## Contents

Editor and Editorial Committee .......................... 1

### ANALYSIS

Fairness in Public-utility Regulation: A Theory .......................... 5
Darryl Biggar

### ARGUMENT

Formulating Policy Responses to Global Warming in the Face of Uncertainty .......................... 33
Harry Clarke

A New Mind-set for Exchange Rates .......................... 55
Stephen Grenville

Reflections on "A Tax System for New Zealand's Future" .......................... 63
John Creedy

### SYMPOSIUM:

**Krugman on Economics in the Light of the Crisis**

How US Economists Got It So Wrong .......................... 77
Ross McLeod

Froth and Bubble: The Inconsistency of Paul Krugman's Macroeconomic Analysis .......................... 83
Don Harding and Jan Libich

Beyond Krugman to Behavioural Keynes .......................... 89
Ian M. McDonald
Krugman on the Malaise of Modern Macro: Critique without Alternative 95
Keith Rankin

What Keynes Missed and Krugman is Missing: The Short/Long Choice 101
David Vines

Beauty ≠ Truth? Thoughts on Krugman's "How did economists get it so wrong?" 113
John Quiggan

RETROSPECT

Yegor Gaidar: Pragmatic Economist or Romantic Revolutionary? 121
Gennadi Kazakevitch

REVIEWS

*Educating for Business, Public Service and the Social Sciences: A History of Economics at the University of Sydney 1920–1999* by Peter Groenewegen, and *Balanced Growth: A History of the Department of Economics, the University of Melbourne* by Ross Williams (ed.), reviewed by Selwyn Cornish 133


*Immigrants: Your Country Needs Them* by Philippe Legrain, reviewed by Mathew Pollock 143
Editor

Dr William Coleman

College of Business and Economics, ANU

Editorial Committee

Professor Harry Clarke
Department of Economics and Finance, La Trobe University

Professor Max Corden
Department of Economics, University of Melbourne

Mr James Cox
Independent Pricing and Regulatory Tribunal of New South Wales, Sydney

Dr Anne Daly
Faculty of Business and Government, University of Canberra

Professor Kevin Davis
Department of Finance, University of Melbourne

Mr Henry Ergas
Canberra, Deloitte Australia

Dr Stephen Grenville
Lowy Institute, Sydney

Dr Arthur Grimes
Motu Research, Wellington

Professor Kim Hawtrey
Department of Economics, Hope College

Professor Doug Irwin
Department of Economics, Dartmouth College

Dr Gary Johns
Public Policy Institute, the Australian Catholic University

Professor Tony Makin
Griffith Business School, Griffith University

Dr Deborah Peterson
Department of Primary Industry, Melbourne

Professor Jeff Pope
School of Economics and Finance, Curtin University of Technology
ANALYSIS
Fairness in Public-utility Regulation: 
A Theory

DARRYL BIGGAR

Abstract
Regulators routinely and systematically depart from policy prescriptions that are soundly based in conventional economic theory. In doing so, they often appeal to notions of fairness, justice, or reasonableness. Economists have historically struggled with these notions, which seem to be separate from, or in conflict with, conventional economic efficiency. This paper identifies five stylised facts about public attitudes to fairness in utility pricing, and argues that these stylised facts can be explained as an implicit attempt to protect the sunk investments of consumers in a natural monopoly’s services. Thus the paper suggests that the notion of fairness is not in conflict with the conventional notion of economic efficiency, but can be seen as consistent with the desire to promote sunk investment by the monopolist and its customers.

Introduction
Open any textbook on microeconomic theory or natural monopoly regulation and you will learn that the primary economic harm from natural monopoly is the allocative efficiency loss. First-year economics students are taught that an unregulated monopolist will produce ‘too little’ and set a price ‘too high’ relative to the theoretically efficient level. The ensuing reduction in economic welfare is known as the ‘deadweight loss’. According to the conventional theory, the primary objective of the economically literate and enlightened regulator is to regulate in such a way as to minimise or eliminate the deadweight loss.

1 Darryl Biggar is an economic consultant to the Australian Competition and Consumer Commission and the Australian Energy Regulator, Darryl.Biggar@accc.gov.au. The views expressed here are those of the author and are not necessarily the views of the ACCC or the AER. The paper has benefited from the comments of an anonymous referee.
During the twentieth century, a large body of economic theory was developed showing how a regulator should set tariffs so as to minimise the deadweight loss under different assumptions and constraints. This theory highlights the benefits of, amongst other things, marginal-cost pricing, various forms of price discrimination, Ramsey pricing, and peak-load pricing. However, although there has been some limited acceptance of these policies, the response of the regulatory community has, on the whole, been cautious and muted. Experience shows that regulators routinely depart from policies which are soundly based in conventional economic notions of efficiency.\footnote{See, for example, the citations in Biggar (2009) and later in this paper, such as Berg and Tschirhart (1995: 324), and Baumol (1986).}

Instead, regulators argue that these conventional economic pricing principles often do not satisfy their moral or legal obligation to take into account notions of ‘fairness’, ‘equity’, ‘justice’, or ‘reasonableness’. Many regulators, especially in the US, are statutorily obliged to set rates which are ‘just and reasonable’ or ‘not unduly discriminatory’. Even in the absence of an explicit legislative requirement that rates be ‘fair and reasonable’, regulators rarely adopt rates which — even though soundly based in conventional economics — might violate notions of fairness or equity.

Economists have historically reacted to the apparent trumping of efficiency concerns with notions of fairness or equity in a variety of ways. As discussed below, one approach has been to dismiss or belittle these other concerns, viewing them as illegitimate deviations from the pursuit of economic welfare. Another approach views fairness and equity concerns as potentially legitimate, but outside the professional sphere of the economist. There have also been several attempts to formalise and/or axiomatise notions of fairness and equity.

This paper argues that key elements of the notion of fairness in public-utility regulation can be explained through the lens of economic efficiency, using a different underlying economic model. Specifically, I argue that fairness concerns in public-utility regulation can, to a large extent, be explained as an explicit or implicit concern on the part of a well-intentioned regulator to protect and thereby promote sunk complementary investments by consumers.

Following the argument in Biggar (2009), I point out that users of a monopoly service must typically take some irreversible action which increases the value of the monopoly service — such as the decision of a large gas consumer to locate close to a gas-transmission pipeline, or the decision of a factory to install electrical wiring on its premises. The value of such investments is contingent on continuing to receive access to the monopoly service at reasonable prices and quality. Such an investment is therefore at risk of expropriation through an
increase in the price or a decrease in the quality of the monopoly service. The fear of such expropriation has a chilling effect on such investment, reducing overall economic welfare.

I suggest that regulators (in common with the rest of the population) are intuitively sensitive to the notion that their pricing policies could threaten the incentives of users to make valuable sunk investments. In rejecting certain pricing policies as ‘unfair’ or ‘inequitable’, well-intentioned regulators implicitly have in mind the impact of those pricing policies on the incentives on users to make sunk complementary investments.

This hypothesis seems to explain the stylised facts about attitudes to fairness in pricing, set out below. Furthermore, this proposition is, at least in principle, empirically testable. For example, I suggest that fairness concerns are important precisely in those sectors in which the value of the complementary investments by consumers are dependent on the price and quality of the service they receive from a specific firm. Conversely, I suggest that fairness concerns are, in practice, diminished in those industries where customers’ complementary investments are not specific to any one firm — that is, those industries in which customers face effective competition.

First, however, section 2 makes the case that there is a need for economists to study and understand fairness. This section also distinguishes notions of fairness which are not directly relevant to this paper and briefly addresses the concern that perhaps fairness is not systematic enough to be a legitimate subject for study. Section 3 surveys the responses of economists to fairness concerns in the past. Section 4 derives some stylised facts on attitudes to fairness in pricing. Section 5 shows how those stylised facts are consistent with a concern to protect sunk complementary investments.

**What is fairness and why is there a need to explain it?**

Economists — especially those at the ‘coalface’ of interaction with the regulatory community — have long been aware that regulators do not always do what conventional economic theory tells them they should do. Even the most well-established pricing policies in conventional regulatory economics are often downplayed or overruled by regulators.

For example, economists have long observed that Ramsey pricing and peak-load pricing, although solidly based in conventional economic theory, have been only partially accepted by regulators. In rejecting these policies, regulators routinely
point to concerns about fairness or equity (see, for example, Kahn 1970; Baumol 1986; Faulhaber and Baumol 1988; Berg and Tschirhart 1995; Raux and Souche 2006).

This focus on fairness or equity does not seem to be a recent innovation. Indeed, Jones and Mann (2001) point out that concern for fairness and justice dates back to the origins of public-utility regulation:

The concept of ‘fairness’ is deeply rooted in the 120-year history of public-utility regulation in the United States. Indeed, the word ‘fair’ appears in numerous regulatory concepts and propositions: for example, a fair rate of return, fair value, fair and reasonable rates, and full and fair evidentiary hearings. … While the idea of fairness is elusive and perceptions differ, it clearly is a potent force in regulation, as indicated by the vehemence with which participants complain when they feel they have been treated unfairly.

Even in Australia it is straightforward to find explicit legislative requirements for the regulator to pursue fairness. For example, in the National Electricity Rules, the terms and conditions for access to negotiated distribution services and negotiated transmission services must be ‘fair and reasonable’ as must the price for connection to an existing network. An ‘efficiency benefit sharing scheme’ must provide for a ‘fair sharing’ of efficiency gains between a network provider and users.

Perhaps, however, these references to fairness merely represent a historical hangover from an earlier and less economically sophisticated time. Perhaps fairness concerns have diminished in recent years, especially following the recent reforms of the public-utility sector, with their focus on promoting competition? This doesn’t seem to be the case. In a survey of US state and federal regulatory commissioners, Jones and Mann (2001) found no indication that fairness is ranked lower than other goals of regulation than it was in the past. All but 4 per cent of the respondents ranked fairness as ‘either more important, or as

---

3 For example, Berg and Tschirhart (1995: 324): ‘…The fact is that most regulators do not view the world as economists view it. In reality, regulatory commission have objectives, motivations, and responsibilities far more complex than “setting price equal to marginal cost subject to a profit constraint”. … Neoclassical economics has focused almost exclusively on efficiency, which is but one of many objectives to most regulators. The justification for the narrow focus is that the other concerns, legitimate as they are, are too difficult to balance against efficiency given our regulatory control mechanisms. Whether or not this is true, it does not seem to dissuade regulators from a strong focus on equity or fairness.’

4 Jones and Mann 2001: 1. Bonbright, in his classic text of regulation, identifies fairness as ‘one of the three primary characteristics of a sound rate structure’ (Bonbright et al. 1988: 385).

5 Clause 6.7.1(9) and clause 6A.9.1(9), National Electricity Rules.

6 Clause 5.3.6, National Electricity Rules.

7 Muir (2001: 3) notes that: ‘Examination of the objectives of a small and fairly random selection of Australian and foreign utility regulators reveals that fairness figures explicitly in almost every case.’
important, as any other goal of public utility regulation’ (Jones and Mann 2001: 7). Sixty per cent reported that in a conflict between the goals of fairness and efficiency, fairness would be considered to be more important.6

**Dimensions of fairness**

There are many dimensions of what might constitute fairness, many of which go beyond the scope of this paper. To begin with, it is common in the literature on fairness to make a fundamental distinction between fairness in process, which is known as ‘procedural fairness’, and fairness in outcomes, which is known as ‘end-state fairness’ (see Konow 2003).

Procedural fairness (also known as ‘natural justice’) relates to the processes by which decisions are made. The principles of natural justice are designed to ensure that the decision-making process is fair and reasonable. I am primarily interested in the perceived acceptability of various pricing outcomes, rather than the process by which those outcomes are reached, so I will put to one side issues of procedural fairness.

There is also a sense of fairness which is linked with the distribution of income. This is sometimes known as ‘distributive equity’. Muir (2001) observes that:

> At the retail level, considerations of a customer’s ability to pay appear to dominate [considerations of fairness]. Underlying this concern appear to be attempts to improve the affordability of the service to lower-income consumers, those who are costly to supply or those with particular ‘merit’ claims (such as educational institutions or community organisations).  

The linkage between income distribution and fairness is confirmed in laboratory experiments which show that providing information on the income levels of participants alters perceptions as to whether or not a given scenario is fair.

Nevertheless, as we will see later, fairness is more than just a surrogate for questions of income distribution. Fairness concerns remain even when participants can be presumed to have a similar income. For the purposes of this paper, I will follow the conventional economic approach of simply putting these questions to one side.10 In addition, for the purposes of this paper I will not make a distinction between the notions of fairness, equity, and justice. Instead, I will follow common usage and treat these terms largely as synonyms.

---

6 Jones and Mann (2001: 10). Zajac (1995: 101) and Baumol (1986) also argue strongly that economists should pay more attention to notions of fairness.
10 The conventional economic justification for this approach is that ‘if one disapproves of the distribution of income, one might better correct it by lump-sum taxes (for example, on rents, income or inheritance) and
Is fairness systematic enough to be studied?

Historically, several authors have suggested we should not be over-optimistic about our ability to pin down just what constitutes fair, as opposed to unfair, pricing decisions. Fairness, it is argued, is an imprecise or elusive concept (Bonbright 1988; Baumol 1988; Fehr, Kirchsteiger and Riddl 1993), subjective (Konow 2006; Young 1994), or merely a cover for self-interest (Zajac 1995; Gielissen et al. 2008; Young 1994).

Nevertheless, a small but growing body of literature, starting with the pioneering work of Kahneman, Knetsch and Thaler (‘KKT’) (1986) has shown that there are empirical regularities in what members of the public perceive as fair or unfair in economic transactions. These regularities are discussed further in section 4. First, let’s review how economists have responded to fairness concerns in the past.

The response of economists to fairness objectives in public-utility regulation

When confronted with the fact that some of their most well-founded policy proposals are rejected by regulators on the grounds of fairness, the economics profession has reacted in different ways. It is possible to identify four different possible responses:

• The ‘conflict’ response, which asserts that the objective of conventional economic efficiency is paramount over all other objectives, and that any deviation reflects ignorance, incompetence, or capture by special interests.

• The ‘apartheid’ response, which asserts that objectives such as fairness or equity have a legitimate ‘separate but equal’ status to conventional economic efficiency, and the economist in his professional capacity has nothing to say about these other ‘non-economic’ objectives.

• The ‘assimilation’ response, which seeks to use economic tools and techniques to formalise or systematise the concept of fairness.

money transfers than by departing from the requirements of economically efficient pricing’ (Kahn 1970: Vol I: 68). This line of argument has been heavily criticised as not being economically sound. Nevertheless, it remains convenient for the purposes of this paper to put this aspect of fairness to one side.

11 Young (1994) opens his book on equity with several arguments that equity ‘fails to exist’: ‘The first is that equity is merely a word that hypocritical people use to cloak self-interest. It has no intrinsic meaning and therefore fails to exist. The second argument is that, even if equity does exist in some notional sense, it is so hopelessly subjective that it cannot be analysed scientifically. Thus it fails to exist in an objective sense. The third argument is that, even granting that equity might not be entirely subjective, there is no sensible theory about it, and certainly none that is compatible with modern welfare economics. In short it fails to exist in an academic sense.’
• The ‘integration’ response, which seeks to explain fairness and equity considerations as being derived from economic efficiency notions — possibly by varying the economist’s standard approach to efficiency.

The ‘conflict’ response

To many economists, the goal of economic efficiency (usually defined as allocative efficiency) is self-evidently paramount over all other objectives. Under this view, placing any weight on other objectives at best results in leaving untapped social welfare on the table.\footnote{12 Kaplow and Shavell (2002) go further to argue that placing any weight on a non-economic-welfare objective will result, in some circumstances, in leaving all members of society worse off.} To do so reveals ignorance, incompetence, or capture by one or more interest groups.

Under this view, any departure from efficient policies, ostensibly due to concerns with fairness or equity, is ill-informed, misguided, a cover for self-interest, or at best secondary to the pursuit of allocative or productive efficiency.\footnote{13 See Baumol (1986) and Bonbright \textit{et al.} (1988: 182). The most active contemporary proponents of this view are Louis Kaplow and Steven Shavell of the Harvard Law School. See Kaplow and Shavell (2002).} Zajac (1995), for example, reflecting on his time at Bell Labs designing optimal rate schedules for AT&T, recalls that: ‘Those opposing us, we felt, were either misguided, ill-informed, or just plain evil.’ (Zajac 1995: 4).

Under this view, the task of the economist is to use the powers he has at his disposal to persuade the relevant authorities as to the virtues of maximising economic welfare (as conventionally defined), to expose the conflict with other objectives (see, for example, Bonbright 1988: 192), and to subordinate these other objectives to the efficiency objective.\footnote{14 As an aside, there is of course, an equal-and-opposite view (not usually held by economists) that fairness and equity considerations must be paramount in the mind of the regulator, with allocative efficiency subordinate where it is considered at all. For example, Jones and Mann (2001) report that many regulators view fairness and equity considerations as more important than economic efficiency as conventionally defined.}

The ‘apartheid’ response

A second common reaction of economists is to declare that the objective of maximising conventional economic welfare is separate and distinct from the objective of promoting fairness or equity, and each has its own validity.

According to this approach, the discipline of economics is concerned with the narrow objective of promoting economic welfare (defined as allocative or productive efficiency). There may be other equally legitimate objectives, such as fairness or equity, but the economist has nothing to say about these objectives,
at least not in his/her professional capacity as an economist. If regulators, serving their political masters, want to take into account other non-economic considerations, so be it.

For example, Bonbright asks whether or not the economist should ‘take part, in his or her professional capacity, in controversies about rival standards of fairness’. According to Bonbright, the usual answer is ‘no’, ‘on the ground that the question, being one of ethics, goes beyond their professional competence’ (Bonbright 1988: 192). Baumol is even more explicit:

The natural reaction of members of our discipline is to question the economists’ qualifications to make any pronouncements on fairness … Credulity is really strained by the prospect that an economist witness will tell a Congressional committee or a regulatory commission that he or she is qualified professionally to make some pronouncement on the fairness of some proposal that such a body may be considering. Why should economists ever make the attempt, and on what legitimate basis can they hope to do so? (Baumol 1987: 1)

Under this view, the economic and ‘non-economic’ objectives of a regulator are ‘separate but equal’. This response might be called the ‘apartheid’ approach.

The ‘assimilation’ response

A few economists have recognised the importance of fairness and equity considerations and have responded by attempting to formalise or axiomatise these concepts, so as to give the economist some ground on which to stand when providing advice on fairness issues or when trading-off fairness with conventional objectives.

There are two strands of this research worth highlighting. The first strand seeks to formalise the notion of fairness as equivalent to the ‘absence of envy’. As Konow (2003) explains, this concept was first formally stated by Duncan Foley (1967) and was further developed by Hal Varian (1974), Elisha Pazner and David Schmeidler (1978), William J. Baumol (1986) and others.

Under this approach, an allocation is said to be ‘envy free’ if no participant prefers the bundle of goods and services of another participant. This approach has the strong advantage that it uses many of the micro-economists’ existing tools and techniques, such as the concept of the Edgeworth box and the concept of the ‘core’ of a co-operative game.

Unfortunately, however, this approach has met with limited success in explaining the everyday concept of fairness. As Konow notes:
Absence of envy is an appealing construct and seems like a reasonable goal. The question asked ..., however, is whether it describes allocations people call fair, or whether it is distinct. Robin Broadway and Neil Bruce (1984) are sceptical about equating the two: ‘I might envy a friend’s lucky find in an antique store yet perceive no “unfairness” that he, not I, owns it.’

Konow (2003) provides an example of a scenario in which the participants do not envy each other’s final allocation, but 90 per cent of respondents found the outcome unfair.

A second important strand of research starts from the observation that all tariff structures which yield the regulated utility sufficient revenue to cover its costs are equivalent to some allocation of the common costs — so the question as to what constitutes a fair or just tariff structure is equivalent to the notion as to what constitutes a fair or just cost allocation. This research then seeks to formalise the notion of a fair or just cost allocation.

This approach typically starts by asserting some fundamental principles or axioms that a fair cost allocation must satisfy, and seeks to characterise the class of cost allocations that satisfy this principle. The most well-known example of this approach is the famous work on cross-subsidies by Faulhaber (1975). Faulhaber starts from the principle that the revenue from any service or group of services should not exceed the stand-alone cost of this service. One possible justification for this ‘stand-alone cost ceiling’ is that it is ‘unfair’ for any group of customers to be paying more than what they could pay if they could organise to provide the service for themselves. As Ralph (1992) observes: ‘This seems a compelling minimum criterion for identifying a cross-subsidising group — common sense suggests it is unjust that a group should have to pay more than its stand-alone costs.’

There are important generalisations of this approach. Willig (1979) introduces the notion of ‘anonymous equity’. A cost allocation satisfies anonymous equity if no individual customer or group of customers is charged more than the stand-alone cost for the services that he/she (or the group) consumes. Young (1985) describes several other axiomatic approaches to cost allocation, including the ‘separable costs remaining benefits method’, the Shapley value and

16 Holcombe (1997) similarly rejects equating fairness with absence of envy (although in this case on procedural grounds).
17 In co-operative games a key question is how the gains from co-operation should be shared. The Shapley value, proposed by Lloyd Shapley in 1953, is a formula for sharing the gains from co-operation which satisfies certain key desirable axioms. See, for example, Young (1994), or Sergiu Hart (1989).
the nucleolus. Young suggests several properties that we might like a cost allocation to satisfy (specifically, ‘additivity’, ‘monotonicity’ and ‘consistency’). Not all of these properties can be satisfied simultaneously, but he shows how various combinations of these properties, with various ‘strengths’ lead to different cost-allocation methods.

To be clear, this line of research is independent from, and unrelated to, conventional economic efficiency notions. Zajac, discussing the various tests for cross-subsidisation notes: ‘These tests were … basically motivated by fairness, and we shouldn’t necessarily expect them to be economically efficient, nor that an economically efficient pricing structure will pass all of these tests.’ (Zajac 1995: 211).

But how well does this line of research do as a theory of fairness? Is it the case that, say, a random sample of objective members of the public would find cost-allocation methodologies which satisfy these axioms to be fair, while other cost-allocation methodologies which do not satisfy these axioms are deemed to be unfair?

As we will see in the next section, empirical studies of fairness seem to show that the concept of fairness is not so much related to a particular tariff structure or cost allocation as it is to changes in that tariff structure or cost allocation. It seems that virtually any cost allocation could be considered to be fair if consumers have had a long enough period of adjustment. In other words, attempts to define what constitutes a fair cost allocation may be simply missing the point. As we will see when we survey the empirical literature on fairness, an allocation which satisfies relatively few axioms might be considered fair if customers have had time to adjust. Conversely, a move towards an allocation which satisfies several axioms (such as, say, the Shapley value) might, in fact, be deemed to be unfair — precisely because it raises the price on some group of customers. It seems that these approaches fail to capture key elements of public perceptions of fairness.

---

18 The nucleolus is a particular way of sharing the gains from co-operation introduced by Schmeidler in 1969. The nucleolus has certain desirable properties. For example, it is always unique and lies within the core (the set of all allocations for which no group of players could be better off if they refused to co-operate). See, for example, Young (1994).

19 Young (1994) observes that a ‘difficulty with the axiomatic method is that it can easily become disengaged from the problem that it was intended to solve. The invention of axioms and conditions is a fascinating business. The danger is that the exercise can take on a life of its own.’

20 As suggested later in this paper, the problem with this line of research is possibly that it focuses on too narrow a bargaining game between the firm and its customers. This paper suggests later that fairness is not simply a matter of sharing of the gains from co-operation at a given point in time; rather, fairness is about achieving efficient outcomes in a game in which one or both parties can make investments to increase the value of co-operation.
The ‘integration’ response

The fourth possible response to the persistent reference to fairness in regulation is what might be called the ‘integration’ approach. This approach seeks to explain the patterns of behaviour of regulators — including their description of some rate designs as fair or unfair — using economic models and approaches. This approach inevitably moves beyond the traditional focus on allocative efficiency and asks whether or not there are other economic models or approaches which can better explain the behaviour of regulators.

This paper falls into this fourth category. This paper seeks to identify patterns of regulatory outcomes that might be considered fair or unfair and then seeks to argue that these patterns can, in large part, be explained using economic models.

The next section attempts to identify patterns in attitudes towards fairness. The last section then sets out a model which seeks to explain these patterns in a manner that is consistent with economic efficiency.

What constitutes ‘fairness’ in public utility pricing?

The first step in this process is to identify key regularities in public notions of fairness. What, in fact, are the key characteristics of public attitudes towards fairness? Drawing on the few empirical studies on attitudes towards fairness in pricing more generally, I suggest that public attitudes to fairness in public-utility pricing can be summarised in the following five stylised facts:

Fact 1: It is considered unfair to raise prices above expected or anticipated levels

There appear to be solid grounds for the view that a key element of fairness relates to the stability or predictability of the regulated prices. It is perceived to be unfair to increase regulated rates above a promised, pre-announced, predicted, or anticipated level, except where the rate increase is necessary to offset an increase in long-term costs.

One of the key conclusions of the research by KKT is that the fairness of an economic transaction is measured in part by the change in the terms of the transaction relative to a ‘reference transaction’:

The main findings of this research can be summarized by a principle of dual entitlement, which governs community standards of fairness:
Transactors have an entitlement to the terms of the reference transaction and firms are entitled to their reference profit. A firm is not allowed to increase its profits by arbitrarily violating the entitlement of its transactors to the reference price, rent, or wage. When the reference profit of a firm is threatened, however, it may set new terms that protect its profit at transactors’ expense. (KKT 1986: 729–30)

KKT emphasise that it is the change in the terms of the transaction which is important, not any particular desirability of the original or reference transaction:

It should perhaps be emphasized that the reference transaction provides a basis for fairness judgments because it is normal, not necessarily because it is just. Psychological studies of adaptation suggest that any stable state of affairs tends to become accepted eventually … Terms of exchange that are initially seen as unfair may in time acquire the status of a reference transaction. (KKT 1986: 730–31)

Interestingly, there is ample evidence from within the field of public-utility regulation that a change in regulated prices is often seen as unfair. Baumol (1988: 4) observes that: ‘It is important to recognize that “fairness” in a pricing arrangement depends heavily on consistency with the practices of the past to which people have become habituated.’

The same point is echoed by Kahn (1970, Vol I: 115): ‘As Ben Lewis has put it, “any scheme of compensation is fair provided only that it was reasonably anticipated at the time of the investment”.'

Along these lines, both Bonbright and Kahn go to great lengths to emphasise the importance of predictability and stability in regulated tariffs (Bonbright 1988: 387 and 187). A mail survey of a representative sample of Swiss households by Bruno Frey and Beat Gygi (1988, cited in Konow 2003: 1220) found that: ‘raising price in response to a demand increase is perceived as less unfair if the demand shift occurs at predictable intervals, alternative supplies exist, buyers are previously informed of and able to prepare themselves against the price increase, and sellers do not profit from the price increase’.

This is consistent with research by KKT which reveals that the permanence of the ‘reference transaction’ is a key determinant of the assessment of fairness. Specifically, it is not considered unfair to raise prices to eliminate a temporary discount or to cut wages by cancelling a worker’s bonus. For example, 61 per cent of respondents considered the following scenario unfair: ‘A small company employs several people. The workers’ incomes have been about average for the
community. In recent months, business for the company has not increased as it had before. The owners reduce the workers’ wages by 10 per cent for the next year.’

On the other hand, 80 per cent of respondents considered the following scenario fair:

A small company employs several people. The workers have been receiving a 10 per cent annual bonus each year and their total incomes have been about average for the community. In recent months, business for the company has not increased as it had before. The owners eliminate the workers’ bonus for the year. (KKT 1986: 732)

As an aside, the perceived unfairness of price increases can (of course) lead to situations where prices do not rise to a level that is able to clear the market. KKT argue that this can explain the temporary incidence of rationing that arises from time to time. Several subsequent studies have confirmed that rationing is often perceived as more fair than raising prices to clear the market (see Raux and Souche 2006, and Dickson and Kalapurakal 1994: 431).

