Evidence-based evaluation of social policy

Susan St John & M. Claire Dale[1]

ABSTRACT

Much attention in recent times has been given to the concept of evidence-based policy. One outcome has been a desire for research-based evidence to inform policy design; another is for there to be proper evaluation of policies after their implementation. For policies that attempt to affect behaviour this approach can raise many challenges. The pressures politically may appear to be for speedy evaluation to secure ex-post justification of the policy. This paper reviews the role of evidence-based input and the policy making process with two examples from recent policies designed to affect behaviour: Working for Families and KiwiSaver. How could the policy framework have been used more appropriately?

1. Associate Professor Susan St John, Economics Department, Co-director, Dr M. Claire Dale, Research Fellow, Retirement Policy and Research Centre, University of Auckland. Contact s.stjohn@auckland.ac.nz.

1. INTRODUCTION

One of the most important functions of the discipline of economics is to provide a rational basis for policy making. If the goal of economics is to extract maximum value out of scarce resources, the issues of policy design cannot be ignored by economists, nor can the distributional outcomes of policy be sidelined. Reflecting this concern, much attention in New Zealand has been given to the concept of evidence-based policy, following its promulgation in the UK in the late 1990s.

One of the intents is not only for research-based evidence to inform policy design, but also for proper evaluation of policies after their implementation. Major policies such as Working for Families (WFF) and KiwiSaver are implemented with a large role for evaluation built into their budgets. However, measuring outcomes from attempts to affect behaviour can raise many challenges.

This paper first outlines the background to 'evidence-based' policy, sometimes called 'what works', and suggests a simple framework for policy analysis that highlights the points at which research-based evidence of a variety of disciplines may, could, or should have an impact.

Using this framework to illustrate the process, examples from recent policy in New Zealand are discussed, in particular, the policy development, implementation and evaluation of WFF. Ex-ante theorising, ex-post evaluations, and data gathering of different kinds have all played a part. For WFF the conclusion is that while the evaluation techniques that have been used may appear to be sophisticated, at best, their contribution has been, and indeed could only be modest. Meantime other more helpful evidence including qualitative work, ex-post theorising, and meta analysis is sidelined. The result is that the economic evaluation becomes an endpoint of the process, and there is little broader critical analysis that might suggest improvements, or even a fundamental rethinking of policy. In similar vein, the difficulties of evaluation of KiwiSaver against its objectives may have resulted in a lack of focus on emerging policy design issues.

Could the policy framework have been used more appropriately? Has technical economic evaluation using narrow measurable objectives been used to secure an ex post justification of policy? What is the appropriate role of economic evidence in improving policy design? Who asks whether the broader problems of the policy have been addressed?

> *Critically, if evaluation is to be useful, rather than simply testing if our policies reduce dependence on the state, or if they increase work participation, we must frame the question in a way that reflects the ultimate goals. This broader vision of a better society must include the perceptions and well-being of those who are the focus of our concern.* (St John, 1997, p. 429)

2. EVIDENCE BASED POLICY

2.1 Origins

In the UK social science environment, Nutley, Davies, & Walter (2003, p. 29) noted that in the 1980s and early 1990s "there was a distancing and even dismissal of research in many areas of policy, as the doctrine of 'conviction politics' held sway". By the late 1990s it had become fashionable to make social science and policy making appear more 'objective' by pointing to the role of evidence and analysis.

In 1999, the UK White Paper *Modernising Government* pledged to be "forward-looking in developing policies to deliver outcomes that matter, not simply reacting to short-term pressures" (CM 4310, 1999). The same White Paper proposed that being evidence-based was one of several core features of effective policy making, a theme developed in subsequent UK government publications (Bullock, Mountford J., & Stanley, 2001).

In addressing the "Evidence-based Policy and Practice" (EBPP) agenda in 1999, the UK Government Cabinet Office described evidence as:

> *Expert knowledge; published research; existing statistics; stakeholder consultations; previous policy evaluations; the Internet; outcomes from consultations; costings of policy options; output from economic and statistical modelling.* (Strategic Policy Making Team, 1999)

This broad definition clearly identifies *research*-based statistical evidence as just one source amongst many, and explicitly includes informal knowledge gained from work experience or service use:

> *There is a great deal of critical evidence held in the minds of both front-line staff … and those to whom policy is directed.* (Strategic Policy Making Team, 1999)

In practice, as a 2001 survey of policy-making found, a more limited range of evidence appears to have been used by government departments: domestic and international research and statistics, policy evaluation, economic modelling and expert knowledge (Bullock et al. 2001).

As noted by Nutley et al (2003, pp. 31 - 32), in health particularly, reviews of randomised experiments are placed at the apex of the "hierarchy of evidence" for assessing what works. Observational studies and professional consensus are accorded much lower credibility. Understandably, providing evidence of efficacy or effectiveness of selected treatments in the health field has been highly sought after: proof of efficacy of a drug for example has wide commercial ramifications.[2] Nutley et al (2003) suggest that "bitter experience" provides the second reason for the explicit methodological hierarchy: "much empirical research suggests that biased conclusions may be drawn about treatment effectiveness from the less methodologically rigorous approaches" (Nutley, et al., 2003, pp. 31 - 32).

In contrast to the hierarchical approach in health care, other sectors such as education, criminal justice and social care are fractured by disputes regarding 'appropriate' evidence. Also, in comparison to health care, in these sectors there is relatively little experimentation, and divisions between qualitative and quantitative approaches run deep (Davies et al. 2000). This happens in part because of the more diverse social science underpinnings in these sectors, compared to the natural sciences underpinning in much of health care; and in part because of the multiple and contested nature of the desired outcomes. It follows that knowledge of 'what works' tends to be influenced strongly by the kinds of questions asked, and in any case, is largely provisional and dependent on context.

It is tempting to think of evidence entering the policy process as part of a rational decision-making process, when reality is often far more messy and overtly political. There are many caveats, and these are especially relevant for social policy where it is difficult or impossible to do randomised control experiments. However, adopting 'what works?' as a slogan is dangerous, as Roberts (2005, p. 34) argues:

> *Those of us in the evidence-based arena who have been pushing policy makers and practitioners to adopt evidence-based practice need to be careful that we do not sell the "What works?" agenda as a simple way to solve problems. Social and educational interventions are complex and are capable of doing as much or even more harm than medical ones.*

In social policy, such as 'make work pay' and savings policies where behaviour is to be modified, the role of 'objective' evidence may be even more suspect. Perhaps the 'appropriate' conclusion is that any such evidence needs to be considered, but not regarded as providing more than one lens for policy evaluation.

> *The term "evidence-based" when attached as a modifier to policy or practice has become part of the lexicon of academics, policy people, practitioners and even client groups. Yet such glib terms can obscure the sometimes limited role that evidence can, does, or even should, play. In recognition of this, we would prefer "evidence-influenced", or even just "evidence-aware", to reflect a more realistic view of what can be achieved.* (Nutley et al., 2003, p. 30)

2. For example, many statistical studies have shown that red wine in moderation has benefits for health. This has influenced drinking behaviour and has been good for the wine industry. Interestingly, wider, more qualitative examinations have cast serious doubt on these findings.

2.2 The New Zealand experience

As in the UK, evidence-based approaches to social policy became popular in New Zealand, with two major social policy conferences around this theme in 2003 and 2004. In her opening address to the 2003 *Social Policy Research and Evaluation Conference,* Dame Anne Salmond suggested "biculturalism or Treaty-based policy, New Zealand's aspirations to be a fair and prosperous society, and providing a good life for our children" are three areas of contemporary social life in New Zealand where evidence-based approaches to policy-making could make a huge difference (Salmond, 2003, p. 3). She went on to comment:

> *At present we are facing some formidable social challenges, partly as the result of past ideological adventures…. With insightful questions, research can get to grips with the issues and the data can be managed. Integrity among researchers and those who commission research, high standards of enquiry, and generosity of spirit can work wonders, changing people's lives for the better.* (Salmond, 2003, p. 5)

Later, in 2004 another Social Policy Research and Evaluation conference was held called 'What Works'. The belief that objectivity and hence better policies would result from taking an evidence- based approach was thus further strengthened.

Wylie (2006, p. 8) noted the increased interest in more sophisticated analysis was often not matched by the budget or time frame for the research, including the translation of that work into policy or lay language. She suggested that there were three barriers to the use of research-based evidence. First, an unrealistic desire for statistical "proof" of causality, either for a policy or initiative as a whole, or individual elements of it; second, "not wanting to hear that something is not working as intended"; and third, "not wanting any reference to other studies or parallel endeavours". Perhaps the expectations are just too high:

> *... the reality is that no single statistical study of policy impact or evaluation of an initiative is going to be able to provide definitive proof, or be able to definitively isolate the action of one aspect from all others.* (Wylie, 2006, pp. 8 - 9)

Statistical methods, designed for an idealised world, rely on some assumptions that cannot always be met, particularly in the social world. Also, unless the samples are large, and a large range of factors are included, statistically teasing apart aspects that usually occur together is difficult to do. Very few New Zealand or indeed international studies have such samples, or include the factors needed to develop the deeper understanding that policy makers (and others) need (Wylie, 2006, pp. 8 - 9).

The evidence-based approach has, however, a very wide appeal to social scientists who may wish to have their research perceived as objective and quantitative and therefore more credible:

> *It is difficult to imagine anyone aguing that policy should be based on anything but the best available evidence. The concept of evidence based policy has an intituitive common sense logic, which partly explains how it has become naturualised in a diverse range of policy settings.* (Marston & Watts, 2003)

This paper looks at WFF as a New Zealand case study in using the simple framework developed in section 3 and comments briefly on KiwiSaver as another example where evidence-based evaluation is fraught with difficulties.

