THE UNIVERSITY OF MELBOURNE

FACULTY OF BUSINESS & ECONOMICS

Melbourne Institute Policy Briefs Series

Policy Brief No. 4/15

Beyond Nudge: The Potential of Behavioural Policy

Reuben Finighan

THE MELBOURNE INSTITUTE IS COMMITTED TO INFORMING THE DEBATE

MELBOURNE INSTITUTE®
of Applied Economic and Social Research

Beyond Nudge:
The Potential of Behavioural Policy*

Reuben Finighan
Melbourne Institute of Applied Economic and Social Research
The University of Melbourne

Melbourne Institute Policy Brief No. 4/15

ISSN 2201-5477 (Print)

ISSN 2201-5485 (Online)

ISBN 978-0-7340-4385-6

July 2015

* I thank Deborah Cobb-Clark, Paul Jensen, Guyonne Kalb and Mike Pottenger for useful comments and suggestions. For correspondence, email <reuben.finighan@unimelb.edu.au>.

Melbourne Institute of Applied Economic and Social Research
The University of Melbourne
Victoria 3010 Australia
***Telephone* (03) 8344 2100**
***Fax* (03) 8344 2111**
***Email* melb-inst@unimelb.edu.au**
***WWW Address* http://www.melbourneinstitute.com**

Abstract

Policymakers often mistakenly see behavioural policy as synonymous with "nudging". Yet nudges are only one part of the value of the behavioural revolution—and not even the lion's share. This policy brief looks at the full potential of behavioural policymaking. Using examples in cigarette regulation, retirement savings, and poverty, it makes three arguments:

1. Trade-offs between social welfare and individual freedom of choice affect both nudges and conventional policy tools (like mandates, taxes and transfers). Nudges are not special tools that avoid such trade-offs.
2. Nudges typically sacrifice social welfare in favour of individual choice, and for this reason many well-known nudges will be unattractive in Australia. Some nudges can, however, play a complementary role by fine-tuning conventional policy regimes.
3. Conventional policy tools will continue to be the most powerful tools for countering behavioural biases, and have the most promise for driving major, behaviourally-informed reforms in Australia.

Policymakers should not ask "can we nudge this?", but should instead ask how behavioural evidence changes the way they think about *all* the options in the policy toolkit.

JEL classification: D03, D60

Keywords: Behavioural economics, behavioural policy, public policy, nudges, regulation, paternalism

Behavioural economics is being hotly debated in policymaking circles across the world. In Australia too, it has been welcomed with the blend of enthusiasm and scepticism that is rightly due to any paradigm shift. While the field has won substantial academic credibility—with economists widely accepting that behavioural economics can improve upon the standard analysis in important ways—it faces a new set of contentious questions. What are the field's *practical policy implications?* Does behavioural policymaking have the potential to contribute big ideas and stimulate important reforms, or will it merely tinker at the edges?

In trying to answer these questions, many public servants substitute them with another: "what is the potential of nudges?" This is an error, and a potentially serious one. Nudges are not synonymous with behavioural policymaking, but represent just one part of its value. The close identification of nudges with behavioural economics is largely a historical quirk: Sunstein and Thaler's groundbreaking *Nudge* (2008, Yale University Press) was the first book to popularise behavioural policymaking, and has powerfully framed the discussion.

This policy brief aims to take a more complete look at the potential of behavioural policymaking, paying special attention to policy tools beyond the nudge. Nudges, it argues, largely tinker at the edges. Major reforms, especially in Australia, are more likely to come through using behavioural evidence to inform the use of conventional tools, like taxes, transfers and mandates. Behavioural policymaking does not mean asking, "can we nudge this?" Much like conventional policymaking, it involves weighing each of the options in the policy toolkit to find the tool that best balances the value of individual choice with the magnitude of potential welfare gains. What is new about behavioural policymaking is that it takes the rich body of behavioural evidence into account.