Fact 2: It is not unfair to raise prices to cover an unexpected increase in long-run costs

Although price stability is an important component of fairness, there are circumstances where price increases will be tolerated and considered ‘fair’ — in particular where the price increase is required to cover a change in the long-term costs of the supplier. KKT illustrate this with the following two questions. Seventy-nine per cent of respondents considered the following scenario was fair:

Suppose that, due to a transportation mix-up, there is a local shortage of lettuce and the wholesale price has increased. A local grocer has bought the usual quantity of lettuce at a price that is 30 cents per head higher than normal. The grocer raises the price of lettuce to customers by 30 cents per head.

Similarly, 75 per cent of respondents considered the following scenario was fair:

A landlord owns and rents out a single small house to a tenant who is living on a fixed income. A higher rent would mean the tenant would

---

22 ‘Conventional economic analyses assume as a matter of course that excess demand for a good creates an opportunity for suppliers to raise prices, and that such increases will indeed occur. The profit-seeking adjustments that clear the market are in this view as natural as water finding its level — and as ethically neutral. The lay public does not share this indifference. Community standards of fairness effectively require the firm to absorb an opportunity cost in the presence of excess demand, by charging less than the clearing price or paying more than the clearing wage.’ (KKT 1986: 735)
have to move. Other small rental houses are available. The landlord’s costs have increased substantially over the past year and the landlord raises the rent to cover the cost increases when the tenant’s lease is due for renewal. (KKT 1986: 733)

KKT summarise this result as follows: ‘[I]t is acceptable for firms to protect themselves from losses even when their transactors suffer substantial inconvenience as a result. The rules of fairness that yield such judgments do not correspond to norms of charity and do not reflect distributional concerns’ (KKT 1986: 733).

Interestingly, the previous two stylised facts could possibly be interpreted as suggesting that pricing according to the costs of providing the regulated service is considered fair, whereas pricing according to demand conditions is considered unfair.23 This, in turn, suggests that the common regulatory requirement that prices be ‘cost-based’ reflects an underlying concern for fairness.

Fact 3: It is not unfair for the service provider to retain (or ‘share’) some of the benefits of a cost reduction

Importantly, fairness does not require that any cost reductions are passed on, dollar for dollar, to users. Fairness allows the monopolist to retain a share of the benefits of any cost reductions it achieves.24 KKT (1986: 734) note that 79 per cent of respondents considered the following scenario to be fair: ‘A small factory produces tables and sells all that it can make at $200 each. Because of changes in the price of materials, the cost of making each table has recently decreased by $40. The factory reduces its price for the tables by $20.’

Allowing the regulated firm to keep some of the benefits of cost reductions could, of course, be necessary to induce investment in cost-reducing effort, in line with stylised fact 5 below.

In contrast, the KKT survey suggested that an attempt to retain all of the benefits of a windfall gain is considered unfair. Seventy-nine per cent of respondents thought the following scenario was unfair: ‘A grocery store has several months supply of peanut butter in stock...on the shelves and in the storeroom. The owner hears that the wholesale price of peanut butter has increased and immediately raises the price on the current stock.’

---

23 Dickson and Kalapurakal 1994: 431–2: ‘Overall these studies suggest that there will be a higher perceived fairness for cost-based rules in price determination, especially those that treat cost increases and decreases symmetrically. They imply that the traditional “fair” price is one based on cost, rather than market conditions.’

24 KKT (1986: 734) observe that this is consistent with a ‘dual entitlement view’ which ‘suggests that the firm is only prohibited from increasing its profits by causing a loss to its transactors. Increasing profits by retaining cost reductions does not violate the transactors’ entitlement and may therefore be acceptable.’
Fact 4: Even where it is unfair to raise the out-of-pocket costs of a service, it is not considered unfair to raise the opportunity cost of the same service

In some cases a regulated firm may be able to, in effect, raise the ‘price’ of a service by lowering the price of the alternative. Importantly, it turns out that even where raising the price of particular consumption choice is considered unfair, it may be considered fair to raise the opportunity cost of taking the same decision.

For example, in the research of Raux and Souche (2003), peak-period pricing of TGV rail services or car-parking services was, as noted earlier, ‘overwhelmingly perceived to be unfair’. However, Raux and Souche also explored the possibility of rationing access to scarce capacity by ‘buying off’ those who were prepared to travel at another time (in the case of TGV services) or move to another car park (in the case of car-parking services). While only 10 per cent thought that rationing access to scarce TGV seats by raising the price was fair (whether the scarcity was one-off or recurring), 95 per cent thought that it was perfectly fair to ration access to scarce TGV seats by rewarding passengers holding reservations with a cash payment in exchange for switching to another day. This mechanism is, of course, similar to the practice of offering cash payments to airline customers to induce switching to another flight in the event of overbooking.

Similarly, while only 7–10 per cent thought it was fair to ration access to scarce car-parking spaces by raising the price, this increased to 89–93 per cent when the scarcity of car-parking spaces was rationed by ‘buying off’ customers willing to switch to another car park.

Fact 5: It is considered particularly unfair to raise the price for a service to exploit the good-faith effort or investment of the buyer in that service.

Importantly, the survey results suggested that respondents were particularly sensitive to the notion that the price might depend on certain characteristics specific to the buyer. Strikingly, more than 90 per cent of respondents thought the following scenario was unfair:

A landlord rents out a small house. When the lease is due for renewal, the landlord learns that the tenant has taken a job very close to the house and is therefore unlikely to move. The landlord raises the rent $40 per month more than he was planning to do.

Interestingly, the number of respondents who rated this outcome as unfair was 91 per cent — a higher ‘unfair’ rating than any of the other questions asked
in the survey. Respondents clearly found this particular change in the terms of the transaction — which resulted from the decisions of the tenant — was particularly unfair. KKT observe: ‘The near unanimity of responses to this and similar questions indicates that an action that deliberately exploits the special dependence of a particular individual is exceptionally offensive’ (KKT 1986: 735).

In summary, I suggest that the key characteristics of fairness in pricing can be summarised in the following five stylised facts:

1. It is considered unfair to raise prices above reasonably expected or anticipated levels.
2. It is not unfair to raise prices to cover an unexpected increase in long-run costs.
3. It is not unfair for the service provider to retain (or ‘share’) some of the benefits of a cost reduction.
4. Even where it is unfair to raise the out-of-pocket costs of a service, it is not considered unfair to raise the opportunity cost of the same service.
5. It is considered particularly unfair to raise the price for a service to exploit the good-faith effort or investment of the buyer in that service.

Is there an economic-efficiency rationale underlying these notions of fairness? This is the question examined in the next section.

**Fairness and the promotion of sunk complementary investment**

The final step in the argument of this paper is to show that we can explain the empirical regularities set out in the previous section using a simple economic model. In particular, I suggest that, to a large extent, these patterns of fairness can be understood as an attempt to preserve incentives for valuable sunk investments on the part of both the monopolist and its customers.

**Sunk complementary investment and public utility regulation**

Although economic textbooks have, of course, long recognised that public-utility firms are often required to make a substantial long-lived investment, the possibility of sunk investment on the buyer side of the market has been largely neglected.
Conventional economic theory has treated buyers as largely passive and inert and has, for the most part, assumed that everything we need to know about a buyer is reflected in the demand curve. However, as emphasised in Biggar (2009), buyers are seldom completely inert and passive. Rather, buyers must often make their own sunk complementary investments to extract the maximum value from the monopoly service. The literature on transactions costs groups the different kinds of sunk investments into the following categories:

- The decision where to locate, when that decision will have an impact on the demand for monopoly services (for example, close to a rail spur, or close to a mine mouth, and so on). These are known in the transactions-costs literature as ‘site-specific investments’ (see Crocker and Masten 1996: 8).
- The decision to invest in discovering or developing new skills or knowledge when the value of that knowledge relies on a continuing supply of the monopolist’s product or service. These are known as ‘human capital-specific investments’.
- The decision to invest in sunk customer-premises equipment or other assets which are specialised to the monopolist’s product or service (such as electrical wiring on the customer premises). These are known as ‘physical asset-specific investments’.
- The decision to invest in the development and marketing of new products or services which rely on a continuing supply of the monopolist’s service. We might call these ‘product-specific investments’.

Where the benefits of these sunk complementary investments exceed the cost, it will be socially desirable for the investment to be carried out. However, buyers fear that, once they make these investments, the monopoly service provider will increase its prices, expropriating some or all of the value of that investment. This threat of hold-up has a chilling effect on buyer-side investment. The failure of the buyer to make that investment reduces the economic value of the monopoly service or, in some cases, prevents it being provided at all; in either case, reducing overall economic efficiency.

For example, in the absence of regulation or a long-term contract, a firm may be unwilling to construct an aluminium smelter supplied by a monopoly electricity-transmission grid for fear that the grid will increase its charges in the future. Similarly, a freight forwarder may be reluctant to purchase specialised railcars for fear that the price or quality of access to the rail track network will decline in the future. The failure to make these complementary investments represents a loss of overall economic efficiency.\(^{25}\)

\(^{25}\) In some instances, of course, private arrangements can mitigate this hold-up problem. The clearest example of such an arrangement is a long-term contract between the monopoly service provider and its
Sunk complementary investment and fairness

Can this focus on the sunk investment of buyers explain the attitudes towards fairness in public-utility regulation identified above? We can make the following points:

First, this approach can explain the apparent ethical significance of the ‘status quo’ and the ‘unfairness’ of changes from the status quo. Buyers form expectations about the future path of tariffs and make their investment decisions accordingly. These expectations are typically formed by the current level and structure of tariffs, unless an element of the current level and structure is explicitly identified as temporary. A subsequent adverse movement in price or quality (relative to the expected level) is a potential threat to the value of that investment and therefore a potential threat to carrying out the socially valuable investment in the first place. By this hypothesis, therefore, an unanticipated adverse movement in prices will be judged as unfair. This is, of course, consistent with the first stylised fact identified above.

This link between ‘unfairness’ and sunk investments is highlighted by Meyer and Tye in their discussion of the difficulties associated with rapid rate changes following partial deregulation of the rail industry in the US. They note that where some consumers have:

sunk costs or made commitments that tie them for at least some time to particular vendors … the transition process can involve some aggrieved consumers who perceive themselves as unfairly victimized by the transition. The difficulty is created by an overhang of sunk costs committed under the prior rules of regulation. (Meyer and Tye 1985: 50)

Bonbright also notes an economic argument, based on sunk investments, that can explain the traditional focus on ensuring rate stability:

The argument runs to the effect that the ratepayers were induced to locate their factories, or to abandon their isolated generating plants, or to convert their furnaces from coal to gas burners, in contemplation of the low promotional rates and on the faith that that this rate would remain in effect for the indefinite future. (Bonbright 1988: 187)

At the same time, however, the promotion of rate stability could threaten the incentive of the regulated firm to make its own sunk investment, if the pursuit of rate stability prevented rates rising in response to a permanent increase in customers. However, in the case of many public-utility services the transactions costs of negotiating such contracts with customers render this approach infeasible. Instead, the public-utility regulator takes on the role of creating and enforcing the long-term contract that the parties would have written had they been able to negotiate before either party made any sunk investment. See Goldberg (1976); Biggar (2009).
costs. By this hypothesis, therefore, we would predict that a rate increase would be perceived as fair if it was necessary to cover a permanent increase in costs, consistent with the second stylised fact noted above.

On the other hand, an unanticipated reduction in costs poses no threat to the sunk investment of either the regulated firm or its customers. We might predict that it would not be labelled unfair to not pass on the full impact of a cost reduction to consumers. (Indeed, if the reduction in costs is brought about, at least in part, by cost-reducing effort on the part of the regulated firm, we might predict that it would be considered unfair to not allow the firm to share in some of the benefits of that cost reduction.) Overall, these predictions are consistent with the third stylised fact identified earlier.

Importantly, the mechanism used for rationing scarce supply can have a significant impact on the incentives for investment by buyers of the monopolist’s service. The decision by a consumer whether or not to make a sunk complementary investment involves a trade-off between the cost of the investment and the future increase in consumers’ surplus which that investment brings about. In the event that demand exceeds supply ex post (either due to an increase in demand or a shortfall in supply) the regulated firm must adopt some mechanism for rationing the excess demand. Where this excess demand is rationed by raising the price, the ex-post consumers’ surplus is reduced for all consumers — both those who choose to continue to consume and those who choose not to consume. This has a chilling effect on ex-ante investment. On the other hand, where excess demand is rationed by ‘bribing’ some consumers not to consume, the consumers’ surplus of all consumers is preserved — each consumer either continues to consume at the present price or is ‘bought off’, presumably when the reward for not consuming exceeds the consumers’ surplus from continuing to consume.

In other words, ex-post rationing through price results in a lower ex-ante surplus to the consumer and exposes the consumer to the risk of ex-post congestion compared to ex-post rationing by bribing some consumers not to consume. Unless the monopolist can directly subsidise the investment by the consumer and can make a credible commitment to maintain a given level of congestion, ex-post rationing through price will have a chilling effect on ex-ante complementary investment. By this hypothesis we would expect that unanticipated peak pricing would be considered unfair, while paying some consumers to not consume to relieve unanticipated excess demand would be considered fair, consistent with the fourth stylised fact set out above.

Finally, it is clear that raising the regulated tariff ex post on just those customers who have made a sunk complementary investment will have the effect of deterring further investment and therefore will be considered unfair. This
might arise, for example, if carrying out the sunk investment shifts the customer into a different tariff class with a higher price. By this hypothesis, these tariff structures would be considered to be unfair or 'unduly discriminatory', consistent with the fifth stylised fact above. At the same time, this approach allows us to suggest which forms of price discrimination would be considered fair — where customers of the regulated firm have had sufficient notice to adjust their own complementary investments to the pricing structure and where the tariff structure does not discriminate between customers on the nature or extent of their sunk complementary investments.

Similarly, this approach predicts that peak-load pricing could, under some circumstances, be considered acceptable — particularly when customers have sufficient notice to adjust their own complementary investments, where the frequency of the peak periods is fixed in advance, and where the peak prices themselves are fixed in advance. There is some evidence to support this. Raux and Souche (2006) have shown that people are less averse to pricing in a recurring situation where there is time to adjust to the price fluctuation than in an exceptional situation where pricing serves to ration demand.

In summary, I suggest that the notion of fairness in public-utility regulation — and in public pricing more generally — can be explained as an attempt to protect and thereby promote sunk complementary investment by users and consumers. I propose that regulatory decisions consistent with the hypothetical efficient ex-ante long-term contract between the regulated firm and its customers will be considered fair, while regulatory decisions which are inconsistent with that hypothetical ex-ante long-term contract will be considered unfair.

**Alternative explanations?**

Is there an alternative approach which can better explain the pattern of attitudes to fairness set out above? For example, some economists might argue that claims of fairness are merely a veil for the pursuit of self-interest. Perhaps regulators, in pursuing fairness over traditional economic efficiency, are merely bowing to the demands of powerful interest groups? As noted earlier, some research has indeed found that fairness judgments are biased by self-interest.

But fairness seems to be more than a mere reflection of self-interest. The research of KKT (amongst others) shows that disinterested, objective observers hold consistent notions of fairness even when they have nothing at stake. If fairness is nothing more than a reflection of the self-interest of consumers, why would it be considered fair for service providers to retain some of the benefits of a cost reduction (Fact 3)? If fairness is a reflection of the self-interest of the regulated firm, why would it routinely be held to be unfair to raise prices above expected
or anticipated levels (Fact 1)? The self-interest hypothesis seems to have trouble explaining these aspects of attitudes towards fairness. Fairness, as I have sought to argue, seems to be something more than a mere reflection of self-interest.

What about the axiomatic approaches to fairness discussed earlier? As noted earlier, certain cost-allocation methodologies, such as the Shapley value, can be justified as satisfying certain desirable axioms. But, as noted earlier, this line of research seems to do a poor job of explaining public attitudes to fairness. Specifically, as already noted, fairness seems more closely related to changes in allocations than to the allocations themselves.

We are left with the hypothesis put forward in this paper — that fairness concerns can be explained as an attempt to protect and promote the sunk complementary investments of the regulated firm and its consumers. Does this hypothesis have any testable implications? One possible implication is that concerns of fairness or unfairness would be expected to arise more frequently in those sectors in which customers have made a material sunk investment in a relationship with the supplier. As noted above, sunk investments on the part of customers are reasonably common. However, those sunk investments are not specific to a particular supplier precisely in those markets where the customers can switch easily between suppliers — that is, in those markets which we would normally label competitive. Therefore, an implication of this hypothesis is that we would expect to see concerns of fairness or unfairness arise much more frequently in those sectors in which customers have no choice of supplier (that is, in public-utility industries) than in those sectors which are reasonably competitive (and in which customers make sunk investments specific to an individual supplier). This implication seems broadly consistent with casual observation.  

However, even where customers face some choice of supplier they may be required to make some degree of relationship-specific investments. Where those investments cannot be adequately protected through long-term contracts, we might expect to see fairness concerns, and some of the pricing outcomes discussed here, arising in normal commercial transactions. Blinder et al. (1998), in a study of price stickiness, find that around two-thirds of private, non-farm, unregulated, for-profit US firms responded ‘yes’ when questioned whether firms have an implicit understanding with their customers, who ‘expect the firms to not take advantage of the situation by raising prices when the market is tight’. Interestingly, consistent with the results reported above, Blinder et al. find that

26 Another observation we might make is that both firms and households must make sunk complementary investments as customers of a regulated monopoly firm. Therefore another implication of this thesis is that we should expect to find no material difference in fairness concerns when the primary customers of the monopolist are other firms compared to the case where the customers are households (except where the customer firms are sufficient large and are present before any investment by the monopolist has been sunk, so that the negotiation of long-term contracts is a viable alternative).
these implicit contracts tolerate price increases when costs increase but do not symmetrically insist on price reductions when costs decrease. ‘Furthermore, the notion that prices should be stabilised seems to apply a bit more to demand shocks than to cost shocks’, confirming the suggestion by Okun (1981) that ‘price increases that are based on cost increases are “fair” while those based on demand increases are often viewed as “unfair”’. These findings are consistent with the discussion in this paper. The central claim of this paper — that fairness in public-utility pricing can be understood as an attempt to protect and promote sunk investments — may extend to commercial pricing practices more generally.

Conclusion

Economists have long viewed users and consumers as essentially passive — so that all the relevant behaviour of consumers can be summarised in the demand curve. In this framework, the most important goal of the regulator is to minimise the deadweight loss. In fact, users and consumers can, and often must, make substantial sunk investments to extract the most value from infrastructure assets. The lack of recognition of these investments in mainstream regulatory policy has contributed to a lack of understanding and conflict between the allocative-efficiency aims of economists and the fairness notions of regulators. I suggest that the desire to protect and thereby promote these sunk complementary investments can explain much of the notion of fairness as it applies to public-utility regulation and public pricing more generally. In my view, by taking into account the impact on these sunk investments in their policy advice, economists will go some distance to bridge the remaining gap between economic theory and regulatory reality.

References


Willig, R. D. 1979, ‘Customer Equity and Local Measured Service’ in Joseph A. Baude et al. (eds), Perspectives on Local Measured Service, Kansas City.


ARGUMENT
Formulating Policy Responses to Global Warming in the Face of Uncertainty

HARRY CLARKE

Abstract

The effects of risk and uncertainty on climate-change policy design are analysed. It is argued that while the science of climate change involves a consensus that anthropogenic greenhouse-gas emissions are driving higher temperatures and a changed global climate, there remains uncertainty about the extent of such changes and the risk of catastrophically large temperature increases in the absence of emission mitigation. The possible cost consequences of not addressing climate change are much greater than the costs of addressing it so that, from a range of perspectives, activist mitigation policies make sense. A critical assessment of Australian climate-change policy is provided, emphasising the key specific uncertainties that impact on policy design.

Introduction

The paper is concerned to discuss three aspects of climate-change policy. The paper begins with the role of uncertainty in the background science. A distinction is raised between the role of the explicit scientific uncertainty that forms such an important part of the work of the Intergovernmental Panel on Climate Change (IPCC), and the contrived uncertainty promoted by those that will be referred to here as ‘climate-change delusionists’. The paper then turns to the broad economics of climate change, with a focus on the impact of risk and

---

1 This paper is based on a presentation to a meeting of the Economic Society of Australia, Victorian Branch. I thank Des Moore and those who attended for the ensuing enjoyable, vigorous debate. Without implication, I also thank Editor William Coleman, John Quiggin and two anonymous referees for their comments.

2 La Trobe University, H.Clarke@latrobe.edu.au.
uncertainty on our understanding of climate-change economics. There follows a critical appraisal of Australian climate-change policy, emphasising the role of uncertainty.

Climate science, climate uncertainty and climate delusionism

Scientific consensus

Modern climate science developed in the early nineteenth century from observations in the 1760s by Horace Saussure about how actual greenhouses trap the sun’s heat. In the early nineteenth century it was suggested that components of the atmosphere trapped heat energy analogously to the way a greenhouse did so. These ideas were developed by famous scientists and mathematicians of the day including Joseph Fourier, John Tyndall and others. In important work around 1895, Svante Arrhenius linked the heating effect of the atmosphere to the logarithm of its CO$_2$ concentration. Subsequent work identified other so-called greenhouse gases, while atmospheric physics provided a theoretical basis for understanding the global-warming issue in terms of the vibration rates of certain gas molecules in the atmosphere (IPPC 2007: 103–6).

The predictions of this theoretical work are supported by the evidence. It is now certain that warming is occurring — this is a matter of fact — and, with high probability, that the warming that has occurred since pre-industrial times is due to the accumulation in the atmosphere of anthropogenic ‘greenhouse gas emissions’ (GGEs). These emissions are linked to the burning of carbon-based fossil fuels, such as coal and oil, and also to modern agriculture and land-clearing activities.

The consensus of scientific opinion suggests that, ignoring slower feedbacks, the doubling of CO$_2$ concentrations in the earth’s atmosphere over pre-industrial concentrations will bring about an approximate expected mean global surface temperature increase of about 30°C — the Charney sensitivity. This increase will be more intense in the earth’s Polar Regions and there will be substantial regional variations.

CO$_2$ concentrations have grown from 280 parts per million (ppm) in pre-industrial times to 379 ppm in 2005 (IPPC 2007: 2). Growth rates in global emissions over

---

3 The Charney sensitivity is defined as the global mean surface temperature anomaly response to a doubling of CO$_2$, assuming that surface, ice sheets and atmospheric composition (chemistry and aerosols) stay the same (National Academy of Sciences 1979).
the last two years have been distorted, in a temporary and inessential way, by the effects of the global financial crisis on demands for carbon-based fuels. However, projecting smoothed emission-growth trends, and ignoring the impact of an active global mitigation program, a doubling of CO\textsubscript{2} concentrations over 2005 levels should occur in about 40 years.

Science and uncertainty

Much information has been accumulated about the process by which climate change is occurring. The degree of certainty regarding the anthropogenic-warming hypothesis has strengthened, although (paradoxically) the range of temperature forecasts from climate science remains substantial. The range of forecasts in the third and fourth IPCC Assessment reports has not narrowed. Increased knowledge has heightened awareness of uncertainties as much as certainties.

For example, the Charney temperature sensitivity reflects only a narrow range of short-run simple climatic drivers. This sensitivity assumes that land surface, ice sheet and atmospheric/aerosol compositions do not change in character or chemistry as climate change occurs, thereby providing positive feedback loops that further increase temperatures. Longer-run sensitivity measures such as Hansen’s long-term sensitivity (the ‘Earth System sensitivity’ or ‘ESS’) allow these factors to vary and hence for the feedbacks to help determine climate (Hansen et al. 2008). The ESS suggests that global mean surface temperatures will rise 6°C with a doubled CO\textsubscript{2} concentration but over a period of hundreds of years rather than a few decades. Decisions to include specific feedback effects in a sensitivity measure are based on the timescale and complexity of feedback processes. Until recently, slow feedbacks have been excluded, as have complex feedbacks due, for example, to changed ozone/aerosol concentrations. In principle, the ESS could be estimated from paleoclimatic records but this requires good data on global temperatures as well as on the various climate forcings that operated over extended periods. This data is difficult to obtain.

Science does not have perfect information about how climate sensitivities vary with time horizon and with the various forcings. For example, some feedbacks drive higher sensitivities and aerosol clean-up may ‘unmask’ greater heating. For example, ice-sheet melting, as a consequence of warming, creates stronger long-run heating responses through reduced albedo (heat reflection) effects.

---

4 Aerosols comprise the suspended fine solid particles or liquid droplets in a gas. Examples are smoke, oceanic haze, air pollution and smog. Industrial aerosols absorb and scatter solar radiation and can have heating as well as cooling effects on the atmosphere.
A fear held by many is that the current climatic situation is deteriorating. That 11 of the 12 years over the interval 1995–2006 were among the 12 warmest since 1850 highlights concerns about sensitivity assessment (IPCC 2007: 5). Moreover, the growth in temperatures has continued beyond 2006. BOM data indicate that Australia’s annual mean temperature for 2009 was 0.9°C above the 1961–90 average, making it the nation’s second-warmest year since high-quality records began in 1910 (BOM 2010).

It is worth emphasising that modern climate science makes no attempt to conceal uncertainties in knowledge. The IPCC Assessment reports focus on uncertainty with, for example, extensive discussion of the problematic role of clouds, aerosols, Antarctic sea ice and so on as climate drivers. Indeed, in The Physical Science Basis Report (IPCC 2007: 81–91) there are 43 key uncertainties listed in relation to natural/human climate drivers, six on global projections, two on sea-level changes and three on regional forecasts. None of these issues are minuscule. It is therefore wrong to portray institutions such as the IPCC as comprising propagandist ‘greenies’.

**Delusionism**

The key doubt some sections of the media and certain interest groups seek to promulgate is whether the entire science of climate change is wrong. The science would be wrong if it were not the case that human beings are significantly altering the world’s climate — if the anthropogenic global-warming hypothesis (AGWH) was false because the warming that has been observed is natural.

This question of the believability of the science was examined in the Garnaut Review (2008). In a recent speech Garnaut (2009) commented:

> The Review accepted the views of mainstream science ‘on a balance of probabilities’. There is a chance that it is wrong. But it is just a chance. To heed instead the views of the small minority of genuine sceptics in the relevant scientific communities would be to hide from reality. It would be imprudent beyond the normal limits of human irrationality.

This seems indisputable. It seems sensible to act on the presumption that the mainstream science is correct and imprudent to suppose otherwise.

Oreskes (2009) confirms Garnaut’s presumption that almost all climate scientists support the AGWH. In a bibliographic survey, she found no explicit rejections of the hypothesis in the published scientific literature. Specifically, of 928

---

5 Key uncertainties include the radiative effects of aerosols, reasons for recent changes in CH3 concentrations, land-atmosphere interactions that lead to radiative forcing, disagreements on the role of cloud cover and *in situ* ice cover and trends in glacier ice loss.
abstracts containing the keywords ‘global climate change’ in the database of the Institute for Scientific Information (ISI) from 1993–2003 not a single paper provided scientific data to refute the consensus scientific position on climate change.

There is a group of contrarians but they are not now doing scientific research on this topic. In a further recent study, Doran and Zimmerman (2009) found overwhelming support for the AGWH, with 75 out of 79 self-described ‘climatologists’ — out of 10,257 geoscientists in US federal research facilities — affirming belief in AGWH. Quoting Doran and Zimmerman (2009):

> It seems that the debate on the authenticity of global warming and the role played by human activity is largely nonexistent among those who understand the nuances and scientific basis of long-term climate processes. The challenge, rather, appears to be how to effectively communicate this fact to policymakers and to a public that continues to mistakenly perceive debate among scientists.

