10. Reverse engineering and policy design

R. Kent Weaver

1 INTRODUCTION

A mechanistic perspective on policy analysis and design has been described as focused on “a theory of a system of interlocking parts that transmit causal forces from X to Y” (Beach and Pedersen, 2013, p. 29). Hedström and Ylikoski (2010, p. 53) argue that a “mechanism-based explanation describes the causal process selectively. It does not aim at an exhaustive account of all details but seeks to capture the crucial elements of the process by abstracting away the irrelevant details.” In the approach used in this volume, first-order causal mechanisms are seen as those that “alter the behavior of individuals, groups and structures to achieve a specific outcome” through use of policy activators embedded in government policy (Capano, Howlett and Ramesh, Chapter 1 this volume). Second-order causal mechanisms are the use of knowledge about mechanisms at work in individual and collective behaviors to inform revisions to policy “activators.” This chapter examines the first- and second-order causal mechanisms at work in a policy sector that many governments have tried to influence through use of policy activators: retirement savings by households. It uses that analysis to draw broader implications for the understanding and utilization of causal mechanisms in policy research, and in particular the potential of and limitations on reverse engineering, that is, using an understanding of how causal mechanisms operate to design mechanisms whose predicted outcomes “coincide with the *desirable* outcome” (Maskin, 2008, p. 567; emphasis in original) sought by government. Reverse engineering can thus be seen as one form of second-order mechanism.

In this analysis, government policy activators are one of several factors that shape individual and household decisions on (and ultimately aggregate levels of) retirement savings. In the terminology used in this volume, the causal mechanisms at work in moving from Stage 2 to Stage 3 of the simplified causal model shown in Figure 10.1 – both policy activators and other factors that may influence individual and household retirement savings behavior – are first-order causal mechanisms. Retirement savings behavior in turn affects

Figure 10.1 A model of causal mechanisms in retirement savings policy

aggregate levels of retirement savings, and may also have spillover effects in other policy sectors that may or not have been anticipated and intended when the policy regime was put in place (Stage 4). The impact of aggregate savings and spillover effects that feed back to affect revisions to the policy activators at a later time period are second-order policy mechanisms. While the feedback loop of individual and aggregate retirement savings on objectives and attributes of individual workers is shown in the lower right-hand corner of Figure 10.1, it is not the focus of the analysis in this chapter.

The next section of this chapter briefly introduces the concept of reverse engineering. Section 3 develops a general framework for understanding first-order contextual and activator mechanisms that shape individual behavior and applies it to retirement savings behavior. These attributes include incentive structures, availability of information, cognitive and decision-making biases, and availability or lack of resources. I argue that retirement savings should not be viewed as a single decision but as a series of steps that can be analytically distinguished even when they take place simultaneously; I focus here primarily on the entry stage (decisions to begin saving for retirement) and accumulation rather than decumulation post-retirement. Section 4 focuses on policy activators (instruments and specific settings on those instruments) that have been developed to address those barriers. Section 5 discusses second-order mechanisms: the degree to which and constraints on use of knowledge about the effectiveness of policy activators and context to revise policy activators in later time periods. The concluding section addresses the utility of the retirement

savings case as a vehicle for understanding causal mechanism approaches, and argues that efforts to utilize reverse engineering in policy reform confront a variety of issues in policy adoption, implementation and political sustainability as well as program design.

2 REVERSE ENGINEERING

The primary use of the concept of reverse engineering has, not surprisingly, been in the fields of engineering and software engineering. Perhaps the most famous example of reverse engineering is the Soviet Union's laborious re-creation of the American B-29 strategic bomber as the Tupolev Tu-4 from three interned (and one wrecked) planes that were landed in the Soviet Union during World War II at a time that Russia was still neutral with regard to Japan. The end result was a virtually identical copy, requiring the design and manufacture of 105 000 parts (Gorman, 1998). As Curtis, Harston and Mattson (2011) note, reverse engineering of products is fundamentally concerned with information extraction rather than imitation of those products, though imitation (and possibly improvement) may be the motivation and the result. Indeed, reverse engineering is a common strategy for firms that are "market followers" seeking to overcome the advantages of "first-mover" firms (Fitzpatrick and DiLullo, 2006). A core premise of the reverse engineering approach is the *uniformity and replicability* of causal mechanisms: the performance of a reverse-engineered and copied propeller on a Tu-4 should be very similar to that on the original B-29.

