active or passive and (ii) whether consumers have active ongoing medical care (which makes them less likely to switch).</td>
</tr>
<tr>
<td>Abaluck and Gruber (2013)</td>
<td>Medicare Part D</td>
<td>Documents money left on table in Medicare Part D prescription drug plan choices over time. There is limited consumer learning; consumer forgone savings increase over time, in large part because of changes to plan designs over time combined with consumer inertia.</td>
</tr>
<tr>
<td>Study</td>
<td>Market</td>
<td>Key Results</td>
</tr>
<tr>
<td>--------------------------------------------------</td>
<td>-------------------------------</td>
<td>-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Abaluck and Grubar (2011)</td>
<td>Medicare Part D</td>
<td>Documents money left on table in Medicare Part D prescription drug plan choices. Elders place higher weight on premiums than on other financial characteristics, and place very little weight on aspects of plans that reduce financial risk. Consumers would have been 27 percent better off if all chose rationally, and market remained as observed.</td>
</tr>
<tr>
<td>Ho, Hogan, and Scott Morton (2015)</td>
<td>Medicare Part D</td>
<td>Documents money left on table both in active choices and from inertia in Medicare Part D market in New Jersey. Studies supply-side responses to consumer inertia, and shows that reducing inertia could have substantial impact on competition, and markedly reduce premiums, leading to both increased consumer welfare and government savings.</td>
</tr>
<tr>
<td>Kling et al. (2012)</td>
<td>Medicare Part D</td>
<td>Studies elders making Medicare Part D plan choices. Performs information intervention where elders are given targeted information about which plans might be best for them. Increases switching rate for elders, and improves their plan choices.</td>
</tr>
<tr>
<td>Ericson (2014)</td>
<td>Medicare Part D</td>
<td>Documents persistence in consumer choice in Medicare Part D market, and pricing patterns consistent with “invest then harvest” pricing where insurers take advantage of consumer inertia in pricing.</td>
</tr>
<tr>
<td>Heiss, McFadden, and Winter (2010)</td>
<td>Medicare Part D</td>
<td>Provides evidence on choice in Medicare Part D, documenting consumer attitudes and money left on table in initial, active Medicare Part D choices.</td>
</tr>
<tr>
<td>Ketcham, Lucarelli, and Powers (2015)</td>
<td>Medicare Part D</td>
<td>Shows that 50 percent of consumers were not enrolled in their 2006 drug plans by 2010, and that switchers gained better plan value. Having more choices is correlated with increased switching rates, implying choice overload may not be a problem on the margin in Medicare Part D.</td>
</tr>
<tr>
<td>Polyakova (2014)</td>
<td>Medicare Part D</td>
<td>Investigates switching costs and inertia in the Medicare Part D market, showing that switching costs are large and have important implications for the plans consumers are enrolled in.</td>
</tr>
<tr>
<td>Marton, Yelowitz, and Talbert (2015)</td>
<td>Medicaid Managed Care</td>
<td>Studies a policy where Medicaid enrollees in Kentucky were automatically enrolled in one of three managed-care plans and given 90 days to opt out. Some enrollees were defaulted into plans with their primary care physicians, and others were not (likely a poor option for them). 30 percent of all enrollees remain in matches without their primary care provider over a long time horizon, exhibiting evidence of substantial inertia in presence of default options.</td>
</tr>
<tr>
<td>Ericson and Starc (2013)</td>
<td>Massachusetts Exchange</td>
<td>Studies change in Massachusetts where exchange plans were required to standardize many financial dimensions of insurance products. Consumer valuation of certain attributes change, in manner that conforms more closely to rational valuation models.</td>
</tr>
<tr>
<td>Fang, Keane, and Silverman (2008)</td>
<td>Medigap</td>
<td>Studies choice in Medigap, with key result that consumers with limited cognitive ability may make poor choices, leading to adoption by the healthiest individuals (so-called advantageous selection).</td>
</tr>
</tbody>
</table>
Notes on Additional Topics We Plan to Mention, but Only Briefly Cover

• **End-of-Life Care:** Some interesting potential experimental results to cite here, still limited work overall. Could be added into section on consumer treatment choices if we find enough papers.

• **Provider Behavior:** How do doctors respond to complex incentives, adopt IT, and use information? Related to Kolstad (2013), Handel et al. (2016). Amitabh has a recent survey on these topics. Plan now will be to mostly exclude them, but discuss at end while referring to other papers.

• **Addiction:** Briefly mention work on distinction between rational and behavioral addiction
  • Our current plan is to not carefully cover this topic
Section 2

Behavioral Environmental Economics
Environmental economics also active area of application. Examples:

- Are people attentive to the future cost of appliances (e.g., air conditioners)?
- Do people pay attention to the marginal, as opposed to average, price of energy (Koichiro Ito’s work)?
- Do people account for the cost of gasoline in used car’s price?

- Present on one of the most effective interventions for energy consumption, OPower nudge

Environmental Economics
Brief Opower History

- Opower is great setting to study welfare effects of nudges
- After default effects, social comparisons greatest success
- Also, truly psychology-and-economics success story
  1. Starts in psychology from Bob Cialdini in *Psych Science*
     - Builds on Cialdini’s social comparison studies
     - Research assistants read meters in San Ramon
     - Leave doorhanger smiley or frowney faces
     - Large effects, 10% reduction off mean of 15
     - BUT underpowered (N=287)

  2. Opower implements with several utilities and millions of households
Brief Opower History

3. In economics: First systematic analysis, Allcott (JPubE)
   • Documents precise 2 percent reduction in energy consumption from Opower social comparisons
   • Effect larger for higher users

Fig. 8. Treatment effects by decile of baseline usage.
Brief Opower History

4. Allcott and Rogers (AER): short-run and medium-run effect

- Sizeable short-run reminder effect
- Some backsliding
- But more than half of effect persists after one year
Brief Opower History

5. Allcott (QJE): heterogeneity of Opower treatment

- Analyze treatment effects by site
- Relate to when started (among other things)

- Remarkable consistency (around 2 percent)
- Yet also selection of most successful programs first
Opower Nudge and Welfare