The consensus view in science may be incorrect, as it was historically with respect to Alfred Wegener’s theory of continental drift. But such historical examples only reinforce the need for humility — the scientific consensus may be wrong ⁶ and there may be a need to revise views as knowledge advances — but, as yet, no new knowledge has arisen seriously questioning AGWH.

As noted, doubts about AGWH do not reflect the state of science. There is, however, a mistaken public perception that there is a substantial scientific debate. It is worthwhile trying to understand how this false perception has developed.

The politics of climate change was decisively driven by Revelle and Suess (1957), which identified global heating caused by anthropogenic CO₂ emissions as a significant policy concern. In the words of these authors, human use of fossil fuels amounted to a ‘large scale geophysical experiment’. Through to the early 1990s there was a consensus of international political opinion supporting AGWH that led to the 1992 Climate Convention and to an ongoing role for the IPCC.

From the late 1980s, however, organised groups emerged who denied AGWH and who sought to force a public debate on whether AGWH was believable. These comprise a tiny group of respectable sceptical scientists, such as the atmospheric

---

⁶ It is sometimes argued that a consensus has low informational value if individual scientists irrationally ‘borrow’ each other’s views. For this to have force it must be shown that such irrational borrowing occurs. The evidence suggests independence of view, both on the basis of the range of independent institutions holding the same consensus attitudes and the bibliographic studies cited which reject the major competing alternative hypothesis — that the warming which is occurring is ‘natural’.
physicist Richard Lindzen; a larger group who, for a living, manufacture delusions on many topics; and a much larger group of conservatives who accept ideologically convenient delusions. These groups interact to a considerable extent. I label them as ‘delusionists’. These well-funded groups have worked to create and disseminate uncertainty, ignorance and confusion on climate-change issues. They have been extraordinarily successful, with much of the political right in Australia and almost all in the US being deluded into thinking the climate-change problem is illusory. In Australia, most National Party MPs, many in the Liberal Party, a few in the Labor Party and some academics in universities endorse delusionism.

The delusionist groups are an influential political movement which promotes the rejection of mainstream science. In the US, these groups centre on the George C. Marshall Institute and the Heartland Institute. In the past, these groups have suggested that passive smoking may not cause health damage, that CFCs in the atmosphere did not cause the hole in the ozone layer, and that SO2 did not cause acid rain. None of these claims stand up to scrutiny. Furthermore, it is a matter of public record that these groups have in the past received funding from Exxon-Mobil, and their officials (Fred Singer and Frederick Seitz, for example) have worked as consultants for the tobacco companies and written reports suggesting that secondary tobacco smoke might not be harmful to health.

Australian delusionism draws heavily on British and American delusionism. An important local offshoot is the Lavoisier Group, which has enjoyed success in influencing public opinion beyond its due.

A major activity of delusionist groups is to create the false impression of a substantive debate in climate science. As Oreskes (2007) and Doran and Zimmerman (2009) show, there is none. The false implication that there is a substantive debate is further twisted to suggest that the case for activist climate policies is somehow weakened by the existence of such a debate. The implication that the false doubt undermines the case for activist policy is an invalid conclusion in several senses. Most importantly, with correctly specified uncertainty, things might indeed be ‘worse’ — not ‘better’ — than expected as is discussed below, thereby intensifying the case for activist climate-change policy.

A disturbing feature of delusionist activity is the repetition of discredited or heavily criticised viewpoints. An example is the repeated reputed claim that global warming stopped during the 10 years after 1998. This was rejected using evidence collated during the preparation of the Garnaut Review and by statisticians and climate scientists on numerous occasions since then. The

---

claim is still repeated mostly without any reference to the contrary claims by eminently qualified statisticians and climatologists (Breusch and Vahid 2008; Fawcett and Jones 2008) who show that the claim is unsupported by evidence.

Delusionism is fostered in Australia by spurious notions of ‘press balance’, according to which any view, no matter how far-fetched, has a right to equal treatment in the public domain. Academics, such as the geologist Professor Ian Plimer, promote delusionism in Australia through books such as *Heaven and Earth* (2009), which deny AGWH. According to his academic website, Professor Plimer has never published a refereed scientific paper in the field of climate science. ⁸ Indeed, the analysis of Plimer’s book is inconsistent with science.

As an economist, it is not my right or role to question the veracity of mainstream science. Indeed, I generally refuse to debate the mainstream science of climate, and believe that economists whose policy preferences are secured by rejecting science have lost their bearings. I do, however, read the science, and am familiar with the arguments advanced by delusionists. Rebuttals to their arguments are simple to find in the scientific literature (for example, IPCC 2007). There is no more reason to question the validity of science than to take seriously the crank monetary theories sometimes pushed by physicists. ⁹

**Climate-change economics**

Climate-change economics seeks to determine the economic impacts of unmitigated climate change and the costs of mitigating and/or adapting to it, either at the global level or at the level of an individual project such as, for example, a forest planting. Given these estimates, the case for policy reverts to a cost-benefit analysis based on opportunity costs. What are the costs of pursuing active climate policy compared to the costs of not doing so?

Inputs into this endeavour include information from science-based models that link the emission of greenhouse gases to climate changes and to the damage functions that suggest the costs that will be inflicted by such temperature changes. There is also a need to forecast the human responses, such as climatic adaptations and the induced innovations that will be associated with climatic changes. This is an ambitious task.

---

⁸ See http://www.ecms.adelaide.edu.au/civeng/staff/iplimer01.html#research.
⁹ This line of argument was suggested to me by John Quiggin. An example of a crank-physicist economic theory is Frederick Soddy’s critique of fractional reserve banking (Zency 2009).
Cost benefit analysis

There are problems of valuing the damages climate change will inflict on societies and, with demand and technology uncertainties, the costs of policies designed to offset such damages. But even ignoring such valuation issues, the complexity of climate-policy analysis is driven by its formal character as a particular instance of cost-benefit analysis (CBA). The ingredients of this CBA involve several types of irreversibility coupled with learning processes, nonlinear responses and threshold effects, all thrown together in an intrinsically dynamic setting where uncertainty is an intrinsic ingredient. Such CBA tasks can be analysed in a distinctive ‘real options’ setting (see Pindyck 2006; Dobes 2008).

Analysts are highly uncertain about the extent of likely climate change, of associated costs and hence of the viability of various technological fixes such as carbon-capture and storage (CCS) technology. They have information about these issues but can expect to learn more as the future unfolds. Thus, information about climate evolution and valuations are stochastic processes. These learning dynamics supplement the uncertain intrinsic dynamics associated with the accumulation of the stock GGE pollutants in the atmosphere.

There are also potential sun-cost irreversibilities and non-linearities associated with threshold effects. First, sunk costs associated with investment designed to address climate-change effects will involve irretrievable waste should business-as-usual climate-change costs prove less costly than forecast. These irreversibilities motivate policymakers to delay and de-intensify policy responses until better information becomes available. They are partially offset by the increasing costs of abatement that can arise with delays in responding. For example, incurring the incremental costs of expanding electricity supplies using new carbon-friendly technologies (solar, nuclear) are lower than replacing pre-existing carbon-intensive technologies with those that are carbon-friendly. Delay therefore locks in emissions-intensive infrastructure which constrains future abilities to mitigate. This is especially true for developing countries where incremental costs of energy capacity are much lower than replacement costs (Office of Tony Blair 2009).

There are also irreversibilities associated with potential threshold effects that give rise to catastrophic outcomes. With respect to biodiversity conservation, species extinctions induced by climate change are irreversible costs, while steadily worsening climate change generally might induce non-linear cost responses if ice sheets or the Arctic permafrost melts. Provided that policy actions can address these concerns, the prospect of such phenomena motivates policymakers to seek prompt and intense responses.
The differing sunk-cost and environmental irreversibilities tug policy analysis in opposite directions. The sunk-cost irreversibilities encourage delay in order to learn, while the prospects of catastrophic irreversibilities motivate decisive action now. Pindyck (2006) has sought via numerical simulations to find out which effects dominate — his emphasis is on sunk-cost effects — but, as he has no actual data, his arguments are unpersuasive. Similar comments apply to work by Kolstad (1994 and 1997) emphasising the role of sunk costs.

Some policy insights can be formulated on the relative effect of the two irreversibilities. A policy measure which reduces sunk costs increases the case for immediate, decisive action. For example, sunk-cost irreversibilities are lower if investments provide benefits, even if climate change should prove to be less damaging than expected — this is equivalent to Kolstad's (1994 and 1997) case for using reversible policy tools as a 'no regrets' option. Clean-power technologies provide useful local environmental benefits in terms of reduced air pollution, while enhanced biodiversity-conservation programmes provide valuable spill-over benefits in promoting agricultural sustainability even should they prove to be based on overly pessimistic climate assumptions. Less of the costs in these situations are now sunk.

Reducing the prospects of catastrophic damages on the other hand delivers increased flexibility for policymakers by putting less weight on feared outcomes. Geo-engineering carbon clean-up options even at high cost, or costly captive-breeding programmes of endangered biodiversity to limit species extinctions, provide a backstop technology that gives policymakers extra time to learn about prospects and implications of severe climate change.

As a general theoretical paradigm, policymakers should think of climate-change issues using this 'real options' perspective that accounts for risk and these irreversibilities. This amounts to using CBA under conditions of risk where irreversibility interacts with structural and learning dynamics. It is inappropriate to account for uncertainty by replacing random variables by their expected values since the effects of waiting to learn and of delay must be accounted for.

**Discounting**

What weight should be placed on the welfare of people today relative to that of their children and of future generations? Should market or intergenerational equity-based discount rates be used in valuing future costs and benefits? This

---

10 It is difficult, however, to agree with Kolstad's arguments for temporary rather than permanent tax-increase measures since these will not lead to the types of long-lived investment decisions one would hope energy-supply agencies would make.
is a matter of practical importance. Stern (2007) uses a discount rate of 1.4 per
cent which values a dollar to be received in 32 years at $(1.014)^{-32} = 73$ cents.
Nordhaus (2008) prefers 3 per cent, putting a current valuation on a dollar of
$(1.03)^{-32} = 38$ cents.

It is important to understand that a 3 per cent choice has major intra-generational
implications. Such a choice puts a weight of only 0.38 on my 11-year-old son’s
welfare in 32 years, and that is subjectively less than I think is appropriate.
Hence, a 3 per cent discount rate is too high for me and my intra-generational
objectives. I am more comfortable with low Stern-type discount rates because I
wish to provide a sound environmental future for my children.

The choice of discount rate is an uncertain issue but the convexity of the
discount factor in the discount rate means that, if an expected rate is to be
used, it should be ‘small’ (Pindyck 2007). Thus, suppose the expected present
value of a $100 benefit to be received 100 years from now is to be determined
but that the discount rate is either 0 or 10 per cent each with probability $\frac{1}{2}$. At
the expected discount rate, 5 per cent, $100 has a present value less than $1.
But the expected present value is $\frac{1}{2} \times (100) + \frac{1}{2} \times (0) = 50$ and the discount rate
applied to a $100 benefit to be received in 100 years that yields a present value
of $50 is 0.7 per cent. Thus even though the expected value of the discount rate
is 5 per cent, uncertainty over its value implies an effective discount rate of less
than 1 per cent.

Risky damages

Furthermore, environmental damages themselves are certainly convex in future
temperatures — they increase with higher temperatures at an increasing rate.
Therefore suppose the IPCC forecast of a $3^\circ C$ increase in global mean surface
temperature is associated with moderate economic damage equivalent, for
example, to 5–10 per cent of GDP depending on how environmental damages are
assessed. If the IPCC forecast is a mean forecast it can be interpreted as a forecast
of $2^\circ C$ with 50 per cent probability and $4^\circ C$ with 50 per cent probability, or if
there is greater uncertainty, as a forecast of $0^\circ C$ with 50 per cent probability
and $6^\circ C$ with 50 per cent probability. Without warming there are no damages
but with an equal probability of warming of $6^\circ C$ there would be catastrophic
costs that should be avoided. More risk — in the sense of a mean-preserving
increased variance in forecast temperature changes — implies a stronger case
for effective climate-change policy action whenever catastrophically large costs
are associated with high temperature increases.

This issue can be looked at in a different way. Suppose there are various possible
future climate states contingent on current climate policies but one state where
catastrophic costs (for example, melting of the Greenland ice sheet and hence
drastic sea-level rises) occur with non-negligible probability. Then, rational policymakers who can have any significant impact on global outcomes should act to avoid that state irrespective of discount rates or even strategic concerns in relation to the policies of other countries. The prospect of preventable catastrophes at non-negligible probabilities favour unilateralism and prompt policy action in relation to climate change (Clarke 2010; and Clarke and Reed 2006).

**Knightian uncertainty**

These probabilistic insights depend on the assumption that information about probabilities can be provided by climate science or that at least subjective probabilities can be assigned. Insights can, however, be sought in settings with pure Knightian uncertainty without probability information.

Clarke (2008) suggests use of classical decision rules not involving use of even subjective probability information. Thus the Precautionary Principle (PP) can be understood as an attempt to avoid severe costs by introducing policy to avoid the worst possible outcomes. This is a version of the classical Minimax Principle which seeks to select the least bad possible outcome. This plausible-sounding heuristic is, however, unhelpful unless the policymaker can be certain that costly policy, once it needs to be undertaken, will not fail. This assurance is difficult to provide. Unless a policy is known with complete certainty to be effective, the most costly contingency is undertaking a costly policy which fails, in the event, to deal with climate change. Thus, from this perspective, PP would provide the counterintuitive direction to remain policy-inactive, irrespective of the policy problem or the available policy instruments.

An alternative heuristic that underlies much modern policy thinking and which lends conditional support for policy activism is Minimax Regret. Here, policy interventions are undertaken if unmitigated climate change occurs, involving substantial costs which could have been avoided at much lower cost, thus potentially creating regret. This heuristic provides a probability-free insurance heuristic that many of us doubtless use in everyday life — we insure our homes against fire damage not by using subjective probabilities and cost data but because we see that the costs of a fire will be vast compared to the insurance cost. Inexpensive policy will now be undertaken if it addresses concerns that might have vast cost if left unaddressed; and if climate change proves not to be as serious as thought, the opportunity losses are low. This case for policy disappears if policy costs are only ‘somewhat’ less than the costs avoided without policy but, as is argued below, empirical evidence from various perspectives confirms that in a climate-change setting policy costs are much lower than possible costs of not taking action.
Earlier work along these lines by Arrow and Hurwicz (1972) showed that, under ignorance, choices based on extreme outcomes satisfy plausible rationality criteria. This work was implemented in a climate-change policy-decision context by Woodford and Bishop (1997). Here a ‘panel of experts’ was portrayed as offering divergent opinions and a policymaker had to choose between them without being able to assess which was more likely to be correct. Without allowing for the possibility of policy failure, the work suggested a case for strong immediate control action to address catastrophes followed by a switch to less-intensive policies if evidence eventually suggested small damages. This is broadly consistent with the arguments made above, although inconsistent with those of Kolstad (1994 and 1996).

Finally, there are heuristic reasons for responding to policy uncertainty by both diversifying choice of policy options and by retaining, as an option, the facility to introduce, perhaps expensive, backup policies if all else fails. In climate-policy settings, various policy approaches should be employed without a telescopic committed focus to a particular approach. Where possible a portfolio of policies should be explored and policy choice should be adaptive and sensitive to the unexpected. Backstop insurance options should be researched, and retained as longer-term options, should current policy fail. For example, if mitigation options fail to be effective, geo-engineering approaches involving, for example, direct removal of CO$_2$ from the atmosphere can provide a backstop technology as insurance against disaster. Hansen et al. (2008) conjecture that, at $200/tC, CO_2$ can be directly removed from the atmosphere. The cost of a once-and-for-all removal of 50 ppm of CO$_2$ globally would be $\sim$20 trillion, which is of the order of 150 per cent of US GDP in 2009. Such policies can be unilateral and implemented by a wealthy superpower without the ‘free-rider’ issues that thwart more-conventional mitigation strategies that rely on international cooperation.

With respect to biodiversity conservation, proposals to strengthen existing reserves by expanding their size in isolation might be pursued to minimise the risks arising from unsought species migrations. Other strategies that should also be independently considered involve almost the opposite strategy of linking up existing reserves to facilitate sought migrations (see Dunlop and Brown 2008). If these strategies jointly appear likely to fail, an expensive backstop is a captive-breeding programme.

On secondary energy technologies there are substantial uncertainties on the commercial potential of low-CO$_2$ technology options and fourth-generation nuclear power. There are, for example, specific issues of the economic viability of CCS which would not seem to be economic at carbon prices much less than US$60/ton. At this stage, a portfolio of technologies (renewable, nuclear) should be researched, with a focus on CCS.
Focusing on CCS alone and on renewable technologies such as solar and wind is imprudent given the imperatives Australia will face to cut its GGEs dramatically by 2050. The possibility of a major shift toward nuclear fuels should be entertained. Nuclear fuels have potential economic advantages over coal, gas and wind technologies once environmental costs are accounted for, and nuclear power involves negligible CO$_2$ emissions. In 2004 France generated slightly less than 80 per cent of its electricity using nuclear power and emitted about 9.3 tonnes of CO$_2$ per head in 2003. Australia that year emitted 26.1 tonnes of CO$_2$ per capita (figures from the UN’s Globalis websites). Suppose, then, that the world agrees to reduce its aggregate carbon emissions from 50 mega-tonnes now to less than 20 mega-tonnes by 2050 to avoid more than 2°C of warming. This amounts to 2.2 tonnes per person in 2050 with an estimated world population of 9.1 billion assuming equal per-capita entitlements. To achieve this, France would need to cut its GGE emissions to less than 23.9 per cent of current levels. Australia, however, would need to cut its emissions by about 71 per cent of current levels!

Despite the uncertainties concerning various particular aspects of climate-change economics, there is an economic consensus on these issues that parallels the consensus in climate science. Most economists agree, for example, that while some countries benefit most lose with limited climate change. All lose with substantial change because positive effects on agriculture then disappear (Stern 2006: 62).

In addition, even with substantial uncertainties there is a presumption that the cost of active policy is low relative to the cost of doing nothing. As evidence, note that the Stern Review (2007) estimates the cost of stabilising at 500–550 ppm at approximately 1 per cent of world GDP, whereas the costs of not addressing climate change are 20 per cent of GDP. The IPCC (2007b) estimates the 2030 costs of stabilising at from 445–535 ppm at less than 3 per cent of GDP, with ongoing growth of GDP reduced by less than 0.12 per cent. Nordhaus (2008) targets lower levels of cutback than does Stern (and Nordhaus is very critical of Stern’s pessimism) but nevertheless estimates optimal costs of abatement at only 0.1 per cent of discounted world income. Finally, Weitzman (2009) focuses on possible catastrophic costs of not abating. Gradually ramping up climate-change policies over the next 200 years creates significant global catastrophic risks and huge costs. There is a 5 per cent probability of a mean global surface temperature increase of greater than 10°C and a 1 per cent probability of an increase of greater than 20°C.

These studies employ distinct modelling strategies, but a consensus is that mitigation costs are low relative to the costs of not addressing climate change.
While this is a basis for optimism, GGEs have grown by 70 per cent from 1970–2004. Most growth has come from the global electricity sector (145 per cent) (IPCC 2007). Global energy intensities — energy consumption per unit output — have decreased but not by enough to avoid being influenced by the effects of income and population growth. GGE mitigation policies have not yet worked to stall emissions growth. Moreover, policies as they currently stand will fail. From 2000–2030, GGEs will grow strongly, with three-quarters coming from developing countries. There is no basis for complacency.

**Australian climate-change policy**

Australia will be impacted on heavily by unmitigated climate change because it is a ‘fringe climate’ society. Australia has high gross per-capita emissions but is a ‘small’ country in terms of aggregate GGE emissions. Globally, the US and China provide 50 per cent of total GGEs, another 15 countries provide a further 30 per cent and another 158 provide 20 per cent (Baumart 2005). Therefore, in aggregate, small countries such as Australia are important polluters. In addition, for a viable global climate-change response post-Kyoto, poor and developing countries must begin to mitigate emissions in the medium term. Australia should not provide negative moral suasion by refusing to accept its share of the international policy response (Clarke 2009).

Australia’s policy position can be exposited in terms of its mitigation and adaptation policies as well as its attitude toward forthcoming international climate-change agreements.

**Mitigation**

Governments can employ various policy instruments to encourage GGE cuts. They can, for example, prescribe energy-consumption or energy-production targets. Economic instruments, which rely on price changes to induce changes in emission behaviour, have advantages over prescriptive targets. Pricing GGEs generates broad supply-and-demand responses to climate problems that reflect both energy prices and the costs of altering emission behaviour. Prescriptive target-setting is a more expensive policy because regulators cannot identify the costs different agents face in meeting these policies.

There is a debate about the case for using a GGE tax or a national emission-trading scheme (ETS) based on targeted aggregate emissions. There are costs and benefits with each approach, although choosing either policy is a decisive advance over choosing neither. For the most part, choice between these policies is an issue of second-order importance. Setting carbon prices by imposing a
tax means that the extent of mitigation will be uncertain, which is unattractive if a particular level of emission cuts is sought. Setting emission quotas using an ETS will leave carbon prices buffeted by market conditions, creating price uncertainty and futures markets in carbon quota sales. With an ETS, carbon prices will show significant month-to-month variations that make investment planning in firms more difficult, although longer-term carbon prices that fall during periods of recession and which increase during booms provide an automatic macroeconomic economic stabiliser. The claim that carbon taxes make it difficult for interest group bargaining to secure exemptions is unproven — it is as easy to argue for a tax break as a free emissions permit. If either tax revenues or revenues from auctioned quotas are used to cut other more distorting taxes then further double-dividend advantages may accrue.

An advantage of a global ETS is that improved opportunities to trade permits mean lower mitigation costs: costs are estimated to fall by 20 per cent with a global ETS, as opposed to a set of independent national schemes (Office of Tony Blair 2009).

The proposed Australian carbon-pollution reduction scheme (CPRS) is an ETS which will be phased in with unlimited permits at $10/tCO$_2$ from 2011/12, with full permit auctioning from 2012/13 at what is expected to be around $29/tCO$_2$. This will yield revenues of about $13b in 2012/13 which will be returned to households ($5b), fuel-excise offsets ($2.2b), trade-exposed firms ($3.6b) and to electricity generators ($0.8b) (DCL 2009). Caps will be pre-announced for the first few years but subsequent caps will depend on 2020 targets that will be determined once the rest of the world announces its cutbacks.

The structure of the CPRS is production-based, with some free quotas allocated to trade-exposed firms and generators of non-traded secondary energy and with income compensations for electricity consumers that make consumers respond only to the pure substitution effects of electricity price changes. Indeed electricity prices will rise about 25 per cent in 2012/13, with consumers being partly income-compensated for the higher prices.

There will be significant longer-term impacts of the CPRS on coal, which is both Australia’s main source of electrical energy and its biggest export. There are strong effects on brown-coal generators, although 93 per cent of such capacity will be in place by 2020 (Garnaut 2007: 491). There is the potential for substitution by nuclear power and renewables although there is limited engineering capacity to introduce nuclear technology — there is no school of nuclear engineering at any Australian university. As noted earlier, given the cost uncertainties, a mix of alternative-energy sources ex ante makes sense given that the alternatives have cost advantage at differing carbon prices and interest rates. As Garnaut stresses, there are significant potential payoffs to developing
‘clean coal’ (CCS) in cutting our own emissions and in securing valuable export markets. At present CCS is a technically feasible technology but its commercial potential remains unclear.

The CPRS is a useful start at providing a comprehensive demand-management policy, although it can be improved.

The scheme as it stands is production-based but its free quota distributions give it a partial consumption basis. A better scheme would begin with a CPRS based on consumption so that carbon-intensive exports were exempted from the need to comply with emissions quotas and with all carbon-intensive imports from countries which do not impose comparable carbon-emission controls on a production basis being subject to a border tax. Such BTAs are probably not inconsistent with the rules of the GATT (Tamiotti et al. 2009) since they operate essentially as a consumption tax. Such a scheme would resemble the longer-term ETS proposed in the US by the Waxman-Markey Bill. The consumption base provides incentives for other countries to tax their exports on a production basis and, in the longer-term, when enough do so, Australia too can switch to a production-based tax without either exemptions for exports or border taxes on imports.

Consistent with this position, there should be no free carbon permits to Australian electricity producers. Exporters which use electricity as an input would be eligible for rebates.

In addition, it has been pointed out that voluntary action to reduce GGE emissions will be ineffectual since the effect of such actions will be to reduce the carbon price. This can be most easily addressed by buying back ‘saved’ allowances, by reducing emission caps, or by more-elaborate schemes such as the CEEM’s Additional Action Reserves (Betz 2009).

The policy of assigning firms free carbon quotas has the undesirable consequence that such quotas must be used to be effective. It would be better to make such quotas tradeable and to then use the revenues to fund energy-saving adjustments.

Finally, there is probably no need for explicit renewable targets — and inefficiencies might stem from setting such targets — if the 2020 carbon price is around $50/tonne. Markets will then efficiently determine the viability of non-carbon-based energy technologies with such a carbon price.

If markets worked perfectly, then setting the correct carbon price would be all that is needed. However, market imperfections exist related to the supply of information to reduce uncertainties for consumers and producers and
for developing and transferring new energy-saving and carbon-friendly technologies. These create a case for public support by providing information and through explicit subsidies. Such support will make the CPRS more effective.

A practical difficulty with the CPRS is that it has twice been rejected by the Australian Senate so that, as it stands, it cannot be introduced. Opposition to the CPRS has been led by the Federal Coalition parties who have denounced the CPRS as ‘a big tax’. Paradoxically for a conservative political party, the Coalition has proposed, instead, a ‘Direct Action Plan’ of funding directly selected CO₂ reductions by subsidising the provision of solar panels in homes, schools and communities, by establishing geothermal and tidal-power projects, by planting 20 million trees, and so on (Abbott 2010). The main disadvantage of this non-market-based scheme is that it will cost more than a CPRS tailored to meet similar reductions targets unless its designers have better information than do firms and households operating in energy markets.

**Adaptation**

Australia can pursue a timely, successful agreement to replace the Kyoto Protocol but would be foolish to assume that such an agreement will be either prompt or entirely successful. Plausibly, the world can expect 1.8-2°C warming over pre-industrial temperatures from existing GGE concentrations, so there is a case for policies which seek an adaptation to climate change in agriculture, industry, urban settlements and in the conservation of Australia’s biodiversity resources.

Adaptation investments are not subject to the ‘public goods’ market-failure issues associated with mitigating climate change although, if mitigation is undersupplied globally, private incentives to adapt to many aspects of climate change are enhanced on the basis of private self-interest. Conversely, incentives to mitigate are reduced when active adaptation policies are in place but, if adaptation policies alone are pursued globally, nations are engaged in a race toward collective disaster.

With respect to adaptation, there is therefore greater potential to rely on market-driven responses and on policy responses which seek to encourage market-driven responses. In agriculture, for example, it has been widely observed that farmers are skilful at adapting to climate uncertainties. Indeed, their livelihood often depends on the capacity to make such adaptations. Provided that climate change is not proceeding too quickly, that prices for current outputs are not deteriorating too rapidly, and that they can revise their production plans quickly enough, farmers can make market-based adaptations that will come close to optimising their welfare. The only role for public policy is to provide climate information and to expand the range of technological options farmers have as they experience climate change (Clarke 2009). Public-good market failures mean
that these informational investments to reduce uncertainty, as well as R&D, need to be a focus. Of course, policies which reduce the need to adapt — such as ‘exceptional circumstances’ drought relief — need to be either redesigned, or as the Productivity Commission (2009) has suggested, abolished.

