Intergenerational Contracts and Time Consistency: Implications for Policy Settings and Governance in the Social Welfare System

Lewis Evans and Neil Quigley

Victoria University of Wellington

Acknowledgements

The authors acknowledge helpful comments from members of the The New Zealand Treasury Long-Term Fiscal Planning Group, and referees, Arthur Grimes, Kirsten Jensen, Becky Prebble, and Graham Scott

Abstract

We suggest that public policies implementing Intergenerational contracts can be sustainable if they are interpreted as relational contracts, but will not be sustainable if the historical structure of entitlements created by those policies is regarded as inviolable. Comprehensive public social welfare schemes, whether fully funded or pay-as-you-go, may prove to be unsustainable across generations because of changing technology, demographic structure and openness of the world economy; and the incentives for politically opportunistic contribution holidays and benefit levels associated with full funding. Additionally, in the case of partial funding, there are incentive problems and inefficiency of middle class churn associated with taxation. We do not propose specific policy changes, but we suggest that New Zealand should be exploring whether the current pay-as-you-go social welfare system could be gradually replaced with a mixed system that addresses the goals of past approaches to social welfare rather than locking in current approaches. The new approach might include mandatory investment-based personal accounts associated with at least some of the contingencies that are currently covered by the social welfare.

JEL CLASSIFICATION I38, H23, H55

KEYWORDS intergeneration, time consistency, relational contract, insurance, social welfare

Table of Contents

Summary

Time consistent policies are important for socially desirable performance of the private and public sectors. This is because they provide stability that enables individuals and the state to plan for the future. We argue that time consistency is achievable if intergenerational arrangements between the state and the populace are treated as relational contracts. A relational contract is quite different from a legal contract since the latter typically impose specific constraints on arrangements. A relational contract does not define specific constraints but rather a process for developing and changing rules by which all parties agree to abide. The few institutional checks and balances of New Zealand administrative governance, requires particularly that the rules or policies balance the interests of groups in the population. The issue is present for all policies but is particularly important for long-lived policies and issues such as state-provided insurance and social welfare. For these, beneficiaries typically differ from the funders across time, and (for government schemes) across generations. This means that time consistent actions are important for the survival of the relational contract and the economic performance of the schemes themselves.

Time consistent policies admit change that is in accord with the purpose of the policy and its place in the relational contract. The aging and wider diversity of the New Zealand population, more open international labour markets, and the digital revolution are all changes that might predicate, even demand, changes in policies that do not violate time consistency. Public policies that enhance social welfare will be acceptable to the populace if they are interpreted as relational contracts, but will not be sustainable if the historical structure of entitlements created by those policies is regarded as inviolable. The literature suggests that rules brought down to render government administrations subject to particular, usually financial, constraints, are not time consistent because the inter-temporal position of the economy can be represented in various ways under various measures.

Key inter-temporal time-consistency issues arise in public social welfare schemes, whether fully funded or pay-as-you-go. The parameter settings of these policies may prove to be unsustainable across generations for the aforementioned reasons, and interact with incentives for politically opportunistic contribution holidays and benefit levels associated with full funding that may render them time inconsistent. Additionally, in the case of partial funding, there are incentive problems and the inefficiency of middle class churn associated with taxation. We suggest that the current pay-as-you-go social welfare system could be gradually replaced with a mixed system that addresses the goals of past approaches to social welfare rather than locking in current approaches. The new approach would include mandatory investment-based personal accounts associated with at least some of the contingencies that are currently covered by the social welfare. Given recognition of property rights of individuals in such schemes, they will provide more policy stability and more consistency over time than is inherent in present schemes. Tracking individual contributions and entitlements in these schemes is more feasible and cost effective in the modern digital economy than it was at the time that these schemes were first developed. The more sustainable structure is likely to include:

- greater use of social insurance (compulsory private contributions to schemes designed to provide for particular contingencies)
- greater use of deductibles or co-payments
- greater use of earmarked taxation
- greater use of private accounts, either fully funded or notionally funded, given the political unsustainability of fully-funded public schemes

- greater separation of redistributive welfare payments from social insurance provisions, and
- a mix of contingency-related and means-tested benefits, with the former based as closely as possible on individual contributions, and the latter based as closely as possible on actual need (income below a certain threshold).

Each of these areas of policy development will of course need to be the subject of more detailed consideration and application of the framework than we have provided before explicit policy recommendations could be provided.

1 Introduction

Social Welfare systems have a long history, but the scope and the scale of the benefits provided by governments increased rapidly after WWII. The cost of welfare systems accounts for a large part of the increase in government spending as a proportion of GDP that has occurred in most industrial democracies over the period from 1950 to the present.

Becker and Murphy (1988) attribute the existence of social welfare schemes to an intergenerational contract between the old and the young. Specifically, parents provide investments in the human capital of their children and then receive a return on this investment in the form of social security benefits when the children are working and the parents are retired. Because children cannot be parties to a legally enforceable contract, the government needs to provide a mechanism for these transfers to occur.[1]

However, the sustainability of that intergeneration contract has been called into question by a wide range of studies that have focused on the long-term costs of current social welfare schemes given expected changes in the working age population and productivity (Kotlikoff (2008); Feldstein (2005a)). The problem with long-term sustainability has been ascribed to the "selfish generation" of post-war workers who (partially in compensation for the sacrifices made during WWII) used their power in the polling booth to have governments introduce levels of social welfare benefits that are unsustainable (Thompson 1996). Alternatively, the problem has been ascribed to changes in population structure and longevity that could not have been foreseen by the post-WWII generation. The population structure of all of the advanced economies is aging as a result of long-term declines in fertility rates and improvements in mortality rates. In New Zealand, fertility rates remain higher than in many advanced economies, but emigration of young and mid-career adults helps to offset this effect, so these factors in combination with longer life expectancy means that our population is also aging.

The performance of the intergenerational contract between the young and the old is also affected by the changing mobility of individuals that is part of the globalization of world economies. In particular, the contract is difficult to enforce on those from the younger generation who emigrate from New Zealand after having consumed large state subsidies to health care and education. This is illustrated in the Demange, Fenge and Uebelmesser (2013) discussion of the optimal subsidy/fee structure for higher education, where they conclude that closed economy settings will differ from open economy settings and that the mobility of students, as well as skilled workers matters for the existence of the optimal price and subsidy structure. Intergenerational contracts between the young and the old play a lesser role now than do voters' perceptions of the benefits of skilled labour to the economy (Borck, Uebelmesser and Wimbersky (2012)).

[1] Other explanations for the role of the state in social security have been advanced (see for example Laslett (1992)), including the inability of private insurers to deal with adverse selection, but this is less satisfactory because adverse selection can be dealt with by the state requiring the purchase of insurance and does not require that the state actually provide the insurance.

In this paper we explore the question of the sustainability of the intergenerational contract that is represented by the current structure of social welfare. We argue that sustainability and time consistency of social welfare policies could be improved by more explicit recognition of the social welfare system as a relational contract that should be reinterpreted in the light of changes in technology, changes in our understanding of the incentive effects of different approaches to social welfare provision, and changes in society as a whole. We suggest that too much of our social welfare policy is based on approaches developed under the social and economic conditions and technology of the past, and that this is a key source of the potential challenge to the sustainability of current policies.

The analysis provided in this paper suggests to us that sustainability of the current social welfare system will be enhanced by changing key elements of the current model of funding a wide range of social welfare provisions from general taxes on a pay-as-you-go basis. In particular, taking advantage of the low costs that modern technology now provides, we suggest that a move to a "dual system" with underlying needs-based (rather than universal) welfare support together with more widespread use of individual contributory welfare accounts would

- increase transparency and improve governance over welfare issues,
- increase savings and reduce the requirement for future tax increases, and
- dramatically improve incentives associated with our taxation and social welfare system.

A move in this direction is consistent with retention of some base level of benefits, and top-ups to the individual accounts of those below a minimum income threshold, funded from general taxation.

2 Time Consistency in Social Welfare Policies

2.1 Importance of Time Consistency

As will become clear in our discussion of the relational contract between individual members of society and the government we take a broad definition of time consistency.[2] Time consistent policy arises where policy is sustained over time; even during limited periods where the policy is not optimal. An inconsistent policy is one where a policy rule is changed in circumstances that are predictable at the time of policy formation. Time consistency requires incentives to induce maintenance of a policy over time.[3] These may arise with respect to expected future-period circumstances or contemporaneously because of the state of other policies.

The Greek mythology of Ulysses and the Sirens illustrates the issue: Sirens were very alluring singing creatures that lured ships onto nearby rocks. Ulysses had his crew bind him to the mast before arriving in earshot of the Sirens and thereby commit to a policy of ignoring the Sirens. Commitments for a government arise where policy is made more credible by integration into the world economy and by domestic institutions – e.g. an independent central bank - that provide bounds on, particularly short term, policies.