- Opower comparison is one of best-understood nudges
  - Reliably effective
  - Channel part reminder, part behavior/equipment change
- Has good benefit-cost evaluation, when evaluated with
  - Energy savings for household
  - Versus administrative cost

- BUT comparison with neighbors could be costly or aversive
  - Example: social pressure/moral cost
- Consumers may rather prefer not to receive it
- Evaluate welfare using opting out? Low (1 percent), but default effect
Design

- Allcott and Kessler (AEJ Applied 2019) – Use WTP:
  1. Randomize previous experience with the Opower “nudge”
  2. Send survey and elicit WTP to keep Opower going
     - Response rate overall of 24.5%
  3. Validate WTP with qualitative questions
  4. Integrate WTP into full-scale welfare analysis
Main Results

- **Finding 1**: WTP is remarkably positive
  - Mean WTP $2.98
  - Sizeable share in positive tail
  - Not so in negative tail
- **Not obvious!** (cfr. DellaVigna, List Malmendier QJE → Avoidance of charity)

- **Two main concerns:**
  - Noise in measure of WTP → unlikely to affects skew
  - WTP be among non-responders? → Some evidence it could bias (respondents to first survey more positive), but not obvious
Main Results

- Finding 2: WTP is about half of value of private energy saving ($2.98 versus $5.52)
- Energy savings matter, but about half of benefit dissipated in moral/thinking costs

Potential concern:
- This part relies more heavily here on magnitudes
- Estimate of energy saving actually from Allcott and Rogers (AER)
- Also, need to take WTP number literally

Nonetheless important, allows for full welfare evaluation
Section 3

Behavioral Public Economics
Behavioral Public Economics

Government Intervention

- 219a presents more in depth discussions of behavioral implications for public economics
- Let me highlight some threads

1. It is not enough to create incentives for public program, say, EITC, incentives need to be clear and salient
   - Example: Duflo et al. H&R Block paper on IRA match

2. Self-control problems and inertia play a major role
   - Example: retirement savings

3. (Non-)Take up of benefit programs has plenty of room for behavioral factors
   - Inertia, limited attention, shame – e.g., Janet Currie’s survey

- Focus now on impact of nudge interventions in public economics
Benartzi and Thaler (2004)

- Intervention that affects default
- First behavioral paper in JPE since 1991!

Setting:
- Midsize manufacturing company
- 1998 onward
- Company constrained by anti-discrimination rules → Interested in increasing savings

Features of SMT 401(k) plan:
- No current increase in contribution rate
- Increase in contribution rate by 3% per future pay increase
- Can quit plan at any time
Biases targeted:

1. Self-control
   - Desire to Save more
   - Demand for commitment

2. Partial naiveté’
   - Partial Sophistication $\rightarrow$ Demand of commitment
   - Partial Naïveté $\rightarrow$ Procrastination in quitting plan

3. Loss Aversion with respect to nominal wage cuts
   - Hate nominal wage cuts
   - Accept real wage cuts
Solutions and Implementation

**Solutions:**

1. Increase savings in the future (not in present)
2. Set default so that procrastination leads to **more** (not less) savings
3. Schedule increase only at time of pay raise

**Implementation:**

<table>
<thead>
<tr>
<th>Table 1: Participation Data for the First Implementation of SMarT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of plan participants prior to the adoption of the SMarT plan</td>
</tr>
<tr>
<td>Number of plan participants who elected to receive a recommendation from the consultant</td>
</tr>
<tr>
<td>Number of plan participants who implemented the consultant’s recommended saving rate</td>
</tr>
<tr>
<td>Number of plan participants who were offered the SMarT plan as an alternative</td>
</tr>
<tr>
<td>Number of plan participants who accepted the SMarT plan</td>
</tr>
<tr>
<td>Number of plan participants who opted out of the SMarT plan between the first and second pay raises</td>
</tr>
<tr>
<td>Number of plan participants who opted out of the SMarT plan between the second and third pay raises</td>
</tr>
<tr>
<td>Number of plan participants who opted out of the SMarT plan between the third and fourth pay raises</td>
</tr>
<tr>
<td>Overall participation rate prior to the advice</td>
</tr>
<tr>
<td>Overall participation rate shortly after the advice</td>
</tr>
</tbody>
</table>
Results

- Result 1: High demand for commitment device
- Result 2: Phenomenal effects on savings rates

<table>
<thead>
<tr>
<th>TABLE 2</th>
<th>AVERAGE SAVING RATES (%) FOR THE FIRST IMPLEMENTATION OF SMarT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participants Who Did Not Contact the Financial Consultant</td>
<td>Participants Who Accepted the Consultant’s Recommended Saving Rate</td>
</tr>
<tr>
<td>Participants initially choosing each option*</td>
<td>29</td>
</tr>
<tr>
<td>Pre-advice</td>
<td>6.6</td>
</tr>
<tr>
<td>First pay raise</td>
<td>6.5</td>
</tr>
<tr>
<td>Second pay raise</td>
<td>6.8</td>
</tr>
<tr>
<td>Third pay raise</td>
<td>6.6</td>
</tr>
<tr>
<td>Fourth pay raise</td>
<td>6.2</td>
</tr>
</tbody>
</table>

* There is attrition from each group over time. The number of employees who remain by the time of the fourth pay raise is 229.
Second Implementation

- Simple letter sent, no seminar / additional information + 2% increase per year
- Lower take-up rate (as expected), equally high savings increase

| TABLE 3
<table>
<thead>
<tr>
<th>AVERAGE SAVING RATES FOR ISPAT INLAND (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>EMPLOYEES WHO WERE ALREADY SAVING ON MAY 31, 2001</td>
</tr>
<tr>
<td>Joined SMarT (N=615)</td>
</tr>
<tr>
<td>Pre-SMarT (May 2001)</td>
</tr>
<tr>
<td>First pay raise (October 2001)</td>
</tr>
</tbody>
</table>

