Behavioural Economics: What do we really know (and why does it pay to be on guard)

UNSW Behavioural Insights for Business and Policy Research Network Roundtable
November 8, 2016

Andreas Ortmann
School of Economics, UNSW Business School
More on incentives … and behavioral nudges

_Recall_

“…monetary incentives are strong motivators, nonmonetary psychological inducements are moderately effective, and results using behavioural factors are generally consistent with models of social and time preferences.” – Stefano Dellavigna and Devin Pope, “What Motivates Effort? Evidence and Expert Forecasts” (wp 2016)

_Next is a graph succinctly summarizing their 18 treatments …_

- The task: pressing two buttons alternatingly as often as possible in a ten-minute period
- Standard M-Turk payment in the baseline condition
- Horizontal axis show the number of button presses: Note where it starts
- _Pay particular attention to these treatments:_
  - “… you will be paid a bonus … . Your score will not affect your payment.” (gift exchange)
  - “Please try as hard as you can.” (exhortation)
  - “We will show you how well you did relative to other participants.” (social comparison)
  - “Many participants scored more than … points.” (generalized social comparison)
  - “You will be paid an extra (bonus) if you score at least … points.” (conditional gift exchange)
More on incentives … and behavioral nudges (ctd)

Figure 6. Average Button Presses by Treatment and Average Expert Forecasts

Actual and Forecasted Button Presses by Treatment - All Expert Survey Takers

- Gift Exchange
- Exhortation
- Social Comparison
- Generalised Social Comparison
- Conditional Gift Exchange
So, who is mis-behaving?


  *tries to paint himself as the resident bad boy who fought the law – here “standard” economics -- (and won, or so he wants us to believe)*

• Loewenstein & Ubel, in NYT article titled *Economists Behaving Badly* (2010)

  *argue that the “over-reliance on behavioral economics” has become a real problem and that “behavioral economics (BE) should complement, not substitute for, more substantive economic interventions.*

  *If traditional economics suggests that we should have a larger price difference between sugar-free and sugared drinks, BE could suggest whether consumers would respond better to a subsidy on unsweetened drinks or a tax on sugary drinks.”

  *Loewenstein, of course, one of two “early students of misbehaving” that Thaler dedicated his book to. The irony.*
What is BE anyways?


- BE provides “a richer understanding of how people actually think and behave” (World Bank, WDR 2015 (December) Overview: Mind, Society, and Behavior)

- In sum, a more realistic version of economics, so to speak, made more realistic by importing insights from psychology.

- That narrative is the dominant one … … but unfortunately patently wrong.

- There is no monolithic body of evidence in psychology that can be imported into economics hook, line, and sinker. In fact, there are huge controversies in psychology about the reality of cognitive illusions (biases), and the (de)merits of heuristics, and they go back for decades and remain heated (e.g., Smith JPE 1991, AER 2003, Gigerenzer PR 1996, 1996a, Kahneman & Tversky PR 1996, …, Spiliopoulos & Ortmann, in press)

- Also continued controversy between, and amongst, psychologists and economists about how to produce laboratory and field evidence. Does it matter?

- We are in the midst of a deepening replication crisis in psychology, economics, and the sciences generally. About that can be little doubt.
There are huge controversies in psychology about the reality of cognitive illusions (biases) and the (de) merits of heuristics.

- Thaler and WDR 2015 ignore them and continue to allege the reality of endowment effects, sunk costs, mental accounting, income targeting, and so on when in fact each and every of these “anomalies” have been contested and in many cases thoroughly deconstructed.

- The only interesting issue at this point should be, when, and under what conditions, they might appear and what kind of people might be afflicted by them. Claims of categorical existence ought not to be trusted.

- “The Report provides literally hundreds of examples of interventions - … What the Report does not do, unfortunately, is the kind of red teaming that it advocates as ‘one way to overcome the natural limitations on judgment among development professionals … In red teaming, an outside group has the role of challenging the plans, procedures, capabilities, and assumptions of an operational design, with the goal of taking the perspective of potential partners or adversaries. … Group deliberation among people who disagree … increase the odds that the best design will come to light and mitigate the effects of groupthink’ (p. 19 of the WDR 2015) … there is little red teaming on display here. Not that that is a particularly new development. … The troubling omission of contrarian evidence on display in the Report is deplorable … .” (AO, reviewing the WDR 2015, in JoEP 2016)

- Red teaming to the rescue. By all means …
If you want to nudge (benevolently the errant actor...) you better know what you do ... and what you do not know.
How behavioural nudges can go wrong, take 1

The Piano stairs: https://www.youtube.com/watch?v=2lXh2n0aPyw

O.K., they are fun (and their effect, incidentally, can be easily explained by standard economic decision theory).

But … what, when the same person has to take the stairs everyday?
When will the novelty wear off?
And what about all the SMS reminders?
When will they wear off?