In the social sciences, the concept of reverse engineering has been used more as an informal metaphor for efforts to trace and isolate the causal mechanisms and impact of public policies on complex social processes, as in Richard Elmore's (1979) concept of backward mapping in the policy implementation process. (For a more formal effort at reverse engineering in social science research, see King, Pan and Roberts, 2014.) The applicability of the reverse engineering metaphor is likely to be most problematic, however, when policy outcomes are the result of complex interactions between multiple policy activators (P_1 to P_n) and environmental factors (E_1 to E_n) that may change over time, and a heterogeneous set of actors (A_1 to A_n) whose preference rankings (R_1 to R_n) are fungible and intransitive. All these factors may be subject to change over an extended period of time, especially when processes of policy layering and restructuring (or "packaging") are occurring (Capano, Howlett and Ramesh, Chapter 1 this volume). In such a complex setting, what kind of information can be extracted from reverse engineering the decision processes of individuals responding to policies and a variety of other constraints? And can lessons about first-order causal mechanisms in one national context be applied to other national contexts where cultural contexts, spillover effects

of other policies (e.g., tax treatment of non-retirement savings and mortgage interest deductibility) and other influences on individual behavior all differ? Certainly law-like statements of uniform cause-and-effect are likely to be less appropriate than probabilistic statements about a distribution of effects across individuals, given the impossibility of describing all possible combinations of policy activators, environmental conditions, actors and preferences over an extended period, or even the likelihood that any single policy will have a uniform effect on a heterogeneous intended or unintended target population (see Gerring, 2010 on the limitations of mechanistic analysis). The following analysis will address limitations suggested by this latter, probabilistic conception of causal mechanisms and reverse engineering across heterogeneous target populations, and especially across different national contexts.

3 FIRST-ORDER MECHANISMS IN RETIREMENT SAVINGS POLICY

Population aging has caused provision of adequate incomes for people of advanced age to become an increasing concern for governments in both wealthy and developing societies. As expanded public pension programs have reached their fiscal and political limits, many governments have sought to encourage or require individuals and households to save for retirement. Several pension reforms (notably those in Sweden and Germany) have explicitly promoted increased retirement savings as essential to maintain income in retirement as public pension replacement rates are cut. However, these policies need to address multiple barriers to retirement savings – barriers that differ substantially within the target population of working age residents that is heterogeneous on many characteristics, including income, information levels, work histories and peer effects. Increasing retirement savings, in short, requires addressing complex causal mechanisms that create barriers to retirement savings and that affect the efficacy of policies intended to increase those savings. Indeed, what constitutes “retirement savings” is not entirely clear. Many individuals accumulate assets over time without distinguishing between retirement savings and other forms of asset accumulation, and they may strongly resist government policies that place restrictions on those assets to ensure that they are available as retirement income streams, such as forbidding or penalizing withdrawals before a specific age, withdrawing large lump sums, and requirements for annuitization to provide a steady income flow over the course of an expected lifetime. Any effort to mandate “retirement savings” specifically is likely to encounter resistance if those policy proposals involve restrictions on what assets individuals and households can hold, how long they are held, how they are spent down, and so on.

A second problem is that retirement savings behavior is not a single behavior but several. We can distinguish a number of behaviors involved in retirement savings over the course of a working life, divided roughly into three stages: system entry, asset accumulation, and decumulation. Multiple barriers can discourage or diminish accumulation of retirement savings at each of these steps.

Unless retirement savings is a completely individual and ad hoc activity (as with most Individual Retirement Accounts in the United States, individual savers must enroll in some sort of retirement savings scheme sponsored by their employer, a financial institution (e.g., automatic transfers of funds from a checking account to a retirement savings account every month) or government. Once contributions have actually begun, several additional problems may arise. Individuals may, for example, have low earnings and little disposable income early in their work lives, but fail to increase their contributions when they have more disposable income. Other problems that may arise during the accumulation phase include investing in vehicles that do not provide reasonable risk/return trade-off or charge administrative fees that are excessive, failure to adjust investments over their working life, taking early withdrawals, and failure to re-enroll in a retirement savings plan if they change jobs. In the decumulation stage, individuals may take large lump-sum payments and fail to preserve adequate capital to provide a long-term retirement income stream, or fail to annuitize adequately.

Several barriers may affect retirement savings behavior, with many of the barriers operating across multiple stages. This chapter will use the categorization of barriers to behavior change developed by Weaver (2014, 2015), adapting it to focus on the underlying causal mechanisms at work in shaping retirement savings behavior. These barriers can be divided into three broad categories, which include a total of eight subcategories: barriers that affect the incentives for individuals to save for retirement (incentives, monitoring and enforcement), those that affect their perceptions of the desirability of and ease of saving for retirement (information and cognitive/decision-making biases, beliefs and attitudes, and peer and network effects), and those that affect their capacity to save for retirement (resources and autonomy). Each of these barriers suggest a particular set of causal mechanisms at work in the retirement savings decision. Of course, these are not air-tight categories; there is some overlap in the features of specific barriers to retirement savings.