Note.—The sample includes 5,817 employees who are eligible to participate in the 401(k) plan and have remained with the company from May 2001 through October 2001. The sample includes 414 employees who were already saving at the maximum rate of 18 percent, although they were not allowed to join the SMarT program. The reported saving rates represent the equally weighted average of the individual saving rates.
Third Implementation

With Randomization:
- Division A: Invitation to attend an informational seminar (40% do)
- Division O: ‘Required’ to attend information seminar (60% do)
- 2 Control Divisions

Two differences in design:
- Increase in Savings take place on April 1 whether pay increase or not (April 1 is usual date for pay increase)
- Choice of increase in contr. rate (1%, 2%, or 3%) (Default is 2%)
- Increases capped at 10%

Results: Sizeable demand for commitment, and large effects on savings + Some spill-over effects
### TABLE 4
**AVERAGE SAVING RATES (%) FOR PHILIPS ELECTRONICS**

<table>
<thead>
<tr>
<th></th>
<th>Employees Who Were Already Saving in December 2001</th>
<th>Employees Who Were Not Saving in December 2001</th>
<th>All Employees</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Joined SMarT</td>
<td>Did Not Join SMarT</td>
<td></td>
</tr>
<tr>
<td><strong>DATE</strong></td>
<td><strong>Joined</strong></td>
<td><strong>Did Not Join</strong></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>7,405</td>
<td>7,053</td>
<td>14,458</td>
</tr>
<tr>
<td>Pre-SMarT (December 2001)</td>
<td>5.65</td>
<td>.00</td>
<td>2.90</td>
</tr>
<tr>
<td>Post-SMarT (March 2002)</td>
<td>5.76</td>
<td>.70</td>
<td>3.20</td>
</tr>
<tr>
<td><strong>A. Control Group</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>180</td>
<td>339</td>
<td>815</td>
</tr>
<tr>
<td>Pre-SMarT (December 2001)</td>
<td>5.26</td>
<td>.00</td>
<td>3.40</td>
</tr>
<tr>
<td>Post-SMarT (March 2002)</td>
<td>6.83</td>
<td>5.72</td>
<td>4.61</td>
</tr>
<tr>
<td><strong>B. Test Group (Divisions A and O Combined)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>66</td>
<td>190</td>
<td>449</td>
</tr>
<tr>
<td>Pre-SMarT (December 2001)</td>
<td>5.47</td>
<td>.00</td>
<td>3.12</td>
</tr>
<tr>
<td>Post-SMarT (March 2002)</td>
<td>7.32</td>
<td>5.97</td>
<td>4.38</td>
</tr>
<tr>
<td><strong>C. Division A</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>114</td>
<td>149</td>
<td>366</td>
</tr>
<tr>
<td>Pre-SMarT (December 2001)</td>
<td>5.14</td>
<td>.00</td>
<td>3.74</td>
</tr>
<tr>
<td>Post-SMarT (March 2002)</td>
<td>6.55</td>
<td>5.41</td>
<td>4.89</td>
</tr>
<tr>
<td><strong>D. Division O</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Note:** The “test” group consists of individuals at Divisions A and O.
Issues

- Saving too much? Ask people if would like to quit plan

<table>
<thead>
<tr>
<th>TABLE 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEDIAN INCOME REPLACEMENT RATIOS (%)</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>INCOME</td>
</tr>
<tr>
<td>$25,000</td>
</tr>
<tr>
<td>$50,000</td>
</tr>
<tr>
<td>$75,000</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>$25,000</td>
</tr>
<tr>
<td>$50,000</td>
</tr>
<tr>
<td>$75,000</td>
</tr>
</tbody>
</table>

Note.—The table displays the median income replacement ratios for different age and income profiles, using investment advice software by Financial Engines. The projections are based on the following assumptions: no defined-benefit pension, statutory social security benefits, employee saving rate of 4 percent before SmarT and 14 percent thereafter, employer match of 50 cents on the dollar up to 6 percent, portfolio mix of 60 percent stocks and 40 percent bonds, and retirement age of 65.

- General equilibrium effect of increase in savings on returns
- Why didn’t a company offer it? How about teaching people?
Leverage biases to help biased agents
Do not hurt unbiased agents (cautious paternalism)

SMarT Plan is great example:
- From Design of an economist...
- ...to Research Implementation with Natural Experiment and Field Experiment
- ...to Policy Implementation into Law passed in Congress: *Automatic Savings and Pension Protection Act*
However...

SMRT may be a unique example for several reasons:

- Defaults are hard to leverage in many situations
  - How to get people to exercise more?
  - Eat less?
  - Pay more attention to hidden information?

- Saving more is desirable for almost all
  - Interventions on other fronts are more open to criticism

- Company was open to SMRT: Firm happy to increase savings of employees
  - Often firm would often rather exploit biases than counter-act them
  - Example 1: Neglect of mutual fund fees
  - Example 2: Overconfidence in trading
Nudge Agenda

More generally, Nudge agenda (Sunstein and Thaler, 2011)
- Use behavioral interventions w/o financial incentives to induce a given behavior

Second examples, not based on defaults

Bhargava and Manoli (AER 2016): EITC take-up
- One of the earliest large-scale nudge interventions
- Rich design with many arms, on important outcome
- Example to follow
MOTIVATION & BACKGROUND

- **EITC is largest means-tested cash transfer program.** It disburses $58 billion per year to 26 million recipients through income supplement that encourages work

- Fully refundable, supplements earned income by average of 17% which amounts to $2,100. Must file your taxes to claim

- **25% of eligible do not take-up (~6.7m).** Of 25%, 16% do not file taxes, and 9% files taxes (~2.3 m) (Plueger 2010). 9% is focus of this study

- (Many) filing non-claimants receive a reminder notice / claiming worksheet (CP 09 or CP 27) from IRS

- **Policy consequences profound.** Foregone benefits amount to average of 31 days of income, up to ~115 days for some (est. $1,096 benefit, $8,900 income). Health, education, consumption benefits linked to EITC (Hoynes 2011; Dahl and Lochner 2011; Smeeding and Phillips and O’Connor 2001)

- Despite considerable research, incomplete take-up in benefit programs regarded as puzzle to economists (Currie 2006)
EITC BENEFIT SCHEDULE FOR TAX YEAR 2009
RESEARCH STRATEGY