The brief builds this case upon three key insights:

1. **Trade-offs between agency and welfare affect *both* nudges and conventional tools.** Nudges are not special tools that can achieve major welfare gains without interfering with choice. In practice, they are affected by the same trade-offs as conventional tools.
2. **Nudges, which Sunstein and Thaler point out have been in use long before *Nudge,* have generally been best suited to driving modest changes in behaviour at low cost**. They tend to play a complementary role, extending the effect of conventional policy regimes. Only under ideal conditions are they powerful by themselves.
3. **Conventional policy tools like mandates, taxes and transfers have been, and will continue to be, the most powerful tools for countering behavioural biases.** The major, prosperity-enhancing reforms that behavioural economics will drive in the coming decades will, for the most part, employ conventional policy tools.

Theoretical foundations: Behavioural biases and market failure

The behavioural economics literature explains observed behaviour as the interaction of two factors: normative preferences (what the individual actually wants to achieve, or what would actually maximise their welfare), which are then distorted by behavioural biases to form revealed preferences (the actual choices observed, which often fail to maximise welfare). The literature describes hundreds of different behavioural biases, and a multitude of them could apply to a given policy problem. For example, the typical individual holds a normative preference to save enough to have a comfortable retirement. In the absence of a supportive policy regime like superannuation, this preference may be distorted by present bias, analytic errors, forgetfulness, inattention, and many biases beyond these. For a complex, intertemporal choice like savings, this list of biases could go on and on.

Biases are usually organised into three categories: first, *imperfect optimisation*, comprising the biases that prevent us from calculating the consequences of our choices, such as the bias toward forecasting exponential growth as though it were linear; second, *bounded self-control*, including psychological phenomena like procrastination, emotions and impulses, which prevent us from implementing our normative preferences (and guarantee the failure of our New Year's resolutions); and third, *non-standard preferences*, where we simply want things that conventional economists do not expect us to want, such as to avoid experiencing painful losses even if this means winning less overall. Note this last category includes some phenomena that can be genuine desires, like conformity with social norms, although these too are regularly exploited—for example, by marketers paying handsome actors to smoke cigarettes.

When behavioural biases distort normative preferences, the decisions we make may be far from welfare-maximising. Scholars like Brigitte Madrian argue that biases should therefore be regarded as a new class of market failure.[1] Behavioural market failures, like other kinds of market failure, open up potential room for government intervention. Where policy is well-designed, it may improve the efficiency with which the market produces social welfare.

Conventional policy tools typically address market failures, or indeed achieve other policy goals, by shifting monetary incentives or restricting what people can choose. They include taxes and subsidies, mandates or regulations, and various kinds of material transfers. Nudges, on the other hand, eschew monetary incentives or strong restrictions on choice, and instead depend upon small changes in "choice architecture" that gently guide choices in the desired direction. Nudges are only effective because of the ways in which real-world people are different to the

rational, self-interested agent—in Richard Thaler's terms, good nudges will affect the real-world "Humans" of psychology, but not the perfectly rational "Econs" of economics. Note that conventional tools can be strengthened behaviourally, such that an Econ would not notice the change (e.g. reframe an incentive as a loss, more painful for Humans subject to loss aversion).

Importantly, a behavioural policy problem does not necessarily call for a nudge solution. As we shall see, nudges and conventional policy interventions can be used to resolve either behavioural or conventional policy problems. Table 1 illustrates this with a matrix of four possibilities—although note that policy problems are often a *mix* of the conventional and the behavioural, and interventions are often a *mix* of nudges and conventional approaches.

Table 1: *The range of approaches available to the behaviourally-informed policymaker to resolve both conventional and behavioural policy challenges*

	Conventional policy problem	**Behavioural policy problem**
Conventional intervention	E.g. align teacher interests with student learning by providing a pay bonus for better student performance.	E.g. to overcome savings myopia, use a mandate (e.g. superannuation) or a financial incentive to increase savings.
Nudge intervention	E.g. align teacher interests with student learning by providing social rewards, like recognition and thanks.	E.g. to overcome savings myopia, use a nudge-like intervention such as pre-commitment to saving.

