

# Creative Distraction

## Issue IX



### *Contents*

- |    |  |
|----|--|
| 1  | Editor's Note  |
| 2  | Stern and critics on Discounting<br><i>Professor John Quiggin</i>  |
| 9  | The Phenomenon of Business Cycle Synchronisation<br><i>Danniella Tomcznski</i>                                 |
| 23 | The Wage Gender Gap: Are Women Catching Up?<br><i>Annie Nguyen</i>   |
| 33 | Review of <i>A Critique of the Private Health Insurance Regulations</i><br><i>Amy Wagner</i>                   |
| 36 | Review of Will Subsidising Private Health Insurance Help the Public Health System?<br><i>Madeline Veenstra</i> |

## ***Editor's note***

Welcome to the relaunched *Creative Distraction*! Now in its 9<sup>th</sup> issue, *CD* has taken a new focus, aiming to showcase the works of high achieving economics students. A showcase it is, with fierce competition for publication. For the coveted spot of Micro Policy, there were 10 articles submitted! Each new edition of *CD* will include numerous student pieces, as well as a feature by a member of staff. To kick off this new initiative, we have one of Australia's leading economists, Professor John Quiggin from UQ writing about the Stern Report of 2006.

Now published at the start of each semester, *CD* will put forward the top pieces from subjects in the preceding semester. It aims to provide you with interesting papers, and give depth to what you learn in your degree.

In other news, the new executive was elected at the end of last year, and they have been hard at work planning for what is to be an eventful year. The ESA executive for 2007 is:

President: Jonathan Pitchford

Vice President: Fay Chen

Secretary & *CD* Editor: Mitch Riley

Treasurer, website and IT: Ang Li

Social Convenors: Alice Heathcote and Jonathan Humphrey

Sport Convenor: Tom Gole

The new executive is looking forward to a bigger and better year in 2007, with more social events, a reinvigorated careers programme, the new *CD*, staff-student sport and a whole lot more! Make sure you check out the brand new website at [www.uqesa.com](http://www.uqesa.com). If you want to get involved, or have any questions about the society, feel free to email [esa@uqesa.com](mailto:esa@uqesa.com). If you missed us on market day, please email the above address and we will get you signed up. Pencil Tuesday 20 March into your diaries, for the premier event of the year, Careers and Cocktails – it's set to be a big one! The next edition of *CD* will call for submissions immediately after the end of semester 1, 2007 to be published in week 2 of semester 2, so start thinking about ideas. Until then, I hope we have provided you with some *Creative Distraction* from your studies!

Mitch Riley

[m.riley@uqesa.com](mailto:m.riley@uqesa.com)

## STERN AND THE CRITICS ON DISCOUNTING

JOHN QUIGGIN\*

The Stern Review of the economics of climate change has had a substantial impact on public debate over policy responses to climate change. This is in part a matter of timing. Although economists have debated the relevant issues for years, this debate has been overshadowed by debate over scientific issues, initially among climate scientists and then, for much of the last decade, between scientists and political critics of climate science, primarily associated with US-based think tanks. The operations of these groups in relation to climate science and other environmental issues, evolution and 'intelligent design' theory, stem cell research and other issues is discussed in detail in Mooney (2006)

By 2006, the credibility of the anti-science groups was on the verge of collapse, at least as regards climate change. Factors contributing to this outcome included the resolution of remaining scientific controversies, particularly relating to satellite measurements of tropospheric temperatures, the success of the documentary *An Inconvenient Truth* and increased publicity regarding the extent to which prominent skeptics were beneficiaries of funding from fossil fuel industries.<sup>1</sup>

The Stern Review radically changed the terms of the debate by presenting the issues in economic rather than scientific terms. The effects of global warming, previously discussed in qualitative terms were shown to correspond to large losses in economic welfare. The result was to outflank the remaining skeptics. They could either continue denying the results of scientific analysis, or try to salvage the fallback position, undermined by the Stern Review, that although global warming is real, the costs of doing anything significant about it exceed the benefits, at least in the short term. It is this latter position which will be critically examined in the present paper.

Some aspects of the Stern Review have been accepted with little controversy. Most significantly, the estimate that stabilising global CO<sub>2</sub> levels at 500 ppm would reduce world consumption by around 1 per cent has been generally accepted as reasonable, if towards the optimistic end of the range of plausible estimates. Although most serious economic analysis has produced cost estimates for climate stabilisation of 1 to 2 per cent of world consumption, much informal discussion has presupposed far larger estimates.

There has also been little discussion of the Stern Review's projections of climate change, which generally followed those of the IPCC or of estimates of the economic value of damage to natural ecosystems which broadly followed those of Nordhaus and Boyer (2000). Nordhaus and Boyer have been criticised for undervaluing damage to natural environments (Quiggin 2006) and it seems likely that a similar undervaluation will be evident in Stern.

The main focus of discussion of the Stern review has been the way in which future costs and benefits have been converted to present values through discounting. This review will focus on the same question. It begins with an outline of the expected utility model used in the Stern review, then examines 'pure' time preference before considering the Stern review and its critics.

### Expected utility

Stern, and nearly everyone else in the debate so far, uses a model based on expected utility theory. There are very strong reasons to go this way. First, expected utility has the property of dynamic consistency, which means that, if you make a plan, anticipating all possible contingencies, you will want to continue with that plan over time, whichever contingency arises. No other choice model has this property except under special conditions.