Field experiment to test leading causes of low take-up

- Modify tax documents (notice + worksheet + envelope) and distribute to eligible filing non-claimants
- Simultaneously test three hypotheses regarding role of information (benefits, costs, program rules), Informational complexity, and program stigma on response
- Randomize three components independently and distribute in blocks defined by zip code and dependent status

Tax-return data plus micro-data on demographics, EIC claiming history

Survey of perceived incentives. Surveys of ~1200 low to moderate income taxpayers to assess perception of EITC cost/benefit parameters

Psychometric scoring of interventions. Second survey with ~2800 subjects illuminates psychological mechanisms underlying experimental response
EXPERIMENT CONTEXT – ILLUSTRATIVE TIMELINE

2009

Jan to Dec

Earn income, qualify for EITC, (CA only)

2010

Feb

File TY 2009 taxes, neglect to claim EITC

March

IRS reminds you to claim with CP09/27 notice

May

For 41% who return CP, IRS mails check

Nov

Experimental notices mailed to CP non-respondents (CA)
## Table 3

**Experimental Interventions by Mechanism**

<table>
<thead>
<tr>
<th>Mechanism</th>
<th>Intervention</th>
<th>Description</th>
<th>Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Informational Complexity</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Simplicity / Complexity (Design)</td>
<td>1. Simple Notice</td>
<td>Relative to complex (original CP) notice, &quot;simple&quot; single-sided notice has</td>
<td>3,676</td>
</tr>
<tr>
<td></td>
<td></td>
<td>simplified layout and excludes eligibility information repeated in worksheet</td>
<td></td>
</tr>
<tr>
<td>Simplicity / Complexity (Length)</td>
<td>2. Simple Worksheet</td>
<td>Relative to simple worksheet, a complex worksheet includes additional, non-</td>
<td>10,979</td>
</tr>
<tr>
<td></td>
<td></td>
<td>discriminatory, questions regarding eligibility</td>
<td></td>
</tr>
<tr>
<td><strong>Program Information</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Benefit and Cost Information</td>
<td>1. Benefits (Low and High)</td>
<td>Simple notice reports upper bounds of benefit range</td>
<td>6,761</td>
</tr>
<tr>
<td></td>
<td>2. Transaction Costs (Low and High)</td>
<td>Simple notice provides guidance as to worksheet completion time</td>
<td>3,475</td>
</tr>
<tr>
<td>Penalty/Audit Information</td>
<td>1. Indemnity Message</td>
<td>Worksheet with message to indemnify against penalty for unintentional error</td>
<td>17,027</td>
</tr>
<tr>
<td>General Program Information</td>
<td>1. Attention Envelope</td>
<td>Envelope with message indicating enclosed information is &quot;good news&quot;</td>
<td>17,044</td>
</tr>
<tr>
<td></td>
<td>2. Informational Flyer</td>
<td>One page flyer offers program information and trapezoidal benefit schedule</td>
<td>4,019</td>
</tr>
<tr>
<td><strong>Program Stigma</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Personal Stigma</td>
<td>1. Emphasis on Earned Income</td>
<td>Simple notice emphasizes that benefit is reward for hard work</td>
<td>1,844</td>
</tr>
<tr>
<td>Social Stigma</td>
<td>2. Social Influence</td>
<td>Simple notice communicates that similarly situated peers are also claiming</td>
<td>1,753</td>
</tr>
</tbody>
</table>
(A) INFORMATIONAL COMPLEXITY

THEORY

- Poor financial choices due to lack of experience and familiarity with complex documents or low “financial literacy”

- Transfer programs are complicated. EITC has 24 pages of instruction in tax book, 56 pages in separate Publication 596; average length of state FSP application is 12 pages (Bertrand and Mullainathan and Shafir 2006)

- Simplification appears to “improve” choice in many contexts (e.g., Bettinger et al. 2009)

INTERVENTIONS

(1) **Complex Notice**: Tests “design complexity”. Features textually dense design, is two pages, and repeats eligibility information from worksheet. Resembles original CP Notice.

(2) **Complex Worksheet**: Tests “length complexity”. Features additional, “non discriminatory” questions.
COMPLICATED NOTICE (ADAPTED FROM CP)

Summary of the notice and program

Instructions for eligibility worksheet; very exclusionary language

Headline describing purpose of notice

Details of dependent eligibility, next steps, and instructions for further information

COMPLICATED NOTICE (ADAPTED FROM CP)
“BASELINE” NOTICE

- Headline communicates program eligibility.
- Summary explains purpose of letter and program. Tax Year is specified.
- Recipients instructed to complete worksheet to determine eligibility; eligibility criteria not repeated on notice.
- Information on Notice + Worksheet held constant.
SIMPLE WORKSHEET

- Guides reader through determination of eligibility (distinct version for dependent and non-dependents)

- Worksheet checks valid SSN, elicits names of eligible dependents, and instructs recipient to sign and return if eligible

- Original CP worksheet, with alternative formatting and organization, not tested
COMPLEX WORKSHEET

• Same formatting and organization as simple worksheet

• Lengthier than simple worksheet due to additional eligibility criteria questions taken from IRS Pub 596 (in Step 1 for dependents version, and in Step 1 and 2 for non-dependents version)

• Example: “I was not a U.S. citizen (or resident alien) for any part of 2009

• Additional criteria do not have bearing on true eligibility as per administrative records
(B) INFORMATION ON BENEFITS, COSTS, RULES

THEORY

- Individuals optimize with respect to incentives
- Individuals have limited attention, may only respond to perceived or known incentives (Kahneman 1986; Taylor and Fiske 1975)
- Basic information regarding incentives helps optimize behavior (e.g., Liebman and Luttmer 2011)

INTERVENTIONS

1. Benefit Notice: Generic benefit information (high and low)
2. Cost Notice: Information on worksheet claiming time (high and low)
3. Penalty Worksheet: “Indemnification” message on claiming worksheet
4. Informational Flyer: Information on benefits and program on 1 page flyer
5. Messaged Envelope: Persuasion message on envelope
BENEFIT DISPLAY

• Identical to baseline notice in design and content except...