Urban-sector adaptations involve planning water and wastewater infrastructure expansions, and transport network designs need to be configured so climate change is either factored into the expansion decision or options are left to expand eventual capacity at low cost, again to deal with hot states of the world. There is also a role for public planning with respect to concerns regarding disaster preparedness.

Many other issues can also be market-driven, again on the basis of campaigns to increase private-sector awareness of the implications of climate change. Choices such as roof colour, the use of insulation and location of housing in relation to threats posed by changes in sea level are decisions best kept private on the basis of, perhaps, publicly boosted information about climate-change consequences and such things as opportunities for improved energy efficiencies.

Adaptation policies cannot always be primarily private. There are no markets for the supply of biodiversity in Australia, so market-based solutions will not work here. There is a need to manage the national reserve system while promoting conservation and corridors on private land (Clarke 2007). With respect to enhanced conservational investments there are significant ‘no regrets’ benefits, even if climate change proves less extensive than expected.

Policies helping the global response

Australia ratified the Kyoto Protocol on the 3 December 2007. This sent a good ‘moral suasion’ message about Australia’s climate-change policy intentions even if it had limited immediate effect and even though the force of the Protocol will expire in 2012. In 2012 a new co-operative international agreement needs to be signed, and the Australian Government pledged a series of unconditional and conditional GGE reductions at the initial meeting, in Copenhagen in November 2009, to determine such an agreement.

The Australian Government was committed to match a comprehensive agreement in Copenhagen. If there was no agreement in Copenhagen, then Australia would offer an unconditional 5 per cent cut over 2000 GGE levels by 2020. If other countries target a 510–540 ppm GGE target by 2050, then Australia offers a conditional 15 per cent cut over 2000 levels by 2020. Finally, if other countries agree to target a 450 ppm agreement then Australia would offer a conditional 25 per cent cut over 2000 levels by 2020, with up to 5 per cent by deforestation credits.
The mix of conditional and unconditional targets is a reasonable approach to addressing uncertainties in the climate-change policies of other countries, even though the Green movement in Australia argued that still more ambitious targets of 330 ppm or lower should have been pursued. This is not sensible if targeting such low GGE levels limits the potential for achieving any collective agreement at all. It seems more sensible to target emission cuts that are plausibly achievable now and to seek tighter cuts in the future.

To enhance the prospect of gaining cooperation from developing countries, Australia should agree to a common emissions entitlement per person by 2050, as suggested in the Garnaut Review.

**Conclusion**

Economists should base their policy views on the mainstream science of climate change. This science makes explicit a range of significant uncertainties and involves, primarily, belief in the view that human beings are having a significant impact on the global climate. To avoid taking policy action on the basis of the implausible views of a handful of practising scientists who do not endorse this primary view would be folly.

The core economics of climate-change policy amount to comparing the opportunity costs of taking policy actions with the costs of not taking action. This is an ambitious field because there are significant issues of uncertainty, the setting is intrinsically dynamic and there are important non-linearities and irreversibilities. Conceptually, the way to proceed is to use cost-benefit analysis, which correctly accounts for uncertainty when there are significant irreversibilities.

Australian policy has made a credible start to initiating climate-change adaptations and mitigations locally and has a reasonable stance toward international negotiations. Ultimately, Australia needs a more comprehensive and integrated energy-use plan that will provide a closed-loop approach to commercialising non-carbon-polluting secondary-energy technologies. As the Garnaut Review makes clear, Australia gains enormously with CCS technology development and transfer, so this is a sound R&D focus. Another sensible focus is agriculture, which is a significant contributor to GGE emissions but also promises carbon-sequestration synergies by means of sustainable agriculture.

Over the coming decades, Australia and other countries need to think about pursuing stricter GGE targets with the eventual aim of restoring pre-industrial GGE levels. This will eliminate longer-term impacts of GGE emissions on climate.
References


Department of Climate Change 2009, Revised Fiscal Impact of the Carbon Pollution Reduction Scheme (CPRS), Fact Sheet, Department of Climate Change, Canberra, May.


National Academy of Sciences 1979, *Carbon Dioxide and Climate: A Scientific Assessment*, Washington, DC.


A New Mind-set for Exchange Rates

STEPHEN GRENVILLE

The Old Mind-set

Following the breakdown of the Bretton-Woods system of fixed exchange rate in the early 1970s, many countries chose to hand over the exchange-rate determination to the financial markets, with a free float. That a free float would deliver the optimal result was presumed by many practitioners and nearly all academics. Harry Johnson (1972) promised that “a freely flexible exchange rate would tend to remain constant so long as underlying economic conditions (including government policies) remain constant; random deviations from the equilibrium level would be limited by the activities of private speculators.”

Things turned out very differently, with large fluctuations seemingly unconnected to any change in the ‘fundamentals’. The US dollar fluctuated by more than 30 per cent in the mid 1980s. The Japanese yen moved over a range from almost 150/US$ to 80/US$ within a space of three years in the mid 1990s. Smaller currencies such as the Australian dollar fluctuated even more widely. Emerging countries fared worse still, with Latin America in general crisis in the 1980s, Mexico again in 1994, East Asia in 1997 and Russia in 1998. This didn’t seem to fit with the idea that ‘fundamentals’ would change slowly and arbitrageurs would smooth the rate over time. It was not just a matter of short-term volatility (which might not do much harm): as well, there were large and persistent misalignments with sudden readjustments, even for the largest countries with deep and mature financial markets.

Despite this, the faith in the magic of the market — in the free float — remained. Even for emerging countries with immature financial markets, the International Monetary Fund’s message was to float and let the market sort it out, with minimal intervention and no capital controls. Fixed rates against the US dollar were seen as the major cause of the Asian crisis of 1997 and the remedy was to float freely.

1 The Lowy Institute, stephengrenville@hotmail.com.
Two factors have opened up this debate again. First, the Global Financial Crisis has highlighted the imperfections of the market’s price-discovery process.\(^2\) No longer will any debate be stopped in its tracks by making the point: ‘You don’t think you know better than the market, do you?’

Second, there is now a fair bit of accumulated evidence that what is pejoratively called a ‘dirty’ float is not only feasible, but can work pretty well.\(^3\) Few countries have practised the benign neglect implied by a pure free float. A number of countries have run successful exchange-rate regimes which rely on a free market float much of the time, but use market intervention (sometimes quite substantial) when the market doesn’t seem to be delivering a sensible answer. Australia provides one of the clearest examples of this. Of course, it is difficult to convince a sceptical audience that the substantial intervention altered the course of the exchange rate: we don’t know for sure what would have happened without the intervention. But in the first 20 years of the float, the RBA made a profit of more than A$5 billion through its intervention (Becker and Sinclair 2004). Even Milton Friedman (1953) accepted that if intervention was profitable, there is a powerful presumption that it was stabilizing.

Still the IMF staff remained unconvinced. A working paper detailing the Australian experience concluded: ‘The apparent success of intervention in affecting the level of the Australian dollar appears to have come at the expense of increasing volatility of exchange rate movements.’\(^4\)

Australia was not alone in profitable intervention. Japan had similar success (Ito 2002) but, following the adage that ‘it is the spouting whale that is harpooned’, it has kept a low profile on its intervention.

Still, the received academic wisdom is that sterilised intervention has little or no effect. Simple theoretical models often assume perfect substitutability of assets in different currencies. Empirical work often relied on a heterogeneous mixture of examples. Some intervention episodes, such as the 1992 defence of sterling, reflected politics more than economics and were foredoomed. Certainly, there were enough examples of policy failure where the authorities had attempted

---

\(^2\) It is tempting to say that the GFC highlighted serious flaws in the Efficient Markets Hypothesis. The EMH, taken by itself, says no more than that the market price has already taken account of all publicly available information. But its name seemed to promise more than this near-platitude. Combined with the idea that people are rational and that market participants face risk (which can be measured by probabilities) rather than uncertainty, something more loosely styled ‘efficient markets’ seemed to imply that the price-discovery process would be optimal.

\(^3\) See Sarbo and Taylor (2001).

\(^4\) Edison, Cashin and Liang (2003). Of course the volatility this paper is testing is daily volatility, irrelevant to the question of stabilization over the course of the cycle. Seeing the profit made by this intervention, some argue that it needs to be discounted for risk. But the usual measures of risk — short-term volatility — are not relevant for long-term position-taking, where the position will not be unwound to meet some short-term liquidity need. The ‘risk’ for a long-term holder is that the position will be eventually unwound at a loss, and this is fully captured by the profit calculation.
to defend disequilibrium rates. Rather than examine cases where intervention worked well and try to analyse just why it was effective in these cases, these futile interventions were often regressed together with the successes, drawing the general conclusion that intervention was futile at best and harmful at worst.

At the same time that studies were showing intervention to be futile, the best academics also noted something that should have shaken confidence in our understanding of exchange rates. In the early 1980s, Meese and Rogoff (1983) found that no fundamentals-based model of the exchange rate (whether based on purchasing power parity, trade flows, capital flows or interest differentials) could beat the ‘random walk’ (that today’s estimate of the exchange rate is the best estimate of the future rate). In other words, no model could offer any help at all in predicting the path of the exchange rate. Extensive testing in the ensuing quarter-century has left this conclusion unaffected over the short run, although some models fare better at explaining longer-term exchange-rate tendencies (Rogoff 1999).

This amazing result might have taken researchers down either of two paths. It might be that the models were accurately capturing where the equilibrium exchange rate should be and the market was not delivering this answer. Or it could send researchers to develop a theory that could explain the behaviour of the exchange rate. Largely, the latter path was chosen. Researchers worked hard to fit the real world into the EMH, explaining the exchange rate in terms of ‘time-varying risk’ and other *deus ex machina* devices.

A more fruitful line of analysis noted that not all market participants based their trading on economic fundamentals. Frankel and Froot (1986) hypothesised that the market comprised both ‘fundamentalists’ and ‘chartists’, and sometimes the chartists were dominant. But even those who accepted this hypothesis did not ask whether the policymakers could or should intervene when the chartists were in the ascendancy, taking the rate to non-equilibrium levels. The free market was still sacred, even when it was wrong.

A much smaller group (De Grauwe and Grimaldi 2006) worked on the intuitively plausible idea that while no-one had a precise idea of what the equilibrium exchange rate should be at any moment, it was more useful to think about equilibrium as lying within quite a wide range of ‘rational beliefs’ (that is, not inconsistent with the evidence). Not only did this explain why the actual market outcome might not correspond to the fundamental equilibrium, but it also explained why the rate could move sharply in response to trivial ‘news’. When people are uncertain as to the right rate (and their ‘model’ of exchange-rate formation has a wide margin for error), a random event can cause people to shift their view on what is the equilibrium exchange rate quite substantially and suddenly.
But this range of rational beliefs is not enough, in itself, to explain the magnitude of the departures from equilibrium. The more extreme movements can only to be explained in terms of market dynamics and endogenous risk. Once the price starts to move, portfolio and accounting constraints force market participants to cut their positions in currencies, even when they believe that the price is already below the equilibrium and will revert in due course.

Why doesn’t arbitrage fix the problems? In practice, most arbitrageurs cannot actually take much risk. Very few institutions can hold an open FX position for long. Persistence is a problem for disequilibrium rates: as Keynes is said to have remarked, markets can remain irrational for longer than you can remain solvent.

**A New Mind-set**

If we accept that the market price is not always gravitating towards the equilibrium exchange rate, the policy mind-set can and should shift fundamentally, to focus on two questions. What is the equilibrium rate? Can intervention shift the actual rate towards this?

The first question poses a major challenge for policymakers. While a free-float regime does not require the policymakers to have any view at all as to the ‘right’ exchange rate, the new mind-set requires them to have a view on the equilibrium exchange rate. This is not easy (perhaps impossible) to do with precision, but the answer might well be in the form of a range reflecting the array of ‘rational beliefs’ (just as inflation targeting sees the acceptable inflation rate as a band rather than a single point).

The second question requires a similar shift in policy mind-set. Instead of assuming that the market has established the equilibrium rate (and therefore any intervention will take the actual rate away from its equilibrium), if the market rate is sub-optimal, this opens the possibility that successful intervention would take the rate towards equilibrium. Then the policy question is an operational one: does the policymaker have the means to shift the rate in the right direction?

This opens up a range of feasible exchange-rate regimes and policies, other than free floating. One approach to the equilibrium issue is the band/basket/crawl (BBC) regime suggested by Williamson (2008). The authorities announce a target band which defines the range in which they believe the equilibrium lies. This allows variation within this band but requires firm intervention at the edges of the band. If the market accepts the band as credible, speculators will not challenge the band (and might even support it: after all, just as the authorities don’t know for sure what the ‘right’ rate is, neither do the speculators). The
A New Mind-set for Exchange Rates

BBC, of course, is in inflation-adjusted real terms and the ‘crawl’ of the band acknowledges the need for adaptation to the (implicitly slowly changing) fundamentals.

For those who (understandably) worry that countries may be too active in their intervention and will defend a non-equilibrium level for mercantilist reasons, an alternative regime is to establish an internationally endorsed ‘reference range’ and specifically renounce intervention within this range (a genuflection to the market and acknowledgment that the model is uncertain) but allow intervention once the band is breached.

Others (for example, Australia and Japan) have a more ad hoc approach, intervening when the market seems to be clearly misbehaving, usually only when the market has taken the rate to a clearly incorrect level, where the misalignment is more likely to be harmful.

Policymakers will have to consider that the equilibrium might be changing over time. It will also have a cyclical component. Some temporary departure from the longer-term fundamental equilibrium exchange rate (FEER) over the course of the cycle is desirable (and part of the transmission mechanism of monetary policy) as the exchange rate can usefully ‘spill’ excess demand into net imports.

The IMF itself has an internal committee which has been wrestling with the notion of the ‘right’ rate, although its two-fold objectives are carefully defined to avoid endorsing market intervention. The first purpose is to give countries warning if their exchange rate risks causing a crisis (for example, through excessive current-account deficits and too much foreign debt). The second purpose is directed at identifying exchange-rate ‘manipulation’ (that is, rates which were being held away from their equilibrium by policy, usually for mercantilist reasons) rather than providing a starting point for policy action to shift the actual rate towards its equilibrium. China is a case in point, although it has not been formally declared to be a ‘manipulator’.

In the IMF’s exchange-rate analysis there are three measures of ‘equilibrium’ — one akin to Williamson (based on the sustainability of the current account), one based on a macro-model and one depending on the sustainability of foreign debt (closely associated with the Williamson approach). The precise answers are not publicly available, and the routine write-up usually finds that the answers (which often differ significantly between the three methods) are all compatible with the actual rate.

At least at this stage, these attempts lack enough precision and have too many questions about methodology to be the basis of firm policy action, especially of a mechanistic nature. If they are to be the basis of serious discussion, the sensitivities (which at the moment stop the IMF from publishing the exact
exchange rates which their methodology implies) need to be overcome. But this sort of analysis is the starting point of a new and better policy approach to exchange rates.

The doctrinal fixation with persuading countries to adopt a pure free float has distracted us for three decades from this practical operational task of identifying the equilibrium exchange rate (in terms of a range or central tendency) and devising effective intervention strategies. But the Fund position now seems to have shifted significantly. The Fund has abandoned the ‘corner solutions’ argument, which said that countries needed to be either rigidly pegged (like Hong Kong) or free floating. The previously infeasible middle ground is now judged to be viable (Gosh and Ostry 2009).

The Challenge

What is required now is a wider change in the mind-set. We should accept that, while the market is a powerful tool for the discovery of the equilibrium price that may well be right on average and most of the time, at times it goes awry. If this is accepted, then efforts to push the rate towards equilibrium can’t be dismissed on the basis of principle or doctrine. It may be difficult to establish the equilibrium with accuracy, and the authorities may have imperfect instruments to influence the exchange rate (so they would not be aiming to achieve the FEER, only to move towards it). But this is a debate which should be joined, and it should be done in the open, without the inhibiting presumption that intervention is a sign of policy inadequacy.

The conceptual mind-set is simple enough: the policymakers should be trying to mimic the equilibrium exchange rate that a well-functioning market would produce. The theoretical attraction of a free-market floating solution was that the financial markets promised to deliver this. This would guard against policymakers attempting to impose their mistaken preconceptions or their mercantilist export-oriented priors on this. It falls down only because the market has shown itself unable to provide this price discovery, at least in a form that is consistent and stable over time. But the key point is that any intervention should be trying to mimic a well-functioning market, not override it to produce a non-equilibrium exchange rate. ‘Active’ policy must resist the temptation to use the exchange rate for some non-equilibrium purpose — for a mercantilist purpose or to hold the rate at some non-equilibrium level because change is politically unpalatable or harms certain vested interests.
Will the authorities make mistakes in their interventions? Of course they will, just as they make mistakes with the other policy instruments. But principled inaction is not the answer: to claim innocence on the basis of inaction is no defence if a valid and beneficial policy is available.

There remains a further issue. While we noted that there have been very substantial and sustained shifts in exchange rates, even for countries with deep and mature markets such as Japan and Australia, these shifts seem to have caused few problems, at least as far as can be judged by complaints from those who might have been adversely affected. One curious stylised fact that coincides with the floating-rate period is that the pass-through to domestic inflation seems to have been either small or very slow. Thus the authorities have been prepared to see large downward shifts in the exchange rate without feeling the need to support the currency through higher interest rates. While elements in the tradable-good sector complain from time to time, developed economies seem to have adapted to these large swings in the exchange rate. This can’t simply reflect widespread hedging, as the large cumulative external imbalances in both Japan and Australia must have left substantial foreign-exchange exposures. In any case, if exchange-rate changes don’t have any effects, we might need to re-examine some core parts of international trade theory.

Thus the policy implications of this paper for the developed countries are small. Japan and Australia will presumably go on intervening, profiting from this intervention but without doing more then ‘filling the troughs and lopping the peaks’ of the path of the exchange rate over time.

Where this analysis has more practical policy force is with the emerging countries and the advice which the IMF gives to them. Here, the big swings in exchange rates are clearly disruptive. Both Indonesia and South Korea suffered large falls in their exchange rates during the GFC, even though their economies were not directly involved in the crisis. These exchange-rate falls posed a serious threat to financial stability. In both cases, the authorities intervened very substantially, and in both cases the fall was short-lived, justifying the intervention in terms of profitability, at least. Nevertheless, the IMF’s reaction, in the case of Korea, was to persuade the authorities that future intervention should be limited to (irrelevant) ‘smoothing and testing’. The new doctrine, as espoused by Gosh and Ostry, has not yet filtered down to the Fund’s operational staff. It is within that institution that the greatest mind-set change is needed.


References


Reflections on ‘A Tax System for New Zealand’s Future’

JOHN CREEDY

Abstract

‘A Tax System for New Zealand’s Future’, the Report of the Tax Working Group, is reviewed. The Report is judged a model of rational policy analysis, explaining the need for reform, the basic principles used to consider alternative policies, and the advantages and disadvantages of a range of reform proposals. A number of aspects concerning the evaluation of tax structures are considered, and some arguments which not stated explicitly in the Report are clarified.

Introduction

The Report of the Victoria University of Wellington Tax Working Group (TWG), ‘A Tax System for New Zealand’s Future’ (2010) represents a unique collaboration among individuals from the New Zealand Treasury, the Inland Revenue Department and the Victoria University Centre for Accounting, Governance and Taxation Research (CAGTR), along with a range of other acknowledged tax experts. The Report has also been produced with remarkable speed and relatively cheaply. Favourable comparisons may be made with the review carried out by the Australian Treasury (‘Australia’s Future Tax System’) and the Institute for Fiscal Studies (‘Reforming the Tax System for the 21st Century’) in the UK.

1 The University of Melbourne, jcreedy@unimelb.edu.au. In preparing this paper, I have benefited from discussions with Bob Buckle, Norman Gemmell, Denis O’Brien and David White. I should also like to thank the editor and two referees for their constructive suggestions. This paper is based on a Treasury Guest Lecture (jointly sponsored by the New Zealand Treasury, the Centre for Accounting, Governance and Taxation, and the Research Institute for the Study of Competition and Regulation at Victoria University of Wellington) presented at the Auckland Business School.
Taxation is an area where strong disagreement and argument at cross purposes is notoriously present. The major strength of the Report is that it is an exercise in ‘Rational Policy Analysis’, in that it aims to examine the implications of a range of alternative policies, using a variety of measures. Such a procedure involves a clear statement of the criteria against which a tax structure may be compared. It thereby encourages an understanding of why different people may not agree about the best way to proceed, and so allows for rational discussion rather than heated argument. The Report is refreshingly free of the kind of rhetoric which is so often a part of taxation debates.

At times the TWG was unable to resist the pressure to make specific proposals for reform of the NZ tax structure, though in doing so it is careful to point out that its members are not always unanimous. The TWG was, it seems, responding to outside pressures to produce specific recommendations. But rational policy analysis at its best provides the information which enables others, using their own value judgements and views about likely orders of magnitude, to make up their own minds.

After briefly describing the reforms suggested by the TWG, the paper goes on to consider the way in which tax structures are evaluated. Some criticisms regarding the precision of chosen concepts are aired, and suggestions for further analyses are made.

**New Zealand Taxes and Reforms**

Some distinctive features of the New Zealand tax structure may be briefly mentioned. The personal income tax has no tax-free threshold. There are few deductions, and the top marginal rate is 38 per cent. It has a very limited capital-gains tax, a substantial difference between the corporate tax rate (of 30 per cent) and the top personal rate, and a Goods and Services Tax with a very broad base and low rate relative to European countries. The Report explains why changes are needed, arising from domestic and international changes since the previous major reforms of the 1980s, such as the increasing extent of international capital mobility. Motivated by the extent to which the tax rate applied to corporations, trusts and the top personal tax rate have become ‘misaligned’, it then considers a range of alternative methods of taxing corporations. These reforms include: reductions in personal income-tax rates; methods to broaden...
the tax base, including the introduction of a Land Tax and extensions to the Capital Gains Tax; and changes to rules regarding depreciation allowances for property investment. The report stresses the need for changes which place less reliance on those taxes (such as personal income taxes and corporation taxes) where there are efficiency problems and where labour and capital are mobile. The report argues for the need to bring corporation, trust and personal tax rates into closer alignment. In each case the advantages and disadvantages are carefully discussed.

In considering changes, an important feature is that the TWG restricted its attention to reforms which are revenue neutral, so that total tax revenue is estimated to remain unchanged. The restriction to revenue-neutral changes is necessary because any change in revenue involves another unspecified policy, so that alternatives cannot properly be compared. Thus, for example, the argument that concerns about the future adequacy of tax revenue would disappear if only government expenditure were cut would be irresponsible, unless at the same time a detailed set of proposals for cutting expenditure were produced and their implications examined. Expenditure considerations were outside the scope of the review.

Any revenue-neutral change in the tax structure must of course involve some winners and some losers. Hence unanimity cannot be expected and distributional value judgements cannot be avoided. It is therefore important to attempt to make such value judgements explicit, since they are the source of many disputes about tax policy. Given the inevitability of there being some losers from a tax-policy change, a strong desire on the part of policymakers to avoid creating losers clearly creates a bias towards the status quo.

Revenue-neutral changes examined by the TWG assume an absence of behavioural responses to tax changes, such as profit shifting by corporations and labour-supply adjustments by individuals. Such behavioural responses may be substantial. To the extent that a policy change improves incentives and increases taxable incomes (including profits), post-reform tax revenue is likely to be understated.

Evaluating Tax Structures

This section discusses the TWG Report’s approach to the evaluation of tax structures.

---

3 For examples using a behavioural microsimulation model for Australia, see Buddelmeyer, Creedy and Kalb (2008).
It is suggested that an appreciation of these points should help readers to obtain a clearer understanding of the main arguments of the Report.

**Basic Tax Principles**

The TWG Report sets out six ‘principles of a good tax system’. These are: efficiency and growth; equity and fairness; revenue integrity; fiscal cost; compliance and administration cost; and coherence (as part of a multi-tax and benefit structure). Any statement of tax principles inevitably brings to mind Adam Smith’s famous ‘four maxims’, which are frequently discussed and extended. 4 A fundamental problem is that, even if these principles were substantially elaborated, it would not be possible to treat them as axioms and from them derive an implied ‘ideal’ tax structure. The TWG appeared to recognise that even the production of a table, listing tax policies in the rows and the principles in columns, with ticks and crosses in appropriate boxes, would not in the present context provide sufficient information. The Report thus more usefully produced tables listing various advantages and disadvantages of each alternative. These features in fact include further criteria such as ‘certainty regarding future taxes’ (see pp.28 and 64).

However, there are some important principles which are implicit in the Report, but which really need to be made explicit. Thus, only on p.56 is mention made of the fact that ‘value judgements are necessary to determine the priorities and respective weighting of these objectives’. The fact that value judgements play such an important role in tax-policy decisions explains also why there is usually so much disagreement in tax-policy debates. 5 It is seldom clear if opponents disagree because they hold different value judgements, or because they form different views about the way people respond to tax changes (given that complete information is never available). This is precisely how rational policy analysis can help — by indicating the implications of adopting different values, so that readers can make up their own minds.

---

4 Smith’s maxims, stated in *The Wealth of Nations*, were: payment according to benefit received (though there was some ambiguity as to whether this coincides with ability to pay); certainty; convenience; and economy in collection. A longer list was produced by G. W. Norman, whose list included: frugal; computable; simple; constant; divisible; popular; non-interferent; equal; uncorruptive; unvexatious; and unevasible. See O’Brien with Creedy (2009).

5 This point is discussed in detail in Creedy (2010).
The Tax Base

A fundamental judgement taken by the TWG — and not fully articulated in the Report — is that taxes should reflect ‘ability to pay’, rather than the ‘benefits received’ as a result of the subsequent public expenditure. Thus, it is stated that, ‘a good tax system should [ensure that] taxes paid reflect ability to pay’ (p.59). In considering ability to pay, a further important step is taken in suggesting that it ‘should’ be measured by ‘comprehensive income’. Unfortunately the Report never states this explicitly, suggesting for example only that, ‘taxing capital gain “on accrual” would bring the tax system closer to taxing comprehensive income’ (p.48) and referring to a capital gains tax as ‘a more comprehensive option for base broadening’ (p.11).

The term ‘comprehensive income’ means simply that income is defined as the amount that can be consumed in a given period without reducing wealth. This ‘maintaining capital intact’ income concept is indeed widely used in economics, but its adoption as a tax base involves a value judgement. The important point here — not brought out sufficiently clearly in the Report — is that the use of this concept is the fundamental reason for attempting to tax capital gains. Such taxes obviously face problems in dealing with those gains which accrue but are not realised, and the Report recognises that a ‘comprehensive capital gains tax’ is impossible to impose. Nevertheless, there may be scope for extending CGT along the lines used in other countries. 6

In popular debates regarding the introduction of, or extension to, a tax there is perhaps a tendency to judge a proposal against a ‘perfect’ tax, when of course no existing tax is perfect. And, there is always some resistance to a new tax. As the writer of the first English-language full-length treatise on public finance, John Ramsay McCulloch, declared: ‘it is sometimes better “to bear those ills we have, than to fly to others that we know not of”’. 7 It is therefore likely that the TWG’s discussion of extending CGT in New Zealand will meet with some resistance, even among those who may not expect to lose by such a change, but it provides a useful starting point for debate.

---

6 The Report discusses the use of a risk-free rate of return method (RFRM) on property, but is generally not supportive.

7 See McCulloch (1863).
Rules of Thumb

An important position taken by the Report is that, ‘the TWG considers that the broad-base low-rate [BB-LR] option is generally a sound principle to adopt’ (pp.16 and 64). However, the adoption of BB-LR is not in fact a ‘principle’, but a rule of thumb. 8 It is nevertheless a useful rule. It arises from recognition that the efficiency cost (in terms of the excess burden) of a tax is approximately proportional to the square of the tax rate. 9 Hence it is useful, other things being equal, to keep the rate low, and this is achieved by keeping the tax base as broad as possible. But, of course, in evaluating a tax structure there are other considerations (as indeed are indicated in the TWG’s own list of criteria), involving distributional value judgements, along with ‘merit good’ and externality arguments.