3 ECONOMICS AND THE POLICY FRAMEWORK

There are a number of ways to set out an economics framework for policy analysis. One suggested sequence is:

1. Clarify the problem. Measurement of ‘what is’;
2. Set clear objectives (aims) for policy. Involves normative judgements;
3. Make aims measurable or quantifiable and explain trade-offs which may also involve normative judgements;
4. Set out the theories or models that inform policy development. Ex-ante evidence including relevant overseas experiences to back this up;
5. Examine a full range of policies that might, according to theory and empirical evidence, achieve the objectives;
6. Select the best policy on ground of criteria of cost–effectiveness, economic efficiency, equity and administrative simplicity;
7. Implement policy as intended with efficiency and timeliness;
8. Measure outcomes: what is the evidence that the policy is working to achieve the objectives as set out?
9. What are the unintended consequences?
10. Evaluation, including the whether the ‘problem’ has been addressed, how the criteria of cost–effectiveness, economic efficiency, equity and administrative simplicity have been met, and whether the previous steps re the economic theorising or policy design need to be revisited and the policy modified.

There is potential for confusion and a loss of clarity at each of the 10 steps as set out above. At each step different kinds of evidence may be used, not just the conventional statistical and quantitative ones. While the process might be set up to look objective, there is a large element of subjectivity and normative thinking at each stage of the process, leaving huge scope for disagreements. Evidence has a role at all steps in this process, but is itself capable of manipulation and subjectivity. Always, the ‘question’ determines the possible ‘answer’. Thus policy-making is bound to be a messy process, at best. There is also scope for political interference with the process, for electoral advantage perhaps.

One of the points of the process that often gets lost is the first point: asking what the problem actually is. The framework assumes a rational ordered approach to an indentified problem. In practice, social policy analysts are often confronted with an existing policy that may have been in place for some time or have been implemented on a political whim. In order to establish what the actual initial policy problem was, it is often necessary to examine preceding political statements, discussion papers, and legislation. Unfortunately this can throw up confused and contradictory elements. In some cases the policy framework might even need to be used to clarify what the objectives ought to have been.

Because the aims of policy must be measurable for there to be a quantitative evaluation, the underlying social policy outcome, judged on broader considerations may be forgotten, which is why point 10 is of critical importance:

> *… 'what works' depends as much on who asks the question as who answers it. How should a good result be measured? Does it 'work' if it meets the objectives of the policy? Or should it be assessed according to a set of higher order principles capable of transcending political ideologies and good intention?* (Durie, 2004, p. 2)

4 Case Study 1: Working for Families' In Work Tax Credit

New Zealand's In Work Tax Credit (IWTC) was introduced in 2006 as part of the Working for Families (WFF) assistance package as outlined in St John (2006) and for a full review and discussion of the IWTC see St John and Dale (2010 Forthcoming). This case study does not attempt to review the actual policy process in detail. That has been set out in documents obtained in CPAG v the Attorney General (2008) and involved an extensive policy development process from 2002 to 2005.

4.1 Clarify the problem. Measurement of 'what is'

The problem that WFF was to address, apart from a much needed inflation catch-up in child related financial support, was twofold: the perceived lack of workforce participation by beneficiaries; and child poverty. The two problems were not unrelated, as child poverty was perceived to be a problem of lack of workforce participation. The high level objective can be seen to improve the lives of families with children.

4.2 Set clear objectives (aims) for policy. What should be done involves normative judgements

Dalgety, Dorsett, Johnston, & Spier (2010) noted the objectives of WFF set out by the Cabinet Policy Committee (2004) were:

- to make work pay by supporting families with dependent children so work effort is rewarded;

- to ensure income adequacy, with a focus on low and middle income families with dependent children to address issues of poverty, especially child poverty;
- to achieve a social assistance system that supports people into work, by making sure that people get the assistance they are entitled to, when they should, and with delivery that supports them into, and to remain in, employment.

For purposes here, we focus on that part of the WFF package called the ‘In Work Tax Credit’ (IWTC), specifically designed to achieve bullet points one and two. The other aspects of improved childcare subsidies in WFF were also to support these goals of the IWTC.

The focus of the work incentive was to be sole parents. In New Zealand’s tight labour market of 2004-2007, sole parents on the Domestic Purposes Benefit (DPB) were viewed as a source of untapped labour. Dalgety, Dorsett, Spier & Johnston (2009, p. 1), for example, note that WFF was designed to impact on the employment of sole parents “who in recent decades have had a relatively low level of engagement with the labour market”.

As in many developed countries, child poverty elimination had also become a focus of policy attention. After a dramatic rise in child poverty in the 1990s, the New Zealand government pledged to make it a top priority (Ministry of Social Development, 2002). This concern was clearly a factor in designing the IWTC, and is explicitly stated as a goal of the WFF package (Centre for Social Research and Evaluation, 2007b).

So the IWTC part of the WFF policy was supposed to get sole parents back to work and reduce child poverty. Innocuous yet worthy goals?

4.3 Make aims measurable or quantifiable and explain trade-offs which may also involve normative judgements

The choice of the techniques to measure outcomes is highly subjective. The work objective was to be measured by how many sole parents were moved off the DPB. The poverty objective was to be measured by the numbers of children falling below two poverty lines; 60% and 50% of the before housing costs equivalised median income (Perry, 2005). Data for 2004 showed that New Zealand 15% of children fell below the OECD defined 50% before housing costs line. WFF was projected to reduce this by 70% and lift New Zealand to child poverty rates similar to those enjoyed by Scandinavian countries.

Both measures were mere numbers. For example a head count of those falling under any poverty line is a crude measure of success. Moving sole parents off benefits is not a good measure of whether life has improved for them or for their children.

The unacknowledged trade-off was that using one instrument to achieve two goals was going to compromise at least one of the objectives. To ‘make work pay’ using a tool that was designed to meet the costs of children and reduce their poverty inevitably meant that some children would be treated differently to others. In the absence of a commitment to reducing the poverty of all children, some would be 'left behind' (St John & Craig, 2004).

4.4 Set out the theories or models that will inform policy development. Ex-ante evidence including relevant overseas experiences to back this up

The basic theorising was that work provides the route out of poverty but that work incentives were required and they would be effective. Thus a work incentive should 'make work pay', so that an income gap is created between those in paid work and those not in paid work. Internationally, it had been acknowledged that creating such a gap with work-related incentives, usually referred to generically as ' in work benefits' should not make the latter group worse-off as that would simply increase poverty (Pearson & Immervoll, 2008, p. 2).

The theorising was crude as the IWTC was to be a most unusual 'in work benefit' (IWB). It does not reward an extra hour of work, as would for example a lower tax rate, but provides a lumpsum to families who met the work-based criteria. Little attention was paid to any 'income effect' of the IWTC which would unambiguously act to reduce work effort, while the substitution effect towards more work was only operable up to the point of qualification of 20 hours a week for a sole parent and 30 hours for a couple. The incentive was to reach the required hours of work, and no more.

There appeared to be no theory or international basis for applying the IWB only for those with children. Only six countries use a child-related work incentive: New Zealand, Austria, Belgium, Ireland, South Korea, the Netherlands, and the Slovak Republic (Human Rights Tribunal, 2008, p. 10). Although New Zealand's IWTC appeared to have been influenced by the UK tax credit system, a vital difference is that since 2003, Britain's child-related assistance has not differentiated between working and non-working families.

The UK child-related per-week payments comprise a universal child benefit, the Child Tax Credit (CTC), paid for all children on the same basis, regardless of the source of the parents' income (Millar, 2008, p. 23). After the UK's 2003 reforms, the CTC recognized the costs of children in low income families, and the Working Tax Credit (WTC) provided the incentive to work. To qualify for the WTC, a lone parent must work for a minimum of 16 hours a week, and a person without children must work for 30 hours. Lone parents receive a larger payment than two-adult families. The WTC is aimed at the transition to work and abates from a low level at a high rate; is not related to the numbers of children; is available to all low-income workers; and is paid to the worker, not to the principal caregiver of the child as it is in New Zealand.

As in the UK system, the US's Earned Income Tax credit (EITC) is aimed at the transition to work, and is paid to the worker. It is not restricted to those with children, although a stated objective in 1993 was to lift lone parents in full-time, low wage employment out of poverty (Pearson & Immervoll, 2008, p. 31). The EITC subsidises extra hours worked up to a maximum determined by the number and age of children, and then it claws that subsidy back.

The EITC entitlement, assessed on 'earned income', is designed to offer a subsidy to low earnings and operates over three ranges of income. In the first 'phase-in' range, the credit increases as income increases. In the second "plateau" range, the credit remains constant; and over the third 'phase-out' range it is abated (Pearson et al, 2008, p. 24). While the EITC does not distinguish between single parent families and two parent families, unmarried couples can file separately and each receive the tax credit, but earned incomes of married couples are aggregated. If their aggregate income puts them in the phase-out range, the perverse incentive is for couples not to marry, or if married, to separate and receive individual EITC credits (Ellwood, 2000).

Evidence, gathered in buoyant economic times, shows both the UK WTC and the US EITC had a positive effect on the labour force participation rate of single women with children. This was clearly influential in New Zealand policy development even though the IWTC policy is very different in design as Table 1 shows.

Table 1. Comparing In Work Benefits: UK, US, and NZ

UK- WTC	US EITC	NZ- IWTC
Minimum hours worked required, Adult based	No minimum hours, Adult based	Minimum hours worked required, Child-based
Paid to worker	Paid to worker	Paid to carer
Abates from very low income level. Affects transition to work	Phased in over low income and phased out over transition	Abates from level above transition
Abates quickly	Abates moderately quickly	Abates slowly

Source:(St John, Dale, & Littlewood, 2009)

Almost no attention was paid to Australian policy where there has never been a child-related payment to encourage work and where all low income children are treated the same (St John & Craig 2004).

4.5 Examine a full range of policies that might achieve objectives according to theory and empirical evidence

A full examination of the way workforce participation can be encouraged by enhancing returns to labour was not part of the WFF policy development process. This spectrum of possible ways includes generic tax reductions, changes to the level of benefits or their abatement that reduce effective marginal tax rates (EMTRs);[3] minimum wage legislation; "welfare to work" case management; and labour market regulations.

Ignoring the approach in Australia where only minor targeted use has been made of work incentives, (Pearson e al 2008) the choice was to implement an extensive and expensive cash-based, child related targeted tax credit designed to also impact on poverty levels.