• Headline communicates refund may be up to specific amount determined by number of dependents [IRS did not allow exact benefit amounts]

• Indicated range is $457 for those with no dependents, $5,657 for those with 3 or more dependents, and randomized to be either dependent specific, or overall, maximum for 1 dependent ($3,043), and 2 dependents ($5,028)

• Summary reiterates benefit information
Cost Display

- Identical to baseline notice in design and content except...

- Headline communicates that completing worksheet should take less than 60 (or 10) minutes

Important information about the Earned Income Credit
You may be eligible for a refund. Claiming your refund usually takes less than 60 minutes.

Do not discard or overlook this notice because you may be entitled to some additional money.

To claim your benefit, simply complete and return this form. This usually takes less than 60 minutes.

What you need to do

Complete the Earned Income Credit Worksheet on Page 3. If the worksheet confirms that you are eligible for the credit, sign and date the attached worksheet, and mail it to us in the enclosed envelope.

Next steps

If you are eligible for the credit, we will send you a refund check in 6 to 8 weeks. If you owe back taxes or other debts, such as child support which we are required to collect, we will use your credit to reduce or pay off those debts.

Additional information

If you need additional assistance, please call 1-800-829-1040, or visit online at IRS.gov. For tax forms, call 1-800-TAX-FORM (1-800-829-3676).

You can also find forms and other helpful documents which explain the EIC program in greater detail (e.g., Publication 596) at IRS.gov.
INFORMATIONAL FLYER

• One page sheet containing incentive information through a graphical display, and text clarifying confusing aspects of eligibility and requirements

• Graphics generally complicated to digest for those of low financial literacy

• Flyer accompanies select baseline notices
Messaged Envelopes

- Treatment envelopes communicate that contents contain beneficial and important information

- Mail marketing firms estimate that up to 44% of non-personal mail is not opened

- Our surveys indicate that 16% of low to moderate income filers do not open mail from IRS
THEORY

- Stigma may deter participation in means-tested benefit programs (e.g., Weisbrod 1970; Moffit 1983; Currie 2006)
- Stigma due to either social sanction (social) or threat to identity (personal)
- Encourage behavior through social influence (Cialdini et al. 1990)
- Energy use and peer feedback (Costa and Kahn 2010)

INTERVENTIONS

“You may be eligible for a refund. Usually, 4 of every 5 eligible people claim their refunds.”
Notice Headline for Intervention 1

“You may be eligible for a refund due to all your hard work.”
Notice Headline for Intervention 2
RANDOMIZATION

- Notice, worksheets, envelopes independently randomized
- Randomization by blocks defined by zip code and dependent indicator (3,148 blocks)
- Oversampling – Baseline notices 4x sample; salience, 3x sample; complex worksheet, .5x sample
- Balancing checks suggest randomization successful
- Mailed mid November 2010; data collected through May 2011
WHAT IS THE COUNTERFACTUAL RESPONSE?

CA Notice Response since July 2010
(IRS Processing Date)

Experimental Notices Mailed
(mid-November 2010)

Pre-Period Response to CP Notices
(since approx July 2010)
SUMMARY OF OVERALL RESPONSE

- Mere receipt of second notice yields 0.22 response (0.14 control condition)
- Language may be a barrier to response
- Simplification raises response from .14 to .23; Information from .23 to .28; No beneficial effect of lower stigma
- Effects not driven by denial of claims rate
Predicted Response for Benefit and Cost Notices

- Benefit Display (w/o dependents):
  - Baseline: 27%
  - +9%
- Benefit Display (w/ dependents):
  - Baseline: 16%
  - +5%
  - +6%
- Cost Display:
  - Baseline: 23%
  - 60mn: -1%
  - 10mn: -2%
Potential Problems

Potential problems with Nudges

Problem 1. Are we nudging for good?
- Nudges could be used to pursue sinister objectives
- (Companies have used them for decades to increase sales)
- Even when well intentioned, do we know that it is good to induce a given behavior?
  - Savings: What is the right savings rate?
  - Charitable giving: Does it raise welfare? (earlier lecture)

Problem 2. (Related) What is the model?
- A model helps assess the channels, welfare implications

Problem 3. Size of Effects
- Are most nudges as effective as these?
- Evidence on comprehensive impact of nudge, DellaVigna and Linos (2020)
Introduction

- Nudge RTCs adopted in government communication
  - Ideas42 pioneered application of behavioral science
  - Quick expansion of “nudge units”

- In each unit, RCT nudges taken to scale
- Unique opportunity to study RCTs at scale
Reminder of where we started...

- Collaboration with two Nudge Units:
  - SBST/OES, 2015- present
  - BIT North America 2015- present

- Each unit shared its complete record of RCT results
- Remarkable case of transparency in public sector

- Two key features:
  - *(Transparency)* Each trial has a trial report, and many have pre-analysis plans
  - *(Novelty)* 90% of trials unpublished in academic journals
The current paper

1) Measure treatment effect across OES/BIT

2) Measure treatment effect in academic papers

3) Disentangle the gap
   • MDES and sample size
   • Publication bias
   • Heterogeneity

4) Compare to forecasting

5) Simulate the publication bias gap
### Academic Journals: Sample

<table>
<thead>
<tr>
<th>Academic Journals nudges</th>
</tr>
</thead>
<tbody>
<tr>
<td>102 trials</td>
</tr>
<tr>
<td>253 nudges</td>
</tr>
<tr>
<td>1,653,302 participants</td>
</tr>
</tbody>
</table>

**Sample restrictions**

<table>
<thead>
<tr>
<th>RCTs and non-lab studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>52 trials</td>
</tr>
<tr>
<td>115 nudges</td>
</tr>
<tr>
<td>650,517 participants</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>No financial incentives</th>
</tr>
</thead>
<tbody>
<tr>
<td>49 trials</td>
</tr>
<tr>
<td>109 nudges</td>
</tr>
<tr>
<td>642,930 participants</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Targets binary outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>28 trials</td>
</tr>
<tr>
<td>77 nudges</td>
</tr>
<tr>
<td>505,885 participants</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>No default interventions</th>
</tr>
</thead>
<tbody>
<tr>
<td>26 trials</td>
</tr>
<tr>
<td>74 nudges</td>
</tr>
<tr>
<td>505,337 participants</td>
</tr>
</tbody>
</table>