Time-consistent policies are important for the performance of the economy because they provide an environment for planning by investors in the public and private sectors. Time consistent policies are those that will not be revised except in ways that are consistent with the policy framework that was originally developed. The importance of time-consistent policy in social welfare provision is increased by three factors. First, social welfare provision (including in health, old age pension, unemployment, accident and disability support) represents a very large part of GDP. Second, state provision of social welfare has substantial impacts on private choices, particularly relating to employment, saving and retirement, so that radical changes in social welfare provision could create significant problems for sections of the population that had relied on the sustainability of intergenerational welfare policy. Third, social welfare transfers offer the opportunity for policy variation according to interest group pressures.[4] Further the design and governance structures of social welfare policies have an important role to play in determining their long-term sustainability.

Time consistency touches very many policies relevant to individual and household decisions. Unless a policy is viewed as being in accord with time-consistent principles it will not be robust to opportunistic behavior, that in itself has the potential to undermine the sustainability of policy settings. Opportunistic behaviour refers here to agents in the economy responding to incentives directly relevant to their welfare that they see before them in the present and the relevant future as it unfolds. These incentives will discount time inconsistent policies and thereby produce the self-fulfilling result that the policy is ineffective or does not enhance the economy's performance.

---

[2] Kydland and Prescott (1977) argue that time consistency is required for optimal policies and that it can only be achieved by policy that sets rules: not outcomes.

[3] Time consistency policy issues arise in regulating a natural monopoly. If the regulator announces a regulatory framework that the firm knows is not in the regulator's best interest once investment by the firm is in place, the announced policy is time inconsistent. Because the firm knows it is inconsistent it will not invest on the basis of the announced policy and regulation may well not produce desirable outcomes (for an example, see Evans and Guthrie (2012)).

[4] For discussion of characteristics of equilibria resulting from interest group pressures see, MaGee, Brock and Young (1989).

The time consistency of government spending and budgets has received increased attention in economic and political literature from the time that a single currency for the European Union was mooted, and this interest has not waned under the global financial crisis. A key problem with discipline in the government fiscal position and with rules about fiscal discipline is the wide flexibility that governments have to structure combinations of debt and other obligations to achieve identical burdens on the taxpayer but with markedly different impacts on the government accounts and the deficit in particular (Green and Kotlikoff 2007).[5] It means that governments may have, and may even create, asymmetric information in the short run. For example, the government may have an incentive to reveal less than full information about the fiscal costs of policies when it expects that full information will only become available to the population as a whole in during the tenure of some future administration.[6] The impact of time consistent political discipline is relevant to the incentives provided by different approaches to funding government activity.[7]

Time consistent fiscal management of the economy will depend upon such factors; as the political and constitutionally-based constraints on government fiscal positions, transparency, ability to forecast economic outcomes, asymmetrically costly forecast errors, and political consensus about the extent and shape of government. Frankel (2011) argues that forecasting processors matter, and that for many countries officials' forecasts of economic growth have been excessive with the effect that they underwritten excessive government spending, on average but also particularly in times of economic upswing. He gives an example where the forward looking fiscal growth parameters are assigned to a body that is independent of the public service and government, and which has improved fiscal settings.

2.2 New Zealand's Constitutional Framework: Implications for Policy Consistency

The constitutional framework is central to the achievement of time-consistent public policy (Weingast (1989)). New Zealand's policy framework is dominated by institutional and political forces internal to the economy which are not much modified by external institutions.[8] This differs from that of many countries, including the UK from whence New Zealand derives its governance system. The supremacy of New Zealand parliament in law making is distinctive; for this parliament has no checks on policies it transforms into law, nor their implementation, other than those it has chosen to fix by means of (ratified) treaties with other countries and external international bodies.

---

5 As an example, consider the situation in which the government raises debt to fully fund the current value of some welfare obligations. A year later the government switches to a pay as you go basis for that scheme, and repays the debt. The appearance is that the government's debt has been reduced, but the burden on the taxpayer has not changed in aggregate though it has (most likely) changed in intergenerational terms.

6 This mis-information of political opportunism is not considered further in this paper given our focus on the long-term sustainability of institutional arrangements.

7 Ricardian equivalence – that the balance of taxation and government debt does not affect real economic activity, instead activity is affected by the incentives of taxation and other regulations – remains controversial, but the issue is important for time consistent policies. According to Seater (1993, 184) "Empirical success and analytical simplicity make Ricardian equivalence an attractive model of government debt's effects on economic activity". However, Ricardian equivalence, particularly in its pure form, is not widely accepted. An example, Kotlikoff's (2001: sVII) review of Ricardian equivalence, cites theoretical and empirical rejections of necessary conditions for Barro's (1974) finding of Ricardian equivalence to hold. The controversy continues to get theoretical (Werning, (2007)) and empirical (Banzhaf and Oates (2012)) attention.

8 Though some agreements with other countries provide some limitations on policy, including the CER agreement with Australia, other bilateral trade agreements, and international conventions and law.

Most other countries are constrained to some extent by some combination of bicameral systems where an upper house elected on a different basis from that of the lower house (parliament), by federal systems where powerful state/provincial governments constrain the ability of the federal government to make time-inconsistent policy decisions, and external judicial authority. Where there are such checks and balances, policy changes – beneficial or otherwise – are generally subject to more considered scrutiny and are harder to implement because there is no guarantee of the alignment of the political preferences of the different institutions that make up the political structure.[9] New Zealand's position means that its policies are potentially more volatile over time, and that other countries' specific institutions and policies will often not be useful comparators; although political, public service, and private behaviours observed under different institutional settings may be informative. This topic is important (see Persson, Roland and Tabellini, 2007) but will not be considered further in this paper as we take the constitutional structure as given.

2.3 Population Aging and the Time Consistency of the Current Social Welfare System

We have argued that time consistency is necessary for sustainable intergenerational contracts. In this and the following section we consider this relationship. Governments provide insurance and social welfare encompassing response to a wide range of events that include retirement, unemployment, ill-health, disability and loss of income due to accidental and other causes, and the costs of investing in the upbringing of children. These events combine the features of uncertain outcomes and transfers among citizens.

Government provision opens up a wider array of possibilities than the private sector can provide.[10] Government schemes reflect the benefit of statutory specification and all feature mandatory requirements that include compulsory contributions and, (typically) defined benefits that are unrelated to contributions. Government schemes may be of the defined contribution type where individuals are assigned their own accounts into which a government-regulated proportion of income is paid (see Merton 1983) with the purpose of accumulating funds and providing payouts when particular events (e.g. retirement) occur. The fund together with its expected return between contribution and pay out meets the expected obligations of the scheme at all relevant dates into the future. The alternative is a scheme that is not fully funded but where payouts are covered by current contributions: hence the term pay-as-you-go, or pay-go.[11] Compulsion is required for pay-go schemes because individuals are making predictable payouts to other commonly unrelated individuals. Even with risk pooling (as occurs, for example, in the Merton scheme) a fully funded scheme does not have this feature because risk pooling addresses only the incidence of events that are predictable only in the population as a whole and not at the level of the individual.

---

[9] New say this not withstanding that the New Zealand Parliament has recognised the importance of transparency and consistency of government decision-making and enacted internal constraints on the administration of government, such as that provided by Part 2 of the Public Finance Act 1989 and its amendments.

[10] Merton (1983), for example, argues that even if policy design focuses on retirement and ignores other uncertainties and redistribution; government provision of a basic minimum retirement income is suggested by externalities such as economies of scale in managing funds and information acquisition.

[11] Under pay-as-you-go schemes there is a rate return in that for the same the defined benefit of having supported – by payments – earlier generations will grow with the population growth rate when the age cohorts of the population are stable (Samuelson (1958)).

Most of the world's social security systems are financed on a pay-go basis, with few or no assets accumulated to meet future liabilities, and liabilities funded as they arise by taxes. When a pay-go system is first introduced, or when a new set of benefit enhancements is introduced, the generations that are already retired, or are near retirement, receive a windfall gain because they have not paid taxes to fund the benefits that they will receive. Put another way, the return they receive on the taxes that they did pay is very high. But all future generations will pay taxes before receiving benefits, and the more taxes they pay, the lower will be the return that they receive on those taxes. Under a mature pay-go system in which everyone pays taxes over their entire working life and there is a stable distribution of ages in the population, the rate of return on the taxes paid drops to the long-term sustainable rate, which is (approximately) equal to the rate of productivity growth plus the rate of population growth (Samuelson 1958).