The BB-LR rule of thumb leads, in the Report, directly to the suggestion that the top marginal income-tax rate should be reduced. But it is important to recognise that economic ‘efficiency’ criteria alone are not sufficient to determine policy — value judgements play an important role and more detailed information is required. For example, there is considerable heterogeneity in excess burdens, and many individuals below the top marginal rate face high burdens, depending on demographic and other characteristics. The preferred policy depends on the judge’s precise value judgements, including the degree of aversion to inequality.

Some ambivalence is also present in the TWG’s discussion of the ‘alignment’ of the corporation tax rate, the rate applied to trusts and the top personal income-tax rate. It sometimes appears that the Report treats alignment as a basic principle, while recognising that it would involve a loss of degrees of freedom in policy choices. Yet again it is more appropriate to consider it as a rule of thumb based on an assumed ease of shifting between income sources, rather than a principle. Starting from some specified objectives, and given other features (such as regulations regarding allowances and deductions, the costs of income shifting and the precise extent of behavioural responses to tax changes), it would be extremely difficult (as mentioned above) to produce an optimal structure: there is little reason \textit{a priori} to expect identical rates to emerge from such a problem. And of course the relevant rates on trusts and corporations are proportional, while the average personal income-tax rate is below the top

---

8 For further discussion of this rule of thumb, see Creedy (2009). Similarly, references to ‘folk theorems’ were made by Slemrod (2004).
9 For an introduction to the excess-burden concept and explanation of the approximation, see Creedy (2004).
10 On the use of Utilitarian value judgements in tax reform, see Creedy and Hérault (2009).
marginal rate. However, there can be little doubt, as the Report argues, that the rates have now become seriously out of line and some movement towards alignment, if not necessarily to equality, is a high priority.

**Precision Regarding Concepts**

Perhaps surprisingly, one major concept is not defined by the TWG, leading to some potential misunderstanding. A tax structure is progressive if, over the relevant income range (where income is the tax base), the average tax rate increases with income. However, progressivity does not require ‘marginal rate progression’, which refers to an increasing marginal tax rate as income increases. Indeed, a considerable degree of progressivity can be achieved with a proportional tax, combined with a basic income (that is, a fixed and unconditional transfer payment).

This is relevant when considering a change in the tax mix — a shift from personal income taxes towards indirect taxation in the form of the GST. This can of course be achieved in a variety of ways, including reductions in all marginal income-tax rates, or as part of a rate-flattening exercise with reduction in the top personal rate of taxation. A revenue-neutral tax shift must, as suggested earlier, involve losers as well as winners. However, the approach adopted clearly depends on the value judgements regarding income distribution. Where concern is largely with low-income groups, a cut in higher income-tax rates can be combined with an increase in the GST rate, along with a suitable adjustment to benefit levels to maintain their real values. After all, this was done when GST was first introduced.

**Taxes and Elasticities**

The TWG Report suggests that, ‘taxing those bases that are least likely to be subject to significant behavioural change … (inelastic bases) is also a sound principle to adopt when … broadening the tax base’ (p.10). This is an allusion to the efficiency effects of taxes, concerning the distorting effect of non-uniform taxes on behaviour. Reference is made to a result obtained by Ramsey (1927). If the aim is to minimise a measure of the sacrifice arising from taxation, the appropriate tax structure turns out to be one that reduces the consumption of

---

11 On indirect taxes in New Zealand, see Creedy and Sleeman (2006). The Report (p.47) shows that GST as a proportion of total expenditure is similar across income decile groups.
all goods by the same proportion. This implies a higher tax on inelastic goods and a lower tax on elastic goods. However, this result is severely limited because it refers to a single-person economy, and thus excludes any equity concerns.

Importantly, the most widely accepted modern measures of the efficiency effects (excess burdens) arising from taxation are based on ‘compensating variations’ and ‘equivalent variations’. These measures depend not on market price or income elasticities, but on ‘compensated’ elasticities, where the ‘income effects’ of the price changes are excluded. \(^{12}\) This is not merely a pedantic point about definitions. In the context of labour supply and taxation, where income effects of tax changes can be significant, it is possible to observe high marginal welfare costs (excess burdens per dollar of extra revenue), even for those whose labour supply changes very little. \(^{13}\)

Here is the appropriate point to consider how the TWG’s discussion of a Land Tax fits in with their list of tax principles. It is not in fact motivated by the use of a comprehensive income measure of ability to pay. Indeed, the comprehensive-income concept gives no support to any kind of wealth (or net worth) tax — only tax on the income arising from assets (both human and non-human capital). The proposal for a Land Tax relies exclusively on the desire to find an additional tax base and, on the argument that the stock of land is fixed, uses the criterion relating to elasticity. Hence it is concerned only with efficiency, while transition, tax shifting and equity effects are important considerations which need further analysis. \(^{14}\)

**Some Next Steps**

The TWG Report recognises that further analyses need to be carried out. Given the remarkably short time scale within which the Report was produced, and the low level of resources made available to the TWG, this is of course inevitable. One area mentioned concerns the transfer system and its integration with the personal income-tax structure. In particular, the complex system of in-work benefits in New Zealand gives rise to substantial marginal tax rates applying to middle-income groups, and high levels of government expenditure. It also appears that there are unintended consequences whereby some individuals who are not in the ‘target group’ considered by the designers of the system are in

---

\(^{12}\) The compensated elasticity therefore measures only the pure substitution effect of the price change. For an introduction to these concepts, see Creedy (2004).

\(^{13}\) Examples of this kind of result for Australian individuals are reported by Creedy, Hérault and Kalb (2009).

\(^{14}\) Furthermore, the supply of land for sale is not fixed.
receipt of transfer payments. It is suggested here that a comprehensive review of the benefit system, to complement that of the tax side, needs to be given very high priority.

Additional areas where further work is needed, not mentioned explicitly by the Report but undoubtedly supported by it, include behavioural responses. Despite the potential importance of foreign direct investment, and profit-shifting by corporations (both real and that achieved by the exploitation of transfer pricing), very little solid empirical evidence is available regarding companies operating (or previously operating) in NZ. Similarly, there no longer exists a behavioural model in NZ that is capable of producing estimates of labour-supply responses to tax changes and their associated efficiency (welfare) costs.  

One suggestion made several times in the Report is that there should be some kind of review body, or organisation, charged with, and sufficiently equipped to carry out, the task of conducting regular independent evaluations of the tax structure in NZ. This, it seems, is not intended to propose yet another quasi non-governmental organisation, but presumably the TWG has in mind something like the role played (among other things) by the Institute for Fiscal Studies in London. The aim is to inform public ‘rational debate’, by investigating the detailed implications of tax-policy changes. It is suggested that this could impose some constraints on politicians, with a propensity to tinker with a tax structure in ways that have insufficient regard for the integrity of the complete system. It could also help to make the system less vulnerable to the inevitable special pleading by interest groups. The question of the funding and location of such a body was not discussed in the Report.

Conclusions

In any discussion of taxation, it is important to recognise the inevitability of disagreement and a lack of consensus. There are simply too many value judgements involved and too many areas where only informed guesses must be made in the absence of direct information. It is certainly no weakness of the TWG Report that it does not report unanimous support for various policy reform proposals. However, the TWG members managed to agree on a broad framework for discussing reforms, and the need for some action. A Report which simply set out a dogmatic agenda for reform, using the kind of rhetoric that is familiar from the commentaries of special-interest groups, would have little value and would be read only by like-minded people.

15 For a description of the only large-scale New Zealand microsimulation model, and examples of policy simulations, see Buddelmeyer, Creedy and Kalb (2008), Chapter 12.

16 For further elaboration of the reasons for this comment, see Creedy (2010).
The influence of the TWG on the tax debate, indeed in making taxation itself more prominent in public debate, is clear from even the most casual examination of the media in New Zealand. The suggestion that there are substantial revenue risks and that the distorting effects of some forms of taxation have growth effects as well as the less visible — but nonetheless substantial — excess burdens, has clearly influenced the government in its thinking about changes to the tax structure.

As mentioned earlier, the strength of the report is in its attempt to contribute to rational policy debate by rehearsing the various arguments in a clear and dispassionate manner, so that those on different sides of the debate can come to understand just why they differ. That a disparate group of individuals from a range of backgrounds have established some common ground in a way of thinking about taxes is itself sufficient cause for praise. The Report can be read with interest and profit by all those interested in tax policy.

References


The Global Financial Crisis of 2008 occasioned a flash of antipathy towards economics that was far more predictable than it was informed. A serious attempt at the censure of certain intellectual tendencies in economic thought was advanced by Paul Krugman in ‘How did economists get it so wrong?’ (The New York Times, 6th September, 2009). Agenda here publishes some reactions to Krugman’s contentions.
How *US* Economists Got It So Wrong

ROSS MCLEOD

Paul Krugman recently asked: ‘How did economists get it so wrong?’ I have great sympathy with his answer: ‘Economists…mistook beauty, clad in impressive-looking mathematics, for truth.’ But having read the essay I could not help wondering: Is this really a critique of economists in general, or just those in the US? Krugman mentions no fewer than two dozen economists, all but one of whom (with the exception of the long-dead Adam Smith and John Maynard Keynes) are American—or at least based in the US.

It strikes me that in some ways being an academic economist is not unlike being a movie star. The US dominates the English-speaking market for movies. The practical impact of this is that Australian actors such as Cate Blanchett, Toni Collette, Nicole Kidman, Russell Crowe, Anthony LaPaglia, Naomi Watts, Guy Pearce and Heath Ledger need to master American accents and take themselves to Hollywood if they aspire to winning an Oscar. Likewise, the US dominates the English-speaking market for academic economics journals, and nearly all of what are regarded as the world’s top universities for teaching economics are in the US. Non-US economists need to present their work in a style consistent with the dominant US approach if they hope to be published in those journals or to obtain positions in those universities. In particular, they experience considerable difficulty if they do not conform to the US obsession with mathematical elegance and econometric (statistical) sophistication. Of course, many US economists face the same dilemma. ²

I find it disconcerting that in the last two decades there were only two years in which the US did not provide a Nobel-prize winner in economics; and in each of the last 10 years at least one winner has come from the US academic environment. ³ How does the selection process work? A committee of the central bank of

---

1 The Australian National University, ross.mcleod@anu.edu.au. The writer is Editor of the *Bulletin of Indonesian Economic Studies*, published by the Indonesia Project, The Australian National University, Canberra.


3 Although widely known as the Nobel Prize in Economics, the correct designation is ‘The Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel’.
Sweden chooses the winners, but prior to that a number of candidates need to be nominated, and presumably the committee relies heavily on the supporting documentation supplied to it by those who do the nominating. The nominators are said to be chosen in such a way that as many countries and universities as possible are represented over time, but the identity of such nominators does not seem to be made public. My guess is that US economists are heavily represented in the lists of those eligible to nominate candidates — not least because these include previous recipients of the prize — and that the nominations themselves are also US-dominated. One hears stories of discussions at informal gatherings of top US economists along the lines of ‘Whom from among us shall we put forward this year?’. When nearly all of the world’s top economics journals and top economics universities are American, it is perhaps not surprising that the distinction between ‘us’ and ‘US’ becomes blurred.

One of the unfortunate consequences of US hegemony is that American ways of thinking and doing things tend to overwhelm all else. American best practice comes to be seen as world’s best practice, whereas if the global markets for cultural and academic output were more balanced, the world might well be a more interesting place, and better ways of understanding reality might emerge.

Thus Krugman’s essay comes down to discussion about contending views over the global financial crisis not within the economics profession, but among economists in the US. The way Krugman frames this debate is as one between what he calls ‘freshwater’ and ‘saltwater’ economists (FWEs and SWEs). In this geography, economists outside the US are almost totally ignored. According to Krugman, the FWEs have great faith in the ability of markets to get things right, and little faith in the capacity of fiscal policy to moderate the business cycle, while the SWEs have much less faith in markets and considerably more faith in fiscal policy. He argues that these debates had been largely dormant for many years, but erupted again with the onset of the GFC.

Krugman is so strongly focused on the civil war between American economists that he fails to consider the possibility that they are all missing the point: that government failure, rather than market failure, is to blame for the severe downturn we have been witnessing in the US and many other countries. One therefore gains the impression that the two major schools of thought on macroeconomic policy in the US are so preoccupied with defeating each other on the intellectual battleground that the Keynesian/neo-Keynesian (SWE) versus Chicago School (FWE) debates have simply overlooked certain crucially important matters.
In fact, governments have made two major policy errors: one of commission and one of omission. In neither case, however, are the underlying economic issues needful of, or readily susceptible to, mathematical analysis, so they have been largely ignored by economists.

In contrast with the voluminous criticism of bankers for the obvious recklessness with which they managed their institutions, comparatively little has been said about the role of the moral hazard that is generated by the practice of governments bailing out banks (and, more recently, other financial institutions) on the grounds that they are ‘too big to fail’. This is not to suggest that there is no awareness of the problem, but simply that there has been little attempt to come to grips with it. By comparison with the reams of discussion about whether markets are efficient and whether fiscal policy can stabilise the economy, there has been virtually none on how to deal with the widespread perception among bankers that they can virtually ignore risk without incurring the wrath of their shareholders or depositors, secure in the knowledge that the government will not allow their institutions to fail.

When the US authorities decided to let Lehman Brothers go to the wall, I naïvely imagined that they had finally got the message, for there is probably no better way to get the attention of bank shareholders and depositors than by allowing some of them to lose their shirts. But when Lehman is mentioned these days this policy decision is portrayed, absurdly, as an error that triggered chaos in America’s financial sector, the shockwaves from which spread rapidly around the world. The message now seems to be: ‘Oops! That was a big mistake. We will never do that again!’ But that is equivalent to saying to the financial community: ‘Go right ahead. Take all the risks you want. We will back you up with taxpayer money if anything goes wrong.’

How else can we explain the fact that US banks were lending the full purchase price of houses to buyers who were not even employed? That they focused on the relatively recent rapid increase in house prices rather than the fact that they were by now well above their long-term trend? That other institutions were prepared to buy securities backed only by these kinds of assets?

Do those now critical of the Lehman decision seriously believe the same kind of thing will never happen again because older and wiser regulators will do their job better in the future? Do they seriously believe that bank behaviour can be significantly altered by regulating the payment of bonuses to their executives? Policy for the future needs a much firmer foundation than this kind of wishful thinking, and one of the ways in which economists have let the world down is undoubtedly their failure to focus on the moral-hazard problem.
The second key error of governments has been their failure to devote sufficient attention to one of their most important roles in the economy: providing an efficient legal system. The private sector can run banks but it cannot run a legal system, because it lacks the coercive power of government to settle civil disputes and to punish those who commit crimes.

Of crucial importance in the current context is insolvency law: the legal processes put in train when companies — including banks — become insolvent. In effect, broadly speaking, when a company becomes insolvent, shareholders lose their investment, and creditors share proportionately in the remaining value of the assets — whether the company is liquidated or revived by way of injection of new equity. Unfortunately, however, so little attention has been paid by governments the world over to the design of these processes that they have become extraordinarily drawn out. Creditors often cannot get hold of the residual value of their claims for several years.

It never seems to occur to the authorities, or to economists, that a better designed set of legal arrangements might be able to deal with insolvencies within days rather than years — and that this is precisely what is needed in the case of banks. What makes banks special from the point of view of macroeconomic policy is that they provide the payments mechanism for the entire economy. If depositors cannot get access to their money they cannot spend it, and without spending, economic activity grinds to a halt. This is the reason why central banks act as lenders of last resort and governments create deposit insurance and deposit guarantee schemes — namely, to allow depositors continued access to their funds, even if their bank becomes insolvent.

In a frictionless legal system, bank insolvencies would result in the immediate cancellation of existing shares and the writing-off of a sufficient proportion of deposits and other liabilities so as to return capital to a satisfactory level, thus allowing the bank to continue to operate and depositors to access the remainder of their deposits. One of the key challenges facing the economics profession is to design an insolvency process for banks that comes as close to this frictionless ideal as possible.

The central feature of such a scheme would be a bank regulatory authority with power to freeze the operations of any bank immediately there are plausible concerns about its solvency. Following this, there would be a ‘quick and dirty’, highly conservative, evaluation of the bank’s assets by suitably skilled assessors. Assuming the bank is indeed insolvent on this basis, the existing shares would be cancelled and new equity would be injected by means of writing down the value of all liabilities, proportionately, to an extent sufficient to allow the bank to resume operations. Depositors and other creditors of the bank would thus become its new shareholders.
Such a process could be implemented within a few days, thus minimising disruption to the payments system. There is now no longer any reason for a run on the bank. Depositors cannot avoid incurring losses because it is already too late. If actual asset values turn out to be higher than the conservative valuations produced initially the newly issued shares will turn out to be more valuable, such that creditors will lose less than first thought. Indeed, if actual asset values turn out to exceed total liabilities, the original shareholders could also recoup some of their investment.

The bottom line is that disruption to the payments system — and therefore to the economy as a whole — is minimised, while losses incurred by the bank are borne, first, by the shareholders, and, second, by creditors (mainly, depositors). There is no cost to the general public, so moral hazard disappears; and banks would be much more prudently managed as a result. This is precisely how the capitalist system is supposed to work when firms become insolvent.

Unfortunately, putting flesh on the bones of proposals such as that outlined here will require assistance from the legal profession and other bankruptcy specialists who may instead be inclined to oppose them, since their large incomes are the direct manifestation of the gross inefficiency of the legal processes relating to insolvency. Moreover, the dominant and highly insular US economics fraternity unwittingly portrayed by Krugman, with its human capital heavily invested in mathematical and econometric virtuosity, seems too strongly inbred with its own ideas to be thinking about extending the debate in the directions indicated, despite its now heightened concerns about the evidently parlous state of macroeconomics.
Froth and Bubble: The Inconsistency of Paul Krugman’s Macroeconomic Analysis

DON HARDING AND JAN LIBICH

Consistency is one of the touchstones used to evaluate not only arguments but also the people that put forward the arguments. In assessing the person advocating an argument, it is natural to look for coherence over time in their arguments and, secondly, whether the person offers a convincing explanation for a change of view. We apply this framework to evaluate some of Paul Krugman’s macroeconomic analysis.

Froth and Bubble in 2002

In the New York Times in August 2002, Paul Krugman writes: ‘To fight this recession the Fed needs more than a snapback; it needs soaring household spending to offset moribund business investment. And to do that, as Paul McCulley of Pimco put it, Alan Greenspan needs to create a housing bubble to replace the Nasdaq bubble’ (Krugman 2002).

In blog and newspaper pieces around that time Krugman is encouraging the Federal Reserve to achieve such stimulus through looser monetary policy. For example, he argues: ‘It’s still not clear that Mr. Greenspan has caught up with the curve — let’s have at least one more rate cut, please.’ (Krugman 2001). Yet half a decade later he argues: ‘If there are two guys that I blame for this crisis it would be, in order, Alan Greenspan and Phil Gramm.’ (Krugman 2008a)

It is reasonable enough to change one’s mind. But to change one’s mind without explanation, and to compound the aggravation by damning those who implemented the very policy that one favoured at the time, is extraordinary.

1 La Trobe University. J.Libich@latrobe.edu.au.
Not content to leave this well alone Krugman disingenuously explains in regard to his 2002 piece quoted above, that: ‘It wasn’t a piece of policy advocacy, it was just economic analysis.’ (Krugman 2009a)

**Which Paul Krugman Was Right About Policy in 2002?**

The recession of 2001 was one of the shortest (March–November) and mildest in US history. The unemployment rate rose from the December 2000 low of 3.9 per cent to 5.7 per cent a year later; a rate it roughly held throughout 2002. Employment performed similarly well. Inflation reached its low of 1.1 per cent in January and February 2002, and then started rising steadily (with a small temporary dip in June/July 2002). GDP actually grew in the first and second quarters of 2002 by 1.6 per cent and 1.5 per cent (annual rate), respectively. ²

There are many more things that one might want to take into account but even so it is hard to find evidence of sufficient importance to justify Krugman’s extremely pessimistic assessment of the situation in mid-2002.

Moreover, there are well-respected alternative views of policy in that recession. For example, according to Taylor’s (2007) estimates, the traditional Taylor rule (that had described the Federal Reserve’s behaviour quite well up until that time) prescribed increases of the Fed Funds rate to start already at the end-2001, and to continue at a steady pace to reach 4 per cent at the end of 2003 (as opposed to the rate of 1 per cent observed at the time). ³

There have been some revisions in the original data available to policymakers in real time, but these do not change the overall picture: in August 2002 there was little reason to panic and call for dramatic policy measures, as Paul Krugman did. This is especially true considering the fact that the world economy (except perhaps for Western Europe) was performing fairly well, and that the expansionary effects of the 2000–01 interest-rate cuts were still largely in the pipeline.

---

² Note that the level of GDP never recorded a sustained decline, and thus identifying this period of time as recession was not uncontroversial. Furthermore, the NBER dating committee indicated that in the absence of the September 11 terrorist attacks: ‘…it is possible that the decline in the economy would have been too mild to qualify as a recession’ (see Hall et al. 2001).

³ Bernanke (2010) disputed Taylor's conclusions, arguing that monetary policy was not excessively expansionary during the early 2000s if one uses contemporaneous (not forecast) values of inflation and output. Papell (2010), however, shows that Bernanke's results arise from using two different measures of inflation (headline versus core), rather than the contemporaneous versus forecast inflation distinction.
Deficit Trouble in 2003

By 2003 debt and deficits were the issue. In March 2003 Krugman writes:

But what’s really scary — what makes a fixed-rate mortgage seem like such a good idea — is the looming threat to the federal government’s solvency.’

That may sound alarmist: right now the deficit, while huge in absolute terms, is only 2 — make that 3, O.K., maybe 4 — per cent of GDP. But that misses the point.

Think of the federal government as a gigantic insurance company (with a sideline business in national defense and homeland security), which does its accounting on a cash basis, only counting premiums and payouts as they go in and out the door. An insurance company with cash accounting…is an accident waiting to happen.

So says the Treasury under-secretary Peter Fisher; his point is that because of the future liabilities of Social Security and Medicare, the true budget picture is much worse than the conventional deficit numbers suggest.

Of course, Mr. Fisher isn’t allowed to draw the obvious implication: that his boss’s push for big permanent tax cuts is completely crazy. But the conclusion is inescapable. Without the Bush tax cuts, it would have been difficult to cope with the fiscal implications of an ageing population. With those tax cuts, the task is simply impossible. The accident — the fiscal train wreck — is already under way. (Krugman 2003)

This is pretty much a mainstream view that many economists would understand. It’s a point of view that John Maynard Keynes, who unlike many of his followers was not a proponent of over-reliance on deficit spending or tax cuts, would agree with. But the issue again is the consistency of Krugman’s arguments. As Lee (2009) observes:

The problem is that everything Mr. Krugman now writes entirely contradicts his 2003 article, despite the fact that every fundamental problem the economy faced six years ago is now much worse. Mr. Krugman has no issues with Barack Obama and Ben Bernanke committing the same atrocities the previous administration committed. President Obama has ramped up every budget, including the military budget,  

---

4 Many modern-day Keynesians gloss over the fact that in How to Pay for the War, published in 1940, Keynes argues for higher taxation to fund the spending.
while Bernanke runs the presses faster than Greenspan ever did. Mr. Krugman has consistently stated throughout 2009/2010 that there is no danger of interest rates rising in the future and that the budget deficit is not disastrous by comparison to 1940s United States and 1990s Japan.

The Twist and Turns of Economic History

Nobody can be like Keynes and no book can be like the General Theory. It was a book that solved the problem that nearly brought down Western civilization.’ (Krugman 2008b)

There is no doubt that Keynes was one of the most influential economists of the twentieth century, but the claim that he ‘solved the problem that nearly brought down Western civilization’ is not supported by the evidence.

The Means to Prosperity, published in 1933, recommended increased public spending, though not necessarily deficits, as a means to reduce unemployment. But there is little evidence that it had a direct effect on macroeconomic policy during the Great Depression. Skidelsky (2003), Keynes’ biographer, reports that Keynes’ policy suggestions only began to influence US economic policy after 1939, by which time the depression had ended.

It is true that the Swedes did employ some Keynesian ideas in the Great Depression (Skidelsky 2003). It is also true that German public spending in the Depression had some of the positive economic effects foreseen by Keynes. But, Keynes never claimed either of these examples as support for his ideas. Sweden was too small and the other German policies too abhorrent.

Truth and Beauty

So what are the lessons that Paul Krugman has learnt from the global financial crisis of 2007–09? Among others, in his highly publicised article (Krugman 2009b) he expressed the view that the state of macro ‘is not good’. His argument is that the profession has mistaken ‘beauty for truth’ in terms of hi-tech models that do not reflect the economic reality.

Many others have dissected his arguments on this (see, for example, Cochrane 2009 or Kocherlakota 2009). We have documented the changing state and consistency

---

5 As quoted in his Collected Works (volume 21, 94), Keynes stated: ‘I am sure the Premiers’ Plan last year saved the economic structure of Australia’, referring to the Plan of April/May 1931 which recommended a reduction in the Commonwealth deficit from £39 million to £11 million, to be secured by a £13 million reduction in outlays, £12 million increase in taxes, and £3 million from reduced interest.
of Paul Krugman’s macroeconomic policy analysis. The contrast between his assessment of 2002/2003 and 2008/2009 is so large and the justification for the changed view so ephemeral that we feel his policy recommendations no longer have the required consistency and coherency. Rather than being the substance that nourishes public debate they seem to conceal the substance.

References


Krugman, P. 2008a, Interview in Gregory One on One (video), 22 September, Available at: http://www.youtube.com/watch?v=YwqcLbZJ4HA.

---

6 This does not in any way detract from the importance of Krugman’s contributions to new trade theory for which he was awarded the Economics Nobel Prize.

Krugman, P. 2009a, ‘And I was on the grassy knoll, too’ (blog post). Available at: http://krugman.blogs.nytimes.com/2009/06/17/and-i-was-on-the-grassy-knell-too.


Beyond Krugman to Behavioural Keynes

IAN M. MCDONALD

Paul Krugman’s critique of financial and macroeconomics, ‘How did economists get it so wrong?’, is apt. Whilst the financial sector is the biggest threat to macroeconomic stability, the study of the financial sector by academic economists — a major endeavour occupying the resources of some of the brightest people in the economics profession — has not reduced this threat. Indeed some argue that financial innovations based on financial economics have increased financial instability. In spite of this huge research effort the financial sector has been reduced to a shambles reliant on government bail-outs, and the macroeconomies in the US and Western Europe are in deep recession. It is not unreasonable to conclude that there has been a major misallocation of resources in the science and policy of financial economics and macroeconomics. But how could that have happened?

Whilst the financial sector is far from untouched by government intervention, it is a sector in which demand and supply dominates the setting of prices, and in which participants, because of the size of their transactions, have every incentive to make sensible decisions. Homo economicus would be untrammelled and motivated. But instead of the sensible decision-making one would expect from Homo economicus, we observe rashness and short-sightedness. For Krugman, this suggests that the model of Homo economicus is inappropriate for the financial sector. Krugman advocates a different model, that of behavioural finance, in which individuals are strongly influenced by the behaviour of others and suffer from irrational exuberance, unwarranted panic, loss aversion, and ill-considered extrapolations from small samples.

In an emotional response to Krugman’s critique, John Cochrane appears to concede the important pillars of Krugman’s views. Cochrane agrees that asset prices are excessively volatile. He also suggests that this excessive volatility is perhaps due to the vulnerability of people to bursts of irrational optimism.