3.1 Incentives

The incentives approach suggests that the core causal mechanism in retirement saving is individuals' rational calculation of individual, household or intergenerational self-interest (for a discussion of rationality, see Elster, 2007, Chapter 12). Individuals respond to public policies as well as other constraints and

opportunities in a way that maximizes their welfare. If the incentive structure offered by government policies is altered, individuals and households should adjust their behavior accordingly. Thus, policies that provide tax advantages for retirement savings often exempt earnings deposited in retirement savings accounts from taxation at the time they are earned and while they accumulated, with taxation only after retirement in a presumably lower tax bracket. Incentives that favor retirement savings over consumption and other forms of asset accumulation (e.g., incentives for home ownership) are likely to encourage retirement savings through those vehicles. Heavy penalties for early withdrawal from retirement savings have ambiguous effects: on the one hand, such policies make it less likely that retirement savings decisions, once made, will be reversed. On the other hand, they are likely to make it less likely that individuals will utilize retirement savings accounts in the first place, since it reduces their ability to deal with important income and health shocks relative to other potential investment vehicles.

Other features of program design may also affect the incentives for retirement savings. In Chile, for example, many citizens believe that income from their retirement savings accounts is unlikely to surpass the benefit available from the pension guarantee given to workers with more than 20 years of contributions to the retirement accounts system; they thus have an incentive to evade contributions entirely after reaching that years-of-contribution threshold, and to understate their contributions prior to that time.

3.2 Monitoring and Enforcement

Two further potential barriers to retirement savings suggest more complex incentive-focused causal mechanisms at work. As the Chilean example suggests, where government policies designed to mandate government policies are not consistent with individual preferences, failure to monitor and enforce that behavior is likely to lead to evasion of that policy. Successful evasion can be affected by characteristics of individuals – for example, the self-employed and those working in the informal labor market are most likely to be able to evade mandatory savings markets most readily (Ross, 2011, p. 195). The actions of these workers may be abetted or even required by employers as a condition of employment. Poor government capacity, such as weak data systems to track employment and deter informal employment, exacerbate problems of monitoring and enforcement.

3.3 Information, Cognition and Decision-making Biases

Decisions on retirement savings are also affected by the information that individuals have or lack that would help them to make more advantageous deci-

sions, as well as by cognitive and decision-making biases (for an overview, see Benartzi and Thaler, 2007; see also Tapia and Yermo, 2007), and by cultural attitudes and the behavior of people around them. As a recent New Zealand Financial Markets Authority (2016, p. 7) report noted, making financial judgments and decisions is difficult because those decisions “are complex, require consumers to assess risk and uncertainty, require making trade-offs between the present and the future, can be emotional, [and] are done infrequently so don’t provide opportunities to do them well.”

Evidence from a number of countries suggests that many individuals lack important information that would help them to make more advantageous decisions. Levels of financial literacy and engagement in retirement savings planning affects preparedness for retirement in a variety of ways. People who are more financially literate are more likely to avoid high cost debt, diversify investment risk, engage in retirement planning, and be aware of pension fund management fees (see, for example, Lusardi and Mitchell, 2011a).

A number of recent studies from many OECD countries suggest that there are major gaps in individuals’ financial literacy. These gaps are likely to be particularly large for the less educated, for women and for the young and old. Historically disadvantaged racial and religious groups (for example, African-Americans and Hispanics in the United States, Māori in New Zealand, Muslims in the Netherlands) are also likely to have lower levels of financial literacy (Lusardi, Mitchell and Curto, 2010; Alessie, van Rooij and Lusardi, 2011; Almenberg and Sävje-Söderbergh, 2011; Bucher-Koenen, 2011; Crossan, Feslier and Hurnard, 2011; Lusardi and Mitchell, 2011b).

Individuals also have cognitive biases in the way that they acquire, process and apply information and the way that they make decisions. They procrastinate in making plans and decisions, especially when the consequences of their decisions are far off in the future. They are myopic, valuing immediate over delayed gratification. They are loss-averse, weighing a loss from the status quo more than equivalent gain. They may satisfice in considering options rather than continuing to search for optimal outcomes. They may act on impulse with inadequate information, for example in making investment decisions. And they discount risks – for example, of extreme longevity.

Obviously, many of these cognitive and decision-making constraints apply in retirement savings behavior. Individuals do not know how long they will live. Most have little understanding of what level of retirement savings is needed to produce adequate income streams in retirement, understanding of how multiple retirement streams (e.g., from public pensions, employer plans, and personal savings) fit together, or an adequate understanding of how important early saving can be to building up an adequate “nest egg.” Acquiring adequate information in investing retirement savings is costly – in time, management fees, or both. Individuals may simply put off making retirement savings to a later date

– and never get around to it. Procrastination is exacerbated by most consumers' lack of understanding of the compounding of savings and investment early in their working lives, which in turn lowers the perceived costs of delaying savings for retirement (McKenzie and Liersch, 2011).