When a mature pay-go social welfare system is combined with an aging population, implicit rates of return have to fall to reflect the decline in the rate of population growth (Weaver 1994). This will impact on payments for pensions, but it will also impact on payments for health care where the impact is self-reinforcing: living longer means that people consume more health care, and the higher level of health care helps them to live longer.

Although the scope for reducing the substantial deadweight losses of the pure pay-as-you-go system and for increasing the present value of all future consumption should provide a strong incentive for a change in policy, they are often not the reason that has driven the political process in countries to move from pure pay-as-you-go to a mixed system or to consider such a change. In some the primary driving force is the recognition that the increasing age of the population will require a very large tax increase or benefit cut if nothing is done to change the existing system. This is not a temporary effect of the baby boom generation reaching retirement age but a permanent result of the trend to increased longevity and declining birth rate of many countries. This demographic change is significant not only because it drives the political process but also because it increases the potential social gain of making a policy change.[12]

2.4 Problems With Pay-As-You-Go Funding

It is well recognized that pay-as-you-go funding results in a reduction in saving and in the present value of consumption, and so produces a lower current and future growth path for the economy than a system in which all provision for the future was via investment today. This reduces the sustainability of pay-as-you-go funding by making the tax burden required to fund it less affordable.

[12] McHale (2001) considers benefit rule changes in the G7 countries to gauge the political risk of pay-as-you-go systems. He finds that projections of rising costs presaged change in benefit systems: typically in a way that protects the existing older generations at the expense of lower prospective benefit of the schemes for younger generations. This balance may well be affected by stronger, relative, voting power of the older generations as the population ages. McHale suggests that such scheme changes may be credible because the younger generations fear future additional cuts in benefits if the taxes on future working cohorts are relatively high.

In addition, Feldstein (2005a) has pointed out two additional problems with pay-as-you-go-funding:

(i) Distortion of labour supply and of the form in which compensation is paid, because of the increase in the marginal tax rate associated with pay-as-you-go funding of social welfare. The relevant marginal tax rate is the statutory rate net of the anticipated increase in the actuarial value of benefits. By comparison, in a system of compulsory private social welfare savings accounts the actuarial present value of an individual's benefits would be equal to the social welfare saving that they carry out. The only labor supply distortion would result if the mandatory nature of the scheme required some individuals to do more saving for retirement than they preferred.

(ii) The incentive to retire resulting from a mandatory pay-as-you-go public pension paid on retirement at a certain standard retirement age depends on the details of the scheme. The New Zealand public pension scheme has the free option of continuing to work beyond the age of eligibility and thus does not affect the proclivity of individuals to work excepting via the income effect of its provision of pension income, but it cannot be taken as a lump sum or form part of an estate. The costs of the New Zealand scheme are thus increased by payments to those who continue to work (have not retired) specifically, and payments that may start at an earlier age than is justified by current standards of living and life expectancy at age 65. Given New Zealand's free option and aging population the existing scheme, while not materially inhibiting working beyond the age of entitlement, will require increasing tax rates to maintain its existing benefit levels.

Social preferences and opportunity costs, affect generational allocation decisions (Feldstein (2005b)) and hence fiscal stance. Dobrescue, Kotlikoff, and Motta (2008) report that while savings rates vary enormously across countries there was a trend, at least to 2008 across developed countries, of social preferences increasingly favouring present rather than future generations. It draws on Green and Kotlikoff (2007) to propose a model of rational economic agents that do not distinguish among public and private activity and thereby ignore any distinction between public and private property rights. It reconciles public and private choice by means of a social welfare function that has uncertain future time preferences. Choices are made in an inter-temporal equilibrium model embodying uncertainty in two variants: in one the same decision-makers are present throughout and in the second, future and current decision makers differ allowing there to be time inconsistency. The models are estimated for the U.S., France and Italy and in all six variants it is found that social time preferences increase over time, thereby indicating an increasing preference for early gratification by society.

3 Intergenerational Social Welfare as a Relational Contract

The welfare state is sometimes said to exist on the basis of a "social contract" wherein policies are accepted by individuals on the understanding that their treatment under the contract that is the welfare state will in fact be implemented; and thus they accept the state's treatment of others, including transfer payments. To take a narrow example, consider a social contract that an individual is guaranteed government provision of health care at no financial user cost no matter whether the demand for health care coincides in time or in quantum with that individual's financial contribution to the health system. Under this contract elderly citizens most demand services from the system at the time they are paying little or no tax, and have short life-time horizons. Time consistency is critically important for the existence of the contract: if there is prospect of health care requiring a significant fee contribution by retired persons, working persons with low health care demand will reduce their support for the government health scheme and thereby place this social health contract at risk. These types of social contracts therefore embody arrangements for temporal choices and outcomes and to include intergenerational contracts.

In our view, intergenerational contracts are akin to a relational contract (see Baker, Gibbons and Murphy (2002) and Goldberg (1998)) in which the institution of parliament responds to the demands of its constituents in making provisions that the constituents find generally acceptable in the context of the commonly held expectations, and the particular circumstances of the time. Relational contracts anticipate a long-term relationship between the parties, but create options for periodic renegotiation of the terms of the relationship. They are designed to allow parties to a contract to benefit from building new information into the contract at discrete intervals. It is the characteristic of relational contracts that they depend for their viability on ongoing repeated interaction, high-level goals understood at a level of principle, and flexibility to vary arrangements. They are orthogonal to a long-term legal contract in that it is their purpose to provide options for change (Goldberg (1980)), whereas legal contracts seek to limit change.

Relational contracts arise from four fundamental aspects of the environment in which many contracts are negotiated:

1 People are not omniscient, and information is costly to acquire, so it is difficult to write binding contracts about the future

2 Opportunism is present in contracting, that is, parties to the contract will act in their own best interests consistent with the terms of the contract.

3 Aspects of the contract as well as changes in the state of the world may create unforseen opportunities to act opportunistically which the parties will want to control in the long-term, and

4 It will be less efficient for outsiders (judges, arbitrators, regulators) to be required to enforce the contract than to allow relational negotiation between the parties.

Following early attempts to establish the basis for relational contracting such as Goldberg (1980; 1998) and Scott (1990) the relational contracts literature has explored the terms under which it is optimal not to explicitly specify the terms of a contract. The focus of the literature has been on explaining why parties would want to specify renegotiation at specific times or in the event of specific contingencies rather than setting explicit terms in a long-term contract. This literature suggests that parties prefer to renegotiate because:

1 Changes in the state of the world may move the details of past contracts a long way away from the contemporary efficient position.

2 If the past contracts are out of line with the efficient arrangements under contemporary circumstances, one party will have an incentive to act in ways which, while technically not a breach, will reduce the value of the relationship to both parties.

3 Parties may not have outside options. The more reliant are some groups on the social welfare system, the more important it will be to them to have the option to treat intergenerational contracts as relational contracts. In other words, the poor benefit most from treating intergenerational contracts as relational contracts; for these provide stable entitlements.

In the context of intergenerational social contracts, we view the only plausible interpretation of the contract as one that is "relational". We say this because the world has changed considerably since the current social welfare system was first designed, most importantly in that demographic characteristics of the population have changed substantially (as discussed above), standards of living have increased, New Zealand society has become more heterogeneous and has different priorities than it had 50 years ago, technology has changed considerably, and the mobility of people has increased along with globalisation.

For example, the set of feasible approaches to social security has been expanded by changes in the technology of financial administration made possible by the introduction of computers. When Social Security was created, there would have been very high transaction costs associated with individually controlled personal retirement accounts. Today, with the help of computers, the creation of a system of individual investment-based accounts is relatively easy as has been demonstrated by uptake of Kiwisaver.

In addition, an important change has been in the economic profession's understanding of how fiscal incentives affect individual behavior (Feldstein 2005a). In the 1930s, many policies assumed that individuals were so unresponsive to taxes and benefits that any behavioral response could simply be ignored. If social welfare schemes were being created now the programmes would likely be significantly different from those in current law simply because of our understanding of the impact of these programmes on individual decision-making.

New Zealand has an example illustrated by the unprecedented widespread policy changes of the 4th Labour Government. The preparedness of much of the populace for change resulted in comprehensive reforms that left no area of the economy untouched.[13] Evans and Richardson (2002) argue that the voter consensus for reform was aided by the breadth of the reform programme which enabled changes that upset existing special interest group equilibria but were politically saleable because all groups in New Zealand were affected.