Second, expected utility theory allows a single utility function that simultaneously determines attitudes to intertemporal wealth transfers, interpersonal redistribution and risk reduction (transfers of income and consumption between states of nature). With the plausible technical assumption of constant relative risk aversion, (almost) everything is determined by a single parameter (called  $\eta$  in the Stern report), which measures the (reciprocal of the) elasticity of substitution for consumption.

The central idea of expected utility theory is that people are not concerned ultimately with money income

---

\* Australian Research Council Federation Fellow, School of Economics and School of Political Science and International Studies University of Queensland. I thank Nancy Wallace for helpful comments and criticisms. This research was supported by an Australian Research Council Federation Fellowship.

<sup>1</sup> Even more damaging was the disclosure that several prominent skeptics had earlier worked for the tobacco industry, challenging scientific analysis of the dangers of passive smoking.

but with the utility derived from consumption. The assumption of diminishing marginal utility means that a dollar of extra income or consumption is worth less if you're rich than if you're poor. So we derive the conclusions

- (i) assuming rising incomes, a dollar of extra income is worth less in the future than it is today
- (ii) under uncertainty, a dollar of extra income in a bad state of nature is worth more than a dollar in a good state of nature
- (iii) transferring income from rich people to poor people improves aggregate welfare

The first of these assumptions mean that in the presence of technological progress, allowing rising incomes and consumption, we would expect a positive discount rate. The second means that a risky asset (more precisely, a risky asset with returns that are correlated with aggregate consumption) should be worth less than a riskless asset. The third means that even when redistributive taxes and international aid are costly (for practical purposes, nearly always) they can improve welfare.

Assuming that the combination of the expected utility model and inherent discounting (discussed below) captures all the issues under consideration, the riskless discount rate is determined by a simple formula

$$r = \delta + \eta * g$$

where  $\eta$  is the reciprocal of the elasticity of substitution for consumption,  $g$  is the rate of growth of consumption per person, and  $\delta$  is the inherent discount rate. A similar, slightly more complex formula can be used to derive the rate of return for a risky asset, based on its correlation with aggregate consumption.

The parameter  $\eta$  represents the extent to which the marginal utility of consumption is reduced as consumption increases, and the choice of  $\eta$  is central to the debate over discounting. The larger is  $\eta$ , the stronger all of effects (i)-(iii) become. So high  $\eta$  means a high preference for current consumption, high aversion to risk and large benefits from redistribution.

Even economists familiar with the mathematical derivation of  $\eta$  often have problems understanding the implications of different choices of  $\eta$ , particularly when time, uncertainty and interpersonal redistribution interact. So it may be useful to consider a particular example, which happens to be that used in the Stern Review.

#### *The special case $\eta = 1$*

Expected utility theory was first developed to analyze choice under uncertainty. In discussions focused on risk, the most common single choice for  $\eta$  is  $\eta=1$ , corresponding to a logarithmic utility function. This is a particularly tractable function, and seems to fit the data reasonably well. It also gains popularity from tradition having been proposed as a utility function for money by Daniel Bernoulli back in 1738.

There is a natural way of interpreting  $\eta=1$ , that is, logarithmic utility, in the intertemporal context. With this specification (and ignoring inherent discounting as discussed below) one per cent of consumption now has the same value as one per cent of consumption at any time in the future. So, for example, a policy that reduced consumption (not the growth rate of consumption!) by one percentage point from 2000 to 2050, relative to some baseline, then increased consumption by one per cent relative to the same baseline until 2100, would come out exactly neutral. Logarithmic utility is implicit in much discussion of intergenerational equity, notably the intergenerational accounting analysis popularized by Kotlikoff and Summers (2004), which focuses on the proportion of income paid in tax by each generation.

At this point a numerical example might be useful. The world's mean income per person is currently around \$US7000, though the majority of people get much less and the billion or so in rich countries much more. Under the projections used in the Stern Review, average world income in 2100 is estimated at about \$US 100 000. Consumption is approximately equal to income, and in what follows, these numbers will be treated as consumption levels.

Using  $\eta = 1$ , a sacrifice of \$70 per person (1 per cent of consumption) today would be justified if (and only if) it increased the consumption of our great-grandchildren in 2100 by at least \$1 000. If this trade-off appears reasonable, then a value of  $\eta = 1$  is appropriate. If the future payoff required is higher (or lower) then so is the preferred value of  $\eta$ .

Intuition about the far future tends to be cloudy, so it is worth observing that, under expected utility theory, exactly the same arguments apply to redistribution within the current generation. To illustrate, it is useful to turn around the direction of redistribution. Consider a redistributive program that takes \$1000 from the well-off (in this example households with income of \$100 000 per person) and uses the proceeds to benefit the poor (those with \$7000 per person). (Alternatively to keep the focus on redistribution from the poor to the rich, suppose that such a program already exists, and consider scrapping it).

Such redistribution always involves a range of costs including administration, compliance, efforts at

avoidance and evasion and incentive costs. Suppose that, in a particular case, these costs amount to 93 cents per dollar initially taxed, so that for a net loss of \$1000 to the well-off, the net benefit to the poor is \$70. If such a program is exactly marginal, so that any program with a larger net benefit is acceptable, and any program with a smaller net benefit is unacceptable, then the implied social preferences have  $\eta = 1$ . If the minimum acceptable net benefit is larger (smaller) then we require  $\eta$  to be less than (greater than one).

The choice of  $\eta = 1$  also seems to give a good match for calibrated macroeconomic models designed to match growth and business cycle facts. On the other hand, studies of aggregate consumption seem to imply values of  $\eta$  close to zero, while studies of asset prices for equity seem to imply values well above one (as discussed below in relation to the equity premium puzzle)

### *Problems with expected utility*

The expected utility model is neat, logically compelling and tractable, but it suffers from two big problems.