3.4 Attitudes and Beliefs

Broad cultural beliefs (as well as factors related to subgroups within nations) may shape retirement savings behavior in multiple ways, such as through the perceived importance of saving for retirement, sense of obligation to care for one's parents, and trust in pension providers (see, for example, Hershey, Henkens and van Dalen, 2007). More short-term attitudinal factors may also affect retirement savings behavior. Some authors have suggested that the Great Recession in the last decade may have led to "greater public mistrust of financial institutions in the private sector because of the large losses recently sustained and the unreliability of private investment vehicles (Ross, 2011, p. 194).

3.5 Peer and Network Effects

Closely related to cultural effects are what can be called peer and network effects (see, for example, Hedström and Swedberg, 1996). A substantial literature in social psychology suggests that individuals are affected both by what they perceive to be appropriate behavior (injunctive norms) but also by the behavior that they actually observe occurring (descriptive norms). "Network effects" involve *interaction* between different associated individuals as well as observation of their behavior: for example, in a randomized control trial of faculty in several university departments, Duflo and Saez (2003) found that attendance at retirement savings information fairs and subsequent enrollment in the tax deferred retirement account was higher in programs where some faculty members received individual invitations and incentives to attend the event, even for those faculty members who did not themselves receive the invitation and incentive.

3.6 Resources

Individuals' decisions on retirement savings may be influenced by factors that affect their capacity to engage in retirement savings, specifically their resources (defined here to exclude information resources, which were discussed separately above) and autonomy. Most obviously, individuals with very low earnings are unlikely to be able to save substantial sums for retirement in the absence of government policies that subsidize those accounts and restrict early withdrawals.

3.7 Autonomy

Individual autonomy on retirement income decisions refers to factors that limit the range of decisions that workers can make because it requires the cooperation of others. Most notably, policies that make retirement savings mandatory and withhold at source, as in Australia and Sweden, lower the ability of workers to avoid dedicated savings. These can be reinforced by policies that forbid withdrawals from accounts before an individual reaches retirement age to ensure that they are still available at retirement age. On the other hand, employees in the United States who wish to have retirement saving contributions withheld from their paychecks but whose employers do not participate in 401k or employer-sponsored IRA plans are not able to do so. Only a minority of employees of small (less than 100 employees) businesses in the United States have access to an employee-sponsored retirement savings plan (US General Accounting Office, 2013).

4 POLICY ACTIVATORS FOR INFLUENCING BEHAVIOR

As the discussion above suggests, governments can use a variety of policy activators to promote retirement savings, focused on different barriers and with varying degrees of coerciveness (see, for example, Howlett, 2011). Broadly speaking, these instruments can be divided into information, admonition, choice architecture, positive incentives, negative incentives, and requirements and sanctions (Weaver, 2015). Governments can also vary the intensity of the settings on instruments – for example, weak versus strong incentives, admonitions, and enforcement behavior. While potential policy instruments do not map one-to-one on specific steps and barriers, some instruments are more plausible candidates for changing behavior at particular steps than others.

4.1 Information

At the least coercive end of the spectrum, governments can attempt to address informational and cognitive barriers to retirement savings. For example, they can inform the public about longevity risks, and improve information about their likely income streams in retirement. Sweden, for example, has made strides in providing consolidated information about multiple income streams from both public employers and private sources (see, for example, Larsson, Paulsson and Sundén, 2011). Of course, such information is no protection against a public that is not receptive to that information or poorly prepared to make good use of it.

A variety of strategies have been pursued in recent years to try to improve financial literacy, with the Organisation for Economic Co-operation and Development taking a strong role in encouraging cross-national diffusion of effective practices. These campaigns have used a variety of institutional vehicles and venues, including schools, trade unions, pension agencies, pension fund associations (see, for example, Orton, 2007; Atkinson, 2008; Atkinson et al., 2012; Grifoni and Messy, 2012). Despite some success in efforts to increase financial literacy, and increased participation in defined contribution retirement schemes, engagement of citizens in retirement planning remains low in most OECD countries, especially among young workers and those with low levels of education (Lusardi et al., 2010). Experience of a negative income or wealth shock, on the other hand, increases engagement in retirement planning (Lusardi and Mitchell, 2011a).

Tailoring messages to specific groups of retirement savers can be facilitated by using internet-based information vehicles that are user-initiated and allow the user to provide substantial information about their personal financial situation and preferences, such as tolerance for risk. However, those who are most in need of information are probably not those who are most likely to use these vehicles, precisely because they require user initiative.

4.2 Admonition

Admonition instruments do not just provide target populations with information, they also provide direction on what the target population should do with that information – for example, start retirement savings early, adjusting their investment portfolios periodically, and avoiding contribution holidays where they are permitted. Government-provided admonition for retirement savings is particularly fraught, because (1) target populations are heterogeneous, and advice suitable for some might not be suitable for all, yet (2) given low levels of citizen engagement in retirement savings, messages need to be relatively simple if they are going to be heard and acted upon at all. But keeping individuals engaged in retirement planning is just as challenging as getting individuals engaged in the first place – and just as important, since portfolios that are appropriate for young workers may not be appropriate for older workers. Governments also need to think through more specific engagement strategies for particular sectors of the labor force who are least likely to do so on their own, and differentiated messages for different age groups.