1 University of Melbourne, i.mcdonald@unimelb.edu.au. I thank William Coleman and a referee for comments on a previous draft.
and pessimism. To my mind, this supports the view that *Homo economicus* is inappropriate and that behavioural finance should be playing a more central role in the research and teaching of financial economics.

The theories of economic decision-making developed by behavioural economics offer better insights into human behaviour than do those of the standard economic paradigm based on *Homo economicus*. The individuals in behavioural economics are social animals, influenced by context, especially the activities and outcomes of other individuals. One result of this is herding behaviour, a commonly observed phenomenon in asset markets. The decision-makers of behavioural economics are also subject to biases, especially present bias and self-serving bias, and to emotional distortions that override and/or prevent attempts to behave according to the cool, long-sighted calculations of *Homo economicus*.

*Homo economicus* would not lend to highly leveraged financial institutions, since leverage encourages excessive risk-taking by the borrower. *Homo economicus* would not lend to financial institutions that operate incentive schemes that encourage short-run behaviour by their staff that jeopardises the wealth of creditors. *Homo economicus* would not ‘search for yield’; that is, take on excessive risk in rash attempts to achieve a historically determined reference rate of return.²

Krugman views Keynesian economics as the best framework for macroeconomics, saying ‘a more or less Keynesian view is the only plausible game in town’. Krugman argues for the importance of fiscal policy in tackling cyclical unemployment, especially when interest rates have got to zero. Krugman’s advocacy of a Keynesian approach is based on his view that it is inadequate aggregate demand that drives recessions, not confusion about relative prices, nor lapses in technical progress, nor voluntary shifts to leisure during times of low real wages. While rejecting these microeconomic explanations, it is noteworthy that Krugman offers no clues about the microeconomic foundation that would support the sustained impact on activity of the decline in aggregate demand.

This absence of a microeconomic foundation for macroeconomics is also apparent in an earlier paper, Krugman (2000). In that paper, Krugman argued for the importance of a Keynesian approach to macro-policy issues. He advocated a Keynesian/IS-LM/Mundell-Fleming approach supported by ad hoc microeconomic foundations, in particular an adaptive expectations Phillips curve and regressive expectations for the real exchange rate. For Krugman, this Keynesian/IS-LM approach had the advantage of realism and simplicity. However, it’s disconnect from economic decision-making by individuals —

---

² For discussion of these points, see McDonald (2009).
that is, its lack of micro-foundations — suggests a worrying disconnect from academic research. As a stop-gap or temporary solution, this may be fine, but as a programme for macroeconomics it is unsatisfactory.

The microeconomic-foundations revolution of 1968 attempted to extract a macroeconomic model consistent with Keynes’ aggregate-demand hypothesis from the narrow conception of human behaviour as embodied in *Homo economicus*. In that revolution, job-search behaviour played the key role in the determination of wages. However, with no concern for wage relativities, there was no role for the sticky wage story put forward by Keynes. The result was a brittle theory of how inadequate aggregate demand influences the level of activity, brittle because in those natural-rate models the real effect of aggregate demand depended upon a persistent overestimate by workers of the level of prices.

The result of this brittleness was and remains a puzzle to me. It killed Keynesian economics rather than stifling at birth the unrealistic modelling of wage determination on which it was based. The reliance of the theory on a persistent overestimation of the level of prices was vulnerable to the concept of rational expectations. As soon as rational expectations was considered as an alternative to adaptive expectations, the ability of the theory to explain the impact of aggregate demand on the level of activity evaporated. For me, this exposed the underlying theory of job search as inadequate for an explanation of wages. But the mainstream took a different view and decided instead to jettison the Keynesian theory of aggregate demand.

There may have been a problem of timing. In 1968, the sophisticated understanding of the role of relativities in human behaviour as described by prospect theory, with the associated concept of loss aversion, was not available. It only appeared a decade later in Kahneman and Tversky (1979). Had this theory has been around in the 1960s, when Keynes’ influence was stronger, then the micro-foundations revolution may have avoided an excessive reliance on *Homo economicus*. Prospect theory has emerged as the natural way to model concerns with relativities. Using prospect theory to model Keynes’ concerns about wage relativities shows how wage rigidity is consistent with individual behaviour, as shown by the pioneering paper of Baskar (1990). This implication follows from the concept of loss aversion. The extreme dislike felt by workers for cuts in wages below the reference level is an example of loss aversion.

Keynes himself showed prescience by anticipating the concept of loss aversion. Not only did Keynes emphasise the role of relativities but he also regarded this role as being influenced by what would later be called loss aversion. Thus on the page following the introduction of the wage-relativity effect, Keynes went on to say: ‘Every trade union will put up some resistance to a cut in money wages,
however small” (Keynes 1936: 15; emphasis added). ‘However small’ implies a discontinuity in the marginal utility of wages at the current level of the wage, which is the implication of loss aversion. 3

Thus behavioural economics can make a valuable contribution, not just to financial economics but to macroeconomics. 4 It is to be hoped that the turn made in 1968 to the narrow conception of Homo economicus — a turn somewhat similar to the Paretian turn taken (according to Bruni and Sugden, 2007) by economics in the nineteenth century — will be reversed and the important insights of Keynes taken up in the application of prospect theory to labour-market behaviour and macroeconomics.

References


Bruni, L. and Sugden, R. 2007, ‘The road not taken: How psychology was removed from economics and how it might be brought back’, Economic Journal 117(516): 146–73.

3 Some argue that Keynes theory did not rely on any special modelling of wage determination. However, Keynes’ own efforts to explain resistance to cuts in nominal wages even when unemployment is at a high level belies this idea. Furthermore, Keynes contrasted this behaviour with the flex-price behaviour of classical theory by using the expression ‘fall without limit’ to describe classical flex-price behaviour when the economy is at less than full employment (see, for example, Keynes 1936: 253 and 303–4). The decelerationist hypothesis of natural-rate theory implies ‘fall without limit’. Keynes regarded ‘fall without limit’ as empirically irrelevant for an economy at less than full employment and thus implicitly rejected the deceleration hypothesis and natural-rate theory.

4 Ideas from behavioural economics can also contribute to the theory of the determination of aggregate demand. The importance of present bias in the determination of saving rates (see, for example, Thaler and Bernatzi 2004) is a good explanation for why Ricardian equivalence doesn’t hold and why consumption is strongly influenced by income. Bowman, Minehart and Rabin (1999) show how the application of prospect theory to an inter-temporal model of consumption and saving can explain the commonly observed asymmetry in the consumption–income relation. Akerlof (2007) argues that the social norms of managers of firms can explain the sensitivity of investment to cash-flow. Akerlof and Shiller (2009) argue that confidence, sometimes aided by ‘new era stories’ in which historic changes are described to propel the economy into a brand new era, have strong influences on the level of aggregate demand.
Cochrane, J. 2009, ‘How did Paul Krugman get it so wrong?’ Available at: http://faculty.chicagobooth.edu/john.cochrane/research/Papers/#news.


Paul Krugman (2009a) argues that the global financial crisis represents a collective failure of an economics profession which had become complacent, believing that financial markets therefore could not, in themselves, cause non-trivial recessions.

Krugman argues cogently — though not without hyperbole — of economists’ blindness to sources of instability arising from within financial markets. While economists have failed to meet their own scientific criteria — to assemble a body of knowledge that may be used to make important predictions — Krugman observes ‘freshwater’ macroeconomists unable to engage fruitfully with their ‘saltwater’ critics such as himself. This touches a philosophical division — idealism versus realism — that goes back as far as Plato and Aristotle.

A profession that retreats into sects that disparage each other has to be a less efficient producer of knowledge than it might otherwise be. Nevertheless, contestability of ideas — albeit with insults — is a better knowledge factory than complacency. The essence of the problem, Krugman believes, is the bankruptcy of the foundation that both ‘freshwater’ and ‘saltwater’ had come to accept; the efficient-markets hypothesis.

Krugman sees the seductive mathematical attraction of the neoclassical synthesis — especially in its modern financial form, the efficient-markets hypothesis — as the discipline’s Achilles heel. Financial economists want to believe in efficient markets because they are beautiful. Further, the attraction of the hypothesis expanded because, by generating a critical mass of the faithful, it offered lucrative career opportunities. Beauty and money are an attractive combination in matters of career choice as well as in marriage partners. Not only was this kind of economics ‘elegant, convenient and lucrative’, but it appeared to be empirically validated, at least in as much as it predicted the prices of financial assets relative to each other.

1 Unitec Institute of Technology, krankin@unitec.ac.nz.
The seductive beauty of the synthesis of Robert Lucas’ new classical economics with Eugene Fama’s theory of efficient markets is based on the elegance of the price mechanism in ensuring: that unperturbed markets always clear in the short long-run; that the ensuing price signals lead rational buyers and sellers to allocate resources efficiently; that efficient financial markets rule out substantial economic shocks from the financial sector; and that macro-economic imbalances self-adjust in the face of a consistent and credible monetary policy. This appeared to be the solution to the holy grail of macroeconomics; deriving a set of macro-truths from micro-foundations.

Nevertheless, beauty is not substance, as Krugman notes. Krugman eschews the whole idea of drawing useful real-world knowledge from an elegant deductivist model drawn from a set of plausible premises about human behaviour. In reality, economic man falls between the narrow rationality upon which the failed theories depended and irrationality, Krugman argues. This means that economists can make some useful predictive statements, especially by incorporating the insights of behavioural finance, and by pursuing regulatory policies that counter the collectively adverse consequences of much financial decision-making.

Buyers and sellers might not be ‘rational’ in the sense that neoclassical economists use that term. Behavioural economics shows that people are less calculating than the models assume (or at least they seek to minimise calculation costs). Further, those who do calculate commonly miscalculate. We have biases hard-wired into our pre-historical brains that lead us, for example, to naively follow the herd, to avert losses by not selling assets they paid too much for if that means reducing the price, and to accumulate claims to wealth as if they were actual stores of food or other durable goods. But Krugman falls for the latter misapprehension when he says, ‘US households have seen $13 trillion in wealth evaporate.’ A $13 trillion financial bubble never was $13 trillion of tangible wealth.

Thus while Krugman touches on important issues he remains far from producing an alternative analysis. He offers insights into a deeper critique of the new classical model, but is unable to extend his message beyond the combined insights of Keynes, and the findings of the new behavioural economists. So, irrational practices in the financial sector — including ‘institutions that run amok’ — become Krugman’s bogey. This is consistent with the banks having become the miscreant of populist analysis, and the view that substantially greater regulation of financial institutions will save us from future events as serious as (or worse than) 2008.

My reflection suggests to me that the neoclassical emphasis on the price mechanism — with its implicit Platonist idealism — is not as central to the workings of free-market capitalism as most economists, including Krugman, believe. There continue to be elegant pathways within the free-market tradition.
that have been little traversed. I do not believe that Krugman has the necessary vision to overturn the efficient-markets hypothesis by exploring such a new pathway.

In some markets, prices rise or fall in the short run but are restored in the long run at greater or lesser quantities bought and sold; the neoclassical theory of the firm. It’s not clear that price variation was necessary to achieve such resource reallocations. In other markets, such as housing, the opposite would appear to be true; when demand falls, quantity adjustment predominates in the short run. Price-adjustment is rapid in most financial markets; but the mechanism is far from efficient; it repeatedly over- and under-shoots. Other markets — for new durable goods — seem to self-correct with a minimal role for price signalling. The ubiquity of belief in the price mechanism does not necessarily make that mechanism an exclusive micro-foundation for macroeconomics. When addressing economic crises, ‘freshwater’ critics of Krugman (for example, Mulligan 2009) continue their reluctance to look beyond distortions in relative prices. Despite a substantial growth of inequality preceding recent panics, such economists continue to claim that wages are too high.

An important part of the problem as I see it is our failure to persevere with circular-flow models — the 1950s’ ‘hydraulic Keynesianism’ of Bill Philips notwithstanding — and to integrate them with balance-of-payments analyses. Krugman himself noted: ‘As a sheer matter of accounting, the balance of payments always balances...A country that runs a surplus on capital account must run a deficit on current account’ (2009b: 43). Insights gleaned from the classical critique of mercantilism can enhance our understanding of economic decision-making at the micro level.

While Krugman’s historical point of reference is the 1930s’ Great Depression, he barely acknowledges that this was an international event, and (Krugman 2009a) likewise fails to globalise his perspective on the 2008 crisis. As Krugman notes, it is not really credible anymore to claim that better US monetary policies in 1929–31 could have prevented the Great Depression. Nor is it credible that we can claim any global financial crisis is just a domestic crisis that overspilt the Unites States’ borders.

The most plausible explanations of the Great Depression — developed, for example, by Eichengreen (1992) — stem from balance-of-payments analysis. It was the refusal of creditor countries to run trade deficits in the context of

—

2 Indeed Bill Philips is better known for his Philips Curve, in which the changing price of labour plays a key determining role, than for the hydraulic Moniac machines he built in the 1950s, through which he simulated many of the variations of the circular-flow processes. [I first encountered the delightful expression ‘hydraulic Keynesianism’ in Alex Millmow’s (2007) essay on Douglas Copland. Needless to say, Copland, though with Keynesian sympathies, was not a hydraulic Keynesian.]

3 Krugman (2009b) offers a more globally oriented account.
the restored gold-standard currency regime that prevented debtor countries from running surpluses. (We might note that a system of floating interest rates — adopted from the 1970s as a price-based global stabilisation mechanism — probably could not have prevented the Great Depression. Financial flows between nations negate the requisite exchange-rate adjustments. Indeed, in the 2000s under floating rates, it was the debtor/deficit countries that consistently had appreciating currencies.)

Economic historians have documented the mercantilist world of the sixteenth to eighteenth centuries, and the writings that supported national strategies of producing ongoing current-account surpluses. The mercantilist mindset has never disappeared — it is commonplace among Krugman’s ‘policy entrepreneurs’ (MacFarquhar 2010). Indeed the parable of the failed baby-sitting circle he cites (as a model recession) might be interpreted as a mercantilist system in miniature. Contrary to Krugman, the printing of the correct amount of scrip might not have been a sufficient solution to the problem.

Ask yourself: what happens to microeconomics if our individual actors are hard-wired 4 to operate according to the same form of rationality as mercantilist governments did (and continue to do)? Such behavioural assumptions could give a new micro-foundational explanation of macroeconomic instability. Combine this with Keynes’ insights into the importance of the liquidity constraint and we see a macroeconomic order that periodically veers from over-consumption (the predominant bias in a period in which bank lending expands because bank deposits form an increasing percentage share of the money supply) to under-consumption.

A tendency towards under-consumption (deriving from habitual saving as per the mercantilist mindset) in the 2000s was offset by the marketing by financial intermediaries of secured and unsecured debt to households; an enterprise global in scale. This offset was strong enough in most of the decade to maintain high rates of consumer-led economic growth. (Inflation only threatened in 2007 as demand for food and oil outstripped supply.) The price of maintaining aggregate demand though increasingly unsustainable household debt was the creation of a huge bubble of financial claims on households and countries that had to consume more than they produced in order to accommodate those savers who habitually produced more than they consumed.

4 Hard-wiring to store and conserve food — like the squirrel of folklore — and other life-preserving items must have been an important facet of human evolution. That saving bias has been transferred from storing actual wealth to storing claims on wealth, and indeed ‘claims on claims’ (as Krugman calls derivatives).
A second periodic over-consumptionist bias is revealed by a careful inspection of the basic circular-flow model. The link labelled ‘household saving’ is actually a netting of four separable flows: ‘Saving’ = gross saving – expenditure funded by past saving – lending to households + loan repayments by households.

Net saving can easily become negative when expenditure funded by past saving occurs in unusually high quantities — due either to panic, population ageing, or simply the sheer bulk of financial assets in circulation. What happens if there is an unsustainable surge to realise historical claims as expenditure on goods and services? This — a crisis of over-consumption — is likely to be the nature of the biggest macroeconomic crisis this century, and we have no clear precedent for it. By developing the circular-flow approach, we can both explain the recessionary events at the end of the 2000s, and predict the quite different events that can be expected this century when the great pension funds are unleashed upon a supply-inelastic global economy.

To summarise: Paul Krugman is right to question the capability of economics developed under the auspices of the dominant neoclassical paradigm to predict business cycle events. A more constructive critique is required, however. The approach suggested here, which changes the analysis of markets away from its neoclassical emphasis on relative prices, may give many of the additional insights that we will require if we are to navigate the new global environment successfully. The new circumstances include heightened inequalities (especially within developed nations) and indications that the global economy as a whole may be much less supply-elastic — due to both natural and human resource constraints — than it proved to be last century.

References


What Keynes Missed and Krugman is Missing: The Short/Long Choice

DAVID VINES

Introduction

Until three years ago we were living in the ‘Great Moderation’. Macroeconomic outcomes were good. And New Keynesian macroeconomics — the theoretical apparatus which underpinned macroeconomic policymaking — appeared to be in good shape. Olivier Blanchard, now Chief Economist at the International Monetary Fund, wrote that ‘the state of macro is good … The battles of yesteryear… are over, and there has been … a broad convergence of vision.’

Then things fell apart.

How did economists get it so wrong? Paul Krugman’s answer to this question is profoundly misleading (Krugman 2009). As ever with Krugman, it is also profoundly thought-provoking. But I have ended up with a quite different answer.

The Krugman story

Three steps

Firstly, Krugman laments the fact that, over the last 50 years,

… economists fell back in love with the old idealised vision of an economy [which was widely believed before the Great Depression] in which rational individuals interact in perfect markets … This influenced their view of policy … Lucas [a Nobel Prize winner of Chicago] says
the Obama administration’s stimulus plans are ‘schlock economics,’ and his Chicago colleague John Cochrane says they’re based on discredited ‘fairy tales.’

Second, Krugman implies that economists fell back in love with this old idealised version because, like Keats, they mistook beauty for truth. Krugman is saying — although never quite explicitly — that only the ‘old idealised vision’ can be represented in the ‘all-encompassing, intellectually elegant’ kind of analysis that economists like to produce.

Third, Krugman suggests that ‘economists need to abandon the neat but wrong solution of assuming that everyone is rational and markets work perfectly’. As a result,

economists will…[then] learn to live with messiness, will…acknowledge the importance of irrational and often unpredictable behaviour…[and will] face up to the often idiosyncratic imperfections of markets. … In practical terms this will translate into more cautious policy advice — and a reduced willingness to dismantle economic safeguards in the faith that markets will solve all problems.

Evaluation

I think that Krugman’s answer is actually too optimistic, and that things are worse than he suggests.

Most of us have never believed in the idealised vision which Krugman described. I myself studied economics at Melbourne University in the late 1960s because I thought — correctly — that economic policy is necessary to correct the many failings of the free market. And I have spent the rest of my life trying to teach my students what the necessary policy should look like. ³

And those who have led the subject in my generation have certainly never believed in a ‘perfect markets’ vision. Take Stan Fischer, perhaps the best macroeconomist of our time, once Head of the foremost Economics Department in the world (at MIT), then First Deputy Managing Director of the IMF (one level up from the post that Olivier Blanchard now holds), and currently Governor of the Central Bank of Israel. The great practical work of Fischer’s life was achieved during the Asian Financial Crisis of 1997–98. Fischer led a team at the IMF which laboured, using the best tools available, to avert a global financial meltdown.

³ James Meade, my mentor when I was a young researcher at Cambridge, went to Oxford in the late 1920s to study Classics. But Meade gave this up because he saw unemployment everywhere around him and wanted to do something about it. As a result, he studied economics, and then went to Cambridge to work as a graduate student with Maynard Keynes. Many of us are still like Meade.
Now Olivier Blanchard has achieved the same thing in the present financial crisis. Neither Fischer nor Blanchard has been working with silly Panglossian theories. They have both made use of the best available analysis of what has gone wrong, and of what to do to fix it.

The trouble with all of us was that, even although we were fully aware that things could go wrong, we did not see this particular crisis coming. Our trouble was not that we believed that markets are perfect. Rather, we just did not understand how the financial system works. This might seem odd for macroeconomists, but it was true. Our problem was not one of ideological bias. It was one of poor understanding.

Krugman on the future of macroeconomics

At the end of his paper, Krugman suggests that economists ‘should recognise that Keynesian economics remains the best framework we have for making sense of recessions and depressions’.

What is Krugman’s ‘Keynesian framework’? To answer this, we need to remind ourselves of the key insights in Keynes’ General Theory of Employment Interest and Money. There are three.

First, Keynes argued that the level of demand for goods can influence the position of the economy: an economy can end up producing less than what is possible. Second, Keynes maintained that serious falls in demand can be caused by shocks in the financial system. Third, he thought that economies might actually be unstable if they were not actively managed by fiscal policy and/or monetary policy. If there is a large negative shock to demand, and a recession, and prices start to fall, Keynes thought this combination of factors might cause ‘debt deflation’. As prices fall, what debtors owe becomes more and more onerous, leading to further reductions in what is spent. Consequently, the economy may never recover under its own steam. The much-neglected Chapter 19 of Keynes’ General Theory makes this point very clearly. All three of these points really matter.

---

4 Actually many of us thought that we saw coming a crisis arising from a collapse of the dollar — which has not happened. Even Krugman made this mistake.
5 Modern textbooks abstract from this point — it is too inconvenient for them. And so, generations of students have been taught in macroeconomics courses that if only prices were somehow more flexible, then recessions could be quickly averted, reducing the need for active policy intervention. Keynes did not believe this.
What is missing from Keynesian economics

But Keynes’ *General Theory* does not contain a theory of financial crisis. While Keynes’s popular writing (for example, Keynes 1932) contain remarkable insights about the financial crash which happened during the Great Depression, it is not true that, lurking beneath the surface of the *General Theory*, there lies a theory of financial crisis, which economists have somehow stupidly forgotten. Yet somehow the myth of a hidden Keynesian solution has persisted.

Robert Skidelsky implies as much in his recent popular book about Keynes; even the title — *The Return of the Master* — suggests this idea. But the fact is that Keynes could not cut the mustard when it came to building a theory which is relevant to the present financial crisis.

In his discussion of finance in the *General Theory*, Keynes focused on the choice by investors between interest-bearing and non-interest-bearing money. He thought that the interest rate might be forced to rise if investors want more liquidity, and that the resulting rise in the interest rates might cause a fall in the demand for goods, and so might cause a recession. If the fall in demand was large enough, it might even — because of Keynes’ third point described above — cause an unstable collapse of the economy. That is clearly an issue.

But in the world in which we now live this issue has been comprehensively dealt with — partly as a result of what Keynes has taught us. In our modern economy the interest rate is set by the central bank, so as to make sure that inflation is under control, and to dampen down fluctuations in demand. In doing this the central bank ensures that our economy is not troubled by the kind of fluctuations in the demand for liquidity which concerned Keynes. If financial institutions want more liquidity, the central bank will provide it, without letting the short-term interest rate go up too much, if at all. Such central banks will also cut interest rates to offset the effects of any falls in demand. And if a recession does happen, the central bank will cut the interest rate enough to ensure that debt deflation does not occur. In fact, this modern approach to monetary policy deals with all three of Keynes’ points.

Until two years ago, we believed — in a self-congratulatory sort of way — that this kind of policy was all that was needed. And we were right — up to a point. Providing that the inflation-targeting approach worked all right, we thought, financial intermediation would work efficiently. As a result, the financial sector would provide no impediment to the functioning of the economy. A competitive financial system would drive risk premia on private debt down to low levels, so that — effectively — all short-term interest rates, including those on private debt, would mimic rates set by the central bank. And the prices of longer-dated assets would be set by efficient ‘inter-temporal arbitrage’. As a result, the return
on these assets would be governed by the interest rate which the central bank would be expected to set in the future, as it successfully operated its inflation-targeting policy. The central bank would thus effectively set those longer-term interest rates too.

But such a theory was silent on how balance-sheet problems of financial intermediaries might lead to very large increases in the risk premium attached to holdings of longer-dated assets, to a collapse in the price of these assets, to a rise in long-term interest rates, and to a difficulty for longer-term borrowing and investing. This theory was therefore completely silent on how our financial system led our whole economy to crisis.

Getting Finance Properly into Macroeconomics: ‘the short-long choice’

What the financial system does — at its heart — is to transform the maturity of lending. Households make savings decisions and — through financial institutions — lend their savings to others in the economy who invest them. The savings decision is — at least in many cases — a short-term decision. That is, savers need to be able to get their money out easily. Investment is, by contrast, a long-term decision, since investment projects take time to bear fruit.

The important thing in analysing the financial system is to focus on what we can call ‘the short-long choice’. How can we find enough people, or institutions, who take short-term savings and make them available for long-term investment projects? This will be risky. The interest rates charged on long-term loans — or the returns on long term equity investment — have to be large enough to compensate investors for the risk in going long. If the financial system becomes less able to bear this risk, then the interest rate charged to investors will rise, and the price that investors will get by issuing shares will fall (since fewer people will be able to buy them). This is what brought down the world economy two years ago.

What happened was this.

During the early 2000s, many investors engaged in a ‘search for yield’. They did this because in this period interest rates on short-term lending were very low. A good way for a financial institution to increase yield was to ‘leverage’; that is, to borrow money and then to invest the borrowings alongside the firm’s own capital. During the early 2000s, highly leveraged financial institutions (HLFIs) took their own capital, or shareholders’ funds, and supplemented it with large short-term borrowing from elsewhere at low interest rates, and then invested
these funds in longer-term assets paying a higher rate of return. This increased the expected return on their capital — which was good at a time when short-term interest rates were low — but, of course, it made that return much more risky.

An example of what can be achieved by leverage is the following simplified illustration. Imagine an investing institution with $100 in capital. Suppose that, if this was invested long term, it would earn $3, or 3 per cent. Suppose that, at the same time, the investing institution borrowed $900 short term, at a lower interest rate of, say, 2.5 per cent, and then invested the overall sum of $1000 in long-dated securities having a return of 3 per cent. Then the net earnings of the portfolio, after paying the interest due on the short-term borrowing, would rise from $3 to 0.03 x 100 + (0.03 – 0.025) x $900 = $7.50. Leverage would have raised the fund manager’s return from 3 per cent to 7.5 per cent.

And to see why leverage increases the risk faced by an investor at the same time as it increases prospective return, consider what happens to the above portfolio when the price of long-dated assets falls by 1 per cent. This would mean that the value of the investor’s portfolio would fall from $1000 to $990. But one part of the portfolio’s liability structure is unchanged — the value of the outstanding loans which the fund manager has incurred as a result of the borrowing used to finance the leverage. This borrowing remains unchanged, at $900. The other component of liabilities — the value of the portfolio to those who originally invested in it — must take up the slack, falling by $10 to $90. Thus a 1 per cent fall in the value of the long-term assets held by the portfolio will cause a 10 per cent fall in the balance-sheet value of the assets of the portfolio. A leverage ratio of 10 — that is, a ratio of assets-held-in-investment to own-capital of 10 — has increased the proportionate variation in the balance-sheet value of the portfolio by a factor of 10.

In the early 2000s the business model of HLFIs was very simple. Increase the leverage ratio as much as possible in order to increase returns. But only do as much of this as the risk managers in the firm will allow you to do. Leverage ratios well above 10 — up to 30 or 40 — have been common over the past few years.

A ‘Financial Multiplier’

This business model of HLFIs led directly to the financial crisis, as follows.

---

6 Shin (2009) describes how ‘value at risk’ calculations can be used to work out such a maximum leverage ratio.
Suppose — as imagined above — that the value of the HLFI’s portfolio falls by 1 per cent, from $1000 to $990. This initial effect will get magnified by a ‘financial multiplier’. Recall that the initial leverage ratio was 10 ($1000 in investments divided by the $100 of capital). We have already noted that a fall in the value of long-term assets from $1000 to $990 will cause the value of capital to fall from $100 to $90. But this means that the leverage ratio will have actually risen from 10 to 11. If we suppose — for now — that the initial leverage ratio of 10 really was the maximum that the firm’s risk managers would allow, then the HLFI will be required to contract its balance sheet by selling long-dated assets.