Legal contracts have very limited applicability in **New Zealand**'s governance arrangements for social welfare, and where they do (ACC is perhaps an example) it has significant transaction costs**.** These arises from the fact that detailed aspects of the implementation of legal contracts are capable of interpretation and enforcement by the courts, whereas it is much more difficult for a third party to enforce a relational contract (indeed the inability of relational contracts to be enforced by third parties is one of their attractions)**.** It is therefore unfortunate if politicians promote interpretations of current aspects of the intergenerational contract as inviolable when they are not[14] (the retirement age debate is a good example) as opposed to an undertaking to be fair and reasonable in the context of modern economic and social conditions**.** In other words, sustainability implies the potential for changes in the specific terms of any policy when that policy is interpreted as a relational contract.

---

[13] See, for example, Evans, Grimes, Wilkinson and Teece (1996).

[14] Even if one political administration did adopt a rule, it cannot commit other administrations in other parliamentary terms to the rule.

4 Some Taxonomies of Social Welfare Scheme Design and Funding

4.1 Introduction

In this section we discuss key characteristics of combined social welfare and insurance schemes in order that the differences between organisational forms of schemes can be more sharply set out. This facilitates more precisely linking scheme features to time consistency and governance, which we do in Section 5.

The provision of social welfare benefits can be categorized according to the following taxonomy:

- Contributory benefits are those that are available only to those individuals who have contributed at a certain minimum level.
- Means-tested benefits are those available only to those whose incomes and/or wealth fall below a certain threshold.
- Contingent benefits are those which are available based on a certain contingency such as reaching retirement age, being unemployed, being ill, having an accident etc.
- Discretionary benefits are those benefits available on the judgement of providers with delegated responsibility to assess contingent events.

In practice, most aspects of the social welfare system are contingent, and the largest difference between different types of benefits and the approaches adopted in different countries is in the use of contributory requirements and means testing. However, the approach to the contingency threshold varies on the basis of whether it is treated as social welfare or social insurance.

The design of social welfare schemes can also be categorized according to the ways in which the schemes are governed and funded. Here the taxonomy of possibilities includes:

- Public provision from unfunded schemes
- Pubic provision from fully funded schemes
- Private provision from unfunded schemes
- Private provision from fully funded schemes
- Combinations of the above, particularly or public and private provision.

In this section we consider these different issues of taxonomy. We begin by considering the differences between social welfare and social insurance, and then outline the different funding and governance options.

4.2 Insurance

Uncertainty is a fundamental feature of life because individuals are unable to predict the timing and magnitude of events that may have a profound effect on their wellbeing. Insurance may be broadly defined to include all the actions people undertake to mitigate the effects of uncertainty. Modifications of individual behaviour may reduce uncertainty but since there are many events that occur with a frequency that is unrelated to the behaviour of each individual, insurance contracts provide a vehicle by which uncertainty about outcomes for

each individual is transferred (pooled) and the costs of uncertainty are spread across a large group of individuals.

The essence of the insurance contract is the payment of a premium by the insured in return for a promise from the insurer to pay a certain sum in the event that the contingency insured against actually occurs within the contract period. Insurance companies commonly also provide services following the occurrence of a claim, including the payment of claims, the provision of health and rehabilitation benefits etc., but it is the provision of the insurance against the contingency, rather than the provision of services associated with the claim, which is the fundamental and unique feature of the insurance industry. This is seen most clearly in the case of an insured party who does not claim on the policy; this party has benefited from the transfer of risk arising from the insurance even though there has been no claim made.

The benefits of an insurance contract are derived from two elements of insurance: the transfer of risk and the pooling of risk. The individual policyholder transfers the risk, in that a loss is payable by the insurer. The individual policyholder shares in the risk borne by other policyholders in that each policyholder contributes, through the premium paid in advance, to the cost of meeting the claims of the policyholders as a whole.

Policyholders pay a premium sufficient to cover the expected cost of claims and cost of management of the scheme because of the benefit that they obtain in being insured against the small but non-zero probability of the occurrence of some contingency that may impose high or catastrophic financial costs on the individual policyholder. Because policyholders are risk averse by definition, they will be willing to pay individual premiums that are in excess of the expected payout for the population as a whole, and this margin funds the costs of the insurance scheme. In other words, insurance schemes are built on the principle that outcomes for a large sample of policyholders are predictable whereas for the individual they are not.[15]

4.3 Insurance and Asymmetric information

A standard feature of insurance markets is the presence of asymmetric information: in particular, that the insured party has more information relevant to the insurance contract than the insurer. It is not possible for insurance companies to obtain full information about the insured party, or to write a contract that includes provisions addressing all potentially risk-enhancing actions that the insured party might take. The theory of insurance distinguishes between adverse selection (information known only to the insured party before the contract is written) and moral hazard (actions taken by the insured party after the contract is written). Adverse selection means that because the insured party has more information about the likelihood that the contingency insured against will occur, insurance will be most attractive to those in high risk categories whereas those with low risk may self-insure by not taking out insurance with a third party insurer. Moral hazard means that once the insurance contract is written the probability of a claim for the contingency insured against is increased. Moral hazard recognises the incentives that provide both:

- the potential for changes in behaviour that will increase the risk for the insured party; and
- the potential for loss of good faith that may result in opportunistic claims against the insurer for events that fall outside the insurance cover.

[15] In the theory of statistics this is known as the law of large numbers. It requires that the outcomes have some degree of independence.

Private sector insurers address the problems created by adverse selection and moral hazard through a range of mechanisms that include:

- adjusting premiums based on both observable characteristics of the insured party ex ante, and the frequency with which claims are made (experience rating and no claims discounts);
- the sharing of the loss between the insurer and the insured party (co-insurance) through deductibles and limitations on the total value insured;
- restrictive covenants in contracts that make the insurance void in specified cases of moral hazard and adverse selection (such as self-injury and failure to disclose relevant information); and
- accident prevention and information sharing programmes designed to promote low risk behaviour in certain circumstances.

The adjustment of premiums on the basis of observed characteristics involves what is known as establishing risk pools.

The sharing of the loss between the insured party and the insurer is a common feature of insurance because of the extreme incentive problems that may arise. There is also extreme differentiation of incentive effects among different types of workers. This proved to be the fundamental problem with the sustainability of the social insurance schemes founded on the basis of "Beveridge" principles,16 among which New Zealand's ACC scheme can be counted. The problem in the latter case is the difficulty of differentiating between accidental and other causes of disability, the need to make benefits high enough to compensate for a no fault accident scheme, and the need to determine over what period workers are unable to work. Workers who do not expect to be promoted in the future may find it attractive to receive long-term ACC benefits if they can then receive an ongoing income of 80% of their current income, whereas for workers who expect to be promoted in the future to accept 80% of their current income might involve a very substantial sacrifice over the income that they expect to earn in five years' time. Therefore, where insurance companies can identify workers who have more limited prospects for future promotion they may reasonably require higher levels of co-insurance than they would for workers with strong professional qualifications or long-term career prospects.

4.4 Distinguishing Between Social Welfare and Insurance

There are six factors that underlie the distinction between social welfare and insurance.

4.4.1 Ex Ante vs. ex post redistribution

Insurance contracts set premiums that fairly reflect the likelihood of the occurrence of the contingency against which the insurance is purchased. In this sense there is no ex ante redistribution of wealth implied by an insurance contract. Insurance contracts provide for redistribution of wealth among individuals, but only ex post facto, and only on the basis of the (more or less random) occurrence of the contingency insured against. Social welfare schemes, in contrast, are designed to achieve ex ante redistribution; that is, they are partly or wholly designed to redistribute income from one group in the population to another on the

[16] The Beveridge Report (1942) proposed that social insurance should mean that the government would provide adequate income support for all, regardless of their source of income loss using a unified comprehensive system of flat rate benefits linked to flat rate contributions, and in which there was only a minimal role for means testing.

basis of characteristics that are observable ex ante. For example, the social welfare system may redistribute income from rich to poor, and thereby change the expected income stream of individuals.

4.4.2 Compulsory vs. voluntary participation

A key element of social welfare schemes is that they are compulsory. This is because the ex ante nature of the redistribution that occurs in social welfare schemes means that individuals who expect to be net contributors to the scheme may not participate voluntarily, whereas all those who expect to be net beneficiaries will. Compulsory participation is, however, a necessary but not sufficient condition for the identification of social welfare. Insurance is often purchased voluntarily, but there may also be market or institutional failures (such as those resulting from people free-riding on the social welfare net or externalities for third parties) that may make it desirable for governments to legislate for mandatory insurance coverage.[17] Such compulsion does not, in itself, transmute insurance into social welfare, nor does it suggest that private underwriting and delivery of the insurance is inefficient.