Insight One: Both nudges and conventional tools are subject to trade-offs between agency and welfare

Nudge-like policies are not new. Policymakers have long intuited that people are not perfectly rational—that framing is important, that we follow social norms, that we tend to stick with defaults, and so on. What recent advances in behavioural economics have done is dramatically enriched our understanding of how nudges work and how they should be designed—and, notably, increased their salience. As importantly, or perhaps even more so, behavioural economics has improved our understanding of when conventional tools should be used and how they should be designed. Why, then, has all the emphasis been on the nudge?

The answer is partly political. Sunstein and Thaler wrote *Nudge* within the unique political context of the United States, where the libertarianism of political entities like the Tea Party is a powerful force to be reckoned with. They promote the political philosophy of "libertarian

paternalism" as a way of resolving the tension between libertarians and progressives—on the one hand, the desire to preserve individuals' agency in being able to choose what they prefer, and on the other, the growing recognition that individuals' choices are strongly shaped by cognitive limitations and contextual factors, with sometimes severe consequences for welfare.[2] Nudges are supposed to *influence choices* in ways that increase welfare but that *avoid coercing choices*, letting policymakers steer a safe course between these warring tribes.

If the idea of avoiding politically charged trade-offs between welfare and agency sounds too good to be true, unfortunately this is because it *is* too good to be true. The distinction between nudges and conventional tools turns out to be exaggerated, with nudges subject to much the same trade-offs. Take two examples: first, the *Save More Tomorrow* nudge scheme, which has been too non-coercive to be effective; and second, efforts to reduce smoking, an example of nudges that are effective precisely because they are coercive.

Savings behaviour is strongly affected by a range of biases, and the celebrated *Save More Tomorrow* scheme uses a set of cleverly designed nudges to counter them. The scheme's most important innovation is a default automatic escalation in savings over time, so individuals will save more in future years without experiencing a loss in income today. In an early trial, average savings reached almost 14%. Yet actual implementation has been less impressive: a majority of employers offered automatic escalation by 2011, but only 11% of employees chose to enrol in the plan. The escalation default was 3% over 3 years, with no further increases thereafter. The US-wide savings rate increased by a feeble 0.33%,[3] a tiny fraction of the effect achieved by schemes like Australia's superannuation—with enormous implications for welfare.

The problem with agency-preserving nudges like *Save More Tomorrow* is that they continue to rely upon the discredited rational model. People must choose to participate and must choose the right settings, but these are choices that the behavioural evidence clearly shows we cannot reliably make. Yet if a nudge scheme like *Save More Tomorrow* was strengthened enough to be effective—say, all employers are required to place employees in opt-out savings escalation schemes, with a high default savings rate, and penalties for those who do opt-out—then it effectively becomes a coercive mandate with a costly escape clause. Evidence from the behavioural sciences suggests that even this scheme could be welfare reducing, because people who choose to opt-out often do so irrationally, with sometimes enormous welfare costs.

An opposite case is illustrated by cigarette smoking nudges, which are substantially more coercive than the *Save More Tomorrow* scheme. Plain packaging rules make cigarettes

unattractive; the Quit campaign provides services that prompt smokers to plan ahead and make commitments; removing cigarettes from view and selling them at a separate counter makes them harder to access; health warnings placed at the point of sale make for a timely and emotionally powerful reminder; and social stigma pits smokers against the powerful forces of social conformity. People are still free to exercise the choice to smoke, but every nudge in the book is pointing in the other direction. Though this is a nudge regime, it is clearly highly coercive—and it is precisely the high costs imposed on smokers that makes the regime effective.

Insight Two: Nudges are best suited to driving modest changes in behaviour at low cost, and can enhance the effects of conventional tools

Nudges are not special tools that evade agency-welfare trade-offs. They are, nonetheless, very useful under the right conditions. This policy brief will not attempt to improve upon the many excellent overviews that already exist. A more important question for the brief is this: given nudges sometimes completely shift aggregate behaviour, and on other occasions are entirely ineffective, when should policymakers be particularly enthusiastic about the potential of nudges?