First, at least some of the time, most people don't behave in a way that is consistent with the expected utility model. For example, people simultaneously gamble at unfavorable odds and take out costly insurance, which violates the predictions of EU with respect to uncertainty. Similarly people often apply a high discount rate to trades between the present and the near future, but a low discount rate for trades between the near and far future. This is called hyperbolic discounting. Large branches of modern economic theory, commonly referred to as generalized expected utility theory and behavioral economics, attempt to address this problem as discussed below.

The second problem is that observed market outcomes are not consistent with EU theory as it is commonly applied. This problem is partly because people don't act in accordance with EU and partly because markets don't work in the smooth and frictionless way assumed in standard finance-theory models.

The most important problem in this respect is the 'equity premium puzzle', and the closely-related 'risk-free rate puzzle'. The equity premium puzzle is that for plausible choices of  $\eta$ , the real bond rate should be somewhat higher than it is, and the rate of return to equity much lower.

Historically, real returns to investors from the purchases of U.S. government bonds have been estimated at one percent per year, while real returns from stock ("equity") in U.S. companies have been estimated at seven percent per year, a difference of six percentage points. By contrast, for reasonable choices of  $\eta$ , the difference should be no more than half a percentage point. The equity premium puzzle can be resolved by assuming very high values of  $\eta$  since risk aversion increases the premium. But high values of  $\eta$  imply a high discount rate, so the risk-free rate puzzle is made worse.

The inconsistencies between the EU model and observed choices and market outcomes mean that for any possible choice of parameters, it's possible to present hypothetical choices for which most people will reject the implications of the model, or to point to market outcomes inconsistent with the proposed parameters.

### **Inherent discounting and weighting**

Although expected utility provides a complete theory of allocation of consumption across individuals, time periods and states of nature, it is often supplemented by some sort of weighting scheme. This is true whether expected utility is used positively, to model actual behavior, or normatively as a guide to rational individual decision-making and ethical social-decision-making.

In the case of allocating consumption over time we need to consider whether we should discount future consumption simply because it is in the future, even with the same marginal utility (call this inherent discounting). Similar issues arise in relation to the weighting of individuals in income redistribution and states of nature in choice under uncertainty, and these will be considered first.

### *Probability weighting and uncertainty*

In decisions under uncertainty, individuals often seem to put more weight on low-probability extreme outcomes than would be implied by expected utility (Kahneman and Tversky 1979). Although various responses to this observation have been proposed, the most popular has been to use a rank-dependent weighting scheme, as proposed by Quiggin (1981, 1982, 1993) and incorporated in the cumulative prospect theory model of Kahneman and Tversky (1992).

Application of rank-dependent weighting schemes is controversial in a normative sense, since it leads to violations of dynamic consistency, except where decision-making starts from a particular point and holds relative weights fixed thereafter. On the other hand, if individuals make decisions using probability weights, it is necessary to incorporate this fact in public policy. For example, if individuals overweight the possibility of a nuclear power-plant accident, and take expensive precautions against it, the construction of the plant may reduce welfare, even if an expected-utility calculation, assuming that everyone followed expected-

utility, would yield a net benefit.

### *Interpersonal weighting*

The idea that all individuals' utility should have equal weight is clearly not valid as part of a descriptive model of individual behaviour and has been controversial as a normative basis for social welfare. However, the deviations from equal weighting tend to go in opposite directions.

As individuals, we almost invariably place more weight on our own welfare and that of our immediate family than on the welfare of friends, neighbours and more distant relations. In turn, we value the welfare of fellow-citizens in general less than that of those close to us, and the welfare of foreigners lower still.

As members of a political unit, we normally agree to give each member equal weight, since there is no basis on which one person or a small group could claim to have an inherently greater entitlement than the rest and have this claim accepted. However, critics of utilitarianism such as Rawls have argued that we can justify more redistribution to the worst-off than would be implied by expected utility theory, while maintaining neutrality in an *ex ante* sense.

It turns out that the implementation of this idea in a formal sense is identical to that of rank-dependent utility under uncertainty and was first put forward around the same time, by Weymark (1981). As Ebert (...) observes, using rank-dependent weightings allows for a reconciliation between Rawls and Bentham.

### *Inherent discounting*

One of the longest-running controversies in welfare economics has concerned the appropriateness of applying different weights to people in different generations, and, more generally of discounting future utility whoever receives it.<sup>2</sup> Ramsey (...), whose work is the starting point for formal analysis of intertemporal choices, rejected inherent discounting as ethically unjustified, and this viewpoint is shared by most philosophical advocates of utilitarianism. On the other hand, a good deal of evidence suggests that individuals tend to discount their own future consumption.

Before discussing inherent discounting, it is worth observing that standard expected utility suggests one reason for discounting future consumption; namely the possibility that we will not be around to enjoy it. As individuals, we face a typical annual mortality risk of around 1 per cent<sup>3</sup>, and it makes sense to discount future utility by this amount. But at least some of the time people (most notably teenagers) discount the future much more than this.

For society as a whole, there is a comparable risk arising from the possibility of nuclear annihilation, a killer meteor and so on. The risk need not involve a total extinction of the species; it is sufficient that the disaster be great enough that 'all bets are off' in terms of calculations about the future.

With this point addressed, there remains the question of whether we do and should, discount future utility. The evidence on individual behavior is far from clear. On the one hand, there is a lot of evidence to support the idea of 'hyperbolic discounting' ... However, this is offset by a notion of 'mental accounts'. Individuals may allocate resources between activities and follow inconsistent rules in different activities. For example, the same person may allocate money to an automatic saving scheme offering low or even negative real returns, while displaying hyperbolic discounting with respect to the remaining cash flow.