- We then re-code all interventions from the original papers
- What are the characteristics of Academic Journals?
Academic Journals: Effect Size

- **Unweighted treatment effects**

<table>
<thead>
<tr>
<th></th>
<th>Academic Journals</th>
<th>Nudge Units</th>
<th>Published Nudge Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average treatment effect</td>
<td>8.682</td>
<td>0.499</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.467)</td>
<td>(0.110)</td>
<td></td>
</tr>
<tr>
<td>Nudges</td>
<td>74</td>
<td>74</td>
<td></td>
</tr>
<tr>
<td>Trials</td>
<td>26</td>
<td>26</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>505,337</td>
<td>505,337</td>
<td></td>
</tr>
<tr>
<td>25th pctl trt. effect</td>
<td>1.05</td>
<td>0.12</td>
<td></td>
</tr>
<tr>
<td>Median trt. effect</td>
<td>4.12</td>
<td>0.32</td>
<td></td>
</tr>
<tr>
<td>75th pctl trt. effect</td>
<td>12.00</td>
<td>0.69</td>
<td></td>
</tr>
<tr>
<td>Avg. control take-up</td>
<td>25.97</td>
<td>25.97</td>
<td></td>
</tr>
<tr>
<td>Median MDE</td>
<td>6.30</td>
<td>0.49</td>
<td></td>
</tr>
</tbody>
</table>

This table shows the average treatment effect of nudges. Standard errors clustered by trial are shown in parentheses. p.p. refers to percentage point. Minimum detectable effect (MDE) calculated at power 0.8.

- **Average nudge increases take up by:**
  - 8.7 percentage points
  - 0.50 in log odds ratio
  - Large effects!
Published Nudges: Effect Size

Active decision in 401k enrollment (Carroll et al., 2009)
H&R Block FAFSA experiment (Bettinger et al., 2012)
Changing menu order in buffet line for healthier food consumption (Wansink & Hanks, 2013)

Sample: 69 nudges (24 trials)
3 nudges with treatment effects >40 p.p. excluded
95% confidence intervals and quadratic fit shown
Nudge Units: Sample

<table>
<thead>
<tr>
<th>Universe of OES &amp; BIT nudges</th>
<th>165 trials</th>
<th>349 nudges</th>
<th>37,020,243 participants</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sample restrictions</strong></td>
<td>152 trials</td>
<td>332 nudges</td>
<td>36,907,315 participants</td>
</tr>
<tr>
<td>RCTs and non-lab studies</td>
<td>137 trials</td>
<td>263 nudges</td>
<td>24,946,504 participants</td>
</tr>
<tr>
<td>Designates clear control group</td>
<td>134 trials</td>
<td>256 nudges</td>
<td>24,925,529 participants</td>
</tr>
<tr>
<td>No financial incentives</td>
<td>126 trials</td>
<td>245 nudges</td>
<td>24,884,187 participants</td>
</tr>
<tr>
<td>Targets binary outcomes</td>
<td>126 trials</td>
<td>243 nudges</td>
<td>23,584,187 participants</td>
</tr>
</tbody>
</table>

- Notice: Only 12 of 126 trials are published or working paper
Nudge Units: OES

GROUP A ROTH TSP: SMARTDOCS for January 2, 2015

Subject: Important! Your Action Needed in January to Continue Your Roth TSP Election

As a Roth TSP participant, your window to submit new contribution elections is here. You may submit your new Roth TSP elections based on percentages of basic pay, special pay, incentive pays and bonuses any time through Jan. 31, 2015, to avoid any interruption in your retirement investment plans.

Your elections may be submitted quickly and securely using myPay. You may also use the revised TSP-U-1 form available at www.tsp.gov. Forms must be submitted to your finance office to be applied to your military pay account.

We will send you reminders throughout January to make sure you have the information, worksheets and time to get your Roth TSP elections completed within the allotted time.

Election submissions received after Jan. 31, 2015, will result in a lapse in Roth TSP contributions.

For more information on the change to percentage-of-pay selections and how you can make sure your investment plans continue, visit www.dfas.mil/TSP_AC.html.

My POC for this effort is Matthew Stack at matthew.stack@dfas.mil.

Bruce S. Smith
Director, ESS Military Pay

Control communication for OES trial

“Increasing Servicemember Roth TSP Re-Enrollment”
Nudge Units: OES

Treatment communication for OES trial
“Increasing Servicemember Roth TSP Re-Enrollment”
Control
BITNA trial “Increasing Online License Plate Renewals”

Treatment
# Nudge Units: Effect Size

<table>
<thead>
<tr>
<th></th>
<th>Academic Journals</th>
<th>Nudge Units</th>
<th>Published Nudge Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nudge Units: Effect Size</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1) p.p.</td>
<td>(2) log odds ratio</td>
<td>(3) p.p.</td>
</tr>
<tr>
<td>Average treatment effect</td>
<td>8.682 (2.467)</td>
<td>0.499 (0.110)</td>
<td>1.381 (0.302)</td>
</tr>
<tr>
<td>Nudges</td>
<td>74</td>
<td>74</td>
<td>243</td>
</tr>
<tr>
<td>Trials</td>
<td>26</td>
<td>26</td>
<td>126</td>
</tr>
<tr>
<td>Observations</td>
<td>505,337</td>
<td>505,337</td>
<td>23,584,187</td>
</tr>
<tr>
<td>25th pctile trt. effect</td>
<td>1.05</td>
<td>0.12</td>
<td>0.04</td>
</tr>
<tr>
<td>Median trt. effect</td>
<td>4.12</td>
<td>0.32</td>
<td>0.50</td>
</tr>
<tr>
<td>75th pctile trt. effect</td>
<td>12.00</td>
<td>0.69</td>
<td>1.40</td>
</tr>
<tr>
<td>Avg. control take-up</td>
<td>25.97</td>
<td>25.97</td>
<td>17.44</td>
</tr>
<tr>
<td>Median MDE</td>
<td>6.30</td>
<td>0.49</td>
<td>0.78</td>
</tr>
</tbody>
</table>

This table shows the average treatment effect of nudges. Standard errors clustered by trial are shown in parentheses. p.p. refers to percentage point. Minimum detectable effect (MDE) calculated at power 0.8.