One of the political attractions of work incentives is that they appear able to achieve both employment and distributional goals at the same time (Pearson et al, 2008, p 2). Also, there may be a stronger political constituency for the fiscal cost of supporting people in paid work as opposed to supporting them in the welfare system. These were influences on the IWTC policy development as was an increasing focus on paid work as the way, indeed it seemed to the policy makers, the only way out of poverty.

It is possible, and appears to have been the case in New Zealand, to be trapped by circular thinking as the following illustrates:

> *[IWBs] provide additional benefits to low-income families, so reduce the incidence of poverty among those families with children. They also increase the incentive to work, and as poverty rates among those in work are lower than those out of work, any increase in the number of parents moving into work would reduce child poverty.* (Pearson et al, 2008, p. 13)

Is *work* itself the way out of poverty or *only* because the state *makes* it pay with sufficient subsidies? Sometimes the justification is the belief that any attachment to the labour force creates future opportunities for well paid work that will eventually provide an unsubsidised path out of poverty. Recent research in Canada (Card & Hyslop, 2005), and Minnesota (Gennetian, Miller, & Smith, 2005) has thrown into doubt the value of work incentives for achieving any such long-term benefit, but this evidence was not adduced in the IWTC development

Nor was attention paid to the obvious: if work incentives are used as a major and direct method of addressing poverty, there is a risk of creating and perpetuating an underclass who cannot access the incentive and who thus must remain, by definition, in poverty.

A further problem may arise in times of recession:

> *Because severe economic downturns can have marked effects on the earnings distribution, policymakers should review whether the eligibility conditions and payment profiles of existing IWBs are appropriate or should be adapted in order to exploit their potential as a measure that cushions income losses during a recession.* (Immervoll et al., 2009, p. 46)

3. The EMTR is the combined loss from tax and benefit abatement when an extra dollar is earned. A poverty trap for low income people may occur when earning extra income does not materially increase disposable income. The nature of New Zealand's welfare payments are described in St John & Rankin (2009).

This stage of the policy development was informed by 'evidence' from simulation exercises but does not appear to have been thoughtfully used to moderate the policy approach.

A March 2004 paper to Government Ministers on the expected impacts of WFF on employment reached the following conclusions:

> *As a result of the reforms there may be a small increase in labour market participation amongst both beneficiaries and non-beneficiaries who decide to enter the labour force..., this may lead to a small increase in employment… In the event of an economic downturn, employers are less likely to absorb any increase in labour supply generated as a result of improved work incentives. Families already engaged in work are not generally responsive to financial work incentives and may, depending on the structure of assistance, reduce the work effort of second earners in dual-income households.* (Ministry of Social Development, 2004, paras 89-92)

A micro-simulation exercise found an increase in the probability of working was largest for one-child families because the increase in IWTC over the previous tax credit was largest for the 1-child family (Kalb, Cai, & Tuckwell, 2005). For couples, only 8% in the sample observations worked less than 30 hours a week and for these, higher incomes could induce a reduction in labour supply of one or both partners. The labour supply of couples was predicted to decrease, and: "For married women, the most popular choice is to reduce labour supply to zero" (Kalb, Cai, & Tuckwell, 2005, p. 13).

As a result of the WFF package, around 1.8% of lone parents were expected to enter the labour force; about 2.4% were expected to work less; and about 1.9% were predicted to prefer longer working hours, with an overall average increase of less than 1 hour a week. The authors note "the effect is probably largest for low-wage families where the price of childcare might otherwise cancel out nearly all additional earnings to be obtained from additional working hours" (Kalb, Cai, & Tuckwell, 2005, p. 24).

Analysis by Nolan (2004), St John & Craig (2004) and Dwyer (2005) also concluded that the WFF package was unlikely to have any noticeable net positive effect on aggregate employment; and provided no encouragement for secondary income earners to seek employment where one parent was already working 30 hours a week. Once a family qualifies, there is no extra incentive to work more but there is a disincentive to work above the threshold for abatement (Nolan, 2002). The MSD expected only a modest increase in labour market participation, and expressed concern that it may have consequences for other labour market policies:

> *Should there be an increase in labour supply, any downward pressure on wages would reduce the returns to work for people without children and make these workers increasingly reliant on minimum wage provisions.* (Ministry of Social Development, 2004, paras 89-92)

In the meantime, before the policy was implemented, the government announced a greatly expanded role for the IWTC, raising the threshold for abatement of WFF tax credits from $27,500 to $35,000, and reducing the abatement rate to 20%. A priori, this was expected to provide a strong incentive to increase work effort to the new threshold. For those above the new threshold the effect was more ambiguous.

4.6 Select the best policy on ground of criteria of cost–effectiveness, economic efficiency, equity and administrative simplicity

The IWTC was the chosen centrepiece as part of WFF's broad package of tax credits as described in Table 2. The IWTC provides $60 a week for families 'in work' with up to three children, plus $15 a week for the fourth and subsequent children. It is lump-sum in nature, requiring the double test of eligibility: off benefit' and 20 hours a week of paid work for a lone parent, 30 hours for a couple.

Table 2. Working for Families Tax Credits

Tax Credit	Abbreviation	Nature of payment
Family Tax Credit	FTC	Child-related weekly supplement
In-Work Tax Credit	IWTC	Child-related weekly supplement with work requirement
Minimum Family Tax Credit	MFTC	Minimum Family income top- up with 100% abatement
Parental Tax Credit	PTC	Paid NZ $150 a week for 8 weeks for new child

Source: IRD, https://interact1.ird.govt.nz/forms/famcalc2008/

The original package announced in the 2004 budget government was criticised as leaving out the poorest children (St John & Craig, 2004), but the government claimed to have no money left to help beneficiary families. Then in a surprise move pre-election in 2005, an additional NZ$500 million per year was found to augment the WFF package. The threshold for the WFF's joint parental income test was raised significantly and the abatement rate was reduced from 30% to 20%. As the IWTC abated last, higher income families suddenly became eligible. By the time the complete WFF package was fully implemented in 2007, with parameters as set out in Table 3, there was no clear connection between the original rationale for the IWTC and its final form.

All low income families with children are entitled to the per-week per-child FTC as set out in Table 3. If they are in work, lone parents at least 20 hours per week, and couples at least 30 hours per week, and not in receipt of a benefit, they may also be entitled to the IWTC and perhaps the MFTC as well. These three tax credits are paid as one sum to the caregiver. The amount paid depends on the combined annual income of the family, and the number and age of dependent children, and is reconciled at the end of the year. However, if more joint income has been earned than anticipated, some of the WFF tax credits may have to be repaid.

Table 3. WFF weekly and annual payment from 1 October 2008

Weekly payments: Age and number of children	**FTC Rate from 1 October 08 NZ$**	**Lump sum, end of year payments Annual rate 2009/2010 NZ$**
First child if under 16	$86	$4487
First child if 16, 17 or 18	$99	$5198
Subsequent rate if child under 13	$60	$3119
Subsequent rate if child 13 to 15	$68	$3557
Subsequent rate if child 16 or over	$89	$4651
"Working Families" Numbers of children (aged under 13 years)*	**IWTC Rate from 1 October 08**	**IWTC + FTC**
1	$60	$7,607
2	$60	$10,726
3	$60	$13,845
4	$75	$17,744
5	$90	$21,643

* Higher amounts are paid for older children. WFF abates from $36,827, so a family with 5 children can still access some part of the IWTC up to an income of $145,042. (IRD http://www.ird.govt.nz/wff-tax-credits/entitlement/

The different tax credits are treated differently for abatement purposes. The MFTC is a top-up payment to guarantee a minimum level of income for families in work, and the level is not child-related. In 2009/10 it ensures that families have a minimum income of NZ$395 a week after tax or $20,540 net (approximately 51% of the net average wage) and abates at 100% for every extra dollar earned. The FTC and the IWTC are the means by which the costs of extra children are recognised, over and above the MFTC income floor (see Table 3). They are added together and abate from a household income of $37,000 at a rate of 20%, with the IWTC being the last to abate. The PTC is also paid for a small number of families with a new baby.

Eligibility for the IWTC requires minimum hours of worked per week as outlined above. However, income from such hours of work may be insufficient to encourage an exit from the benefit system. To encourage such an exit, incomes may be topped up to a guaranteed minimum with the MFTC.

Lacking at the development stage or the point of adoption of the IWTC was a clear critique of how the policy achieved the objectives and fulfilled the criteria of cost-effectiveness, equity efficiency and simplicity.

4.7 Implement policy as intended with efficiency and timeliness

Various evaluation reports have concentrated on implementation and largely judged it to be a success. The 2007 *Receipt of the Working for Families Package*, for example, reported:

> *The evaluation findings confirm the successful communications and delivery of the WFF package. Original forecasts have been met or exceeded, and there are high levels of*

> *awareness and receipt of WFF. In particular, there have been dramatic increases in the number of working families receiving WFF components…. The 2005/2006 tax year has exceeded the forecast of 260,000 families benefiting from WFF. There is high public awareness of WFF (88%) and recognition of its advertisements (91%) among families. Overall, the number of people receiving WFF continues to increase. The increase among non-beneficiaries is especially pronounced….* (Centre for Social Research and Evaluation & Inland Revenue Te Tari Taake, 2007, p. 1)

However, some of the criticisms as noted above by Dale et al. (2010) were acknowledged:

> *From survey results, we know that potentially eligible families who are not receiving their WFF entitlements say that they need more information about eligibility and how to apply. MSD and IRD are continuing to actively engage with families (eg through outbound calling), have made evening appointments available, and have promoted WFF in public places and at worksites… Some families participating in the communications evaluation survey said that they have concerns about overpayment of WFF Tax Credits. Families have the option of receiving full or part payment at the end of the tax year. One of the initial goals of WFF is that families receive payments throughout the year, thereby ensuring that people receive the payments they are entitled to when they need it. MSD and IRD are continuing to work on reducing the level of overpayments.* (Centre for Social Research and Evaluation & Inland Revenue Te Tari Taake, 2007, p. 2)

4.8 Measure outcomes: what is the evidence that the policy is working to achieve the objectives as set out

It is difficult in a real world analysis to disentangle the effect of any single measure such as the IWTC, when numerous factors, including minimum wage legislation and the economic environment, may change simultaneously. Policies designed to incentivise work effort may appear to work well in buoyant times when labour is scarce, but appear ineffectual when jobs are scarce.