Leaving such phenomena to one side, the evidence for high inherent rates of discount is not strong. The most obvious market measure to use in assessing intertemporal tradeoffs is the real rate of interest on low-risk bonds (government or AAA corporate). This rate has generally been between one and two per cent and is currently around two per cent. Given that the rate of growth of average consumption per person is between one and two per cent, this is consistent with zero discounting and  $\eta=1$ .

Even if individuals do display inherent discounting, that does not necessarily mean that this is appropriate as a basis for social decisions. Future individuals presumably will not share the view that utility in our time is inherently more valuable than utility in theirs. In fact, as individuals, introspection and casual observation suggests that we generally regret decisions made in the past on the basis of inherent discounting. Such decisions represent selfishness on the part of our past selves at the expense of our current selves, analogous to individual selfishness with respect to others.

### *Implications for parametric choices*

Broadly speaking, high weighting on low-end outcomes and high values of  $\eta$  are substitutes. That is, the higher the weighting on low-end outcomes, the lower the value of  $\eta$  required to match given observations

<sup>2</sup> The similarity between this idea and probability weighting is discussed by Quiggin & Horowitz

<sup>3</sup> In the environment in which we evolved, the rate would have been higher. And an ev psych analysis suggests that what matters most is the probability of the end of reproductive life, which is higher still.

or intuitions. Consider the income redistribution example, and suppose that the poor are given a weight twice that of the rich. Then if the redistributive project in the example, taking 1 per cent of income from the rich, and raising the consumption of the poor by 1 per cent, is just marginally acceptable, the required value of  $\eta$  is not 1, as before, but 0.5.

Similarly, a high rate of inherent time discounting implies that to match any observed pattern of market rates, a lower value of  $\eta$  is required.

### Stern and the critics

The analysis in the Stern Review follows the general approach set out above. The value of  $\eta$  is set to 1, which is, as noted above, the most common single choice for this parameter. The value of  $\delta$  is set to 0.1 per cent, reflecting a rejection of inherent discounting, except insofar as it reflects the possibility of extinction. Similarly, there are no interpersonal weights, but changes in the consumption of low-income individuals and countries are weighted more highly at the margin because of the assumed diminishing marginal utility of consumption.

In Stern's analysis  $g$  is derived from the economic scenarios. Typical values of  $g$  are between 1.5 per cent and 2 per cent, so the corresponding values of  $r$  are between 1.6 per cent and 2.1 per cent.

The effect of choosing  $\eta = 1$  and  $\delta$  near 0 is that concern for future generations extends more or less indefinitely into the future, when changes in welfare are expressed in terms of percentages of income or consumption. On the other hand, the discounted value of payments expressed in monetary terms declines quite fast. At a rate of 2.1 per cent, a dollar of (constant price) income received in 2100 is worth approximately 12 cents today. A income stream of a dollar a year, received for a million years into the future is worth a little under \$50.

The controversy over Stern's approach has raised many concerns not all of which have been expressed clearly.. It is hard to disentangle all of these concerns, but I will try to deal with them in the following order.

- (i) erroneous criticisms based on misunderstandings of the discounting procedure
- (ii) criticism of the choice of  $\delta$ ;
- (iii) criticism of the choice of  $\eta$
- (iv) claims of inconsistency with observed market rates
- (v) claims of inconsistency with general practice
- (vi) claims of internal inconsistency
- (vii) criticism of the time horizon used in discounting

### Erroneous criticisms

One part of the debate over  $\delta$  can be dismissed pretty easily. Many of the critics on this point have confused  $\delta$  and  $r$ , apparently assuming that  $\delta$  is a discount rate, rather than a subsidiary factor determining the discount rate.

Examples of this confusion include Kling, Lomborg, and McArdle.

### Criticisms of $\delta$

Among the more serious critics, both Nordhaus and Yohe focus on the sensitivity of the results to changes in the value of  $\delta$ , but do not give any specific argument for inherent discounting. Yohe does not present any argument for a high value of  $\delta$ , simply observing that others have used high rates.

Nordhaus and Boyer (and hence also Nordhaus 2006) appear to backed out  $\delta$  as a residual. For reasons of technical tractability, Nordhaus and Boyer want to set  $\eta = 1$ . On the other hand, they want  $r$  to be at least 4 per cent to match their interpretation of market data. The only way to do this is to choose high values of  $\delta$ . Hence, it is probably more correct to classify Nordhaus' criticism as being based on deviation from observed market rates rather than specifically on the choice of  $\delta$ .

Nevertheless, it may be worthwhile restating DeLong's response to suggestions for a large value of  $\delta$  (numerical values changed)

A  $d$  of 2% per year is unconscionable--it means that somebody born in 1960 "counts" for twice as much as somebody born in 1995, who in turn "counts" for twice as much as somebody born in 2020; somebody born in 1960 "counts" for 256 times as much as somebody born in 2160. That's not utilitarianism

Nor is it sensible in terms of individual decisionmaking. For someone facing a zero real interest rate for savings (not an unreasonable assumption in many cases), the combination  $\delta=2$  per cent,  $\eta = 1$  implies a

consumption path that declines at about 2 per cent per year. To adapt Brad's example, a person beginning such a plan at age 30 would plan to halve their consumption by age 65, and halve it again by age 90. Even allowing for the caution above about EU models and actual choices, this makes no sense.