4.4.3 Choice

Insurance schemes provide not only a choice around whether or not to participate, but also provide choice about the nature of the benefits purchased. High income individuals may choose to pay higher premiums to obtain higher benefits, and other individuals may choose higher co-payment provisions to reduce the premiums that they must pay for each level of insurance. In contrast, social welfare systems rarely provide choice of this type. Social welfare benefits are generally in a standard form, and universally applicable except in situations where income or asset tests make the benefits unavailable to those who are assessed as not requiring the support provided by the welfare benefits.

4.4.4 Funding

Social welfare is generally funded from taxation rather than premiums, because premiums will not be paid voluntarily when there is explicit redistribution of wealth.[18] Funding a scheme through taxes based on income ensures that high income individuals contribute more in premiums, even though all individuals receive benefit levels that are the same or only based on income and premiums to a limited extent. In contrast, insurance contracts offer variable benefit levels that are tightly linked to the magnitude of the premiums paid, thus avoiding ex ante redistribution except that associated with efficient pooling.

4.4.5 Incentive Design

A key component of insurance contracts is the design of incentive compatible premium and payment schedules that address the moral hazard and adverse selection problems pervasive in insurance markets. Some common elements of insurance contracts attempting to address these problems include use of deductibles, co-insurance, and experience rating. Social welfare schemes, in contrast, use compulsory participation as a means of addressing the extreme adverse selection associated with ex ante redistribution and the presence of other welfare programmes. Relatively low benefit levels and monitoring serve as blunt instruments

[17] The relative efficiency of different taxes is relevant to this issue, but we do not consider it further here.

[18] The extent and incidence of social welfare transfers are not obvious when the totality of taxes and subsidies is to be assessed. For example, Feldstein (2005b) suggests that U.S social security is more user pays than might be inferred at first glance because wealthier persons tend to live relatively longer.

for controlling moral hazard. Social welfare schemes will be most appropriate in markets where the type and magnitude of the adverse selection problem requires compulsory coverage, and where the moral hazard associated with the provision of benefits is small relative to the social costs arising from the contingency covered by the scheme. For example, the rate of unemployment in the population as a whole varies across the business cycle in a manner that is unrelated to the effort and actions of individual workers, and therefore, if the social costs (for example poverty and crime) associated with widespread unemployment are high, social welfare may be more appropriate than private provision of insurance in this case.

4.4.6 Recompense for losses vs minimum acceptable living standards

Social welfare schemes normally provide minimum levels of support regardless of past income, whereas insurance schemes typically provide income-related benefits based on the concept of recompense for losses incurred as a result of an injury. This means that insurance is designed to allow the potential to address the opportunity cost of some contingency not the maintenance of minimum socially acceptable standards of living following some contingency.

In practice, many problems will contain elements of both social welfare and insurance. Where the insurance component is large, the social welfare elements of the problem may be dealt with in full by making the purchase of some minimum level of insurance benefits mandatory or placing reliance on a safety net. It is this compulsory requirement to purchase insurance at a minimum level that we call “social insurance”. Social insurance schemes allow considerable flexibility in the design of premium and benefit levels, and may in particular allow the combination of subsidised minimum levels of insurance cover with the ability for individuals to obtain higher levels of cover based on income or choice.

4.5 Funding and Accounting for Intergenerational Benefits

We can distinguish conceptually between four different approaches to funding social welfare and social insurance schemes.

4.5.1 Full Funding

Full funding arises in a social welfare or social insurance scheme when premiums are set at a level that will fully fund all expected liabilities associated with the claims in the period covered by the premium. If a social insurance or social welfare scheme operating on such a basis was to cease operations at any time, it would be expected to have a pool of funds sufficient to meet the expected costs of benefits to existing claimants for as long as they claims may last. Full funding also arises in defined contribution social welfare or social insurance schemes, where the benefits obtain are determined by the fund that has been accumulated by past contributions. As these two possible ways of thinking about full funding imply, there are two possible ways in which to assess the governance structures associated with full funding:

a Full funding in a scheme managed by government implies that the funds accumulated are managed by government. Only the government has ownership rights in those funds, because they support a public scheme and are not assigned to individual ownership, which allows the government a range of discretion in how it invests the funds and even in whether full funding will continue.

b Full funding in compulsory social welfare schemes may alternatively be consistent with individual choice where the funds accumulated are invested, and competitive provision of investment services for those funds. Compulsory individual retirement accounts such as required under the Australian pension scheme represent an obvious example of this approach.

4.5.2 Pay-as-you-go funding

Premiums or levies in any given year are set at levels that will meet the annual costs of the payments made by the scheme in that year, regardless when the accident upon which each claim is based occurred. A key feature of a pay-as-you-go scheme is that it entails defined benefits and requires the ability to tax, in order to ensure that individuals meet the current period costs of the scheme even if this requires a premium rate that is greater than the expected value of the current period insurance benefits received. A second key feature of pay-as-you-go funding is that even in the presence of the ability to tax, the sustainability of the scheme requires that the tax burden required to fund the scheme will be regarded as reasonable by future generations as well as by the current generation of tax payers.[19]

4.5.3 Partial Funding

Premium rates are set at levels that provide less than the full present value of liabilities arising from the claims expected to be made in the current period, but sufficient to fund some of the future liabilities arising from these claims. Partial funding can be viewed as the provision of a capital sum the investment income from which offsets a full pay-as-you-go premium to some degree.

4.5.4 Notional Funding

Under this approach, individual accounts are maintained, but no investment fund is actually accumulated. In other words, the benefits available to individuals are determined by an account of the contributions that they have made, but the contributions are actually made by government and the benefits, which relate to level of contribution, are provided by government on a pay-as-you-go basis. This approach has the advantage that the full extent of the liabilities associated with any scheme can be calculated, and contributions can be determined on an actuarially fair basis, but there is no need for a government agency to invest a large segregated fund. It also has the advantage of providing a basis on which the benefits provided to individuals can be tied to their contributions without a fully funded account: in other words it is a pay-as-you-go defined contribution scheme.[20]

4.6 Sustainability of Different Funding Options

Efficient insurance contracts link actuarially fair premiums to expected benefits as tightly as is possible given transaction costs and the presence of moral hazard and adverse selection problems. The efficiency of social welfare schemes is affected by the extent to which the link between an individual's premium contributions and benefits is weakened or broken. If a social welfare scheme can provide actuarially fair premiums then other things equal it will be sustainable.

---

[19] This issue is discussed by McHale op cit.

[20] The absence of a invested fund over which an individual has ownership may make entitlements in this scheme more subject to being affected by public policy changes than would be the case if there was an invested fund.

The premiums of pay-as-you-go social welfare schemes are not actuarially fair because they do not equal the present value of expected benefits. Because premiums that are not actuarially fair will be endemic in pay-as-you-go schemes, they are sustainable only because participation in these schemes is compulsory and relies on the ability to tax. Government sponsorship is normally required for these schemes because only the government can make membership and funding contributions compulsory. The compulsion is a reflection of the weak link between premiums and expected benefits and is an element of inefficiency of pay-as-you-go schemes unless the scheme is plainly sustainable as part of a credible relational contract. In such a setting the redistributional element of the pay-as-you-go scheme would be widely accepted and the scheme sustainable at its settings. As we have argued, a world-wide heterogeneous labour market and changing demographic structure make a consensus sufficient for an relatively efficient, pay-as-you-go scheme difficult.

Partial funding of the expected costs of social welfare is an intermediate case. Partial funding may provide for greater stability in premiums than a pay-as-you-go scheme, because the partial funding may be used as a buffer against unexpected changes in current benefit payments that would be reflected in premiums under a pure pay-as-you-go scheme. However, partially-funded schemes have in common with pay-as-you-go schemes the problem that premiums depart from the efficient (actuarially fair) level, and that the unfunded liability must be covered with premium income at some time.[21] Thus, partially funded schemes raise problems of efficiency and transfer between different cohorts of levy payers, and also require government support to compel the payment of premiums to cover the unfunded liability. Partial funding may also reduce efficiency because in the absence of a specific funding in excess of the pay-as-you-go premiums may be the subject of political opportunism (for example, through increases in benefit levels funded from scheme reserves or contribution or government contribution holidays).

Full funding is based on the fundamental principle that what we pay in premiums today should equal the full costs of benefits that we expect today to incur in the future[22], and that future generations will also pay the costs relevant to the benefits for their society. The reasons why this principle is important are:

- In the absence of full-funding there is scope for non-transparent transfers between different groups or cohorts of levy payers. Claimants can obtain expected benefits in excess of the value of the premiums that they have paid only if another group is paying premiums that are less than the expected value of the benefits. The danger is that in an "immature" insurance scheme where there are very long-term costs associated with accidents but premiums are not fully-funded, each generation will be tempted to provide itself with benefits that would not be affordable on a fully funded basis by underfunding the insurance liabilities and failing to compensate for this in the overall government fiscal position. In this circumstance, the fact that the benefits are not affordable will become apparent only in the future, as the stock of claims that must be funded on a pay-as-you-go basis increases to the point where the burden on a future generation becomes too heavy and benefits have to be reduced.
- Without full funding the premiums charged may change in any year in ways that are unrelated to current expected outcomes.