- Statistically and economically significant impact
- Smaller than Academic Journals nudges
Nudge Unit Effect: Effect Size

- Raw data on effect size

Sample: 237 nudges (123 trials)
95% confidence intervals and quadratic fit shown
Why are the effect sizes so different in the two samples?

- Channel 1: Different Characteristics of Nudges (e.g., email vs in person)
- Channel 2: Statistical Power of Studies
- Channel 3: Publication Bias
- Channel 4: Residual going “to scale” factors
Characteristics I

Policy area
- Revenue & debt: Academic journals 17.57%, Nudge units 28.81%
- Benefits & programs: Academic journals 19.81%, Nudge units 22.22%
- Workforce & education: Academic journals 9.46%, Nudge units 18.52%
- Health: Academic journals 13.17%, Nudge units 20.38%
- Registration & regulation compliance: Academic journals 8.64%, Nudge units 12.16%
- Community engagement: Academic journals 4.05%, Nudge units 7.82%
- Environment: Academic journals 0.82%, Nudge units 13.51%
- Consumer behavior: Academic journals 0%, Nudge units 4.05%

Control communication
- No communication: Academic journals 43.24%, Nudge units 60.91%
- Some communication: Academic journals 39.09%, Nudge units 56.76%

Frequency (%)
Statistical Power

- How well powered are the trials? MDE at 80% in p.p.
- Much higher power for Nudge Units vs. Acad. Jour.

Nudge Units sample: 243 nudges, 126 trials
Academic Journals sample: 74 nudges, 26 trials
Test 2 for publication bias: Result

- Plot of $t$ statistics for evidence on $p$-hacking

All treatments (74 nudges, 26 trials)

- No obvious bias – but multiple results per trial
Test 2 for publication bias: Result

- Plot of $t$ statistics for evidence on $p$-hacking
- Most significant treatment

Most significant treatments only (25 nudges, 25 trials)
Nudge Units: Pub. Bias

- Is there publication bias in this sample? No
- Test 2. $t$-stat distribution

All treatments
(243 nudges, 126 trials)

Most sig. treatment
(124 trials)
Decomposing the Effect Size Gap

• Why do Academic Journals nudges have such a higher ATE compared to those from Nudge Units? (8.7 vs. 1.4 p.p.)
• Decompose the difference due to precision, publication bias, and nudge characteristics
• Method 1: Regression decomposition

<table>
<thead>
<tr>
<th>Dep. Var.: Treatment effect (p.p.)</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Academic Journals sample</td>
<td>7.301</td>
<td>2.148</td>
<td>5.720</td>
<td>1.443</td>
<td>2.244</td>
<td>0.827</td>
</tr>
<tr>
<td></td>
<td>(2.449)</td>
<td>(1.449)</td>
<td>(2.263)</td>
<td>(1.193)</td>
<td>(1.567)</td>
<td>(1.330)</td>
</tr>
<tr>
<td>Min. detectable effect</td>
<td>0.843</td>
<td>0.770</td>
<td>0.441</td>
<td>0.276</td>
<td>0.277</td>
<td>0.203</td>
</tr>
<tr>
<td></td>
<td>(0.276)</td>
<td>(0.277)</td>
<td>(0.203)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1/MDE</td>
<td>0.179</td>
<td>0.127</td>
<td>0.001</td>
<td>0.131</td>
<td>0.141</td>
<td>0.132</td>
</tr>
<tr>
<td></td>
<td>(0.131)</td>
<td>(0.141)</td>
<td>(0.132)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>1.381</td>
<td>-0.430</td>
<td>1.381</td>
<td>-0.201</td>
<td>1.301</td>
<td>1.337</td>
</tr>
<tr>
<td></td>
<td>(0.302)</td>
<td>(0.723)</td>
<td>(0.302)</td>
<td>(0.758)</td>
<td>(1.522)</td>
<td>(1.750)</td>
</tr>
<tr>
<td>Nudges</td>
<td>317</td>
<td>317</td>
<td>317</td>
<td>317</td>
<td>317</td>
<td>317</td>
</tr>
<tr>
<td>Trials</td>
<td>152</td>
<td>152</td>
<td>152</td>
<td>152</td>
<td>152</td>
<td>152</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.182</td>
<td>0.352</td>
<td>0.133</td>
<td>0.322</td>
<td>0.449</td>
<td>0.439</td>
</tr>
<tr>
<td>MDE &amp; 1/MDE</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Publication bias weight</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Nudge characteristics controls</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

Standard errors clustered by trial are shown in parentheses. Coefficient on Academic Journals sample is the estimated average difference in percentage point (p.p.) treatment effects between the Academic Journals and Nudge Units samples. MDE (minimum detectable effect) is calculated in p.p. at power 0.8. Weighting for publication bias assigns significant trials a relative weight of .22 compared to insignificant trials in the Academic Journals sample. Nudge characteristics controls include the control take-up in % and its squared value, policy area, control communication category, medium, and mechanism. The early vs. late indicator is not included as a control, as the threshold differs between the two samples. A dummy for the 4 nudges (2 trials) missing control take-up data is included with the nudge characteristics controls.
Decomposing the Effect Size Gap

- Why do Academic Journals nudges have such a higher ATE compared to those from Nudge Units? (8.7 vs. 1.4 p.p.)
- Decompose the difference due to precision, publication bias, and nudge characteristics
- Method 2: Re-weighting