Yet Nordhaus and Boyer propose an even higher rate of 3 per cent, which is tantamount to saying that the future (certainly anyone more than two generations away from us) can go to hell for all we care, since the welfare of our great-grandchildren has about a tenth the weight we accord the current generation. Not surprisingly, this translates into a 'do nothing now' approach to global warming.

In the absence of any convincing justification for inherent discounting, the case for a low rate such as that chosen by Stern seems overwhelming. Hence, if there is a problem with the ultimate outcome it is necessary to look elsewhere in the analysis. From here on, the value of  $\delta = 0.1$  will be assumed, and discussion of the implications of other choices is conditional on this.

### *The intertemporal elasticity of substitution*

A more plausible criticism concerns the choice of  $\eta$ , the intertemporal elasticity of substitution. With these preliminaries out of the way, we can look at some of the problems that have been made of Stern's choice of  $\eta = 1$ , and suggestions that higher values should have been considered.

The most direct criticism is that, in a growing economy, a low value of  $\eta$  underweights the welfare of the current generation, at the expense of succeeding generations who will be much richer. This point is made most directly by Dasgupta, who considers the case when society has available an unlimited supply of projects yielding a riskless rate of return of 4 per cent. As Dasgupta shows, with  $\eta = 1$ , the implied policy recommendation is that the vast majority of current income (around 97.5 per cent) should be saved in order to allow for greater consumption in the future.

The underlying problem is observed by DeLong. Looking at current savings rates and rates of economic growth, Dasgupta's estimated rate of return to marginal investment of 4 per cent seems conservative for a classical growth model based on factor accumulation. To achieve 1 per cent growth in consumption per person in such a model, it would be necessary to generate net additions to the capital stock equal to 25 per cent of total income each year (since 4 per cent of 25 per cent is 1 per cent). So the fact that we see more rapid growth with lower rates of net saving seems to imply that there must exist many projects with rates of return greater than or equal to 4 per cent.

However, once technical progress, generated either exogenously or through the existence of increasing returns to scale in knowledge, is taken into account, the picture changes radically. In an economy where most growth in consumption arises from technical progress, the optimal rate of saving is far lower than that derived by Dasgupta.

A more direct way of refuting Dasgupta's argument is to observe that the major premise must be false. If there existed an infinite supply of projects with riskless returns of 4 per cent, the rate of return on riskless bonds would have to equal 4 per cent, rather than the 1 to 2 per cent observed in practice. Although this difference may appear small, it is critical in practice

### *claims of inconsistency with observed market rates*

A number of critics, notably including Nordhaus (..) have argued that the discount rate implied by Stern's procedure is inconsistent with observed market data. As has been observed several times above, this claim is incorrect with respect to the most obviously relevant data point, the real rate of return on low-risk bonds.

On the other hand, application of the expected utility model with parameters like those of Stern produces a severe underestimate of the expected rate of return to equity investments, and this is Nordhaus' main concern. In effect, this is a restatement of the equity premium puzzle, along with an implicit presumption that the market rate of return to equity, and not the market rate of return to debt, is the 'correct' measure of the social rate of time preference.

Although defensible, this seems implausible. Most resolutions of the equity premium puzzle imply that the puzzle mainly reflects an excessive risk premium for equity rather than an inadequate rate of time preference.

However, the problems observed here reflect the underlying difficulties of the expected utility model. No single choice of  $\eta$  and  $\delta$  produces results consistent with actual observations on choices over time and risk, and intuitions about redistributions between individuals. One manifestation of this is the equity premium puzzle. Given plausible estimates of risk attitudes, either the real bond rate is much too low (it's typically between 1 and 2 per cent) or the real rate of return to equity is much too high (it's typically 6 to 8 per cent). The model suggests that the two should differ by no more than half a percentage point.

Logically, any proposition can be deduced from a contradiction. So it is easy, in the present context, to combine plausible empirical propositions with standard EU analysis and derive absurd results. In general,



this is not a useful way to proceed.

*claims of inconsistency with general practice in benefit-cost analysis*

Tol makes the observation that British government generally uses discount rates for benefit-cost analysis that are much higher than those proposed by Stern. This is a neat debating point, but has little practical relevance. The political structure of project appraisal is that estimates of costs and cash flows rely on inputs from project proponents that are almost always over-optimistic. Treasury controls the choice of discount rate and uses it to adjust for downside risk as well as for discounting. Hence, the official discount rate is substantially higher than the true social rate of discount.

*(vi) claims of internal inconsistency*

Tol also observes that whereas Stern uses  $\eta = 1$  for discounting, estimates of the cost of disaster are derived from Nordhaus and Boyer who use  $\eta = 4$  for this purposes. It appears that this inconsistency did not originate with Stern, but with Nordhaus and Boyer. Nevertheless, it would seem appropriate to address it.

*Criticism of the time horizon used in discounting*

The most plausible criticism of low rates of discount is that they require us to take account of developments more than 100 years into the future about which we can in practice, know very little. This is a reasonable criticism, but its main effect is to point up the limitations of utilitarian benefit-cost analysis for a problem like global warming.

We know that the effects of global warming will be felt far into the future. We can either mitigate these effects, at very modest cost to ourselves, or leave the problem future generations, whose technological capacities are unknown, but presumably greater than our own. Perhaps our descendants will be able, at very low cost, to resuscitate species we have driven to extinction and restore ecosystems we have destroyed. Perhaps not. There is no easy way of getting useful probability and cost numbers here.

One partial solution might be to end the analysis at, say 2050 or 2100, with future effects being measured as a diminution in the capital stock (including natural capital). Although logically equivalent to the discounting procedure employed by Stern and his critics, this might turn out to be more tractable and intuitively comprehensible.