Important related question: Do passive decisions crowd out active decisions?
(Carroll et al. QJE 2009 but see also Chetty et al. QJE 2014)
How behavioural nudges can go wrong, take 2

“Retirement Savings: A Tale of Decisions and Defaults”

- How do members of an industry-wide pension fund deal both with the prevalence of defaults and their impact on retirement savings.
- Empirically (using admin data), preferences, demographic characteristics and labor mobility matter greatly when it comes to the extent of active plan choice (or high defaulting) and overall level of pension benefits.
- Using a structural dynamic life-cycle model, we show the ability of these empirically motivated factors to explain retirement savings patterns.
- In our model, individuals decide how much to save in a setting that combines irreversible automatic enrolment with an active decision regime.
- After automatic enrolment, there is an initial choice between two pension plans (DB vs. DC), and then members decide (possibly by default) their voluntary contributions and type of investment allocation.
- Our simulation results show that default settings strongly influence wealth accumulation. For example, changing the default plan from DB to DC leads to a 10% (18%) net increases in total pension wealth for men (women), respectively.
- Settings are highly persistent, both over time and across decisions.
How behavioural nudges can go wrong, take 3

BORIS research dashboard, as announced in recent UNSW research news:

"UNSW is introducing a new research dashboard that brings together a wide range of metrics. The user-friendly, web-enabled dashboard highlights various metrics such as research income, outputs, collaboration, h-index, HDR supervision and completion, FWCI, SNIP, and SJR. Representing a single source of aggregated information that currently exists in multiple places, the new dashboard removes the requirement to collate this manually. With drill-down and reporting functionality, BORIS will also provide a visualisation of an academic’s position relative to others in their school cohort."

In other words, the next step of kpi-sation.

Big data, behavioral nudges (social comparison). What could possibly go wrong?
How behavioural nudges can go wrong, take 3 (ctd)

*Big data, behavioral nudges (social comparison). What could possibly go wrong?*

First, in any multi-task environment, the important question arises how you weigh someone's contributions to research, research income, collaborations, impact, HDR supervision and completion, teaching, engagement, and what not.

Second, in any such environment you have to determine what the window is you want to look at, and how the performance along these dimensions was accomplished relative to opportunity. How will you take into account that someone has, possibly for decades, been research only? How will you take into account that someone had maternity leaves, etc.?

Third, if we accept for a second that an index could be constructed, then you better also understand that the system will be played in predictable ways. This is basic Holmstroem & Milgrom(1990) and Baker (1992) hard core incentives stuff that has plenty of empirical support (see Prendergast 1999, p. 9): "In essence, complex jobs will typically not be evaluated through explicit contracts." In other words, just because you have big data … (you might not want to use it.)

Last but not least, BORIS's promised "visualisation of an academic's position relative to others in their school cohort" has to address the three points above and seems to try to use alleged insights from BE to extract more effort. Unlikely to work, and quite possibly will have just the opposite effect, even if the system is not riddled with mistakes (which most likely it will be).
What does Zuckerberg (the king of big data and nudges) know that others don’t?

**Facebook Employees Are Insanely Happy With Their Jobs**
Mark Zuckerberg must be doing something right.

Emily Peck
Executive Editor, Business and Technology,
The Huffington Post
What does Zuckerberg (the king of big data and nudges) know that others don’t? 2

Facebook Employees Are Insanely Happy With Their Jobs

Facebook Has The Least Stressed, Most Satisfied Employees

- Facebook: 96%
- Salesforce: 44%
- Google: 60%
- Adobe: 70%
- eBay: 80%
- Cisco: 80%
- SpaceX: 60%
- Apple: 70%

Source: PayScale. Icons: Facebook

The Huffington Post
What does Zuckerberg (the king of big data and nudges) know that others don’t?
Take-home lessons

• The (BE) evidence base is rather shaky, not withstanding some people’s claims.
• The effects of behavioural interventions in particular are often rather poorly understood.

• Best not to rely on the results of one study, our intuition, only.
• Best to canvass the evidence base as completely, and as systematically, as possible.

• *Red teaming* is highly recommended to avoid counterproductive design and implementation choices. Transparency and openness seem imperative.

• Setting defaults and nudges is a science; intuition might not to be a good guide.

• Modern economics (the one that wins people Nobel prizes) has a lot to offer …
Thank you!
Discussion points

• What nudges and defaults have you found useful?
• How do you assess the usefulness of nudges and defaults?
  (What is the role of theory and Big Data in your assessment?)
• Do you have formal or informal red-teaming processes in place?
  (If so, how are they informed by theory and Big Data?)
• What barriers have you encountered in implementing nudges and defaults?
• If we take the FB employee’s (and researcher’s) assessment of why FB is a success, what lessons are there for your organization?
  (Recall she said: “(it is) the transparent and open learning environment across different teams and organizations, as well as its endless effort of being honest and transparent about difficult conversations … “)