Three methodologies for evaluating the impact of IWBs on the work incentive goal were set out in Pearson et al (2008, p 25) as follows:

- Experiential. Pilot studies allow for a control group not affected by policy;
- Simulation. Projection of likely effects before policy, based on population characteristics;
- Statistics approach. Difference-in-Differences method eliminates other effects by use of statistics and is most generally used.

Points one and two are not relevant because New Zealand has no regional or state differences in policy, and the IWTC is a national programme without a prior pilot study. The third ex-post evaluation methodology is discussed below.

The first official evaluation of WFF acknowledged the difficulty of assessing the work incentive aspect of the IWTC. The qualitative finding that "nearly two thirds of families agreed that the In-Work Tax Credit is a good incentive to stay off a benefit" (Centre for Social Research and Evaluation, 2007, p. 38), can be dismissed as simply serving to illustrate the gap between those who are able to work and those who cannot.

In their 2007 report, the Centre for Social Research and Evaluation stated:

> *..since WFF has been implemented, New Zealand has experienced the largest fall in numbers receiving DPB since the benefit was introduced in 1973 – the number of families receiving the DPB has fallen by 12,500 (from 109,700 at August 2004 to 97,200 at August 2007).* (Centre for Social Research and Evaluation, 2007b, p. 40)

While exits from the DPB were indeed occurring at a faster pace between 2004 and 2007, this is not evidence of the efficacy of the IWTC because:

- The labour market was exceptionally tight and unemployment itself fell rapidly until 2007;
- The WFF threshold was higher and abatement lower;
- Demand for unskilled labour was high;
- Child care subsidies and supply increased markedly;
- Minimum wage was increased from $9 to $12 or 33% between 2004 and 2008;
- Case management can be expected to have helped;
- Exits from benefits may not have entailed a significant increase in hours of work;
- The role of the MFTC is ignored.

While claiming that the IWTC "appears to be contributing to the decrease in the numbers of DPB recipients" (Centre for Social Research and Evaluation, 2007b, p. 36), the Report acknowledged that the fall in numbers was greatest among those people with an existing attachment to the labour market.

Table 4. Number of DPB recipients [lone parents] with/without income in addition to benefit (NZ$)

DPB Recipients at end of August	No other income	With other income ($ per week)					Total
		All with other income	$1-80	$81-180	$181-300	>$300	
2005	80,000	26,400	8,900	9,000	5,900	2,600	106,400
2006	78,000	23,600	8,200	8,200	5,200	2,000	101,500
2007	75,600	21,600	7,300	7,500	4,800	2,000	97,100
Change							
2005-2006	-2,000	-2,800	-700	-700	-700	-600	-4,800
	-3%	-11%	-8%	-8%	-12%	-25%	-5%
2006-2007	-2,400	-2,000	-900	-700	-300	0	-4,400
	-3%	-8%	-11%	-8%	-7%	-1%	-4%
2005-2007	-4,400	-4,800	-1,600	-1,400	-1,100	-700	-9,200
	-6%	-18%	-18%	-16%	-18%	-25%	-9%

Source: MSD Information Analysis Platform (IAP).
Note: Numbers may not sum due to rounding. Percent change calculated on unrounded data.

Source: Centre for Social Research and Evaluation, 2007, p. 40.

Between 2005 and 2006 there was a 25% fall in numbers for those earning over $300 a week (see Table 4 above). Of these, “some would have been already working sufficient hours (20 a week) to qualify” (Centre for Social Research and Evaluation, 2007b, p. 40).[4]

The IWTC was introduced in 2006. Between 2006 and 2007 only 2,400 of the reduction of 4,400 were not working at all and most of these would not have gone into full-time work. Of those who did go into full-time work, the higher threshold and reduced abatement on the amount of their WFF tax credits entitlement may have provided a more valuable incentive than the IWTC.

The official evaluations ignored the fact that a significant number of those who moved off a part benefit on to the IWTC needed the income top-up provided by the MFTC, as Table 5 confirms. From 2006 to 2007, lone parent numbers on the MFTC jumped from 621 to 2,167. By 2008, the average entitlement was NZ$2,801. The MFTC can in fact result in lone parents receiving more from the state when they are off the DPB than when they are on a part benefit (Gray, 2008).

Table 5. Data for the receipt of the Minimum Family Tax Credit (NZ$)

	2004	**2005**	**2006**	**2007**	**2008**
Numbers entitled to MFTC (year)	1,412	1,007	863	2,727	2,397
% Couple	29%	30%	28%	21%	16%
% lone parent	71%	70%	72%	79%	84%
entitlement Mean	**$2,274**	**$2,298**	**$2,325**	**$2,748**	**$2,801**
25th percentile	$625	$827	$786	$985	$881
75th percentile	$3,418	$3,427	$3,493	$3,945	$4,139
Total WFF tax credits **Mean**	**$7,618**	**$7,650**	**$9,632**	**$10,914**	**$12,025**

Data from Inland Revenue, personal communication (2008)

As far as the poverty objective is concerned, evaluation identified a fall in child poverty rates on the 60% line from the whole WFF package (Perry 2009), but the predictions of reduced child poverty based on the 50% line did not materialise. No attempt was made query the role of the IWTC, or how those excluded from it, largely children in ineligible families, remained in poverty. Evidence of severe poverty problems were identified in June 2007 (Centre for Social Research and Evaluation, 2007a) but this did not feed into the analysis of the effectiveness of the IWTC and the poverty objective.

Evidence-based statistical evaluations

Some preliminary research on WFFs using a difference-in-differences approach found some small impacts on employment (Fitzgerald, Maloney, & Pacheco, 2008).

4. A confounding factor not addressed in any analysis to date is the role of Child Support (payment from the non-custodial parent), which may have profound work incentive implications. Child Support payments offset the DPB and many lone parents prefer a part-benefit even when working so that their income is effectively secured. Child Support problems can be expected to worsen in the recession.

> *… we provide some evidence of employment increases and more solid evidence of work hour increases for those working due the family assistance policy [tax credits] changes. Evidence on partnering is more elusive but there are certainly no large impacts currently.* (Fitzgerald, Maloney, & Pacheco, 2008, p. 48)

It would appear that the statistical approach employed by Fitzgerald et al (2008), while problematic, largely confirmed the earlier results from micro-simulation. However, they did show that partnered women increased their hours of work, an outcome that may reflect the positive incentive effect of the reduced abatement of family assistance and the higher than originally intended threshold introduced with the pre-election 2005 changes.

A further difference-in-differences study (Dalgety, et al., 2010) for the 2004-2007 period was based on data from the Household Labour Force Survey (HLFS) and Income Survey (June Quarters), and a Survival analysis of recipients of the lone parent benefit (DPB), MSD and IRD linked data. The methodology took lone parents as the ‘treatment’ group and single people aged 18 to 65 years as the non-treatment group and examined their respective employment rates over the period in which WFF was introduced.

Two-thirds of the estimated 9.4 percentage points (pp) increase in lone parents employed was attributed to the effect of WFF, with the remaining one-third ‘most likely capturing the strength of the economy over this time’ (p 27). This increase occurred throughout the 3 year period as shown in Table 6 but the results were only statistically significant (6.0 ±4.4 pp) if 2007, the year following the introduction of the IWTC, was included.

Table 6 Impact of WFF on employment of lone parents

	Comparison between 2004 and		
	2005	**2006**	**2007**
Actual change since 2004 (pp)	3.1	5.8	9.4
Impact estimate (pp)	2.6	3.9	6.0
p-value	0.250	0.097	0.007
Sample Size	13,197	12,734	13,629

Source (Dalgety, et al., 2009).Shaded cells are statistically significant (at p<0.05) Difference–in-Differences estimates. Source: Calculations based on unpublished Household Labour Force Survey data, June quarters

These results were for numbers of lone parents in paid employment for least one hour a week, however the work incentive provided by the IWTC did not come into play until 2006 and required lone parents to be working 20 hours at least. Dalgety et al (2010, p. 28) found:

> *the percentage of [lone] parents meeting the eligibility threshold for the in-work tax credit requirement increased from 35.9% in June 2004 to 47.5% in June 2007, an increase of 11.6 percentage points. Difference-in-Differences analysis suggests that around three-quarters (9.2 pp ± 4.5 pp) of this increase was due to effect of WFF.* (Dalgety et al, 2010, p28)

While the increase in lone parents working 20 or more hours by 2007 was 'statistically significant' it is not all attributable to the IWTC which did not apply until 2006. Nor should 'statistical significance' be inferred to indicate 'social significance' (McCloskey & Ziliak, 2008).

More fundamentally, the methodology can be questioned.

There are occasions when a difference in difference approach may be valid, but there are huge cautions: the difference in difference approach typically considers the policy reform itself as an experiment and tries to find a naturally occurring comparison group that can mimic the properties of the control group in the properly designed experimental context.

> *[It relies] on the two critically important assumptions of common time effects across groups and no composition changes within each group. Together, these assumptions make choosing a comparison group extremely difficult. For example, in their heavily cited evaluation study of the impact of Earned Income Tax Credit reforms on the employment of single mothers in the US, Eissa and Liebman (1996) use single women without children as a control group. However, this comparison can be criticised for not capturing differential macro effects. In particular, this control group is already working to a very high level of participation in the US labour market (around 95%) and therefore cannot be expected to increase its level of participation in response to the economy coming out of a recession. In this case, all the expansion in labour market participation in the group of single women with children will be attributed to the reform itself.* (Blundell & Costa Dias, 2000, p. 429)

The Dalgety et al (2010) 'difference in differences' approach compares the growth in employment of single persons aged 18 to 65 with that of lone parents. The choice of this comparator in a time of the already very high overall employment of this group is questionable. In addition single unemployed people face a much higher abatement of their earnings from part-time work (91% after 6.4 hours) than lone parents, a fact made worse as the minimum wage increases. The purpose is to discourage the choice of part-time work and a part unemployment benefit in favour of full-time work. Thus these two groups are very dissimilar yet may be competing in the same part of the labour market.