### **Concluding comments**

Criticism of Stern has focused on the claim that the parameters used in discounting are extremely low, yielding implausible results. In fact, the choice of  $\eta = 1$  is standard, and both lower and higher values are commonly considered in sensitivity analysis. Stern's choice of  $\delta = 0.1$  per cent is primarily the result of applying the standard utilitarian view that all people count equally. If this view is accepted, the pure rate of time discount, reflecting the probability of social extinction, must be close to zero, and there is nothing remarkable about the parametric value  $\delta = 0.1$  per cent.

The real difficulty here is that we are pushing economic analysis to its limits, in an area where fundamental problems, such as the equity premium puzzle remain unresolved. Economists can help to define the issues, but it is unlikely that economics can provide a final answer.

# THE PHENOMENON OF BUSINESS CYCLE SYNCHRONISATION: THE ROLE OF GLOBALISATION AND ECONOMIC INTERDEPENDENCE

DANNIELLA TOMCZYNSKI\*

---

## Abstract

This paper discusses the possible causes, consequences and dynamics of business cycle synchronisation. This paper analyses the Australian experience and attempts to explain those [factors/forces] which specifically drive the patterns observed. In general, and specific to Australia's case, there does appear to be some consensus as to the influential forces, however, empirical results are heavily mixed as to the exact [nature/role] of these forces.

*Key words:* Business cycle, Synchronisation, Interdependence, Globalisation.

---

## 1. Introduction

“The business cycle is usually defined as a regular and oscillatory movement in economic output within a specified range of periodicities,” [Cotis and Coppel (2005)]. Each country can be assumed to exhibit their own ‘typical’ business cycle characterised by a more or less regular pattern of expansion (growth or recovery) and contraction (recession). “The term... ‘international’ business cycle refers to the existence of common elements in aggregate cyclical behaviour across countries,” [Canova and Dellas (1991)]. What drives business cycle synchronisation in the global economy, and the impact these correlations have across country pairs, is of great importance to economists and policy makers alike. Similarities in bilateral economic performance can be accounted for by two distinct factors: the first being notable international economic interdependence, and the second factor, the common experiences of exogenous disturbances (i.e. similar internal or external shocks and similar economic policies) [Canova and Dellas (1991)]. Cotis and Coppel (2005) suggest that a group of countries seems to have nearly ‘extinguished’ the business cycle, as a result ‘classical’ business cycles tend no longer to be observed in OECD economies as they have tended towards reduced fluctuations in GDP output gaps. The literature therefore tends to deal with the alternative, ‘growth’ business cycle theory.

As is evident from the gamut of recent empirical literature, the phenomenon of cross-country business cycle synchronisation still remains a very much unsettled issue. The significance of economic interdependence, globalisation, and the various shock transmission channels remain open to debate and open to further investigation, as do the temporal, dynamic and changing behaviours of these features.

The empirical literature has extended our insight into the concept of cross-country business cycle synchronisation, analysing a variety of questions that are of concern to policymakers. Canova and Dellas (1991), and Dellas (1986) centre their attention on the source of shocks that generate business cycle co-movement as opposed to their transmission between economies. Other authors focus on discerning the transmission mechanisms, such as trade and economic policies, important in transferring shocks between integrated economies. Studies on the effect of trade as a means for transmission can be split into a number of areas of sub-interest; Kose, Prasad and Terrones (2003) analyse how business cycle synchronisation is affected by globalisation, trade, and financial integration. Heathcote and Perri (2003) also investigate the significance of financial integration while Imbs (2003 and 2004) investigates the elaborate linkages between trade in goods and services, financial openness, specialisation and business cycle correlations. With trade partnerships on the one hand, providing the fundamental basis for discerning the significance of international interdependencies and integration, the arguable implications of trade as a determinant of business cycle co-movement remain open to debate and are examined by Crosby (2002), Frankel and Romer (1999), Rodriguez and Rodrik (2000), and numerable others, while Cotis and Coppel (2005) investigate the role for similar and improved monetary policies.

This paper tries to illuminate at least some of these points of interest by examining the empirical literature on business cycle synchronisation within the integrated world economy and discriminating between the

---

\* Written for ECON2040 *Macroeconomic Policy*, Sem 2 2006.

mixed supporting evidence. We discuss the empirical literature outlining the correlatory characteristics observed between the samples: industrialised and developing economies, English-speaking and non-English speaking economies, the OECD economies considering what these implications are for Australia. We then analyse more specifically, the significance of Australia's bilateral relationships with the likes of the US, New Zealand, and the growing economy of China.

## 2. Factors Shaping the Business Cycle

“The recent period of widespread weakness has brought to the forefront questions about how closely economic activity is synchronised across national borders and what transmission channels serve to foster synchronisation. Most OECD economies were characterised by steady expansion and relative macroeconomic stability prior to the global oil crises and failure of domestic policy in the 1970's. A shift in policy framework amongst OECD countries – with an especial focus on monetary policy – in the 1980's has, as some see it, been a key factor in reducing output gaps. Dalsgaard, Elmskov and Park (2002) note a global shift towards more service-based economies and better inventory management facilitated by the increased adoption of Information and Communications Technology (ICT) – the ‘New Economy’. It is important to emphasise the difficulty in discerning the ongoing role of these contemporary factors given their temporal tendencies to change with time. The literature on this matter tends to reveal a general consensus of interest as to whether trade has become more or less important over time in cushioning fluctuations in domestic demand.