It will need to do this to enable its investments to fall in line with the fall in the value of its capital. To do this in such a way as to bring its leverage ratio back to 10 requires the HLFI to sell $90 of assets.

If such a process happens to any one HLFI on its own — facing a fixed price of financial assets — then the outcome would not be remarkable. But if this process is macroeconomic — that is to say if the position of our representative HLFI is symptomatic of what is happening in the economy as a whole — then the problem is serious. That is because the sale of $90 in long-dated assets will force the price of these assets to fall further, beyond the initial fall in price. But that will cause a further contraction in the value of the HLFI’s balance sheet, requiring it to make further sales of assets, causing further falls in asset prices, etc, etc — a true fire sale. In other words, a ‘financial multiplier’ process will come into play. Clearly, because the leverage ratio is much larger than unity — 10, in our example — there is a risk that this multiplier process causes the price of long-dated financial assets to implode, leading to a collapse of the whole financial system.

How bad can the Crash get?

Whether the process outlined above happens clearly depends on whether there is anyone else to take up the supply of long-dated assets. The question is: how much will this asset sale drive down the price? That depends on the price elasticity of demand of the other holders of long-dated securities. If the general public will absorb a large quantity of risky assets without much fall in price — that is, if there are many final investors like Warren Buffett — then asset prices will not fall much in the face of the shock which we have described.  

It will also try to reduce its own borrowing. To analyse the effects of this would complicate our story. But it does not change the overall picture.

My understanding of how a financial multiplier can create large fluctuations in asset prices owes much to a paper by none other than Paul Krugman himself (see Krugman 2008). But in this paper Krugman assumes that there are lots of people like Warren Buffett.
But if there are no others besides HLFIs who might hold these assets, then the
price will fall a lot, because each fall in price will cause HLFIs to dump more and
more securities on the market so as to maintain their preferred leverage ratio. In
reality, in the past two years, the crisis revealed that the holdings of securitized
assets, which (to ensure diversification of risk) were supposed to be distributed
widely beyond the banking system, had in fact largely remained either within
the banking system, or in off-balance-sheet financial vehicles which were still
ultimately part of the banking system (Bean 2009; Frexias 2010). This suggests
that the demand by the non-HLFI public for long-dated financial assets may
have been quite inelastic.  

In the limiting case in which the HLFIs are the only holders of long-dated
securities, the outcome depends on how far the price of these long-dated
securities has to fall, to ensure that the supply of leverage increases enough
(that is, the leverage ratio increases enough) to create a floor under the price of
these assets.  

To continue with our example from above, let us suppose that the
initial fall in the value of long-dated assets from $1000 to $990 represents the
true long-run fall in the value of these assets. As HLFIs offload these assets to
prevent their leverage ratio from rising above its desired level, the value of these
assets will go on falling way below their true value: $990, $980, $970, $960…. But as this happens, there will be an increase in the prospective returns to
holders of these assets, assuming that the value of these assets will, eventually,
return to the true long-run level of $990. This will lead HLFIs to increase their
desired leverage ratio. The further the price falls, the more this desired leverage
ratio will go on rising, because the larger is the fall in price, the larger are the
prospective capital gains which the purchasers of these assets expect to make
after they have bought them. Consider what happens in our example when
the value of long-dated assets falls to $950. At this point the value of invested
capital will have fallen to $50 ($950 – $900 = $50), and the leverage ratio will
have risen to 950/50; that is, all the way to 19. Let us suppose that, at this point,
the capital gain which the HLFIs expect to earn, as the assets rise in value back
to their long-run value of $990, is large enough to induce them to increase their
desired leverage ratio to 19. At that point the value of the long-dated securities
will stop falling, because investors will be prepared to hold them.

---

9 One reason for this may have been that many of the securitised assets were very complex. The seller has
superior information about the assets than the buyer, and they have an incentive to sell the lowest-quality
assets to the buyer, as in the classic lemon case, which, of course, makes these assets less attractive to the
non-HLFI general public. The current law case against Goldman Sachs has revealed just how much this was
going on.

10 My understanding of how there can be very large fluctuations in financial asset prices when there are
not lots of people like Warren Buffett, owes much to the tradition of work on the ‘financial accelerator’ started
by Bernanke, Gertler and Gilchrist (1999). In particular, a recent paper by Gertler and Keradi (2009) is very
important in this area.

11 In our simple arithmetic example we are not actually letting the value of these assets fall very much
before people are supposed to start buying them. This is because if the value of the assets fell by too much
How will recovery get under way — and will it be sustained?

After such a collapse in asset prices, how can an economy ever recover? What does our new way of looking at the financial system tell us about this?

What our story suggests is that the HLFIs will need to build up their capital again, to make up for the capital that was lost when the value of their long-dated assets fell. When the value of our HLFI’s assets has returned to $990, the value of its capital will have returned to $90. Suppose that, when assets have stopped rising in price, the leverage ratio desired by our HLFI will also return to 10. To support asset holdings of $990 with a leverage ratio of 10, the HLFI will have need to have built up its capital, back to $99. It thus needs to find an extra $9 in capital. To obtain that will require that there be a period of time during which there is a large gap between the return on long-dated assets and the short-term interest rate, so that our HLFI can generate large profits and build up its capital again.

All this suggests that, in the real world, HLFIs now need to make big profits, to rebuild their capital! Only when they have done this will they be able to start lending again at sensible low-interest rates. Such high profits are exactly what banks are now earning right around the world. And this has caused much annoyance. It is right that people should be annoyed. This is an outcome which seems very wrong indeed, since it was the behaviour of the HLFIs which caused the crisis in the first place.

But can this be avoided? One way would be for new banks to raise new capital and to enter the industry with it, stimulating the ability of the industry to provide loans. That does not seem likely at present. Another way would be to increase the public ownership of banks — at least for a time. Governments would inject more capital into existing banks, in exchange for extra shares that would be owned by the state. This would dilute the ownership of those who already own the bank — something which has already happened in the US and, especially, in the UK, where two of the three major banks are now more than half-owned by the state. Of course a further move would be wildly unpopular amongst the existing bank shareholders. And it would also be difficult at present, given the fiscal crisis of the state in so many countries. So — at least in the US and the UK — we may be stuck with an extended period during which there is a slow recovery, but during which banks earn high profits.

then all of the capital of the HLFI would be wiped out and it would go bankrupt. That will happen in our example if the fall in the value of the long-dated assets is 10 per cent. Of course, such bankruptcies actually happened in 2008 — think of Lehman Brothers. But we do not want to make our simple story too complicated by including that possibility.
Of course, in the longer term, we will need to regulate the financial sector to prevent this happening all over again. But that is another story.

**How did economists get it so wrong?**

To conclude, let us return to Paul Krugman’s question. Why did economists not see this coming?

In this article I have argued that the prices of financial assets can overshoot wildly in an economy with a highly leveraged financial sector. Economists missed this. Everybody — well, almost everybody 12 — thought that financial institutions could diversify away their risks by holding a mixture of assets. Few economists understood the way in which a sale of assets by some institutions could mean that all institutions end up having to sell assets, leading to a generalised fire sale, and causing crisis. In addition, there was little understanding of other factors which made the crisis worse — matters which I have not had the space to discuss here. 13 These include an understanding of how short-term inter-bank lending might dry up — which it did — or of how domino effects could operate within the banking system, so that the failure of one bank could bring down other banks — which happened.

In sum, it was not a case of mistaking beauty for truth. It was just that economists did not understand what truth looks like.

**References**


---

12 William White, at the Bank for International Settlements in Basle, and Nouriel Roubini, at Roubini Global Economics, were spectacular exceptions.

13 See Haldane (2009); May, Levin and Sugihara (2007); and May (2010) for valuable discussions of these questions.


It will be some years before the global economy recovers from the crisis that engulfed global financial markets in the course of 2008. Unemployment in the eurozone and the US is close to 10 per cent, and there is little to suggest that the situation is going to improve rapidly. The effort to stave off total economic collapse has left governments burdened with massive debt that will take years of painful effort to work off.

In one respect, though, recovery seems almost complete. The economics profession, which was briefly shamefaced by the failure to predict the crisis, or to reach any general agreement on its causes and the appropriate policy response, seems to have recovered its self-esteem. Quite a few economists seem ready to dismiss the entire crisis as a ‘transitory volatility blip’ (Coibion and Gorodnichenko 2010).

The policy prescriptions of market liberalism, including deregulation, privatisation and regressive tax ‘reform’, are being advanced with seemingly undiminished confidence. On the most charitable view possible, the adoption of these policy prescriptions did nothing to protect national economies from the worst impacts of the crisis. In fact, several of the hardest-hit, including Iceland, Ireland and the Baltic States were previously the subject of glowing praise for their embrace of market liberalism.

Among the exceptions to this complacent attitude, Paul Krugman has been notable for his willingness to consider the failures in economic analysis exposed by the crisis. Unsurprisingly, much of his attention has focused on the response of economists associated with the ‘freshwater’ or ‘Chicago school’ of macroeconomics.

1 University of Queensland, johnquiggin1@me.com.
However, Krugman has also pointed out the inadequacies the ‘saltwater’ or ‘New Keynesian’ group with which he is commonly associated. As he notes, this group embraced the spurious claim that the US and world economies were enjoying a ‘Great Moderation’ in which the business cycle had been tamed. They differed from their freshwater rivals only in the extent to which they regarded this happy outcome as the result of the stabilisation policies of central banks.

Both groups, Krugman says, mistook beauty, clad in impressive-looking mathematics, for truth.

Krugman cites Olivier Blanchard’s bald statement that ‘the state of macro is good’. Rather more insight can be obtained from the following, literally poetic, and somewhat ironic, metaphor:

A macroeconomic article today often follows strict, haiku-like, rules: It starts from a general equilibrium structure, in which individuals maximize the expected present value of utility, firms maximize their value, and markets clear. Then, it introduces a twist, be it an imperfection or the closing of a particular set of markets, and works out the general equilibrium implications. It then performs a numerical simulation, based on calibration, showing that the model performs well. It ends with a welfare assessment.

The class of models to which Blanchard refers encompasses the intellectual successors of both the Real Business Cycle models of the 1980s and the New Keynesian challenges to these models. They may be described as dynamic stochastic general equilibrium (DSGE) models.

As Blanchard observes:

Such articles can be great, and the best ones indeed are. But, more often than not, they suffer from some of the flaws I just discussed in the context of DSGEs: Introduction of an additional ingredient in a benchmark model already loaded with questionable assumptions. And little or no independent validation for the added ingredient.

Blanchard’s description brings out the central role of microeconomic foundations in the DSGE framework, and illustrates both the strengths and the weaknesses of the approach. On the one hand, DSGE models were able to represent a wide range of economic phenomena, such as unemployment and asset price bubbles, while remaining within the classical general equilibrium framework. On the other hand, precisely because the analysis remained within the general
equilibrium framework, it did not allow for the possibility of a breakdown of classical equilibrium, which was precisely the possibility Keynes had sought to capture in his general theory.

The requirement to stay within a step or two of the standard general-equilibrium solution yielded obvious benefits in terms of tractability. Since the properties of general-equilibrium solutions have been analysed in detail for decades, modelling ‘general equilibrium with a twist’ is a problem of exactly the right degree of difficulty for academic economists — hard enough to require, and exhibit, the skills valued by the profession, but not so hard as to make the problem insoluble, or soluble only with the abandonment of the underlying framework of individual maximisation.

The DSGE approach has failed to generate a truly progressive scientific-research programme. A study of some new problem such as the incentive effects of executive pay would typically begin with the standard general-equilibrium model, disregarding the modifications made to that model in previous work examining other ways in which the real economy deviated from the modelled ideal.

By contrast, a scientifically progressive programme would require a cumulative approach, in which empirically valid adjustments to the optimal general-equilibrium framework were incorporated into the standard model taken as the starting point for research. Such an approach would imply the development of a model that moved steadily further and further away from the standard general-equilibrium framework, and therefore became less and less amenable to the standard techniques of analysis associated with that model.

As Krugman says, DSGE modelling was beautiful (at least to economists) and illuminated some aspects of the truth, but beauty came first. An approach based on putting truth first would have incorporated multiple deviations from the standard general-equilibrium model and then attempted to work out how they fitted together. In many cases, the only way of doing this would probably be to incorporate ad hoc descriptions of aggregate relationships that fitted observed outcomes, even if it could not be related directly to individual optimisation.

These criticisms apply in particular to New Keynesian macroeconomics. New Keynesianism was ideally suited to the theoretical, ideological and policy needs of the Great Moderation. On the one hand, and unlike New Classical theory, it justified a significant role for monetary policy, a conclusion in line with the actual policy practice of the period. On the other hand, as Krugman observes, by remaining within the general-equilibrium framework the New Keynesian
school implicitly supported the central empirical inference drawn from the observed decline in volatility; namely, that major macroeconomic fluctuations were a thing of the past.

The failure of DGSE models to predict the crisis or to provide useful policy guidance requires a return to the basic Keynesian insight that the economy can be far away from general equilibrium for long periods. This does not mean abandoning all the work of the past 30 years and returning to old-style Keynesianism. But it does mean starting from the traditional Keynesian perspective that a general macroeconomic theory must encompass the reality of booms and slumps and, particularly, of sustained periods of high unemployment that cannot be treated as marginal and temporary deviations from general equilibrium. We must model a world where people display multiple and substantial violations of the rationality assumptions of microeconomic theory and where markets depend not only on prices, preferences and profits but on complicated and poorly understood phenomena like trust and perceived fairness.

First, the programme needs more realistic micro-foundations. As Akerlof and Shiller (2009) observe, we need to look at how people actually behave, and how this behaviour contributes to the performance of the economy as a whole.

Second, we need to reconsider the concept of equilibrium. The whole point of Keynes’ General Theory was that the market-clearing equilibrium analysed by the classical economists, and central to DSGE models, was not the only possible equilibrium. An economy can settle for long periods in a low-output, high-unemployment state that may not meet the neoclassical definition of equilibrium, but does match the original concept, borrowed from physics, of a state in which the system tends to remain and to which it tends to return. More importantly, perhaps, we need a theory which encompasses crises, and rapid jumps between one kind of equilibrium and another. Ideally, this will combine ‘old Keynesian’ analysis of economic imbalances with a Minsky-style focus on financial instability.

Between these two levels, we need to consider the fact that the economy is not a simple machine for aggregating consumer preferences, and allocating resources accordingly. The economy is embedded in a complex social structure, and there is a continuous interaction between the economic system and society as a whole. Phenomena like ‘trust’ and ‘confidence’ are primarily social, but they affect, and are affected by, the performance of the economic system.

An approach to economics that has been dominant for more than three decades will not go away simply because its predictions are inconsistent with the facts.
It is necessary to provide an alternative. The key requirements can be summed up by three simple propositions. In the twenty-first century, economics should focus:

- More on realism, less on rigor
- More on equity, less on efficiency
- More on humility, less on hubris.

The prevailing emphasis on logical rigor has given economics an internal consistency that is missing in other social sciences. But there is little value in being consistently wrong. Economics must move on from the infinitely rational, farsighted and asocial beings whose decisions have been the central topic of analysis in recent decades.

Three decades in which market liberals have pushed policies based on ideas of efficiency and claims about the efficiency of financial markets have not produced much in the way of improved economic performance, but they have led to drastic increases in inequality, particularly in the English-speaking world. Economists need to return their attention to policies that will generate a more equitable distribution of income.

Finally, with the collapse of yet another economic ‘New Era’ it is time for the economics profession to display some humility. More than two centuries after Adam Smith, economists have to admit the force of Socrates’ observation that ‘The wisest man is he who knows that he knows nothing’. While knowledge in the sense of absolute certainty may be unattainable, economists can still contribute to a better understanding of the strengths and weaknesses of markets, firms and other forms of economic organisation, and the possibilities for policy action to yield improved economic and social outcomes.

Every crisis is an opportunity. The global financial crisis gives the economics profession the chance to rethink ideas that led the world into crisis, and to produce more realistic, humble and, above all, socially useful bodies of thought.

References


RETROSPECT
Yegor Gaidar: Pragmatic Economist or Romantic Revolutionary?

GENNADI KAZAKEVITCH

The Controversy

Immediately after the life of any prominent person comes to an end, unalloyed tributes normally flow. Not surprisingly this happened for Yegor Gaidar, both in the Western media and on Russian internet sites of all political persuasions. But there was one difference: the obituaries, both in English and in Russian, reveal how divided public opinion was on the man and his influence.

“If a new Managing Director took over a large sluggish company and caused its value to shrink drastically, you’d call them a bungler, or at best unlucky. So why is Gaidar viewed as a hero?”

“I admire such strong personalities as Yegor Gaidar was — [a man] who is able to perform a great deed and not be afraid of the responsibility.”

These are from a Western source. In Russia, a former mayor of Moscow, Gavriil Popov, and the current mayor, Yury Luzhkov, claimed that Yeltsin appointed Gaidar as Acting Prime Minister in 1992 only under pressure from the United States, and a promise of tens of billions of dollars in aid (Popov and Luzhkov 2010). They added: “The failure of Gaidar’s model [of reform] was inevitable because it was based on a wrong theory, and did not correspond to the fundamental reality predetermined by socialism in Russia.”

In response, one of the most prominent Russian liberal journalists and dissidents Valeria Novodvorskaya (2010) writes: “Gaidar [is] our personal Saviour who opened the door for us to the capitalist Paradise.”

---

1 Monash University, Gennadi.Kazakevitch@buseco.monash.edu.au. The author is grateful to William Coleman for his encouragement in writing this article, and to the insightful suggestions and stimulating comments of the anonymous referee, Anna Taitslin and Luba Torlina.

2 These two quotes are drawn from The Economist, 17 December, 2009, at: http://www.economist.com/node/15125467/comments.
The fierce debate, almost 20 years after Gaidar’s short term as Prime Minister of the first post-communist government, over his contribution to both the building of the new Russian statehood and to the theory and economic history of post-communist transformation, only reflects how important Gaidar was to Russia, both as a politician and an academic.

This article attempts to place Gaidar’s role in contemporary Russia beyond partisan debate. It is predominantly based on Gaidar’s own key publications (Gaidar 2000–2007) and also reflects the author’s personal recollection of Gaidar’s intellectual and professional environment, from the 1970s to the early 1990s.  

The Life

Yegor Gaidar was born in 1956. The history of Soviet Russia and Gaidar’s family heritage were closely interrelated. Both his grandfathers, Arkadiy Gaidar and Pavel Bazhov, fought with the Red Army in the Civil War. Both became very famous — indeed, iconic — Soviet writers, whose books have been widely read by several generations of children and adolescents. Both lived through Stalin’s industrialisation, the Great Terror and the Second World War. Arkadiy Gaidar fell in the war against the Nazis, as a member of the poorly equipped home guard that was hurriedly recruited to defend Moscow in the autumn of 1941. Pavel Bazhov, a school teacher, miraculously avoided imminent arrest in 1938 (Gaidar 2000: Chapter 1), and thereafter lived a life of seclusion as a home-bound writer.

Gaidar’s father, Timur, was a prominent military journalist and held the military rank of a ‘one star’ admiral, a very high, but not unprecedented, rank for Pravda’s military correspondent. Since Yegor’s death, speculation has resurfaced that his father’s ‘real’ occupation was in the intelligence service. Though such speculation cannot be proved nor disproved, considering Timur Gaidar’s military rank and his array of postings to ‘hot spots’, it is perhaps fair to assume that, if he was affiliated to any intelligence service, then his real master most probably was the GRU (Military Intelligence Directorate), rather than the KGB. Military journalist or military/naval attaché was the usual cover of a GRU officer. GRU personnel used any opportunity to distance themselves from the political spying activities of the KGB, and regarded themselves as a high-level professional intelligence community. They were indeed more liberal-

---

3 The author knew Yegor Gaidar from his student days at Moscow State University, and throughout the 1970s–1980s in the circles of professional economists in Moscow. He had a lengthy conversation with Gaidar during the latter’s visit to Australia in 1996.
minded and more outspoken. Reinforcing the understanding of Yegor Gaidar’s upbringing within the elite Soviet establishment, this is also an indicator of relatively non-orthodox critical thinking within his family circle.

As a child Yegor travelled to Cuba, where his father was posted, and was living there during the Missile Crisis of 1962. He later wrote of the ‘revolutionary and enthusiastic’ Havana of that time (Gaidar 2000). Like most Russian children of his generation, he believed the Soviet Union to be the best country in the world. That is why he was surprised when he arrived in Yugoslavia in 1966, on another of his father’s postings, to discover a socialist country with a market economy, open public discourse and fully stocked shops. That was the time Gaidar began to follow economic news, and to ask questions about Yugoslavian economic reforms. In addition, his father, who noticed his son’s aptitude with numbers, entrusted the 10-year-old with keeping monthly business expenditure reports. That was the beginning of Gaidar the economist.

As for many youngsters of that time, Gaidar’s comfortable childhood in a privileged family and a beautiful country was broken in August 1968 when Warsaw Pact countries, headed by the Soviet Union, invaded Czechoslovakia to suppress the Prague Spring movement. This caused Gaidar to question what kind of truth it was that needed to be imposed on people with the help of tanks (Gaidar 2000).

The Scholar

The generation of young Russian intellectuals of the late 1960s — who had to grapple with the Yugoslavian deviation from the Soviet model, Euro-socialism, the Prague Spring and, finally, radical economic reform in Hungary — did not lose their belief in ‘socialism proper’. Not surprisingly, they were initially seeking answers through comprehensive studies of the Marxist literature. This was also facilitated by the liberal and creative atmosphere — despite strict ideological control — in the best educational institutions in the capital and major provincial cities. Paradoxically, the elite layer of the educational system, established mostly for the children of the privileged members of Soviet society, created the generation of people who ultimately rejected the very principles of the Soviet system.

Yegor Gaidar graduated from the Economics Faculty of Moscow State University, which was by all accounts the best school of economics and one of the sources of liberal economic thought in the Soviet Russia of the 1960s–1980s. This was only possible because the controversial and eclectic structure of the curriculum included both orthodox Marxian theory, as interpreted by Soviet ideological apparatus, and contemporary Western economics. Along with studying —
almost by heart — the works of Marx, Engels, Lenin, and the Communist Party of the Soviet Union, students were exposed to a variety of courses in modern conventional economics screened by titles that were inoffensive to ‘ideological controllers’. Thus, under the title ‘Critics of Western Economic Thought’, a decent programme in contemporary economic thought was taught; under ‘Mathematical Economics’, there was contemporary neoclassical microeconomics; under ‘Advanced Macro-econometrics’ there was contemporary monetary economics; ‘Socialist Planning and Forecasting’ taught the principles of macroeconomics, and furthermore, why the Soviet planning system did not work. In the early 1970s Professor Shatalin, who lectured on the latter course (and who later became the mentor of both Gaidar and the author of this article), dared to explain to his students behind closed doors that the Soviet economy was about to collapse, and the only reason it had not done so already was the revenue it received from its oil exports.

Meanwhile, even the most liberal economic tendency in Russia originated from the ‘intelligentsia’ educated within the Marxian ideological environment and for this reason tried to explain the failure of the Soviet system and find a way to improve it, initially, within the Marxian paradigm. That is why the concepts of the socialist economy or socialist planning were essential for the liberal economic discourse. Furthermore, Gaidar and his fellow young reformers were educated in the ‘Marxist’ belief that it was possible to reform the economy by revolutionary socio-economic engineering. In fact, the community of progressive economists was implicitly divided between two major schools of thought (both of them, of course, considered to be ‘heretical’ by the orthodox Marxists). One, founded by the prominent mathematician and Nobel Prize winner Leonid Kantorovich, was essentially socio-economic engineering on account of both its philosophy and its proposed technology of planning. It believed that a large-scale optimal-planning model could be seen as a prototype of an optimal national plan; and the price variables of that model as the prototypes of regulated prices. Therefore, the national planning process could be formalised as the optimal solution of a large-scale optimisation problem, making a centrally planned economy as efficient as a market economy (Kantorovich 1959). As utopian as it may sound, this approach nevertheless opened the door to studies in Western microeconomic theory, and so to an understanding of the infeasibility of optimal planning. And so the way was opened to the second, alternative school of thought, which favoured the introduction (and, once again, socio-economic engineering) of some elements of a market mechanism, without dismantling its socialist foundations. And this is where Gaidar, like many of his fellow students, started to ask himself further difficult questions about the feasibility of market socialism.

For Gaidar, those questions were even more difficult and important because of his experience with Yugoslavia; how to ensure that firms under the management
of workers would create new jobs; how to ensure that capital would be appropriately re-distributed to new, more efficient, ventures; how to solve the problem of sharply differentiated wages; and whether it would be possible to avoid accelerating the unemployment and inflation that had been endemic to Yugoslavia for decades. Though Gaidar was fluent in English, Spanish and Slovenian, he was seemingly unaware of the theory of the labour-managed firm, well known among Western economists. Such firms maximise profit per employee, but not the volume of profit. Consequently, if they prevail in an economy, as was the case in Yugoslavia, they restrict employment, accelerate inflation, and also are a major impediment to economic growth (Ward 1958). Despite being unaware of this gloomy theoretical prediction, Gaidar came to the general conclusion that any form of socialism was imminently destined to fail: while state socialism was a base for the absolute power of bureaucracy, market socialism obviously demonstrated its inefficiency (Gaidar 2000: Chapter 2). Paradoxically, however, as a scholar Gaidar never broke with Marxian methodology. Even during his time after politics as one of the greatest reformers of his era, who rejected the Marxian criticism of capitalism and adopted the transition to a market economy for his country, he maintained the Marxian deterministic explanation of the genesis of economic systems by the development of productive forces, despite the recent historic experience and contemporary literature (Shlyapentokh 2005).

The Reformer

As a young university graduate, Gaidar was fortunate to be able to start earning a living as a full-time researcher at the Institute of System Research — an elite think-tank in Moscow — in the area that interested him most. He continued to research into liberalising economic reforms in the socialist countries. However, it would be incorrect to think that from a young age Gaidar was developing into a dissident, anticommunist, pro-market economist. On the contrary, for a decade and a half he was an establishment insider and a member of the Communist Party until its dissolution in August 1991. Furthermore, during the time of Gorbachev’s liberalisation, Gaidar was working for the party’s official daily, Pravda, and for the thrice-weekly Communist.

Together with many younger and progressive members of the Party, Gaidar embraced the changes that became possible during Gorbachev’s time, when radical economic reforms moved out of liberal academic discussion circles to be placed on the agenda of the government of the day. He found himself in the right place at the right time when a commission was established, under the umbrella of the Politburo of the central Committee of the Communist Party, to look at how to improve the management of the economy. The scientific support of the Commission was entrusted to the Institute of System Research and, more
specifically, to the department Gaidar was working with. Together with his colleagues, Gaidar became not just a witness to, but instrumental in facilitating, the gradual introduction into the official language of such terms as ‘radical economic reform’, ‘market’, ‘individual entrepreneurship’, ‘cooperative’, ‘joint venture’, as well as ‘unemployment’, ‘inflation’, ‘poverty’, ‘social stratification’ and ‘budget deficit’.

At the same time, the practical official decision-making was incompatible with economic reforms, of which the acceleration of capital investments and growth were declared to be the key strategies. Gaidar was trying hard to promote the idea that liberal reform and financial stabilisation — rather than growth or ambitious new investment projects that would lead to further destabilisation — should be the first priority. He wrote one submission after another and met Gorbachev several times. Unfortunately, while Gorbachev was looking favourably at numerous proposals from liberal academics, he was also under ideological pressure from his conservative Party apparatus, and yielded to compromised, unworkable solutions. This led to an inevitable disaster.