A well-designed nudge may be a particularly powerful policy option where:

1. individual underlying preferences are relatively homogeneous and are already aligned, or only weakly misaligned, with the desired behaviour;
2. the cognitive biases that distort underlying preferences can be easily overcome, or used to the policymaker's advantage; and
3. either individual choice cannot be avoided, or agency is valuable and the benefits of preserving agency are not outweighed by welfare costs—in either of these cases, the use of stronger policy options is not possible.

Organ donation nudges are the poster child for such success: (1) a large majority of people hold normative preferences that are aligned with the socially optimal decision to donate; and (2) the decision to donate is not fixed by powerful biases, but is largely the result of a default bias that policymakers can easily harness. These conditions allow a nudge—a switch from an opt-in to an opt-out default—to completely transform outcomes, such that a majority rather than a minority choose to donate. On point (3), a nudge is superior to a mandate here because agency is crucial: people cannot reasonably be forced to donate their organs, and the benefits of preserving choice are judged to outweigh the costs.

A more modest example is a program aimed at improving drug prescription charts to reduce medical errors in hospitals: (1) doctors and nurses hold a normative preference to enter, and read, the right drug and prescribed dosage from the prescription paperwork; (2) the problem of illegible handwriting can easily be overcome with better designed forms; and (3) this is a case where the behaviour is already mandated to the hilt—medical practitioners face strong incentives to deliver the correct drugs and dosages. Nudges can help staff comply with the mandate (and, for most practitioners, their own preferences). Indeed, the new charts virtually eliminated errors caused by confusion between milligrams, or "mg", and micrograms, or "mcg".[4]

In many cases these conditions are not met, and nudges will produce relatively small gains at the margin—albeit usually at very low cost. They will be a useful, but limited, part of a wider policy mix.

Insight Three: Conventional policy tools will continue to be the most important tools for countering many behavioural biases

Policymakers do not typically think of conventional policy instruments as central to behavioural policymaking, yet they should. This sub-section provides examples where a behavioural analysis of a policy problem favours the use of a conventional policy tool, with examples provided under mandates, standard incentives, and transfers.

Mandates and regulation

Mandates have been the tool of choice for many important welfare optimisation problems—seatbelts, bicycle helmets, cigarette advertising, rental property standards, fair mortgage contracts, water fluoridation, milk pasteurisation, superannuation, minimum wages, and much beyond. Regulations are the most powerful instruments for shaping behaviour, with the level of interference increasing as they move from requirements or restrictions imposed on marketing (e.g. payday loan advertising), to product attributes (e.g. unfair interest rates), and finally behaviours or whole products (e.g. payday lending).

Their strength gives them high potential for both benefit and harm, so it is unsurprising that they are often controversial. Benefit is most likely where normative preferences are relatively homogeneous so that a one-size-fits-all mandate can produce benefits for most people; where choice errors are common and costly in the absence of a mandate; and where agency is not too

highly valued. Harm is likely where normative preferences are heterogeneous, and individuals are blocked from undertaking behaviours that genuinely enhance their welfare.

Mandates can improve welfare in cases of bounded self-control or imperfect optimisation. Beshears et al. (2014) provide an example of how mandates can address *bounded self-control*: they study a model of retirement savings where individuals are subject to different degrees of present bias, and therefore different levels of likelihood for withdrawing from their savings too early. What level of penalty on early withdrawals, to discourage the behaviour, would maximise welfare? Their analysis finds the welfare-maximising penalty is 100%—that is, withdrawing should be impossible. A lower penalty does allow a small number of people to withdraw early in a rational way, but their welfare gains are completely overwhelmed by the welfare losses experienced by those who irrationally withdraw due to present bias.[5]

This is one weakness of nudge-based saving schemes that allow opt-outs, like *Save More Tomorrow*: the people who opt-out of such schemes may believe they can do better without it, but this is typically not the case. On average, they experience significant welfare losses. Indeed, the US Internal Revenue Service imposes a 10% penalty on early savings withdrawal from 401(k) plans, but even so, for every $1 that is saved, $0.40 is withdrawn prematurely at substantial cost.