With it now widely accepted that the actions of any one economy may impact the experiences of another, the experiences of larger economies like the US, China, the UK etc seem able to influence smaller interdependent economies like Australia, via a number of spillover effects. The recent slowdown in the ICT sector in the US, for example, was accompanied by a major adjustment to stock prices in many economies affected by the US experience [OECD (2002)]. In general, these transmitted effects can be both positive and negative and it is therefore important that Australia's policymakers understand the nature (and structure) of economic interdependence, both in general and specific to Australia's interrelatedness with the rest of the world.

### 2.1 Globalisation and International Economic Interdependence

As the international economy has continued to mature it has inarguably become more consistent with the concept of globalisation – progression towards a single, unified global economy – and realisation of this paradigm is becoming increasingly accepted as an economic (and social) truth. The economic influence that other countries exert on their typically smaller neighbouring economies has cultivated the synchronised fluctuations in business cycles observed amongst economy pairs and ‘economic communities’.<sup>1</sup>

The self-fulfilling concept of globalisation has made the far reaches of the international economy more accessible, presenting greater opportunities for more trade and freer trade, with more partners. The ‘opening up’ of the global economy is an upshot of individual economies opening themselves to the global economy and its internationally competitive forces, in turn decreasing the need to discriminate amongst trading partners on the basis economic distance. Growing exposure to the competitive pressures of an increasingly more open international trade system has led to arguable benefits of higher efficiency in production on the basis of the “specialize and trade” nature of comparative advantage [Thompson (2006)]. Freer trade has led to increased international specialisation with economies gradually switching to production in goods for which they have lower opportunity costs and a comparative advantage in producing internationally. Globalisation has, stimulated by the rapid diffusion of technology, led to economies producing some – but not all – goods required for domestic consumption (i.e. specialising in production), necessitating the formation of trade partnerships between economies.

Obviously trade does not occur between just any pair of economies, and is dependent upon both economic and other non-economic factors. It could be generally presumed that trade is, for the most part, carried out between country pairs who find it both economically advantageous and culturally (and politically) agreeable to do so, hence, leading to the ever-increasing integration and economic interdependencies observed within

---

<sup>1</sup> Cotis and Coppel (2005) and Wacziarg and Welch (2003) examine broad cross-country samples finding that groups of economies who have similar country characteristics, such as: language, legal structures, and policy frameworks, tend to exhibit more highly synchronised business cycles with each other, within what this author terms, ‘economic communities’.

the global economy.<sup>2</sup> Fischer (2003) highlights the point that economies seeking strong growth require appropriate policy frameworks which clearly include “an orientation toward integration into the global economy” while Kose, Prasad & Terrones (2003) find at best, limited support for the line of contention that globalisation is a major source of increased business cycle synchronisation.

## 2.2 International Linkages and Business Cycle Synchronisation

The core rationale of globalisation is flush with the idea that the economic performance of economies from abroad is a partial determinant of growth performance domestically. Whether the Australian economy experiences expansion (positive growth) or a recession (negative growth) is particularly interrelated with the economic performance of its major trading partners, and whether they are performing strongly or not. How does Australia fare in a globalised economy? The effect that its larger neighbours may have on Australia’s relatively small, open economy indicates a vital need for Australia’s policy-makers to possess a wholly comprehensive understanding of the precise nature of its economic interdependencies with the rest of the world.<sup>3</sup>

A simple indication of the countries most likely to be important in exerting some level of influence over Australia can be seen by noting who Australia principally trades with. This can be observed in Table 1 where the level of imported and exported trade with specific trading partners is recorded as a percentage of Australia’s total importing and export activities for 1990 and 2000. The table points to Japan and the United States as Australia’s most significant trading partners in both periods, showing that the sum of trade with Japan and the US constituted trade flows of around 40 per cent in the first sample period, while decreasing to about 30 per cent in 2000. While these two trading partners still hold their rank at the top of the list, a growing shift towards trade with China and the ASEAN countries is observed in the second period. Of its two largest trading partners however, it is the linkage between the Australian and US business cycles which arguably receives most attention. Gruen and Shuetrim (1994) point out that it is now a well-known feature of the Australian economy that its domestic business cycle is strongly correlated with the activities of the OECD, in particular, the US. The two economies exhibit broadly similar industrial structures with about ‘two-thirds of Australian output found to be linked to that of the United States’ [De Brouwer and Romalis (1996)].

The correlation between growth cycles of real GDP for Australia and the selected countries is detailed in Table 2, while the correlation between growth cycles of real GDP for the Australia-US specific relationship can be seen in Figure 1, a simple confirmation that the business cycles of the two economies appear to be relatively highly correlated, and even more so in the last two decades (the period of interest).<sup>4</sup>

It is widely accepted that the US is the world’s largest economy – the leading export market for many foreign economies [Chriszt (2001)]. “For much of the 1990s the United States was the engine for world growth, helping pull the international economy through several negative episodes,” and as such it is often characterised as being the pioneer of the global economy. It’s influence over the Australian economy is such that the pairing is the fifth-highest cross-country correlation [Otto, Voss and Willard (2001)]. The notion that when the ‘US sneezes, the world catches a cold’ is epitomised in numerable catch-phrases, and indicative of the fact that Australia is not the only economy that needs to be concerned with the exertion of US influences nor the economic ramifications of business cycle co-movement with this economic power.<sup>5</sup> Of especial interest to this paper is the transmission of influence from the US to Australia, but also the universal nature of transmission channels.

<sup>2</sup> The observed integration of the global economy is in some means *segregated*, on the premise that individual countries have a discriminatory choice in who they trade with.