4.3 Choice Architecture and Policy Defaults

Governments can also try to affect the choice architecture environment (see, for example, Thaler and Sunstein, 2008) that may influence retirement saving

decisions. Choice architecture narrowly conceived involves manipulating the options or relative visibility of options available to a target population without significantly affecting the payoffs to each option. For example, governments may allow – or even require – employers to offer their employees at the time of employment the option of increasing their retirement savings contribution rate in employer-sponsors plans automatically in future years unless they elect in the future to lower that rate. Such a policy can help to protect against status quo bias and procrastination for employees who are financially pressed at the beginning of their working careers.

One of the most powerful choice architecture tools involves setting the default – that is, what happens if the policy target takes no action. Most notably, research strongly supports the idea that when enrollment in an employment-based retirement savings scheme is the default rather than requiring employees to opt in dramatically increases enrollment rates in such schemes (forced choice, in which employees are required to either opt in or opt out, produces outcomes in the middle). It is not entirely clear whether this effect results from the fact that it is both cognitively and procedurally easier or because an opt-out is perceived as a recommendation from a source with superior expertise: the answer is almost certainly some of each.

Well-designed defaults can be useful to address additional challenges in retirement savings schemes, notably inertia- and procrastination-induced failure to increase savings rates as income increases. In the well-known Save More Tomorrow experiment, for example, Benartzi and Thaler (2004) showed that individuals who are offered the option of pre-committing to put part of future salary increases into higher retirement savings contributions (with the option of later reneging) agree to do so, and that most of them maintain that commitment through several rounds of salary increases.

Defaults can also be used to establish fund allocations for those who do not choose to make an active choice of funds, as well as to shift allocation of accumulated assets toward less volatile investments as retirement nears. Setting defaults in this area is likely to be more controversial, however, because there is no single undisputed criterion for how workers' funds should be invested. One criterion is that investors' principal should not be subject to high risk. In New Zealand's KiwiSaver, participants who do not make an active choice are assigned to one of several privately offered default funds; these default funds are required to follow conservative principles, investing only 15 to 25 percent of their assets in growth assets. So-called "life-cycle" or "generation funds" that gradually shift assets as their owners age without requiring actions by the asset owner are another approach: Sweden uses this approach in the default fund for its mandatory Premium Pension (see Weaver and Willén, 2014).

4.4 Positive Incentives

Government and employer incentives for retirement savings can take a number of different forms. Government incentives for retirement savings often take the form of preferential tax regimes – for example, exempting earnings contributed to retirement savings accounts from taxation at the time they are earned and returns on those accounts while they are in the accounts, taxing them only when they are withdrawn (commonly known as Exempt-Exempt-Taxed, or EET) after retirement, when the individual is presumably in a lower tax bracket than when he or she was working (for an overview of tax incentives, see Marriott and Mackenzie, 2010). But there are many variations on tax advantages for retirement savings. The size of the incentives also varies depending on whether and how incentives exist for other savings vehicles. One problem frequently noted with tax concessions for EET retirement savings plans is that they tend to be regressive: they give the highest benefits to high-income individuals who are able to save more for longer and enjoy higher benefits from lower marginal tax rates at retirement. There are often income limits on contributions to tax-advantaged retirement savings accounts to prevent upper-income earners from reaping disproportionate benefits from those schemes. Government subsidies for KiwiSaver accounts in New Zealand attempt to address the regressivity problem a different way: matching 50 percent subsidies are made on workers' KiwiSaver contributions of up to NZD1043 (benefits were higher prior to July 2011). Homebuyers are further incentivized to participate in KiwiSaver by the availability of grants to first-time home-buyers who have met minimum contribution requirements to KiwiSaver; they are also allowed to withdraw some of their KiwiSaver account for the first-time home purchase.

4.5 Negative Incentives

In addition to offering positive incentives to encourage (but not require) specific behaviors, governments can also put in place negative incentives to discourage (but not prohibit) behaviors. One area where negative incentives have been employed in retirement savings policy is in policies to discourage early withdrawals from accounts. In the United States, withdrawals from IRAs and 401k plans prior to age 59.5 are subject to a tax penalty, though individuals are allowed to borrow from those accounts without penalty.

Given that individuals tend to be more attentive to negative than to positive information, and to be more sensitive to perceived losses than to gains, policies that are framed as negative incentives may be more efficacious in changing behavior than an equal investment in positive incentives. They may also be harder to adopt and sustain, however, as affected interests resist loss imposition.