By the late 1980s the Soviet Union was facing deep economic crisis and collapse. For quite a long time before that the economy had been able to survive primarily through exporting energy resources. However, world prices for crude oil, which had been sharply increasing in the late 1970s, fell substantially between 1981 and 1987. This created the threat of a total deficit of final consumer goods, as Russia was largely dependent on energy exports for importing some basic consumer needs, such as food.

Moreover, Russia’s situation could not be salvaged by its gold reserves. The communist regime had inherited about 1300 tons of gold from the pre-revolutionary Russian government. In spite of massive investments in the production of gold, the first post-communist government of 1991 found itself the inheritor of just 286 tons. Nor did the US$63 billion of Western credit help. But it did triple the level of foreign debt. Starting from the second half of 1991, grocery shops became empty. Even the queues disappeared — there was nothing to queue for. Famine seemed inevitable.

The Politician

This was the situation in the country when Gaidar embarked on his short but historically influential career as a top public servant. In 1991 he joined Yeltsin’s government, as First Vice-Premier of the Russian Government and Minister of

4 Here and below, unless otherwise stated, the data or estimates are from, or are based on, the Goskomstat website: www.gks.ru.
Economics from 1991 until 1992, and Minister of Finance from February–April 1992. He became Acting Prime Minister for a few months in 1992. However, the position was not confirmed by the anti-reform majority of the Congress of People’s Deputies, and that was the beginning of the end of his career as a politician. Subsequently, he was active in advising the government, was a member of the right-wing political movement, and, for some time, was a member of the Duma. Mostly, though, he concentrated on academic work as the founding director of the Institute for the Economy in Transition and was a prolific writer on the theory of reform and Russia’s economic history.

Gaidar’s ‘failure’ to become a career politician was perhaps attributable to the fact that he was not a politician by nature. From his student years he was in many ways a typical scholar. Though he was softly spoken and quite introverted, he was confident in his beliefs. He could impress those who valued knowledge and logic but he lacked the charisma required to convince a broader audience. He spoke the language of economics and political science but could rarely translate it into slogans and spin. He was definitely an academic leader, but not a leader of a mass political movement that would be able to embrace the majority of the nation. Though he was born into and served the official establishment, he was a stranger to it, and was rejected by it for his incompatibility with either the prevailing system or prevailing persuasion of the day.

The Transition

In order to understand the much-debated issue of Gaidar’s role in the revolutionary economic transformation of early 1990s, let us have a look at the facts. The imminent catastrophe of 1992 could be addressed in only one of two possible ways — by introducing traditional rationing supported by the introduction of martial law, or by radically liberalising the economy. Neither of the options was absolutely free of risk of mass unrest and the loss of human life. Gaidar chose liberalisation. He considered this a lesser risk, and the only way to prevent civil war, riots and the possibility of a return to Stalinism. Indeed, his choice meant that the country was saved from collapse, but Gaidar never received the praise that was his due from the majority of the Russian people. At the same time, Russia was plunged into an inevitable transitional depression, and Gaidar has been blamed for this ever since.

Meanwhile, even among Gaidar’s strongest supporters, the accent on his polices and achievements is somewhat biased. Firstly, he is both most praised and most blamed for price liberalisation. However, this measure represented the consensus of the day among all political forces, including the communists, and was approved by an absolute majority at the Congress of People’s Deputies, in October 1991.
Prices would have been liberalised anyway, with or without Gaidar. Secondly, he is blamed for the devaluation that reduced household savings to almost non-existence. Indeed, while price liberalisation contributed to the devaluation, the fact of the matter is that the most considerable devaluation (by 150 per cent) occurred in 1991 — before Gaidar joined the Government.

At the same time, it is almost forgotten that Gaidar rebuilt the country’s fiscal system from scratch — replacing the collapsed Soviet-era system — a contribution which Gaidar himself saw as his government’s major achievement. Within the space of a few weeks, the new package of legislation was drafted and approved by the legislature, and the budget was saved from imminent collapse. However, paradoxically, the right-wing liberal economist Gaidar is now criticised for his soft budgetary policy. Gaidar himself explained that too radical spending cuts would undermine already fragile social stability.

Gaidar is blamed for ‘shock therapy’ that caused a deep plunge in the Russian economy. Deregulation of most prices, as well as sharp measures towards opening the economy and the convertibility of the Russian currency, allowed for rapid elimination of shortages and queues. At the same time, the economy experienced severe post-communist stagflation. In the period 1992–96, GDP decreased by 40 per cent and industrial production by 50 per cent. The price level increased up to 26 times, and increases in nominal wages were much slower than increases in the CPI. This caused a 26 per cent drop in real income, which was accompanied by a considerable stratification within society (Kazakevitch and Smyth 2005).

In fact, Gaidar had time only to save the country from collapse by imposing his macro-financial shock, but was not given the chance or time to undertake the subsequent therapy by restructuring the economy at the micro level. The subsequent privatisation was designed on principles that Gaidar did not entirely agree with, and was conducted mostly after his time in government. It led to an ownership structure whereby at the end of 1990s 70–75 per cent of capital was owned by a mixture of private and government entities. However, only 30 per cent was owned purely by private persons or independent corporate institutions; the government had total ownership of up to 25 per cent of capital assets and part ownership of another 45 per cent (Kazakevitch and Smyth 2005). Under Putin, government control of the key revenue-earning sectors was consolidated much further.

In order to appreciate how radical the reforms have been and how the market has been created, the nature and structure of the privatised assets in the industrial sector should be noted. The pre-reform Soviet Union used to be an industrialised nation with a rather well-developed production apparatus. This apparatus mostly consisted of large and inefficient industrial giants, which
had been established on the orders of the strictly centralised planning system with the aim of serving the needs of the military machine rather than the final consumer.

The corporatisation and privatisation of large industrial enterprises alone could not solve the problem of efficiency and competition. A considerable physical restructuring of assets was necessary to meet the demands of the emerging market. The large government share of ownership, together with the lack of both domestic and foreign investors’ trust in political stability and the future of the market economy in Russia, did not help in efforts to make the manufacturing sector competitive in the domestic or international markets. The level of foreign investment today remains disproportionately low, and the outflow of domestic finances abroad continues. The Russian trade balance and the federal budget continues to rely on the export of fuel and some other mineral resources, and, as such, are strongly dependent on fluctuating world oil and gas prices.

In conclusion, almost 20 years after Gaidar’s reform, non-democratic Russia is once again faced with an inefficient economy and an uncertain future. This is in spite of Gaidar’s endeavours, not the result of them.

References


**REVIEWED BY SELWYN CORNISH**¹

The study of economics in Australian universities commenced when the universities of Sydney and Melbourne were established in the early 1850s. These two books trace the history of the discipline at the two universities from its beginning to the present. Each account is authoritative and comprehensive, drawing upon extensive historical research. Each contradicts the common assertion that contemporary economists have little or no interest in the history of their discipline.

While the two institutions have their own particular stories to tell, the broad outline of the development of economics at Sydney and Melbourne is similar. For the first 50 or 60 years of their existence, courses in political economy were taught ephemerally at both universities. It was not until after the First World War that separate faculties were established. In Sydney, a Faculty of Economics was created in 1920, and in 1924 a Faculty of Commerce was formed at the University of Melbourne. At Melbourne, one of the professors at the foundation of the university in 1853 — W. E. Hearn — was appointed to a chair of History and Political Economy. However, J. S. Elkington, Hearn’s successor, was more interested in history, and as a result economics tended to languish. Even so, two prominent Melbourne graduates before the First World War later held chairs in economics at Australian universities: R. C. Mills (Dean of the Faculty of Economics at Sydney from 1923 to 1945) and A. G. B. Fisher (the successor to Edward Shann as Professor of Economics at the University of Western Australia in the 1930s).

---

¹ The Australian National University, Selwyn.Cornish@anu.edu.au.
The decisive breakthrough at Melbourne came in 1924 when Douglas Copland was appointed Professor of Commerce and Dean of the university’s new Faculty of Commerce. Perhaps no one did more than Copland to set the study of economics on the map at Australian universities. In the early 1920s he had been instrumental in expanding economics at the University of Tasmania, persuading the authorities to create a chair in economics and a Faculty of Commerce. Later, in 1948, he was to become the ANU’s inaugural Vice-Chancellor and was responsible for establishing economics as one of the university’s strongest disciplines. It was Copland, too, who encouraged Australia’s greatest economist — L. F. Giblin — to apply for the Ritchie Research Chair in Economics at Melbourne. Copland was also responsible for the appointment of Trevor Swan to the first chair in economics at the ANU. Ross Williams, the editor of the Melbourne book, rightly states that:

Economics was fortunate that the university appointed Douglas Berry Copland to the foundation chair of commerce and he in turn encouraged Lyndhurst Falkiner Giblin to take the Ritchie research chair in economics. These two eminent economists created the ‘Melbourne School’ over the period 1925–39 and had a profound influence on Australian economic policy.’ (p.6)

At the University of Sydney, which had accepted its first students in 1852, the Professor of Classics (John Woolley) and, later, the Professor of Philosophy (Francis Anderson), presented occasional classes in economics. The foundation Professor of Mathematics, Morris Birbeck Pell, took classes in economics outside the university, as did a later Professor of Classics, Walter Scott. By the early 1900s business organisations, chambers of commerce, professional associations of bankers and accountants and the public service in Sydney were successful in calling for economics to be taught over three years at the university. In 1912, a Department of Economics was established, to which R. F. Irvine, a graduate of Canterbury University College in New Zealand, was appointed professor (Copland was also a graduate of Canterbury — both he and Irvine had been pupils of the remarkable James Hight). The Faculty of Economics at Sydney, the subject of Groenewegen’s book, was created in 1920 with Irvine as Dean. R. C Mills replaced Irvine on his resignation in 1922.

Under the leadership of Copland in Melbourne and Mills in Sydney, economics as a subject for university study took off in Australia. Student numbers rose considerably during the interwar years and the number of courses offered in economics increased. Staff appointments expanded; a professional association of economists — the Economic Society of Australia and New Zealand (now the Economic Society of Australia) — was founded on Copland’s initiative following his move to Melbourne from Hobart. A journal — the Economic Record — was established, with Copland as its Editor-in-Chief. In the 1930s both Copland
and Mills — like many of their economist colleagues — became advisers to government and were to occupy key positions in Canberra during the Second World War.

Each of the books highlights the explosive growth in the number of students studying economics after the Second World War. Staff numbers and course offerings grew accordingly. Opportunities for postgraduate studies also began to expand after the war, and there was a desperate search for new accommodation to meet staff and teaching needs. Multiple departments and additional chairs were created in the two faculties. In the 1970s the introduction of more democratic processes is highlighted: elected Deans, the creation of faculty advisory committees, and more representative selection committees. The reaction against these democratic processes in the 1980s and 1990s is similarly noted and with it the growing adoption of North American academic structures and management systems. There was greater emphasis on research capacity and achievement in the appointment and promotion of staff; on the acquisition of outside research funding; and on the appointment of staff with postgraduate qualifications acquired from overseas, particularly in the United States. More rigorous approaches to teaching are also documented. A premium began to be placed on theoretical work, rather than upon policy analysis and applied economics. Economic history and the history of economic thought — which had once been major fields of teaching and research — were downgraded but not entirely eliminated (though separate Departments of Economic History were abolished). On the other hand, business studies, including accounting, management, marketing and finance, began to challenge economics for pre-eminence in the faculties of both institutions, and the viability of economics departments became heavily dependent on compulsory first-year economics courses taken by students enrolled in business or commerce degrees. Income from these large first-year economics classes heavily subsidised smaller enrolments in optional economics subjects in later years.

These are common themes. Yet the two books differ in important ways. The Sydney story, for example, is written by a single author, the distinguished historian of economic thought, Emeritus Professor Peter Groenewegen, a member of the Faculty of Economics at Sydney for more than 50 years. The Melbourne book, in contrast, was composed by multiple authors. Its 10 chapters were written by nine different authors, all of whom are, or were, members of the Department of Economics. One chapter was written by John Freebairn, who discusses the contributions made by Melbourne staff to applied economics and economic policy. Another chapter, written by Robert Dixon, examines the contributions of staff to economic theory. The other authors of the Melbourne book are Ross Williams, the editor (who writes an introductory chapter on ‘Trends and Cycles: 1925–2008’); Marjorie Harper (who covers economics at
Melbourne from 1855 to 1944); Joe Isaac writes on the 1940s and 1950s; Peter Drake on the 1960s; Neville Norman on the 1970s; Peter Lloyd on the 1980s; and Jeff Borland on the 1990s to the present.

While both books devote considerable space to members of staff — their appointments and departures, teaching, research and external contributions (including policy advice, consultancies and public addresses) — Groenewegen has much more to say than the Melbourne authors about students and alumni, their academic results and post-graduation careers. Here, however, it must be said that, while it may be of interest to read of the impressive achievements of many of the Sydney graduates, it becomes rather fatiguing. For example, from page 195 to page 209 there is a long section headed ‘Graduates and Students, 1985–99’. We are told how many students each year completed undergraduate and graduate degree courses; there is information about the class of honours achieved and medals won; the names of the more distinguished alumni are rehearsed and details are provided about their various careers after graduation. The amount of detail is certainly impressive, and the information will no doubt be convenient for university administrators, but readers may find it difficult to sustain their concentration. Here it might have been better to have included the material in an appendix.

How do Melbourne and Sydney compare in the economics league tables? In his introductory chapter, Ross Williams notes that in ‘a worldwide survey of heads of economics departments in 2006, Melbourne was ranked first in Australia, using international academic standing as the criterion’. This may have been so by 2006 but it was not always the case. The Political Economy dispute at Sydney, which Groenewegen discusses objectively, created great turmoil in the Faculty of Economics and its reputation was tarnished. While this was happening, Melbourne was making a number of decisive appointments, which led to a much-improved research output, especially in economic theory. Postgraduate studies based on coursework were expanded and well-endowed programmes were introduced for the purpose of attracting top international visitors, including Nobel Laureates, to Melbourne. For all these reasons, it seems that Melbourne moved decisively ahead of Sydney.

Is there anything of great significance that the authors have missed? More space might have been devoted to outstanding research conducted by particular members of academic staff. One of the highlights of the Sydney book, for example, is the discussion of the work and achievements of Ray Chambers, the foundation Professor of Accounting. A similar treatment might have been given to the work of Mills, especially his contributions as a member of the 1936–37 Royal Commission on the Monetary and Banking Systems of Australia and his path-breaking work as Chairman of the Committee on Uniform Taxation in 1942. Similarly, an assessment of the academic work of Syd Butlin and Ronald Walker
would have been illuminating. In the case of the Melbourne book, Copland’s research and his outstanding contributions to policy discussion in the late 1920s and early 1930s might have been assessed in more detail. Also, not much is said about the work of the various incumbents of the Ritchie chair, particularly of the research contributions of Dick Downing, the longest occupant of the chair.

But these criticisms are mere peccadilloes and do not detract from the high quality of the books. In fact, the authors deserve the gratitude of the Australian economics community for the light they have shed on the development of the discipline in this country. The scholarship is impressive, particularly the immense research that has gone into the writing of the books, based as they are upon exhaustive archival investigation, and interviews and correspondence with former staff and students. The books will become essential reading for those who wish to understand the way economics as a university subject developed in Australia over the past century and a half.
Interest in the activities of central banks declined after the Second World War when priority was assigned to fiscal policy, and direct controls were imposed upon the financial system. But as a result of the Great Inflation of the 1970s and 1980s and the deregulation of the financial system there occurred a revival in monetary policy and a renewed interest in central banking. These books will add to the current interest in central banks.

The books differ considerably in their coverage and structure. The first encompasses the history of central banking in Britain and in the United States, while the other concentrates on the US Federal Reserve. The first covers more than three centuries of central-bank history, while the other is confined largely to the second half of the twentieth century. One author is an academic with some experience in the world of central banking; the other was a central banker for most of his career. And yet both authors follow a somewhat similar approach: each relies, for example, on the method of historical narrative to draw conclusions about the nature of central banking, though one of them uses extant literature as his principal source of information while the other draws largely on personal experience. A particular highlight of each of the books is the significance attached to the role of individual central bankers in the formation of monetary policy. Economists — though usually not economic historians or historians of economic thought — often dismiss the importance of the human dimension in the policy process. The authors of these books do not fall into that trap.

John Wood traces the history of central banking in Britain since the formation of the Bank of England at the end of the seventeenth century, and in the United States from the late eighteenth century when the first Bank of the United States was chartered. The origins of the Bank of England are identified in the government’s need to raise revenue to fund military expenditure. The Bank’s
progress is then tracked as it sought to earn profits and remain solvent in its early years; through the suspension of gold convertibility in the 1790s and the restoration of gold payments after the Napoleonic Wars; living with the Bank Charter Act of 1844; the adoption of lender-of-last-resort and financial-stability responsibilities in the second half of the nineteenth century; the return to gold in 1925; and the Great Depression of the 1930s. The book concludes with the Great Inflation of the second half of the twentieth century and the Bank’s independence in 1998.

The origins of the first and second Banks of the United States are similarly attributed to the government’s need to secure revenue to meet its various commitments. Wood then traces the emergence of lender-of-last-resort facilities by the United States Treasury, the Suffolk Bank and the New York Clearing House following the demise of the second Bank of the United States in the 1830s. He goes on to map out the crash of 1907 and the establishment of the Federal Reserve system in 1913; the boom and crash of the 1920s–early 1930s; the erosion of the Fed’s independence in the 1940s and its restoration in the 1950s; and the Fed’s role in helping to create and then combat the Great Inflation of the 1960s and 1970s.

Wood relies heavily on secondary sources, though there are occasional references to primary material. Given the amount of detail involved, the author does well to maintain control of the discussion, which is generally lucid and unpretentious. It is true that the book breaks little new ground. But that is not its aim. Rather, its objective is to provide essential historical information for the purpose of drawing general conclusions about the nature of central banking in two major countries over long periods of time. The book will provide an excellent text for courses in the history of monetary policy and central banking, and a reference for information about the development of central banking in Britain and the United States.

The author of Inside the Fed, Stephen Axilrod, is a former head of the research and monetary affairs divisions of the Federal Reserve Board in Washington D.C. and secretary of the Federal Open Market Committee. He began his career there in the 1950s and served under four chairmen — William McChesney Martin, Arthur Burns, William Miller and Paul Volcker — before retiring in 1986. The book began as a long essay on a number of diverse topics, the author’s aim being to highlight the significant events of his career for the benefit of his family and friends. One of the topics was an extended meditation on the achievements of Arthur Burns (chairman of the Fed from 1970 to 1978) and Paul Volcker (chairman from 1979 to 1987), and the policy disputes at the time over how to deal with the Great Inflation. The essay was subsequently extended to incorporate Bill Martin’s long tenure as chairman (1951–1970) and Alan Greenspan’s chairmanship (1987–2006), with some brief comments on Bill
Miller’s term (1978–1979), the early Bernanke years and the genesis of the Global Financial Crisis (GFC) of 2007–09. Axilrod has written a fine book, based on recollections of his interactions with Fed chairmen and colleagues with whom he worked. It is beautifully written, though readers should be warned that it is easy to be carried away by the quality of the writing and miss the essence of the discussion.

What are some of the conclusions reached by the authors of the two books? Wood contends that central bankers and economists see the world differently. ‘It has always been clear to me’, he says, ‘at first very disappointingly, then less so — that central bankers do not see the world like economists.’ (p.xv) Central bankers, he asserts, have at critical times given excessive ‘attention to the financial markets at the expense of their macroeconomic responsibilities’ (p.4). He endorses the view, once expressed by Allan Meltzer, that the Federal Reserve’s ‘knowledge of the monetary process is woefully inadequate… dominated by extremely short-run, week-to-week, day-to-day, or hour-to-hour events in the money and credit markets. [T]heir viewpoint is frequently that of a banker rather than that of a regulating authority for the monetary system and the economy’ (p.4). Writing before the GFC, Wood highlights the consensus that had been reached in the approach to monetary policy at the end of the twentieth century:

Our monetary system is unprecedented. After decades of instability, central bankers, governments, and economists have reached a consensus that the appropriate role of a central bank in the prevailing fiat-money regime includes: (1) the clear assignment of the responsibility for inflation to the central bank; (2) agreement that inflation should be low and stable; (3) rejection of price controls as a means of controlling inflation; and (4) acceptance of whatever degree of fluctuation is required in interest rates to achieve the inflation objective. This is at once more ambitious and more modest (realistic) than earlier systems. (p.1)

Naturally, one wonders whether the consensus can be retained in the light of recent economic and financial events.

Perhaps the major conclusion reached by Axilrod is ‘the important role in policy played not by pure economic reasoning or understanding, but by personalities and their responses to the political, social and bureaucratic contexts in which they find themselves.’ His experience at the Fed, he states, is:

…that a great leader for monetary policy is differentiated not especially by economic sophistication, but by his or her ability to perceive when
social and political limits can be pushed to make space for a significant, paradigmatic change in the approach to policy should it be required, as well as by the courage and bureaucratic moxie to pull it off. (p.4)

He goes on:

Fed policymakers, being very well aware that they are part of a government established to be democratically representative of the people, are themselves likely to be constrained in the policies that they find it practical to consider by their sense of what is tolerable to the country...I am convinced that such judgments, or perhaps such feelings, whether expressed ... or recognized, lie deep within the individual policymaker's gut. The policymakers are independent, but they are making decisions from within the government and within what they perceive to be certain societal bounds. (p.10)

Axilrod ranks the four chairmen under whom he served, together with Greenspan and Bernanke, according to their performance as heads of the Federal Reserve System. Volcker comes out clearly on top, followed by Martin and then Greenspan. Burns and Miller are at the bottom. He admits that it is too early to assess Bernanke's performance. He is not impressed with Bernanke's (and Greenspan's) failure to raise interest rates quickly enough to dampen subprime borrowing, or with Bernanke's subsequent failure to adopt expansionary measures rapidly enough to limit the impact of the financial crisis. On the other hand, he is impressed with the scale of the response once the measures were introduced.

Of the Volcker period, which saw the end of the Great Inflation in the United States, Axilrod writes that ‘It was an exciting period...and...it can also be called a glamorous time. ...It was...one of the few instances in my opinion where a dramatic shift in policy approach could be attributed to a particular person's presence rather than mainly to or just to circumstances.’ Volcker, he judges, was:

the essential man for a combination of reasons. He combined great sensitivity to shifting trends in political economy (he could see what the country would now accept) with a willingness to take dramatic action. Moreover, he was technically very competent in the nuts and bolts of monetary policy, which made it much easier for the F[ederal] O[pen] M[arket] C[ommittee] and the chairman himself to feel confident that the new approach, although not risk free, had a reasonably good chance of working. (p.91)

In short, ‘Volcker the actor was in full display. He was totally in command of himself and the subject matter. He spoke with force and conviction. He responded to questions from Congress and the public with certainty...’ (pp.92–3)
Philipppe Legrain, *Immigrants: Your Country Needs Them*  
(Little Brown Book Group, 2006)

REVIEWED BY MATTHEW POLLOCK¹

The central premise of this book is that the economic, political and social benefits of immigration from developing countries to developed countries outweigh the potential costs. The author, Philippe Legrain, makes a timely case for immigration in Britain, with the recent terrorist attacks in London a not-too-distant memory and labour disputes concerning the Lindsey Oil Refinery on the front page of the newspapers at the time of publication. ²

If, however, you were expecting a book purely on the economics of immigration you would be disappointed, as Legrain is determined to put a human face on his argument: ‘Whenever people talk in the abstract about the pros and cons of immigration, one should not forget that immigrants are individual human beings’ (p.8). This is a refreshing change from the economic norm of referring to individuals as either units of consumption or labour-market participants, and allows the book to appeal to a wider audience. The idea of humanising the issue is systemic and embodied in the text through a series of exampled individual personal experiences. These experiences vary dramatically, from Lasso Kourouma, a former car salesman from Cote d’Ivoire, who was caught and detained trying to enter Spain by boat from North Africa; to Michelle Leboeuf, a field agent for the US Border Patrol from El Paso, Texas; to Leonid Dinevich, a former Soviet General and specialist in atmospheric physics who immigrated to Israel under the guise of Israel’s policy of ‘right to return’ for all of Jewish descent.

Legrain is very clear on his assumption that immigrant labour under a free migration programme would be in the majority, constituting a complementary labour input to the existing domestic labour market. This contention is repeated in the text a number of times and is backed up by as many empirical examples on the composition of immigrants as complements: ‘Many come to service the clusters of global professionals and other residents in places like London

---

1 University of Canberra, matthewpollock111@gmail.com.
2 Editor’s note: In this dispute of 2009 local contractors of the Lindsey Oil Refinery went on strike after the employment of several hundred Italian and Portuguese contractors.
and Silicon Valley: preparing their lunchtime sandwiches, chauffeuring them around, cleaning their houses, etc.’ (p.16). The view of immigrants as substitutes is raised but not discussed in any detail. Rather, it is brushed aside quite conveniently as a non-issue: ‘If immigrants were the same as us, we would have little to gain or lose from letting them in’ (p.67). The second assumption put on the labour market is that it is self-perpetuating in that ‘each person creates work for others, so the more people there are, the more work needs doing’ (p.66). These assumptions are implied to be universal.

Current debate on the subject is rather narrow in its conceptualisation, usually conceived as a ‘cross-border’ perspective. However, as Legrain rightly points out, migration is conducted on a global scale: ‘looking at international migration solely from a national perspective gives a distorted and partial view of what is happening’ because migration is ‘part and parcel of globalisation’ (p.13). Legrain takes this to imply that ‘the case for freer migration follows on logically from the case for freer trade’ (p.20), and poses the question ‘Why can computers be imported from China duty-free but Chinese people not freely come to make computers here?’ (p.18). Further distinction is made in the discrimination between the free movement of professional people from developed nations and the restrictions placed on the unskilled from developing nations: ‘Why is the door open for American managers to run factories in Honduras but the door slammed shut for Hondurans who want to work in American factories?’ (p.18)

Once again, however, arguments in opposition are conveniently brushed aside. Discussion on free trade almost universally accepts the notion of an optimum tariff and the existence of benefits coming from theories of strategic trade. Further, practices such as predatory dumping — a process whereby a foreign firm prices its product below the market rate to drive out domestic competition — are recognised as potentially damaging to the domestic economy. If the case for freer migration truly follows on logically from the case for free trade, then surely similar theories must apply and affect labour markets in much the same way. Legrain does discuss the effect of freer migration on domestic labour markets in terms of unemployment (indeed, an entire chapter is dedicated to the topic). However, the issue is not just unemployment; it is about wage stability and its effect on consumption spending and the promotion of a fair, competitive and efficient labour market for the existing citizens of the country. There is a case for the benefits of regulation in goods and capital markets, and surely this too ‘must follow on logically’ in the case for labour markets.

While the primary audience of the book is British and American, the argument is by no means lost crossing our borders. Australia is not ignored in the text. Indeed, Australia is identified as having the largest relative immigrant population in the developed world: ‘23.6% of the population was born overseas’ (p.56). Our points-based immigration programme gets several mentions, almost exclusively
in the negative, and is described as ‘not always making sense’ (p.107). This is most certainly true if the assumption of immigrant labour as complements holds. However, the argument against Australian immigration policy (p.109) really falls over when it is vilified in a comparison to the migration programme of Israel, almost to imply that a migration programme based exclusively on the discrimination of race and religion is freer than that of a points-based programme.

What emerges is a comprehensive overview of the contemporary issues in global migration that challenge conventional wisdom, and one that reveals a dynamic relationship between all variables and participants. The text tells a story of the benefits of free migration and finishes with the rallying cry: ‘Let Them In’.