---

21 We take the actuarially fair premium to be efficient because it means that the marginal cost of the insurance to an individual equals its (expected) marginal benefit. In this our actually fair benefit calculation embodies the expected earnings on premiums invested. Put another way it is the expected present value of payouts the insured is entitled to on the future occurrence of the insured event.

22 Actual payments from the fund will be for costs of injuries incurred in the past.

- Full funding also allows for privatization of accounts. For example, if a pension scheme was fully funded, it would be possible for government to pass to the private sector the responsibility of managing, on behalf of all of those individuals, the funds required to meet their claims to pension benefits.[23] However, notional funding also allows this possibility to be contemplated in the sense that the government could issue debt to fully fund liabilities if they knew what those liabilities were and wished to fund them in this way. So long as individuals have a notional welfare account, then the government could elect, or allow individuals to elect, to establish a private account associated with their pension obligations.

The problem with full funding is that, in the hands of government or an agency controlled by government, it is unlikely to be sustainable: i.e. time consistent. The moral hazard associated with fully funded schemes is the scope that they allow the government to buy votes by offering contribution holidays (including lower than sustainable contributions, or higher benefits). The problem is that lower than sustainable contributions in a long-term benefit scheme can actually be sustained over a long period of time – in effect over the full period that it takes to shift from a full funded to a pay-as-you-go basis. The funding variations in the history of the ACC scheme in New Zealand provides a good example, and leads us to the conclusion that full funding in a publicly managed scheme does not promote intergenerational sustainability. In fact it may facilitate considerable politically opportunistic intergenerational instability.

However, these reservations do not apply to either fully-funded individual accounts or notionally fully funded individual accounts. In both these cases individual control over and choice about the allocation of benefits reduces markedly the potential for political opportunism that bedevils publicly funded schemes, while they still retain the merits associated with actuarially fair contributions (primarily because full funding of individual accounts also implies a defined contribution basis for the scheme).

A “notional” defined contribution system would only help a little to reduce the adverse effects of the current pay-as-you-go system. A notional defined contribution system is one in which each individual has an account that is credited with his tax payments and with a notional rate of return on his accumulated balance but in which there is no actual investment in financial assets. The notional rate of return that is feasible in the long-term is the modified Samuelson return, ie, the rate of growth of the tax base.

However, since there is no real capital accumulation in a notional defined benefit scheme, the reduction in the present value of consumption is not changed. The distortion in labour supply and in the form of compensation is reduced (but not eliminated) relative to pay-as-you-go schemes because individuals can more clearly see the link between their taxes and their future benefits. A notional defined contribution pension system, for example, also reduces the distortion in retirement decisions because individuals reduce their future benefits if they retire early and increase them by delayed retirement. But even with this improved transparency, the low implicit pay-as-you-go rate of return leaves a substantial distortion in work and compensation incentives.

[23] A difference between private and government-underwritten fully funded schemes may lie in the financial security of government participation.

4.7 Defined Benefit and Defined Contribution

Most social welfare programmes in New Zealand are defined benefit. The notion of a defined benefit scheme derives from the literature on private pension schemes. Defined benefit schemes were originally designed to increase the incentives provide by remuneration schemes (higher pay during the latter part of the career resulting in a higher benefit). But most private pension schemes have now moved away from defined benefit schemes to defined contribution schemes.

In the social insurance arena, defined benefit schemes have the effect of providing for the poor, but limiting choice and imposing cross-subsidies across the full range of benefits. It would be more transparent to shift as much as possible to shift into defined contribution social insurance schemes. In this case, cross subsidies would be removed, and the issue of minimum income (poverty) addressed separately.

5 Achieving Sustainability

In this section we bring together the characteristics of social welfare and insurance schemes with the requirements of time consistency (sustainability) that we have argued are important for socially desirable performance of these schemes, and suggest policy directions. We examine the possibilities of frameworks that in our view are worth further evaluation. We do not examine any particular policy in sufficient detail to recommend it.

The direction of our approach is that additional transparency and flexibility would enhance the connection of insurance and social welfare policies to a relational contract that would enhance scheme performance and sustainability. And that this might be achieved by enabling, even forcing, scheme participation by individuals as individuals. Such individual participation would clarify what benefits they are getting, or may expect, for their contributions, as well as what amounts support others. Individual scheme participation is quite common in other countries and it is facilitated by modern (digital) accounting processes.

In our view, this approach would reveal better the level and incidence of schemes' costs and benefits, and thus aid transparency in the relational contract. It would better meet the individual preferences of individuals, which would enhance stability of the schemes covered by the relational contract. The defined participation of individuals would mean that the schemes' inputs and outputs would to an extent change with demography: avoiding abrupt changes that might arise under pure state management. Established individual entitlements would likely make more difficult politically motivated scheme changes, which would also contribute to stability.

5.1 Permitting Individual Choice

Individuals differ in their preferences. We do not all have the same risk aversion, the same time preference, the same relative taste for goods and leisure. Letting individuals choose among options in a way that reflects their individual preferences and circumstances should be an important aspect of social insurance design. But allowing choice means that programs should be designed so that choice enhances economic efficiency rather than creating deadweight losses. A good example of such a program redesign was the US reform that introduced the actuarial adjustment for early and delayed retirement in a way that, in principle, will allow individuals to decide when they will start collecting benefits without changing the actuarial present value of their benefits.

5.1.1 Creating Programme Transparency

Social insurance programs involve complex rules about the benefits to be received, the taxes to be paid, and the link if any between them is opaque. These weaken the perceived link between the payroll taxes paid and subsequent benefits. Employees in New Zealand have no idea what portion of their current taxes are used to provide them with different types of social welfare benefits: the balance between actuarially fair contributions to any benefit scheme and cross-subsidies to other individuals in the scheme is not explicit. This increases the disjunction between benefit and perceived marginal tax rate, raising the deadweight loss of the tax. The use of a deliberately opaque system to achieve a redistribution of income among individuals is a key element of pay-as-you-go social welfare benefits funded from general taxation that may prove unsustainable as required taxes rise.

New Zealand's social welfare system lacks transparency primarily because it is a defined benefit system rather than a defined contribution plan of the sort that now characterizes many private pension schemes in other countries. Converting social welfare to have at least some elements of a defined contribution plan – even an unfunded "notional" system such as Sweden and Italy now have – would allow individuals to see the link between their taxes and the resulting benefits.

Although a notional defined contribution plan would remain a pay-as-you-go system, it would clearly link each worker's social welfare tax payment to his or her resulting future benefits. A defined contribution system would provide a tax-benefit link for those (high tax-paying) groups in the population that now receive no extra benefits at all in exchange for their additional taxes.

5.1.2 Private Accounts

A much more substantial reduction in the effective tax rate would be achieved by financing the increased cost of Social Security and Health by a funded system that would permit future benefits to be financed without a large increase in the tax rate. Moreover, to the extent that the additional saving of individuals earns a favourable rate of return, they might not consider it a tax at all.

5.1.3 Separating Social Insurance from Income Redistribution

It is well established that much of the current social welfare system in New Zealand involves "middle class churn",[24] that is, that most of the payments are returned to exactly the same group that paid taxes to fund them. The inefficiency of this approach is well established, especially given the transaction and efficiency costs of government management and the removal of individual choice through these systems.

There is, broad acceptance of the need to support the poorest members of society, and thus of the need to provide transfer payments through the taxation system. To the extent that distributional concerns motivate the design of social insurance, the emphasis could be on eliminating poverty and not on the overall distribution of income or the general extent of inequality, since the latter is a recipe for middle class tax churn. An explicit decision to supplement the contributions of low-income earners in such a defined contribution plan would achieve income redistribution without a loss of transparency.

The preference for redistribution may decline as the population becomes more heterogeneous. For example, Dahlberg et al (2012) show that in Sweden the preference for redistributive social policy, especially among higher income earners, declined during a period of substantial immigration. The population in New Zealand has become much more heterogeneous than it was when our current welfare system was first developed, and immigration in the age groups near retirement does have the potential to create a new group within society who benefit from windfall gains in the social welfare system. While a link between heterogeneity and support for redistribution is not yet clearly established in New Zealand, it may well be relevant to the recent debate about the proportion of immigrants who qualify for pensions or other welfare benefits. Furthermore, populace heterogeneity implies improvements in welfare from schemes that allow some self-selection of benefits and contributions.