4.6 Requirements and Sanctions

Finally, governments can use requirements to influence individuals' retirement savings behavior. The most straightforward applications of requirements in savings schemes is the mandatory savings schemes implemented in Australia and Sweden, with penalties for those who try to evade the requirement. As noted above, monitoring mandatory retirement saving is likely to be easiest (and cheapest) to monitor and enforce when the income is withheld at source. In the absence of mandates, participation rates may decline substantially: in 2016, for example, 1.1 million of KiwiSaver's 2.6 million members were listed as non-contributors (New Zealand Financial Markets Authority, 2016, pp. 3, 13).

5 REVERSE ENGINEERING AND ITS LIMITS

Can "reverse engineering" – extracting information from a more rigorous understanding of interactive social and behavioral and policy processes – lead to improvements in policy activators to achieve the goals set by governments? The case of retirement savings reveals several constraints on these second-order causal mechanisms. One is the extraordinary complexity of retirement decision-making in heterogeneous populations, as well as the high frequency of "non-decisions": engaging in whatever behavior is privileged by defaults and by past patterns and habits of behavior by individual households. Also unclear in many cases is the degree to which the impact of specific first-order causal mechanisms (both policy and environmental) is likely to be uniform and generalizable across different national populations and national subpopulations (e.g., the young and the old, the wealthy and the poor, native speakers and immigrants) and stable over time.

The case of retirement savings policies also suggests that even if policy-makers have a good understanding of the causal mechanisms that pose barriers to achieving government objectives and policy instruments and settings that are likely to be efficacious in addressing those barriers, that knowledge may not succeed in getting those policies adopted or in sustaining them. Constraints on adoption and successful and sustained retirement savings policies can take several forms. Pressures from powerful interest groups, notably the financial services industry, may block policy initiatives that would harm their interests. Path dependence frequently leads to the development of concentrated interests that defend the policy status quo and view any initiatives that would make themselves worse off as both threatening and unjustified. These constraints can restrict policy choices intended to address any of the steps in retirement savings outlined earlier. Individuals (and ideological conservatives) may resist policies that mandate retirement savings as an infringement on freedom of choice. Policies that require financial services advisors to act as fiduciaries for

(that is, in the best interests of) retirement account investors rather than a lower “suitability” standard can help to ensure that savers are not saddled with high transaction and management fees that reduce their account balances. But fiduciary requirements are likely to be resisted by financial services companies that see a fiduciary rule as exposing them to lawsuits and reducing their profits. Policies that restrict or penalize early withdrawal and lump-sum withdrawals of retirement savings account balances are also likely to encounter substantial resistance from citizens, especially in countries like the UK and Australia where such practices have been widespread in private sector pension plans for many years. In policy areas that affect the public in very visible ways, existing policies may also create expectations of entitlement or protection from loss; thus, the creation of KiwiSaver defaults that were low risk made policy-makers reluctant to shift to a life-cycle approach for default funds that might expose them to significant short-term losses early in their working lives.

Target populations’ behavior can also be affected by numerous obstacles in the policy implementation process, such as coordination problems across implementing agencies, opposition from front-line workers, and inadequate budgetary and labor force resources of implementing agencies. Reaching groups who are not actively engaged in retirement savings, for example, can be facilitated by enlisting the help of institutions and civil society organizations with particular constituencies such as seniors, students, ethnic and linguistic minorities. Efforts to do so, however, encounter three major obstacles. First, many organizations have well defined “organizational missions” (e.g., provision of specific kinds of social services to their members) and may be reluctant to undertake new tasks unless they are convinced that it helps them to meet that mission (see Wilson, 1991 and McDonald, 2007). Second, they are likely to perceive themselves as lacking in the expertise needed to provide this service, and believe that it will be costly for them to acquire it in an environment that is almost always resource scarce. Third, they may not see a demand from their members, and believe that providing the service may blur the image that they have cultivated with their members. In short, working with civil society organizations, and even government bureaucracies like schools, can be an important vehicle for improving financial literacy. But it is likely to be a complex and long-term process in which the needs, priorities and resource constraints of those organizations need to be fully taken into account.

6 CONCLUSIONS

This chapter has argued that the causal mechanisms affecting individuals’ retirement savings decisions are extremely complex. They are also likely to be very heterogeneous across individuals even within a specific country’s policy

regime due especially to individual differences in characteristics such as information, income and psychological orientation toward the future or the present.