It is highly likely that some employment of those who are subsidised may be at the expense of those who are not. Displacement may also occur between partnered women and lone parents as the incentives for partnered women to work are reduced (Ministry of Social Development, 2004, para. 132).

Dalgety et al (2010) were strangely silent on the role of the MFTC and its sudden increased use after 2006 (see section 3.5). Single people without children are ineligible for the MFTC and do not have this income top-up if they work only 20 hours. For lone parents, moving off-benefit does not necessarily equate to moving to full-time employment or even significantly increased hours. The MFTC, which is arguably just a benefit in a different guise, provides a very strong work disincentive to those working 20 hours to increase their hours beyond that.

In the case of evaluation of the IWTC, the two groups are ludicrously different. Unemployed people without children aged 18-64 and sole parents on the DPB are vastly different in age and gender structure, yet to some extent compete in the same part of the casualised labour market. By 2010 it was clear that there had been a loss of confidence in this kind of evaluation. The MSD regulatory statement of the future focus policy said:

> *There is no research currently available which accurately quantifies the size of the behavioural response from these changes in policies. This prevents estimates, with the degree of accuracy required, from being made of the number of people who will move from benefit to work over a year, as a result of the proposed changes. The inability to determine firm numbers of people shifting from benefit to work as a result of these changes is due to the difficulty of separating out the effect of the policy changes from the effect of changes in other influences such as economic and labour market settings (e.g. employment growth, minimum wage increases). Some broad estimates of magnitude can be made based on previous experiences of similar policy changes. The Ministry of Social Development will also monitor and evaluate the package of changes to determine, to the extent possible, the impact of these*. (Ministry of Social Development, 2010, p. 1)

Perhaps this means that the MSD now believes that their IWTC evaluation did not actually quantify the size of behavioural change and they have since they lost faith in the difference in differences approach. Nevertheless it remained the only evaluation of WFF that attempted to relate the outcomes to the objectives. Having done this the evaluation team were abandoned and there are apparently no further plans to repeat these studies.

Time-frame issues

A problem of this kind of evaluation is that the results of whether the policy is a success are usually required sooner rather than later, for political justification. The Dalgety evaluation was not of sufficient duration to see whether any of the increase in employment was sustained over time. Table 7 shows that between December 2007 and December 2009, combined benefit numbers rose by around 76,000. Those on the DPB increased sharply, by around 11,000, more than reversing the gains found in the Dalgety study. While these may not be the same people, it suggests that the IWTC has not operated to help sustain employment in the downturn.

Table 7. Numbers of working aged clients receiving main benefits, end of December, 1999 – 2009

End of quarter	Unemployment Benefits	Domestic Purposes Benefits	Sickness Benefits	Invalid's Benefits	Other main benefits	All main benefits
December 1999	161,128	110,285	32,870	52,195	44,937	401,415
December 2000	146,692	109,663	33,560	56,711	45,681	392,307
December 2001	133,732	109,047	36,191	61,369	41,095	381,434
December 2002	118,139	109,290	39,426	65,779	34,825	367,459
December 2003	94,348	111,065	42,357	69,588	31,489	348,847
December 2004	65,969	109,339	45,648	72,543	26,200	319,699
December 2005 WFF introduced	51,426	106,302	46,862	74,500	22,993	302,083
December 2006 IWTC	38,796	100,309	48,650	76,816	22,070	286,641
December 2007 WFF fully implemented	22,748	98,154	49,093	80,082	19,655	269,732
December 2008	30,508	100,282	50,896	83,501	20,989	286,176
December 2009	66,328	109,289	59,158	85,038	25,663	345,476

Source: National Factsheets MSD http://www.msd.govt.nz/about-msd-and-our-work/publications-resources/statistics/benefit/2009-national-benefit-factsheets.html (Ministry of Social Development, 2009)

4.9 What are the unintended consequences?

Regardless of type, IWBs share some fundamental design and trade-off issues. They are supply-side tools that do not directly affect the demand for labour and may reduce the employment of non-targeted groups. As Bryson, Evans, Knight, LaValle, & Vegeris (2006b, p. 9) note, “financial inducements to parents with children to enter the labour market may increase their chances of employment at the expense of adults without children”. They may also actually increase the disincentive to work if/when abated. They are usually difficult to administer and evaluate.

Another of the consequences of the IWTC has been the need for compensating policy changes elsewhere. Workers who got nothing out of WFF were effectively denied a tax cut in 2005 when the programme was extended up the income scale, and the original cost was increased by 50%.

Table 8 gives New Zealand’s income tax scales. WFF tax credits substantially offset income tax for middle and low income families, while others have had very little tax relief.

Table 8 Income tax rates from 1 April 2009 ($NZ)

Annual income	Tax rates
$0 - $14,000	12.5%
$14,001 - $48,000	21%
$48,001 - $70,000	33%
$70,001 and over	38%

The National-led Government elected in late 2008 acknowledged the problems faced by high levels of taxation of some lower-income groups, and the inequities that arise with WFF when New Zealand's economy endures a period of contraction and rising unemployment. In response, they introduced two new compensating policies: the Independent Earner Tax Credit, a tax credit for workers who received nothing from the WFF package; and a package to support newly redundant workers.

The Independent Earner Tax Credit

The Independent Earner Tax Credit (IETC) is a tax credit of NZ$520 a year for 2009-2010 aimed at earners who do not receive an income-tested benefit, New Zealand pension, or WFF assistance. It is available in full for people earning $24,000 per year and abates at 13 cents per dollar over the yearly income of $44,000. Since the threshold is less than the full-time income of someone earning the minimum wage, it was intended to be an incentive to for low income single people (New Zealand Government, 2008).

The EITC however introduces yet more complexity to overcome the perceived anomalies created by the IWTC. It also takes New Zealand further from the path of tax reform adopted in the late 1980s of simple low rate broad base comprehensive income taxation (St John, 2007).

ReStart: the redundancy package

The National-led government's redundancy package, ReStart, comprises ReCover, and changes to accommodation assistance. ReCover pays the IWTC for the first 16 weeks of unemployment to workers with children who are made redundant from full-time work. This effectively acknowledges that the IWTC is a payment to assist with the support of children, and not simply a work incentive.

However, a work incentive payment made to only a selected group of the unemployed is likely to produce a new raft of inequities and inconsistencies; an issue bought into sharp focus in the recession.

Human rights breaches by IWTC alleged

The IWTC has been subjected to sustained attack by child rights advocates who have argued it involved unlawful discrimination in terms of Human Rights legislation. In a challenge to the government, a class action was heard in the Human Rights Review Tribunal in June/July 2008.[5] The plaintiff in this case, CPAG v Attorney General, argued that the IWTC is a part of weekly

5. CPAG claimed the IWTC constitutes unlawful discrimination under Part 1A of the Human Rights Act 1993 and breaches New Zealand's obligations under the United Nations Convention on the Rights of the Child. The history, documents and decision of this protracted and expensive exercise can be accessed at www.cpag.org.nz.

family assistance to help meet the needs of low income children, yet access is determined by the parents' employment status. The outcome is to reduce the poverty of those children whose parents are in work while leaving the children of those who do not qualify further behind.

In a 100 page judgment, the judge upheld the claim of discrimination against 230,000 of the poorest children in New Zealand (22% of all children under 18) who do not benefit from the part of their family assistance called the IWTC. Paragraph 192 stated:

> *We are satisfied that the WFF package as a whole, and the eligibility rules for the IWTC in particular, treats families in receipt of an income-tested benefit less favourably than it does families in work, and that as a result families that were and are dependent on the receipt of an income-tested benefit were and are disadvantaged in a real and substantive way.* (Human Rights Tribunal, 2008, p. 72)

The OECD witnesses for the defendant also questioned the appropriateness of an IWB applying so far up the income scale (Human Rights Tribunal, 2008, pp. 4 footnote, 27). While the Tribunal also expressed concern at the lack of human rights considerations in IWTC policy formation, they found the discrimination to be "of a kind that is justified in a free and democratic society" (Human Rights Tribunal, 2008, p. 4). The decision is being appealed in mid 2010.

Long term social cost

As argued in St John (2006), the In Work Benefit policy for families with children in New Zealand requires a more nuanced approach that might:

> *Pay more attention to the lower capability for work that faces many parents; reflecting the parent's health, the numbers ages and health of children and the availability of stable work and child care.* (St John 2006, p. 314)

Paid childcare has been promoted as preferable to parenting one's own child under the 'making work pay' regime.[6] OECD economists are among those who have advised the New Zealand government, and expressed views such as:

> *Providing some parental leave can encourage women to enter the labour market. Providing too much parental leave can mean that they spend so long outside of the labour market that their skills decline.* (CPAG v Attorney General, 2008)

This contrasts with recognition that firstly, the majority of today's children spend large amounts of time in out-of-home care; and secondly, neuroscientific research is demonstrating that loving,

---

6. The commercialisation of early childhood care also holds fiscal risks especially in a downturn when a government, as in Australia, finds they have to rescue the industry they have highly subsidised.

stable, secure, and stimulating relationships with caregivers in the earliest years of a child are critical for every aspect of their development.

> *Taken together, these two developments confront the public and policymakers in OECD countries with urgent questions.* (Adamson, 2008)

Long term social costs be should included in any evaluation of the work first approach along with qualitative studies of whether lone parents and their children experience an actual increase in their well-being as a result of these policies.

4.10 Evaluation including whether the 'problem' has been addressed and whether the previous steps re the economic theorising or policy design need to be revisited and the policy modified

Policy-makers are keen to know: Was everything implemented as planned? Did the eligible population receive what they were entitled to? In addition:

> *Policymakers also want to know what happened and, if it differed from what was anticipated, why it differed. These questions can be addressed through qualitative research identifying what happened, and what people thought about what was happening during the implementation process. Key informants are usually staff at different levels in administering organisations, contracted providers of services and "clients" with experience of applying for assistance. These questions can also be addressed through quantitative analysis of monitoring data, which establishes whether key indicators of the process are going in the anticipated direction*. (Bryson, Evans, Knight, La Valle, & Vegeris, 2006a, p. 2).