[24] See the analysis of Cox (2001).

5.1.4 Accounting for Governance and the Relational Contract

Evans, Guthrie and Quigley (2012) (EGQ) argue that efficient governance requires tying accountability tightly to the residual claimant of the policy; and that selecting the particular claimant really matters for the performance of the policy. While this is relevant to all policies it is particularly relevant to policies that are constrained to manage benefits and costs that are not temporally matched: plainly these include insurance but they also include intergenerational policies that are implied by the contract. The reasons for this include aligning repeated agency relationships and enabling long term planning and decision-making that make decisions about options that arise over time in the context of the objective of the policy. In these policies the existence of options of delay and change imply that contracts are necessarily incomplete and require decisions – and thence a supporting governance structure – that are aligned with the objectives of the policy as laid down in relevant statute that reflects the relational contract. In effect this means maximizing the opportunity for the individual contributor to have decision rights about the level and allocation of contributions to social welfare schemes, and the terms on which benefits are obtained, will be important for sustainability. In this sense, the current Kiwisaver scheme has the benefits of assigning governance to the individuals who will benefit from the savings while also providing for professional management of those individual saving accounts.

New Zealand did in the past have a specific social-security tax, but this separation of tax payments was removed on the grounds that the revenue it generated was simply credited to the consolidated fund and there was no connection between the payment for services from the fund and its sources of revenue. This decision might now be re-evaluated on the grounds of modern technology, the benefits of informational transparency, the incidence of intergenerational transfers and information for governance under the relational contract.

Earmarked taxes and expenditures are present only for the ACC scheme, and yet digital technology would allow them to be more widely applied[25]: to superannuation and health services for example. Their benefits for governance lie in what they imply for the credibility and operation of budgets for the different services. In the case of superannuation the budget and its forecasts would reflect current and prospective payments and revenues that may incorporate both the New Zealand Superannuation Fund (NZSF) and pay-as-you-go funding drawn from current taxation. The approach would make government provision for retirement and funding for retirement more transparent, and it would enable more concrete assignment of management of funding and provision of superannuation. The transparency of revenue and expenditure would make evaluation of variations of the schemes easier to convey to those that have the right to vote about them and the tax and benefit rates would facilitate discussion of any proposed changes to the scheme. Further, defined tax and benefit rates would make commitment to these rates more credible because it is likely that changes to such rates would require a more considered and publicly comprehensible process than expenditure allocation of the consolidated fund. This earmarked funding and expenditure would facilitate the creation of a more transparent governance structure for government superannuation than is possible under the current approach.

Plainly health combines social welfare, insurance and provision in-kind and so would draw different taxes and specialized governance from that of superannuation: but the implications of earmarking remain.

[25] Such specific taxes need not reduce the nature of the tax system as a whole, or preclude private management of private accounts.

5.2 An Example: Individual Unemployment Insurance Accounts

The scheme principles we have considered apply to all those matters for which insurance and social welfare are elements. As described, they include health, retirement, accident and unemployment states of the human condition. The example of the Unemployment Insurance Saving Account (USIA) suggested by Feldstein and Altman, (1998) sets out and example of how one such scheme would function. Each individual would be required to accumulate funds in an Unemployment Insurance Saving Account until the balance was enough to pay benefits for two spells of six months at 50 percent of the individual's current wage (the existing provision by the state). These funds would be invested and would earn a market rate of return. If a balance remains in the account when the individual reaches retirement age, the funds would be available for the individual to take and spend. An individual who dies before retirement bequeaths the account balance. In short, individuals would regard the funds in the UISA as their own money. For someone who expects to have a positive balance in his account until retirement, the UISA plan would provide the same income protection as the current state unemployment scheme but without distortion.

Individuals who experience so much unemployment that they use up their personal insurance fund would be able to borrow from a government fund to receive the same benefits that they would withdraw if they had a positive account balance. After they return to work, they would again save to repay the loan with interest and to rebuild their insurance balance. For those who expect that they will have a negative balance in their account when they retire this plan would represent no improvement over current law. For them the protection and distortion would be the same as it is with the current unemployment benefit programme. The extent of the gain from introducing unemployment insurance therefore depends on the proportion of the unemployed who expect to retire with negative balances and on the sensitivity of unemployment to the change in incentives. It seems likely that the number of such individuals (representing the long-term unemployed) is relatively small, so designing a scheme to provide improved incentives for the majority would have overall positive welfare implications as well as improving sustainability.[26]

5.3 Does Means Testing Promote Sustainability?

Before we consider transitions to mixed public and private social welfare systems, why not simply replace universal provision of social insurance benefits with means tested benefits? New Zealand has done this for housing. The approach does not lock in individual entitlement commitments as a mixed system would, and aggravates the pressure on sustainability by the incentives it provides individuals. Some rational and farsighted individuals would be induced by a means tested system to act in a way that allows them to qualify for benefits. Doing so would impose tax costs on the rest of the population that could make overall well-being lower than in a universal (ie, not means tested) program.

Consider a simple example of a means tested retirement program (Feldstein, 1987). A rational individual would decide whether to act as if he or she is myopic by comparing the lifetime utilities with optimal positive saving and with no saving. Those with relatively high incomes would not be tempted by the means tested benefit. But others with lower incomes would have higher lifetime utility by increasing their consumption during working years even if the means tested benefits would only provide lower consumption during retirement than optimal

[26] Consideration of the details of the design of such a scheme are beyond the scope of this paper.

saving would allow. Although they would achieve higher lifetime utility through their action, their benefits would sourced from tax-financed transfers that would make others worse off.

There is no way for the government to distinguish between the genuinely myopic and those who are rational utility maximizers gaming the system. The government could in principle set the means tested benefit so low that very few rational individuals would be tempted. But a relatively affluent society may not accept that policy. The means tested benefits would be set at a higher level that would tempt many rational individuals to save nothing.

Even if means testing was a panacea for sustainability in the intergenerational contract (ie, it had no perverse incentive effects of the type described above) there is an important political economy reason to avoid simply advocating means tested programs for unemployment and old age (Feldstein 2005). Elected governments will inevitably seek to create universal benefits to capture political support from the largest possible majority of voters. Means testing will always be limited in its ability to stem the cost, or improve the intergenerational sustainability of social welfare schemes. We therefore conclude that means testing has merit as a component of sustainable intergenerational social welfare schemes, but cannot solve sustainability problems on its own.

5.4 The Transition to a Mixed System

The common view that the transition from a pure pay-as-you-go system to a mixed system must require the transition generation to "pay double" – once to save for their own retirement and once to meet obligations to existing retirees – is not generally correct.[27] To illustrate why this is, Feldstein and Samwick (2002) showed how relatively small contributions to personal welfare accounts could offset relatively large increases in required taxes. The key to the transition is using personal retirement account annuities, which earn compound returns on the funds invested, to supplement the pay-as-you-go benefits. The growth of savings in private welfare accounts, driven both by contributions and compound interest, offsets the slowdown of the pay-as-go-benefits that results from not increasing the tax rate, or age of benefit eligibility, as the population ages. Even a transition in which the initial personal retirement account deposits are financed wholly by government borrowing could eventually raise national saving and the present value of future consumption.

Feldstein (2005a and b) argues that a mixed system offers a substantially lower long-term cost of financing welfare payments than the tax projected for the pay-as-you-go system; a higher expected level of benefits from the combination of pay-as-you-go and the private savings; and a very low probability that the actual level of combined benefits will be less than the pay-as-you-go benefits currently expected (in part because of the floor on welfare payments provided by the pay-as-you-go benefits in the mixed system).

[27] Coleman (2012) examines incidence of benefits and payments across New Zealand population cohorts. He suggests that cohorts born before 1981 can expect to pay half as much as they expect to get in benefits as a result of the changing demographic structure of the New Zealand population.

6 Conclusion

It has been our purpose to set out an overview of the connection between policy and policy sustainability, particularly in the presence of intergenerational transfers. We do not promote particular policy structures; although we suggest directions for consideration.

Intergenerational contracts should be interpreted as relational contracts. If they work in this way, they are sustainable, or time consistent. They generally become unsustainable if they are viewed as unable to change consistent with contemporary conditions.

We have argued that comprehensive public social welfare schemes, whether fully funded or pay-as-you-go, may prove to be unsustainable across generations. In the case of full funding this is because of the incentives for politically opportunistic contribution holidays and benefit levels. In the case of the latter, it is because of the incentive problems and inefficiency of middle class churn raising the cost of the system.