Australia's superannuation policy usefully highlights another property of a successful mandate: popular consent. Because mandates are open to charges of paternalism, they depend heavily upon democratic legitimacy. Mandates are most likely to succeed when they are understood to have been imposed by a community upon itself, without marginalising underrepresented sub-groups; and when they enact most individuals' normative preferences, without eliminating any *cherished* choices. In the case of superannuation, most people do wish to save adequately and over 80% of Australians are in favour of the scheme.

Standard incentives

Incentives are most useful when choices are shaped by *self-control* biases. The literature is thus focused heavily upon policy problems like the promotion of healthy behaviours, such as quitting smoking or going to the gym; educational outcomes, like student performance; and labour market outcomes, like incentives for finding a job or for working harder. Some people recognise that they lack self-control, and will willingly impose costs upon themselves in ways that would be bizarre for a rational agent. For example, studies have found that data entry workers

willingly subject themselves to payment schemes with high potential downside, but no obvious upside—for example, they threaten themselves with substantial monetary loss if they fail to reach a daily desired target, say, 1,000 entries in a day. People choose such schemes as *commitment devices,* and tend to work harder under such self-imposed threats.

Let us return to cigarette addiction as a prime example of individuals struggling to implement their normative preferences. Around 70% of smokers would like to quit at any given moment, but only 3% succeed in a given year. Research suggests that smoking imposes costs on one's future self of around US$35 per pack, a figure that is vastly higher than what people ever pay for a pack. Even the most coercive nudges cannot provide a signal that reflects this cost. Cigarette taxes, on the other hand, can provide a motivational counter to addiction that studies suggest is one of the more effective parts of a policy mix for reducing smoking.[6]

While a tax imposes higher costs on smokers and has been criticised as regressive, behavioural evidence suggests that it may reduce smoking sufficiently that its long-run effect on most individuals is net positive. A cigarette tax can therefore be progressive rather than regressive. Behavioural studies of US and Canadian smokers have demonstrated that increasing taxes on cigarettes can actually *increase* the well-being of cigarette smokers, by assisting them to exert self-control.[7]

Incentives are also used to encourage pro-social behaviour, and behavioural evidence turns out to be crucial for successful design. Incentives are *messages*. The mere presence of an incentive, which attaches a positive or negative signal to a choice, can be more important than the magnitude of the incentive. Individuals interpret this message within a particular social context, and incentives thus interact with social preferences in important ways. Careless use of incentives can backfire by displacing and "crowding-out" beneficial social preferences, while well-designed incentives will instead "crowd-in" and strengthen social preferences.

Whether crowding-out or crowding-in predominates is determined by the *interpretation* of the incentives. Incentives are more likely to backfire when they are seen as manipulative or imposed from above; when they are arbitrary, difficult to act upon, or unfair; when they encourage selfish maximisation behaviour; or when they suggest that an activity is unpleasant and therefore must be compensated. Individuals are more likely to see an incentive as legitimate if they understand and endorse its purpose, and if they perceive the incentive as imposed by their peers for the good of the group as a whole. Pro-social preferences can be strengthened when small incentives are used to reiterate group norms.

Finally, note that incentives tend to be less effective for addressing *imperfect optimisation* biases. An excellent example is the design of an aircraft cockpit: early designs were extremely complex, with hundreds of buttons, levers and instruments. Clearly pilots have every possible incentive to fly well, given their very survival depends upon it—and yet the complexity of the environment led to critical errors. In such a case, further incentives are clearly futile. Optimisation biases are better countered by nudges (such as making the choice problem easier, or the incentives easier to understand), as well as mandates and in some cases transfers, although incentives may be used to suggest a useful strategy for solving a complex problem (e.g. small incentives for investors to diversify their portfolio).

Transfers

The rational model holds that transfers, such as those provided as unemployment benefits, should come with a significant hassle cost in order to disincentivise their use and prevent moral hazard.