But the question as to whether the objectives were achieved in cost-effective way is rarely asked. In the case of WFF and the IWTC formal evaluation process has appeared willing to stop at point 8 of the process. There has been no asking of the questions: Has behaviour actually changed or are people just being shifted from one form of benefit to another? In what way does the IWTC actually affect behaviour? Is the complexity worth it? What policy modifications should be considered: are the required hours of work sensible in a casualised labour market; and how is it working out and being monitored?

Do DPBs really need carrots and sticks or is it more a question of suitable work, adequate transport and child care? What are the gains for the economy from shifting so much unpaid work to a marketised form, seen clearly in the case of childcare, especially when tax-funded subsidies are required?

Moreover there has been no serious attempt to see how the lives of families affected by the policy have changed and what the perceptions of the sole parents actually are. Have the parent's lives and those of their children actually been improved? This required qualitative tools and a much closer analysis of employment and poverty data.

As noted by the OECD, IWBs are not a magic bullet:

> *[They are] costly and must be financed by increased taxes elsewhere or cuts in government spending. So they need to be well targeted and implemented carefully and their interaction with social benefits has to be taken into account. The financial incentive should sufficiently large and the duration of the measure long enough to modify behaviour and improve career prospects.* (Pearson & Immervoll, 2008, p. 3)

If an IWB is not sharply focused on the transition to work, and time-limited, it may well be costly and require higher taxes elsewhere. On the other hand, if it is low-cost and time-limited, it may be ineffective in achieving its broader labour-market attachment goals. It is unfortunately possible to have a very expensive IWB that also fails to achieve a meaningful increase in work effort as, this paper argues, the New Zealand IWTC demonstrates In terms of cost- effectiveness, after the 2005 extension the IWTC was bound to perform badly. Paid well up the income scale and abated last, it pays an IWB to a large number of families who have not need of an incentive to work or to stay off the benefit system. In fairness this was not part of the rational policy making process but reflected a political whim.

In 2003, the Treasury suggested that an appropriate policy solution would distinguish between lone parents with low wage earning potential and costly child-care challenges, and lone parents for whom full-time work is a reasonably attainable goal and remarked that "attempts to heavily subsidise the return from longer hours of low paid work seem relatively pointless and costly" (Hurnard, 2007). Another Treasury document noted that proposals to lift labour force participation amongst young women aged 20-34 should be evaluated carefully because early childhood education is expensive and there is a risk of low value for money from spending in this area (The Treasury, 2005, para. 7).

The estimated annual cost of the IWTC is approximately NZ $500 million with the majority of it going to families on more than the average wage. If the IWTC was initially effective in moving 2,000 beneficiaries into equivalent full–time work, the per beneficiary cost is $250,000. If the extra employment disappears in the recession, as seems to be the case, the cost still remains, and the cost per job becomes infinite.[7]

There were problems with equity including the right to be free from discrimination. Eligibility for the IWTC requires independence from any main social welfare benefit or student allowance. Yet care-givers who meet the paid work requirements, and receive the state age pension, New Zealand Superannuation, may qualify. So, too, may parents receiving earnings-related accident compensation payments, even if they do not work. A partnered woman working 20 hours, whose

[7] This calculation attributes the entire cost of providing income assistance to move people on the DPB to the objective of increasing employment by making work pay, and does not take into account distributional objectives.

partner is unemployed, does not qualify for the IWTC; yet, even though her paid work hours are zero, a woman whose partner works for 30 hours may receive the IWTC.

New Zealand's Māori and Pacific Island populations have experienced a much larger decline in their living standards between 2000 and 2004 than the rest of the population (Ministry of Social Development, 2006). They are disproportionately disadvantaged by their exclusion from the IWTC because they have a younger demographic structure and a lower socio-economic status than the general population (Friesen, Woodward, Fergusson, Horwood, & Chesney, 2008). It could thus be argued that the IWTC in effect treats children of different races differently, discrimination not acceptable under Treaty obligations to the indigenous people, illustrating vulnerability of the IWTC policy.

In terms of vertical equity the IWTC and other changes introduced in the WFF package opened up a bigger gap between families 'in work' and those not 'in work'. The IWTC was designed in a time of economic strength and labour market shortages. In a recession, total family assistance falls rather than rises for those with incomes below the threshold of WFF abatement, and thus it fails to provide the cushion that might be expected from a targeted payment. The loss of the IWTC for the children of the newly unemployed is compounded by very low adult benefit levels available to the parent/s.

Work incentives that are linked to the care of children also exclude the childless, who may be more in need of a work incentive than parents with young children. The fiscal cost is borne in the form of higher taxes or lower government spending.

An OECD report, *Economic Survey of New Zealand*, 2007, acknowledged that the WFF package increased the incentives for some welfare recipients to move towards work or increase their hours of work. They were critical however of the changes that extended the effects of abatement so far up the income scale. The report concluded: "Alternative ways of supporting families without these negative effects on incentives to work could do more to raise living standards and should be investigated further." (OECD, 2007, p. 10)

In terms of meeting poverty objectives, the WFF package was credited with reducing the child poverty rate (Perry, 2008, p. 78).[8] However, official living standards data for 2008 showed that 19% of New Zealand's children still lived in 'serious hardship' a much higher percentage than for the older population (Perry, 2009). These were largely children in families on benefits who did not qualify for the IWTC. The solution of paid employment for their parent or parents, regardless of how inappropriate or infeasible that might be did not seem to 'work' .

8. Using the poverty line of 60% current after-housing costs equivalised median household income, child poverty fell from 30% in 2004 to 22% by 2007.

The criterion of simplicity was largely ignored. The final arrangements were exceedingly complex and fixed weekly hours of work proved difficult to monitor. It is hard for people to understand how each part of the package works, and how they should comply, or change behaviour and thus benefit from the package rather than be captured in an overpayment situation at year's end (Dale, Wynd, St John, & O'Brien, 2010).

Less easily quantified, there is also the social value placed on certainty of income by sole parents. The Deputy Chief Executive of Social Policy, MSD, acknowledged that where work has been precarious, people may choose to stay on a benefit (Gray, 2008, p. 105). Churning and unstable employment can lead to income insecurity with high risk of debt, and may lead some to return to the benefit (Ministry of Social Development, 2004).

5. Case Study 2: KiwiSaver

KiwiSaver is the world's first auto-enrolment opt-out national saving scheme. It runs beside New Zealand Superannuation (NZS), the basic universal Pay As You Go partially prefunded pension. The design, implementation and outcomes of KiwiSaver are of potential interest to countries such as Ireland and the UK that are implementing similar schemes (St John et al 2010).

KiwiSaver was introduced in July 2007 and in 2010, after its first three years, is being hailed in many financial circles in New Zealand as a great success:

> *KiwiSaver has proved to be a huge success, far beyond the Government's expectations. At the end of March [2010] the scheme had a phenomenal 1,369,609 members, net of opt outs and closures, compared with the Treasury's projections of 680,000 members by June 2014.* (Gaynor, 2010)

Other commentators have been similarly adulatory, with few, if any, arguing for a reversal of policy or questioning the fundamental design of KiwiSaver:

> *NZS is simple. KiwiSaver is rather less so, but nonetheless no more intrinsically complicated than voluntary private savings schemes encouraged in some places and compulsory ones mandated in others – and having both PAYG and fully funded approaches operating together is now seen as optimal. The auto-enrolment method adopted for KiwiSaver is arguably more complicated than either the voluntary or compulsory approaches, but preserves an element of choice seen as highly desirable.* (Rashbrooke, 2009)

It is timely to ask however: a success for whom? Is it a numbers game? Has it had the desired effect on the economy? What are the macro implications as opposed to the micro effects of providing for some citizens to have more resources in retirement? Will they in fact have more, or will they reduce other savings to compensate? What are the long term implications for New Zealand's overall pensions framework, particularly the very successful universal state pension?

5.1 Background: identifying the 'problem' to be solved

When KiwiSaver was first announced, the pivotal problem was seen to be one of low national saving. New Zealand is heavily reliant on foreign saving with persistently large current account deficits (CADs) and accumulated overseas debt.

> *[The CAD], and a range of other indices, point to a low level of household savings in New Zealand. We are left highly dependent on foreign capital, which means a substantial proportion of our national income is reclaimed by foreigners as theirs. Hence our Gross National Product is significantly less than our Gross Domestic Product. New Zealanders often bemoan the consequences of low saving, such as high levels of foreign ownership. But, if we are to own, literally, more of our future we must lift our level of savings.* (Cullen 2005, p. 4)

It was not clear that KiwiSaver could lift national saving,[9] and there was little mention of the problem in the Kiwisaver Bill. The purpose is described in the 2006 Act 2006 thus:

> *The purpose of KiwiSaver is to encourage a long-term savings habit and asset accumulation by individuals who are not currently saving enough, with the aim of increasing individuals' well-being and financial independence, particularly in retirement. KiwiSaver is designed to complement New Zealand Superannuation (NZS) for those who wish to have more than a basic standard of living in retirement.* (KiwiSaver Act 2006)

Only on p.36 was there a reference to the hope that national saving will improve:

> *If the behavioural changes flow through into increased domestic saving, then economic growth may increase as more funds may be available to fund domestic investment and reduce New Zealand's reliance on borrowing offshore.*

5.2 Evaluating the policy solution

It is too soon to assess KiwiSaver's impact on either national or household saving or living standards of the retired. Gibson, Hector and Le (2008) provided a preliminary estimate of household saving and show the 'shift' effect, "...[i]t appears that out of every dollar in KiwiSaver accounts, only 9-19 cents is new saving" (Gibson, Hector, & Le, 2008, p. 27). The eight year longitudinal Survey of Family Income and Employment (SoFIE) currently being carried out by Statistics New Zealand, may eventually shed light on KiwiSaver's impact. Every two years, starting in 2004 and continuing until 2010, financial data is collected from SoFIE's participants that allow analysis of households' savings.