The research findings and the analytical framework presented here have broader implications for understanding the limitations on using reverse engineering in choosing policy instruments and the settings on those instruments. In retirement savings policy as in most complex policy sectors, governments face very complex trade-offs regarding coverage, adequacy, intrusiveness and equity, among other values, in deciding whether to use more or less intrusive instruments and settings to influence behavior. They need to consider how strong the barriers to behavioral change are for all components of their very diverse citizenry, not just the median citizen, and they need to ascertain and address the distributional consequences of the causal mechanisms on both the behavior of the intended and unintended target populations and their post-intervention welfare – which may extend over a long period of time. They also need to keep in mind that some elements of the population may face very strong barriers to retirement savings, notably resource barriers. If universal coverage for retirement savings schemes are deemed to be important, relatively non-coercive instruments and settings (e.g., providing information, manipulating policy defaults and even providing financial incentives) are likely to be insufficient (Antolin, Payet and Yermo, 2012). The level of uncertainty faced by future retirees exacerbates constraints on retirement savings. They do not know how long they will live, how healthy they will be, what the future return on their retirement savings will be, or how much their home will be worth in the future if they are current homeowners and plan to sell to meet part of their retirement income needs. Given this uncertainty, there is likely to be substantial variation across individuals in the choices that they make on how much to save for retirement and in their choice of investment vehicles. Some individuals will be excessively cautious, others will take risks that they should not take. Some will acquire lots of information, others will not. And these responses are unlikely to be randomly distributed. Those with higher incomes and assets and higher educational attainment are likely to face both easier choices (it is easier to save at a higher rate when your income is higher) and likely to have better information about the consequences of those choices. In this situation, governments will need to consider whether it is more appropriate to give individuals more choices, or to protect individuals against the consequences of poorly informed and highly constrained choices. But governments' ability to process information about these extraordinarily complex first-order mechanisms is generally very limited. At best, highly simplified models of these mechanisms are likely to be utilized, and groups that have privileged access to policy-makers and resources may bias which elements of the complex mechanisms are included.

“Reverse engineering” of social processes can help to illuminate the complex causal mechanisms at work – including the roles played by public policies – but it cannot by itself lead to “optimal” choices of instruments and settings given the underlying policy trade-offs and political conflicts in policy design. More generally, this chapter suggests that a “mechanisms” approach to understanding governmental efforts to achieve “appropriate” behavioral responses from target populations needs to confront a number of problems. First, behavioral responses are often not a single behavior but many, distinctive behaviors, carried out over a long period of time. An overly simplistic definition of the behavior that is being sought while ignoring other behaviors may result in failure to achieve the policy outcomes that were originally being sought, or in the production of negative spillover effects. Second, consideration of policy mechanisms must avoid overly simplistic causal stories – for example, focusing on a single barrier to achieving the desired behavior(s) when multiple barriers are in play. Third, as suggested earlier, a mechanisms analysis should look at how causal mechanisms (including barriers to behavior adaption) may operate differently for different segments of a target population.

REFERENCES

- Alessie, R., M. van Rooij and A. Lusardi (2011), “Financial literacy and retirement preparation in the Netherlands,” *Journal of Pension Economics & Finance*, 10(4), 527–45.
- Almenberg, J. and J. Säve-Söderbergh (2011), “Financial literacy and retirement planning in Sweden,” *Journal of Pension Economics & Finance*, 10(4), 585–98.
- Antolin, P., S. Payet and J. Yermo (2012), “Coverage of private pension systems: Evidence and policy,” *OECD Working Papers on Finance, Insurance and Private Pensions*, No. 20.
- Atkinson, A. (2008), *Evidence of Impact: An Overview of Financial Education Evaluations*, Consumer Research Report No. 68, London: Financial Services Authority.
- Atkinson, A., D. Harrison, F.-A. Messy and J. Yermo (2012), “Lessons from national pensions communication campaigns,” *OECD Working Papers on Finance, Insurance and Private Pensions*, No. 18.
- Beach, D. and R.B. Pedersen (2013), *Process-tracing Methods: Foundations and Guidelines*, Ann Arbor, MI: University of Michigan Press.
- Benartzi, S. and R.H. Thaler (2004), “Save More Tomorrow™: Using behavioral economics to increase employee saving,” *Journal of Political Economy*, 112, S164–S187.
- Benartzi, S. and R.H. Thaler (2007), “Heuristics and biases in retirement savings behavior,” *The Journal of Economic Perspectives*, 21(3), 81–104.