A cognitive scarcity model suggests something quite different.[8] First, transfers should be evaluated not only for their well-being benefits but also their impact on *cognitive resources*. Cash infusions that liberate Indian farmers from poverty, for example, produce increases in IQ of nearly a full standard deviation—a dramatic effect that would be the envy of any educational system. Second, the hassle cost of meeting the requirements of welfare services can perversely impose a "cognitive tax" that only *adds* to the cognitive challenges of poverty. Policymakers should be careful to design schemes that do not pose excessive burdens on already complicated lives.[9]

Where behavioural evidence and the rational model agree is on how individuals spend such transfers: studies suggest that spending is relatively rational. The number of vice goods purchased as a proportion of total spending tends to decrease when transfers are provided, while spending on policymakers' preferred goods, like nutritious foods, tends to increase.[10]

Utah's homelessness strategy makes for an interesting case study. In 2005, "no-strings attached" housing was first provided to the chronic homeless, without requiring drug tests and with minimal rent. There were fears that the project would "incentivise mooching", as per the rational model. Utah's chronically homeless population has since fallen 90%, and by the end of 2015 "may be virtually gone".[11] Most recipients successfully make rent payments, and costs are

much lower than that of providing services for the chronically homeless. While no behavioural analysis has yet been performed, the project may have powerful cognitive effects: providing housing simplifies the lives of the homeless, provides a buffer against shocks, and frees cognitive resources for challenging tasks like quitting drugs and finding employment.

Conclusion

Behavioural economics has peeled away the axioms of the rational model to reveal the layers of psychological complexity lurking underneath. Policymakers can use this richer understanding of human psychology to improve the design of nudges, mandates, incentives and transfers. These tools are essential parts of the policy mix, and in practice are often complementary. Conventional tools likely provide the most substantial potential for significant welfare-enhancing reforms, while nudges tend to be most helpful for fine-tuning institutions in order to bring efficiency gains at low cost. Whichever tool is employed for a given policy problem, policymakers should seek to balance individual agency with social welfare.

Unlocking this potential will require new practices of policy design and evaluation, focused on generating and applying new evidence. Behaviour departs from the rational model in myriad ways, and there is an equally large array of options for setting behaviour on the right course. Such a wide and complex field of possibility must be navigated with an interdisciplinary mix of methods from the applied behavioural and economic sciences, to identify the causal relations that underlie policy problems and that make interventions work. This is behavioural policymaking—the weighing of all policy options in light of the full body of evidence.

---

[1] Madrian, B (2014) "Applying insights from behavioral economics to policy design", Working Paper no. 20318, National Bureau of Economic Research.
[2] Sunstein, C & Thaler, R (2008) *Nudge*. Yale University Press.
[3] Benartzi, S & Thaler, R (2013) "Behavioral economics and the retirement savings crisis", *Science,* 339(6124), 1152–1153.
[4] King, D, Jabbar, A, Charani, E, Bicknell, C, Wu, Z, Miller, G, Gilchrist, M, Franklin, B & Darzi, A (2014) "Redesigning the 'choice architecture' of hospital prescription charts", *Health services research*, 4(12).
[5] Beshears, J, Choi, J, Clayton, C, Harris, C, Laibson, D & Madrian, B (2014) "Optimal illiquidity", paper for the *Retirement Research Consortium*, September 26.
[6] E.g. see Lewit, E & Coate, D (1982) "The potential for using excise taxes to reduce smoking", *Journal of Health Economics,* 1(2), 121–145.
[7] Gruber, J & Mullainathan, S (2005) "Do cigarette taxes make smokers happier?", *The BE Journal of Economic Analysis & Policy*, 5(1).
[8] Mani, A, Mullainathan, S, Shafir, E & Zhao, J (2013) "Poverty impedes cognitive function", *Science*, 341(6149), 976-980.
[9] Mullainathan, S & Shafir, E (2013) *Scarcity: Why Having Too Little Means So Much*. Macmillan.
[10] Evans, D & Popova, A (2014) "Cash transfers and temptation goods: a review of global evidence", World Bank Policy Research Working Paper 6886.
[11] Rascon, J & NBC News (2015) "Utah's strategy for the homeless: give them homes", *NBC News*, May 3.