However, we do not expect that the rapidly growing fiscal gap predicted by writers such as Kotlikoff (2008) will actually arise. The reason is that under any politically sensible interpretation of the relational contract associated with government social welfare obligations there will be ways to transition the intergenerational contract from its current structure to a more sustainable structure. The more sustainable structure is likely to include:

- greater use of social insurance (compulsory private contributions to schemes designed to provide for particular contingencies)
- greater use of deductibles or co-payments
- greater use of earmarked taxation
- greater use of private accounts, either fully funded or notionally funded, given the political unsustainability of fully-funded public schemes
- greater separation of redistributive welfare payments from social insurance provisions, and
- a mix of contingency-related and means-tested benefits, with the former based as closely as possible on individual contributions, and the latter based as closely as possible on actual need (income below a certain threshold).

Following Feldstein (2005b) and the approach embodied in the New Zealand KiwiSaver scheme (but with compulsory membership added) the current pay-as-you-go social welfare system might be gradually replaced with a mixed system that includes mandatory investment-based personal accounts associated with at least some of the contingencies that are currently covered by the social welfare scheme (unemployment and health care for example). It appears to be feasible to design these in a way that maintains or exceeds the benefits that are projected in current schemes while reducing the long-run cost of achieving those benefits. A mixed system would avoid the political risk that future taxpayers would be unwilling to raise taxes to finance promised benefits, and by providing individual control of the funding in the accounts provide more desirable governance of social welfare than is currently provided by pay-as-you-go or fully funded publicly-managed alternatives.

Change is a matter of designing transition paths to more efficient arrangements that do not disenfranchise particular parties or force others to pay twice. The transition can be achieved by a combination of graduated and fiscal incentives. Kiwisaver suggests an example that (with the addition of compulsory membership) might be applied across a whole range of different aspects of the social welfare system.

References

Baker, G, Gibbons, R, and K J Murphy (2002) “Relational Contracts and the Theory of the Firm” *Quarterly Journal of Economics* 117 (1), pp. 39–84.

Banzhaf, Spencer H. and Wallace E. Oates, (2012) *On Fiscal Illusion And Ricardian Equivalence in Local Public Finance*, Andrew Young School of Policy Studies Research Paper Series No. 12-15. Available at SSRN: http://ssrn.com/abstract=2062473 or http://dx.doi.org/10.2139/ssrn.2062473

Barro, Robert (1974) “Are Government Bonds Net Wealth?” *Journal of Political Economy* 48(6), pp. 1095-1118.

Becker, Gary S., and Kevin M. Murphy (1988) "The Family and the State" *Journal of Law & Economics* 31(1) pp.1-18.

Beveridge, W H (1942) *Social Insurance and Allied Services Cmnd* 6404, London, HMSO.

Borck, Rainald, Silke Uebelmesser and Martin Wimbersky (2012), *The Political Economics of Higher Education Finance for Mobile Individuals*, CESIFO Working Paper in Public Finance, No. 3877, 33p.

Coleman, Andrew, (2012) *Pension Payments and Receipts by New Zealand Birth Cohorts, 1916-1986*, MOTU Economic and Public Policy Research Working Paper 12-11, 29p..

Cox, James, (2001) *Middle Class Welfare*, Published by NZ Business Roundtable, 265p..

Dahlberg, Matz, Karin Edmark, and Heléne Lundqvist (2012) “Ethnic Diversity and Preferences for Redistribution” *Journal of Political Economy*, 120 (1), pp. 41-76.

Demange, Gabrielle, Fenge, Robert and Silke Uebelmesser, (2013), Financing Higher Education in a Mobile World, *Journal of Public Economic Theory*, forthcoming.

Dobrescue, Loretti, Laurence J. Kotlikoff, and Alberto F. Motta (2008) “Why Aren’t Developed Countries Saving?” NBER, Working Paper, 14580, 22p..

Evans, Lewis and Martin Richardson (2002) “Trade reforms in New Zealand: unilateralism at work”, Ch.7, 167-218, of J. Bhagwati (ed.) *Going Alone: Relaxed Reciprocity: Historical and Modern Experience with Unilateral Trade Liberalization*, MIT Press, .

Evans, Lewis and Graeme Guthrie, (2012) “Price Cap Regulation and the Scale and Timing of Investment”,*The Rand Journal of Economics*, 43(3), 537-562.

Evans, Lewis, Graeme Guthrie and Neil Quigley (2012) “Contemporary Microeconomic Foundations for the Structure and Management of the Public Sector” New Zealand Treasury Working Paper, 12/01, May, 77p..

Evans Lewis and Arthur Grimes, Bryce Wilkinson, with David Teece, (1996) "Economic Reform in New Zealand 1984-94: The Pursuit of Efficiency", *Journal of Economic Literature*, vol. XXXIV, pp.1856-1902.

Feldstein, Martin (1987) "Should Social Security be Means Tested?" *Journal of Political Economy*, June Vol. XCV, No. 3, pp. 468-484.

Feldstein, M. (2005a), “Rethinking Social Insurance”, *American Economic Review*, vol. 95, n. 1, March.

Feldstein, M. (2005b), “Structural Reform of Social Security”, *Journal of Economic Perspectives*, 19(2), pp. 33-55.

Feldstein, Martin and Jeffrey Liebman (2002) “Social Security” Chapter 32 in *Handbook of Public Economics* 4 pp. 2245–2324.

Feldstein, Martin and Samwick, Andrew (2002) “Potential Paths of Social Security Reform,” in J. Poterba, ed., *Tax Policy and Economy* , Volume 16, Cambridge, MA: MIT Press, pp. 181-224.

Frankel, Jeffrey A. (2011) *A Solution to Fiscal Procycality: The Structural Budget Solutions Pioneered by Chile*, National Bureau of Economic Research, Cambridge Mas., Working Paper 16945, 48p.

Goldberg, V. P. (1980) “Relational Exchange: Economics and Complex Contracts” *American Behavioural Scientist* 23: pp. 337 – 352.

Goldberg, V P (1998) “Relational Contracting” in Peter Newman (Ed) *The New Palgrave Dictionary of Economics and the Law* (Macmillan), pp. 289 – 294.

Green, Jerry, and Laurence J. Kotlikoff (2007) “On the General Relativity of Fiscal Language” mimeo, 15p..

Kotlikoff, L J (2008) “The True Cost of Social Security” NBER, Working Paper, 14427, pp. 33.

Kotlikoff, L J (2001) “Generational Policy” NBER Working Paper 8163, 30p..

Kydland, Finn, E. and Edward Prescott (1977) “Rules Rather than Discretion: the inconsistency of optimal plans”, *Journal of Political Economy*, 85(3), pp. 473-492.

Laslett, Peter (1992) chapter 1, *Justice Between Age Groups and Generations*, James Fishkin and Peter Laslett eds..

Levin, J (2003) “Relational Incentive Contracts” *American Economic Review* 93:3 pp. 835 – 847.

MaGee, Stephen P., Willaim A. Brock and Leslie Young (1989) *Black hole tariffs and endogenous policy theory: political economy in equilibrium*, Cambridge University Press, 438p..

Merton Robert C. (1983) On Consumption Indexed Public Pension Plans, Ch. 10 of *Financial Aspects of the United States Pension System*, Zvi Bodie and John B. Shoven, editors, University of Chicago Press, NBER Books, pp.259-290.

McHale, John (2001) The Risk of Social Security Benefit Rule Changes: some international evidence, in *Risk Aspects of Investment-based Social Security Reform*, John Campbell and Martin Feldstein, eds. Chicago, University of Chicago Press, pp. 142-50.

Milgrom, Paul and John Roberts (1991) *Economics, Organisation & Management*, Prentice Hall, 621p.

Persson, Torsten, Gerard Roland and Guido Tabellini (2007) "Electoral Rules and Government Spending in Parliamentary Democracies", *Quarterly Journal of Political Science*: Vol. 2:No 2, pp. 155-188.

Samuelson, Paul (1958) “An exact consumption-loan model of interest with or without the social contrivance of money” *Journal of Political Economy* 65 (December), pp. 467 – 482.

Seater, John T. (1993) “Ricardian Equivalence”, *Journal of Economic Literature*, 31(1), pp 142-190.

Thompson, David (1996) *Selfish Generations: How Welfare States Grow Old* (The White Horse Press).

Weaver, Carolyn (1994) “Financial Implications of Aging Populations” in P. Newman et al (eds) *New Palgrave Dictionary of Money and Finance* (Macmillan) Vol. 2, pp. 65 -68.

Weingast, Barry (1989) “Constitutions and Commitment: The Evolution of Institutions Governing Public Choice in 17th Century England" *Journal of Economic History*, 49, pp. 803-32.

Werning, Ivan (2007) “Optimal Fiscal Policy with Redistribution”, *Quarterly Journal of Economics*, 122(3), pp. 925-967.