- Bucher-Koenen, T. (2011), "Financial literacy, Riester pensions, and other private old age provision in Germany," *MEA Discussion Paper 250-2011*, Munich Center for the Economics of Aging at the Max-Planck-Institute for Social Law and Social Policy.
- Crossan, D., D. Feslier and R. Hurnard (2011), "Financial literacy and retirement planning in New Zealand," *Journal of Pension Economics & Finance*, **10**(4), 619–35.
- Curtis, S.K., S.P. Harston and C.A. Mattson (2011), "The fundamentals of barriers to reverse engineering and their implementation into mechanical components," *Research in Engineering Design*, **22**(4), 245–61.
- Duflo, E. and E. Saez (2003), "The role of information and social interactions in retirement plan decisions: Evidence from a randomized experiment," *Quarterly Journal of Economics*, **118**(3), 815–42.
- Elmore, R. (1979), "Backward mapping: Implementation research and policy decisions," *Political Science Quarterly*, **94**(4), 601–16.
- Elster, J. (2007), *Explaining Social Behavior: More Nuts and Bolts for the Social Sciences*, 2nd revised edition, Cambridge, UK: Cambridge University Press.
- Fitzpatrick, W. and S.A. DiLullo (2006), "Attack of the clones: Reverse engineering, R&D and the law," *Competition Forum*, **4**(2), 501–14.
- Gerring, J. (2010), "Causal mechanisms: Yes, but. . ." *Comparative Political Studies*, **43**(11), 1499–526.
- Gorman, G.S. (1998), "The Tu-4: The travails of technology transfer by imitation," *Air Power History*, **45**(1), 16–27.
- Grifoni, A. and F.-A. Messy (2012), "Current status of national strategies for financial education: A comparative analysis and relevant practices," *OECD Working Papers on Finance, Insurance and Private Pensions*, No. 16.
- Hedström, P. and R. Swedberg (1996), "Social mechanisms," *Acta Sociologica*, **39**(3), 281–308.
- Hedström, P. and P. Ylikoski (2010), "Causal mechanisms in the social sciences," *Annual Review of Sociology*, **36**(1), 49–67.
- Hershey, D.A., K. Henkens and H.P. van Dalen (2007), "Mapping the minds of retirement planners: A cross-cultural perspective," *Journal of Cross-Cultural Psychology*, **38**(3), 361–82.
- Howlett, M. (2011), *Designing Public Policies: Principles and Instruments*, London and New York: Routledge.
- King, G., J. Pan and M. Roberts (2014), "Reverse-engineering censorship in China: Randomized experimentation and participant observation," *Science*, **345**(6199), 891.
- Larsson, P., A. Paulsson and A. Sundén (2011), "Customer-oriented services and information," in N. Takayama (ed.), *Priority Challenges in Pension Administration*, Tokyo: Maruzen, pp. 135–52.

- Lusardi, A. and O.S. Mitchell (2011a), "Financial literacy around the world: An overview," *Journal of Pension Economics & Finance*, **10**(4), 497–508.
- Lusardi, A. and O.S. Mitchell (2011b), "Financial literacy and retirement planning in the United States," *Journal of Pension Economics & Finance*, **10**(4), 509–25.
- Lusardi, A., O.S. Mitchell and V. Curto (2010), "Financial literacy among the young," *Journal of Consumer Affairs*, **44**(2), 358–80.
- Marriott, L. and G. Mackenzie (2010), "New Zealand retirement saving and taxation: Lessons from Australia," *Australian Tax Forum*, **25**, 3–27.
- Maskin, E.S. (2008), "Mechanism design: How to implement social goals," *The American Economic Review*, **98**(3), 567–76.
- McDonald, R.E. (2007), "An investigation of innovation in nonprofit organizations: The role of organizational mission," *Nonprofit and Voluntary Sector Quarterly*, **36**(2), 256–81.
- McKenzie, C.R.M. and M.J. Liersch (2011), "Misunderstanding savings growth: Implications for retirement savings behavior," *Journal of Marketing Research*, **48**, S1–S13.
- New Zealand Financial Markets Authority (2016), *KiwiSaver Annual Report 1 July 2015–30 June 2016*, accessed 15 October 2018 at <https://fma.govt.nz/assets/Reports/161004-FMA-KiwiSaver-Report-2016.pdf>.
- Orton, L. (2007), *Financial Literacy: Lessons from International Experience*, Ottawa: Canadian Policy Research Networks.
- Ross, S.G. (2011), "Recent developments in the collection of social contributions," in N. Takayama (ed.), *Priority Challenges in Pension Administration*, Tokyo: Maruzen, pp. 193–208.
- Tapia, W. and J. Yermo (2007), "Implications of behavioural economics for mandatory individual account pension systems," *OECD Working Paper on Insurance and Pensions*, No. 11, July 1.
- Thaler, R. and C. Sunstein (2008), *Nudge: Improving Decisions About Health, Wealth, and Happiness*, New Haven, CT: Yale University Press.
- US General Accounting Office (2013), *Retirement Security: Challenges and Prospects for Employees of Small Businesses*, Report GAO-13-748T.
- Weaver, R.K. (2014), "Compliance regimes and barriers to behavioral change," *Governance*, **27**(2), 243–65.
- Weaver, R.K. (2015), "Getting people to behave: Research lessons for policy-makers," *Public Administration Review*, **75**(6), 806–16.
- Weaver, R.K. and A. Willén (2010), "Lessons from the Swedish pension system after 15 years," *OECD Journal on Budgeting*, **13**, 1–26.
- Wilson, J.Q. (1991), *Bureaucracy: What Government Agencies Do and Why They Do It*, New York: Basic Books.