The first pre-KiwiSaver SoFIE data are now available. Scobie & Henderson (2009) have estimated that, before KiwiSaver started in 2007, New Zealand households saved an average 16% of their gross incomes in the two years 2004-2006. Taking property revaluations out of that

---

9. The best government action to improve national savings is running surpluses during economic upswings.

estimate reduced the saving rate to 5%. When the next tranche of SoFIE's financial data is available from 2008, it might be possible to see if KiwiSaver has affected households' saving patterns. However, separating out the specific impact of KiwiSaver is likely to be problematic, especially in times of changing economic conditions.

Even if there is an impact on household saving, there is no guarantee that national saving (the sum of private and public saving) will improve. KiwiSaver is not the only change since 2007: a combination of reduced contributions to the New Zealand Superannuation Fund; lower income taxes; a rebalancing towards the tax on goods and services; lower tax rates on investment earnings; and the impact of the recession; are but a few of the contemporaneous influences. Also, while some of the rhetoric suggests that more KiwiSaver savings equals more investment and growth, in practice more saving from any source does not 'cause' more or better investment.

The success of KiwiSaver has become measured by numbers: quantifiable but nothing to do with the objectives.

> *KiwiSaver's extraordinary success has already demonstrated that creating the right incentives to encourage New Zealanders to save and invest are positive signals people respond to – and Budget 2010's tax changes are a further important step in that regard.* (Dunne, 2010)

The current design of KiwiSaver does not appear to be the outcome of any research-based policy development process. After three years, it is still unclear what problems KiwiSaver was to address, and how they would be addressed. Where was the evidence that New Zealanders were under-saving for retirement, and KiwiSaver would provide the solution? What impact is there on national saving? If KiwiSaver was to ensure more income in retirement, why was there no attention to decumulation of the lump-sum? How does KiwiSaver help moderate the fiscal pressures of an aging society?

In the meantime the following issues of KiwiSaver design as set out in St John et al. (2010) have remained unaddressed:

Implications for New Zealand Superannuation (NZS)

Demographic change implies that the cost of NZS in its current form will double over the next 40 years, from about 4% of GDP to 8%. The costs of healthcare will also reflect the changing age structures of New Zealand's population. Although there is currently no sign from the government that this is under consideration, given the contribution that taxpayers will be making to the accumulation of KiwiSaver benefits, it would seem logical that a future government might link NZS through a means test to KiwiSaver. A similar link applies in Australia to its equivalent of NZS, the 'Age Pension' (but applies to all assets and income, not just those derived from Australia's compulsory saving scheme).

Distributional concerns

Until KiwiSaver, New Zealand's tax treatment of retirement savings was relatively neutral and had been so since 1990. That changed with KiwiSaver II in 2007 with the introduction of a range of subsidies and tax-exempt, compulsory employer contributions of 4%. Although that significant concession was halved with KiwiSaver III, what remains raises concerns that apply to all pay-based interventions in retirement saving schemes. Those workers who for whatever reason do not belong to KiwiSaver, effectively subsidise those that do belong, both from higher taxes to fund the MTC, kickstart, tax exemption and from wages that grow more slowly over time because of the employer contributions for those who do belong. The gender imbalances in retirement savings are reinforced to the extent that subsidies favour higher paid workers.

Distribution across generations depends on whether the Kickstart and the MTC are maintained in real terms. Once a person has joined KiwiSaver, the generosity of the subsidy appears to matter less, as inertia leads to them maintaining their membership.

Why children?

It seems difficult to justify the eligibility of children for a national, subsidised *retirement* saving scheme. Although care was taken to exclude them from the auto-enrolment conditions that apply from age 18, the payment of the $1,000 kickstart (and previously the $40 a year administration fee) in respect of children, seems anomalous.

Most members under age 18 are not contributing to their accounts. For the 2008/09 year, 6% of the more than 180,000 members who are children contributed through Inland Revenue to their accounts at a total value of $2 million (an average of only $11 each). It is likely large numbers of accounts managed by providers contain nothing more than the $1,000 kickstart payment. Of the contributing children, almost all are contributing through salary or wage deductions (Inland Revenue Department, 2009).

From a financial education point of view, the messages about the point of saving given to children in KiwiSaver may be quite perverse. The lock-in rules are very unattractive to the young and with no fees subsidy, these high cost small accounts may not generate a sufficient return to maintain even the nominal value of the original kickstart. From a distributional perspective, most of the benefits are likely to flow to the children of higher income, more financially literate families.

Too many providers?

Most observers expect that the total number of providers (currently 53, including employer-specific schemes) will reduce substantially. It is difficult to understand why employers would set up KiwiSaver arrangements specific to their own employees, and those are likely to disappear. Of the remaining 35 or so ‘public’ schemes, possibly half will merge with other schemes.

The question is whether the remaining 20 to 25 KiwiSaver providers is too many; or, would the single scheme approach chosen by, for example, the UK and Sweden be preferable? Having more than one provider means that disclosure and regulatory oversight needs greater care and there are emerging problems in New Zealand. At very least a reformed regulatory framework and an informed business press that is capable of assessing competing claims and providing reliable information to the general public is required.

Too little regulation?

The regulatory regime overseeing KiwiSaver schemes is relatively light-handed, relying more on registration and filing than on approval and oversight. KiwiSaver has been slotted into the existing regime that applies to all other ‘superannuation schemes’. Although it has its own legislation, the KiwiSaver Act 2006, it remains to be seen whether this delivers the protection that a government-mandated regime requires. As noted, the Securities Commission has raised issues with the behaviour of some providers and the unsatisfactory regularity environment.

The current environment is best described as self-regulation, but the industry has failed to provide an authentic mechanism whereby schemes’ investment performance, fees and costs can be compared. Full disclosure, such as for commissions, is supposed to provide the requisite member protection. But firms gain financial advantage from sales, and disclosure does not impose an obligation to explain the range of alternatives (Sheather, 2010). The global financial crisis may call into question the robustness of a regulatory regime founded largely on the ‘prudent person’ requirement that trustees to act in the best interests of their beneficiaries. It is clear that different trustees have different views as to how to interpret this provision.

Choice, default and opt-out

A defined contribution environment, where the benefits from a given set of contributions depend on the investment returns, implies that members should have the right to decide who manages that money. But too much choice in a small country can be costly for individuals, providers and regulators. The balance between individual choice and what is sensible and what is cost effective has yet to be reached.

As noted in section 5.5, the rationale for conferring a commercial advantage on the six default KiwiSaver providers is unclear. Equally, it is difficult to see why the government would impose investment restrictions on the default investment option of only the default providers. If there were any justification for such rules, why might they not apply to all KiwiSaver schemes? An auto-enrolment regime requires default providers but not default investment options. This flawed structure, a by-product of the haste and secrecy of KiwiSaver's introduction, requires review.

Auto-enrolment

Most new employees must be joined up to KiwiSaver if they are not already members. 'Soft compulsion' is justified on the grounds that it is relatively easy to opt-out and that it is a means of making sure people have a chance to do what is in their best interests. Exempt schemes undermine the intent of auto-enrolment. Whether it is a 'successful' strategy depends on how employees react and then what happens to their other savings behaviour. KiwiSaver's success cannot be properly measured for one or more decades and then only if it can be established that it has helped employees to achieve what they might not have been otherwise able to do on their own. Such an outcome will be difficult, if not impossible, to demonstrate. That may undermine the justification for auto-enrolment, and all the compliance costs that it entails for the employer.

Auto-enrolment is supposed to nudge people to behave in the 'right' way; in this case, to save more for their retirement. It is impossible to assess whether the 'nudge' has been successful if at the same time there are significant monetary incentives to change behaviour.

Housing

Many New Zealanders still make most of their 'retirement' savings through owning a mortgage-free home by the time they reach retirement age. Requiring someone who is not already a home-owner to save through KiwiSaver, rather than use those savings to purchase and pay off a home, was antithetical to New Zealanders.[10] KiwiSaver accommodated that by the concessions and subsidies for first home purchases, and the now defunct mortgage diversion facility. However, it compromised KiwiSaver's fundamental objective: to increase financial savings for retirement. The lesson here is that a government's intervention needs a clear, unambiguous focus. It would perhaps have been better to restrict the auto-enrolment requirement to employees who were over age 35 or 40, and then to give no concessions to home ownership.

6. Conclusion: what works?

Policy is should be assessed for its cost-effectiveness in meeting the stated objectives and against criteria of efficiency, equity and administrative simplicity. In the case of the two social policies examined, it may be possible to conduct quantitative- evidence based studies but these are likely to be of little use in assessing whether the policy has been worthwhile. Statistical significance is not the same as social significance: the later requires a much broad range of tools and thinking.

10. While employees could opt-out, they then lost the advantage of the significant tax breaks and employer's contributions that were part of KiwiSaver II and remain, in reduced form, in KiwiSaver III. Also, everyone's taxes (including non-homeowners) are higher to pay for the incentives.

We can conclude that the use of statistical evidence of ‘what works’ in policy as the sole evaluation tool has proved to lack credibility, at least on the case of complex social policy. In the end we may be left with nothing more than nice ideas and pious hopes: Making work pay will encourage sole parents to work and solve their and their children’s poverty; KiwiSaver will solve New Zealand’s saving problems while also making retired people better-off. Have we substituted slogans for clear unbiased thought and abandoned a critical evaluation process that asks the broader questions: Are society and the individuals in society actually better off?

References

Adamson, P. (2008). *The child care transition. A league table of early childhood education and care in economically advanced countries* (No. 8). Florence: UNICEF Innocenti Research Centre.

Blundell, R., & Costa Dias, M. (2000). Evaluation methods for non-experimental data. *Fiscal Studies, : , 21*(4), 427 - 468.

Bryson, A., Evans, M., Knight, G., La Valle, I., & Vegeris, S. (2006a). *Methodological considerations for evaluating MSD programmes*. Wellington: Centre for Social Research and Evaluation Te Pokapu Rangahau Arotake Hapori, Ministry of Social Development.

Bryson, A., Evans, M., Knight, G., La Valle, I., & Vegeris, S. (2006b). *New Zealand Working For Families programme: Methodological Considerations in Evaluating Working for Families* (CPG.115.0207/0225). London, UK: Policy Studies Institute.