<table>
<thead>
<tr>
<th>Dep. Var.: Treatment effect (p.p.)</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Academic Journals sample</td>
<td>7.301</td>
<td>1.678</td>
<td>5.720</td>
<td>0.187</td>
<td>3.517</td>
<td>-0.092</td>
</tr>
<tr>
<td>(2.449)</td>
<td>(1.313)</td>
<td>(2.263)</td>
<td>(0.994)</td>
<td>(1.819)</td>
<td>(0.762)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>1.381</td>
<td>1.106</td>
<td>1.381</td>
<td>1.106</td>
<td>1.814</td>
<td>1.119</td>
</tr>
<tr>
<td>(0.302)</td>
<td>(0.390)</td>
<td>(0.302)</td>
<td>(0.390)</td>
<td>(0.547)</td>
<td>(0.361)</td>
<td></td>
</tr>
<tr>
<td>Nudges</td>
<td>317</td>
<td>317</td>
<td>317</td>
<td>317</td>
<td>317</td>
<td>317</td>
</tr>
<tr>
<td>Trials</td>
<td>152</td>
<td>152</td>
<td>152</td>
<td>152</td>
<td>152</td>
<td>152</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.182</td>
<td>0.021</td>
<td>0.133</td>
<td>0.001</td>
<td>0.066</td>
<td>0.000</td>
</tr>
<tr>
<td>Weighted by 1/MDE</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Publication bias weight</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Weighted by P-score from nudge characteristics</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

Standard errors clustered by trial are shown in parentheses. Coefficient on Academic Journals sample is the estimated average difference in percentage point (p.p.) treatment effects between the Academic Journals and Nudge Units samples. MDE (minimum detectable effect) is calculated in p.p. at power 0.8. Weighting for publication bias assigns significant trials a relative weight of .22 compared to insignificant trials in the in the Academic Journals sample. P-score is the propensity score using predicted probabilities from a logit regression that includes the same nudge characteristics controls as in Table 4a. When computing P-score weights, Nudge Unit trials with missing control take-up % data are assigned the Nudge Unit sample average.
Decomposing the Effect Size Gap

- The estimate of nudge effect for Nudge Unit trials is reliably about 1.3-1.7 pp

- The published papers have larger effects due to
  - Publication bias
  - This is compounded by low statistical power

- There is also some different characteristics (part of going at scale)

- Conditional on these factors, not much difference from going “at scale”
Nudge Unit Effect (Prior Expectations)

- How does the effect line up with expectations? (DellaVigna and Pope, 2018a,b; DellaVigna, Vivalt, and Pope, 2019)

- Elicit forecasts:
  - 10-minute survey
  - Present the setting, including 3 examples of nudges
  - Ask to predict average p.p. effect of a nudge
    - In Nudge Unit sample
    - Among published papers from meta-analysis

- 237 respondents
  - 27.9% university faculty and post-docs
  - 24% graduate students
  - 16.9% non-profits and government agencies
  - 15.2% private sector
  - 11.8% nudge practitioners
Nudge Unit Effect (Prior Expectations)

- Median forecast of 4pp in between meta-analysis (8.7pp) and actual impact (1.4pp)
- Forecast of nudge effect for published papers is right on!

Forecasts for nudge units: 237 respondents
Forecasts for published nudges in academic journals: 203 respondents
What is effect of experience with nudges?

211 of 237 respondents report approx. no. of field exp. conducted

Sample includes 28 respondents who work in nudge units

Nudge unit predictors right on! Experience helps with accuracy
Section 4

Concluding Remarks
How to complete a dissertation and be (approximately) happy

1. Know yourself, and put yourself to work
   - What gets you going and excited?
   - What instead stops your progress / makes you procrastinate?
     - Are you afraid of undirected research?
     - Not enough intuition?
     - Not enough technicality?
   - We differ in our gifts:
     - Intuition and creative ideas
     - Ability to organize material
     - Technical ability
   - Work in teams with someone who complements you
2 Economics is about techniques AND about ideas

**Rule 1. Study the techniques**

- Everyone needs a knowledge of:
  - Modelling skills (decisions, game theory, contracts, behavioral models)
  - Econometrics (asymptotics, applied metrics)
  - (At least) one field (methodology, questions, previous research)
Economics is about techniques AND about ideas

Rule 2. Think of interesting ideas

- Start from new idea, not from previous papers. Ex.: Mas-Moretti on Safeway data
- Think of an idea that can fix a broken literature (Levitt). Ex.: Fehr-Goette on cab drivers
- Connect two literatures which were unconnected. Ex.: Eisensee-Stromberg on political economy + behavioral

Rule 3. Explore technique you need for idea

- Ideas often come first
- It will be much easier to learn technique once you have an interesting problem at hand
What are good ideas?

1% of GDP (Glaeser)
New questions (better) or unknown answers
Questions you care about and topics you know about (comparative advantage: List)
Socially important topics (Akerlof)
Good research is always useful, even if not policy-relevant
Look for occasions to learn:

- Attend seminars (including student lunch talks)
- Attend job market talks
- Read literature, but critically: What is missing? Where could I apply this idea?
- Discuss ideas with peers, over lunch, with yourself
- Get started on some data set
- Be curious
It is OK to go on the job market as a behavioral student

- Yes, demand for behavioral students is still relatively limited
- BUT supply is even more limited: mainly Berkeley, Harvard + some Cornell, CalTech, CMU
- Many young behavioral economists have Berkeley training:
  - Harvard (Gautam Rao)
  - MIT (Frank Schilbach)
  - Chicago (Devin Pope + Avner Shlain)
  - Princeton (Anne Karing)
  - LSE (Kristof Madarasz, Matthew Levy)
  - UCSB (Erik Eyster, Youssef Benzarti)
  - CMU (Saurabh Bhargava)
  - Wisconsin (Justin Sydnor),....
- Many students with (quasi-)behavioral paper on the market show lack of behavioral training (eg, development papers with experiments in tow)
- Show off your Berkeley training!
Above all, do not get discouraged...

- Unproductive periods are a fact of life
- Ideas keep getting better (and economics more fun) with exercise
- Work hard
- Keep up the exercise!
Section 5

Teaching Evaluations
Teaching Evaluations

- Video: https://www.youtube.com/watch?v=YQygxKl5z3w