Bullock, H., Mountford J., & Stanley, R. (2001). *Better Policy-Making*. London: Cabinet Office, Centre for Management and Policy Studies, http://www.cpms.gov.uk. (accessed 7 June 2002).

Cabinet Policy Committee. (2004). *Reform of social assistance: Working for Families package* Retrieved from http://www.msd.govt.nz/documents/about-msd-and-our-work/work-programmes/policy-development/working-for-families/rep-04-3-229.pdf.

Card, D., & Hyslop, D. R. (2005). Estimating the effects of a time-limited earnings subsidy for welfare-leavers. *Econometrica, 73*(6), 1723 - 1770.

Centre for Social Research and Evaluation. (2007a). *Pockets of Significant Hardship and Poverty*. Wellington: Ministry of Social Development,.

Centre for Social Research and Evaluation. (2007b). *Receipt of the Working for Families Package: 2007 Update*. Retrieved from http://www.ird.govt.nz/aboutir/reports/research/report-wfftc-update/.

Centre for Social Research and Evaluation, & Inland Revenue Te Tari Taake. (2007). *Receipt of the Working for Families Package*. Wellington: Ministry of Social Development.

CM 4310. (1999). Modernising Government White Paper. Retrieved from http://www.cabinet-office.gov.uk/moderngov/whtpaper

CPAG v Attorney General. (2008). Transcripts of cross examination of the OECD economists in the Human Rights Tribunal.

Dale, M. C., Wynd, D., St John, S., & O'Brien, M. (2010). *Work incentives and the labour market: Has Working for Families delivered?* Auckland: Child Poverty Action Group.

Dalgety, J., Dorsett, R., Johnston, S., & Spier, P. (2010). *Employment incentives for sole parents: Labour market effects of changes to financial incentives and support. Technical report*. Wellington: Inland revenue Department, Ministry of Social Development.

Dalgety, J., Dorsett, R., Spier, P., & Johnston, S. (2009). *Employment incentives for sole parents: Labour market effects of changes to financial incentives and support*. Paper presented at the Victoria University of WellingtonTax Conference.

Dunne, P. (2010). *Post-Budget Address to Johnsonville Rotary Club from Peter Dunne, Minister of Revenue, Monday 14 June*.

Durie, M. (2004). *Race and Ethnicity in Public Policy. Does it work?* Paper presented at the The Social Policy Research and Evaluation Conference 2004. Retrieved from http://careers.massey.ac.nz/massey/fms/Te%20Mata%20O%20Te%20Tau/Publications%

20-%20Mason/M%20Durie%20Race%20and%20ethnicity%20in%20Public%20Policy.pdf

Dwyer, G. (2005). Dissecting the Working for Families Package. Retrieved from http://www.nzbr.org.nz/documents/publications/publications

Ellwood, D. T. (2000). "The impact of the Earned Income Tax Credit and social policy reforms on work, marriage and living arrangements.". *National Tax Journal, 53*(4), 1063-1105.

Fitzgerald, J., Maloney, T., & Pacheco, G. (2008). The Impact of Changes in Family Assistance on Partnering and Women's Employment in New Zealand. *New Zealand Economic Papers 42*(1), 17-57.

Friesen, M. D., Woodward, L. J., Fergusson, D. M., Horwood, L. J., & Chesney, A. (2008). Living Standards and Material Conditions of Young New Zealand Families. *Social Policy Journal of New Zealand, 33*(March), 47 - 69.

Gaynor, B. (2010). KiwiSavers need to be more adventurous. *24th April*. Retrieved from http://www.nzherald.co.nz/business/news/article.cfm?c_id=3&objectid=10640441

Gennetian, L. A., Miller, C., & Smith, J. (2005). *Turning welfare into support: Six-year impacts on parents and children from the Minnesota Family Investment Program*. Minnesota: Manpower Demonstration Research Corporation, available at: www.mdrc.org.

Gibson, J., Hector, C., & Le, T. (2008). *The distributional impact of KiwiSaver* (Working paper): University of Waikato.

Gray, D. (2008). *Statement of evidence: CPAG v Attorney General* Wellington: Human Rights Tribunal.

Human Rights Tribunal. (2008). *Child Poverty Action Group Inc v Attorney General. The Decision*. Wellington: Human Rights Review Tribunal

Hurnard, R. (2007). *Roger's Option*. Wellington, NZ: The Treasury.

Inland Revenue Department. (2009). *Kiwisaver Evaluation: Annual Report 2: July 2008 - June 2009*. Retrieved from http://www.ird.govt.nz/resources/2/3/236969804fea97e9881ceb53c1fd2485/ks-ar-2009.pdf.

Marston, G., & Watts, R. (2003). Tampering with the evidence. *The Drawing Board: An Australaian Review of Public Affairs, 3*(3), 143 - 163.

McCloskey, D. N., & Ziliak, S. T. (2008). *The Cult of Statistical Significance: How the Standard Error Costs Us Jobs, Justice, and Lives* Chicago: University of Michigan Press.

Millar, J. (2008). Making work pay, making tax credits work: An assessment with specific reference to lone-parent employment. *International Social Security Review, 61*(2), 21 - 38.

Ministry of Social Development. (2002). *New Zealand's agenda for children*. Wellington: Ministry of Social Development.

Ministry of Social Development. (2004). *Future Directions: Working for Families Impacts*. Wellington: Ministry of Social Development.

Ministry of Social Development. (2006). *New Zealand Living Standards 2004*. Retrieved from http://www.msd.govt.nz/work-areas/social-research/living-standards/living-standards-2004.html.

Ministry of Social Development. (2009). Benefit factsheets. from http://www.msd.govt.nz/about-msd-and-our-work/publications-resources/statistics/benefit/2009-national-benefit-factsheets.html

Ministry of Social Development. (2010). *Regulatory Impact Statement: Future Focus changes to the welfare system*. Wellington: Ministry of Social Development.

New Zealand Government. (2008). *Taxation (Urgent Measures and Annual rates) Bill*. Retrieved from http://www.legislation.govt.nz/bill/government/2008/0006-1/latest/DLM1764914.html.

Nolan, P. (2002). *New Zealand's family assistance tax credits: Evolution and operation* (Working paper WP 02/16). Wellington: The Treasury.

Nolan, P. (2004). *When work does not pay: family structures and poverty traps in New Zealand's social security system.* Paper presented at the New Zealand Association of Economists Conference, Wellington.

Nutley, S., Davies, H., & Walter, I. (2003). Evidence-Based Policy and Practice: Cross-Sector Lessons from the United Kingdom. *Social Policy Journal Of New Zealand Te Puna Whakaaro, June 2003*(30), 29 - 48.

OECD. (2007). *Economic Survey of New Zealand* (Policy Brief). Paris: OECD.

Pearson, M., & Immervoll, H. (2008). *Statement of evidence: CPAG v The Attorney General*. Wellington: Crown Law.

Perry, B. (2005). Social Report Indicators for Low Incomes and Inequality:Update from the 2004 Household Economic Survey. from http://www.msd.govt.nz/work-areas/cross-sectoral-work/indicators-for-low-incomes-and-inequality.html

Perry, B. (2008). *Household incomes in New Zealand: trends in indicators of inequality and hardship 1982-2007* Wellington: Ministry of Social Development

Perry, B. (2009). *Non-income measures of material wellbeing and hardship:first results from the 2008 New Zealand Living Standards Survey, with international comparisons*. Wellington: Ministry of Social Development.

Rashbrooke, G. (2009). Simple, Effective and (Relatively) Inexpensive: New Zealand Retirement Provision in the International Context. *Social Policy Journal of New Zealand*(36), 97 - 110.

Roberts, H. (2005). What Works? *Social Policy Journal Of New Zealand Te Puna Whakaaro, April 2005*(24), 34 - 54.

Salmond, A. (2003). To the Social Policy Research and Evaluation Conference *Social Policy Journal of New Zealand*(20).

Sheather, B. (2010, 22 May 2010). Ban on commissions is a must. *The New Zealand Herald*. Retrieved from http://www.nzherald.co.nz/business/news/article.cfm?c_id=3&objectid=10646641

St John, S., & Dale, M. C. (2010 Forthcoming). Is New Zealand's 'In Work Tax Credit' a good model for an 'In Work Benefit'? *European Journal of Social Security*.

St John, S., Littlewood, M., & Dale, M. C. (2010). *KiwiSaver. The first three years: lessons for Ireland?* Paper presented at the Pension Policy Research Group, School of Business, Aras an Phiarsaigh, Trinity College Dublin.

St John, S. (1997). The Measure of Success for Beyond Dependency: Aims, Methods and Evaluation *Social Policy Journal of New Zealand, 8*.

St John, S. (2006). New Zealand's Financial Assistance for Poor Children: Are Work Incentives the Answer? *European Journal of Social Security (EJSS), 8*(3).

St John, S. (2007). Farewell to tax neutrality: the implications for an aging population. *Economic and Labour Relations Review, 18* (1), 27-52.

St John, S., & Craig, D. (2004). *Cut price kids: Does the 2004 'Working for Families' budget work for children?* Auckland: Child Poverty Action Group.

St John, S., Dale, M. C., & Littlewood, M. (2009). *An evaluation of 'In Work Benefits' in New Zealand: Use of the In Work Tax Credit to encourage labour supply: implications for an ageing population.* Auckland: Retirement Policy and Research Centre, The University of Auckland.

St John, S., & Rankin, K. (2009). *Escaping the welfare mess*: Working Paper 267, Department of Economics, The University of Auckland.

Strategic Policy Making Team. (1999). *Professional Policy Making for the Twenty First Century*. London: Cabinet Office, http://www.cabinet-office.gov.uk/moderngov/policy/index.htm (accessed 7 June 2002).

The Treasury. (2005). *This Year's Critical Issues: Issues that relate primarily to labour utilisation*. Wellington: The Treasury.

Wylie, C. (2006). *Making the most of research in policy making*. Paper presented at the NZCER Seminar: Making the most of research in policy making. Retrieved from http://www.nzcer.org.nz/default.php?cpath=139_133&products_id=1737