<sup>3</sup> Australia’s ‘openness’ is an issue of some debate; see Guttman and Richards (2004).

<sup>4</sup> Gruen and Shuetrim (1994) note that the correlation between Australian and OECD/US output growth is not only high, but that “domestic output seems to track the path of foreign output over time, indicating that this relationship is persistent and long run”.

<sup>5</sup> The phrase, “US Sneezes, World Catches Cold”, employed by the Federal Reserve Bank of Atlanta (2001). Similar phrasing adopted by Bodman & Crosby (2005) who found over 100 sources coining similar phrases.

**Table 1**  
**Australia's Pattern of Trade with Rest of World, 1990 - 2000**

Country	1990		2000	
	Per cent of total exports	per cent of total imports	Per cent of total exports	per cent of total imports
Japan	28	19	20	13
United States	10	24	10	19
New Zealand	5	5	6	4
China	3	3	6	9
Korea	6	2	7	5
Taiwan	4	4	5	3
Hong Kong	3	2	3	1
United Kingdom	4	7	3	6
Germany	2	7	1	5
Other European Union countries	8	11	7	10
Singapore	5	2	6	3
Other ASEAN countries	6	4	8	12
Other	16	10	18	10
<i>Total</i>	<i>100</i>	<i>100</i>	<i>100</i>	<i>100</i>

*The ASEAN group of countries includes Brunei Darussalam, Cambodia, Indonesia, Laos, Malaysia, Myanmar, the Philippines, Singapore, Thailand and Vietnam.*

*Source: Dornbusch, Bodman et al. (2002)*

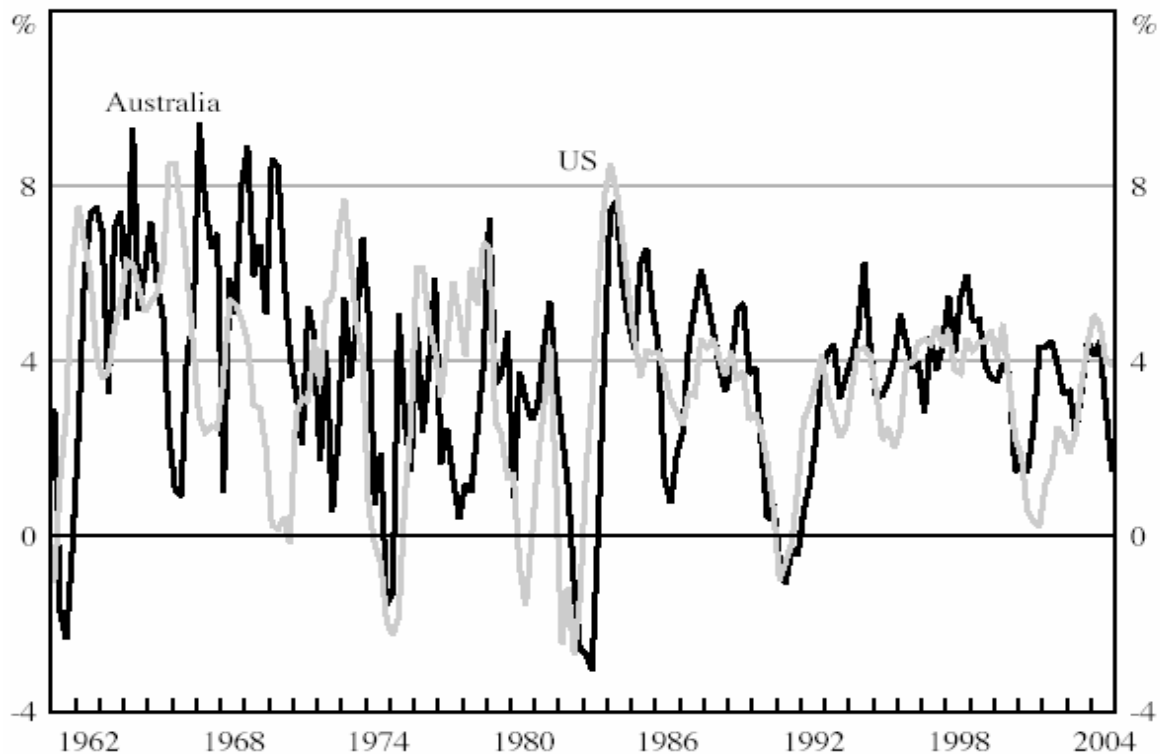
**Table 2**  
**Output Correlations between Australia and Selected Countries**

Australia with:	Correlation of real GDP (year-ended growth rates)			Correlation of GDP cycles (band-pass filtered)		
	1961:Q1 - 2004:Q4(a)	1961:Q1 - 1982:Q4(a)	1983:Q1 - 2004:Q4	1963:Q1 - 2001:Q4(b)	1963:Q1 - 1982:Q4(b)	1983:Q1 - 2001:Q4
	US	0.34	0.32	0.59	0.31	0.02
Japan	0.28	0.41	0.00	0.30	0.49	0.07
New Zealand	0.29	-	0.21	0.38	-	0.32
Canada	0.51	0.38	0.66	0.58	0.25	0.83
UK	0.27	0.19	0.50	0.28	0.06	0.60
Euro area	0.34	0.23	0.59	0.31	0.02	0.82

(a) Sample for Canada starts 1962:Q1, for the Euro area in 1971:Q1 and for New Zealand in 1978:Q1

(b) Sample for Canada starts 1964:Q1, for the Euro area in 1973:Q1 and for New Zealand in 1980:Q1

*Source: Andrews and Kohler (2005)*



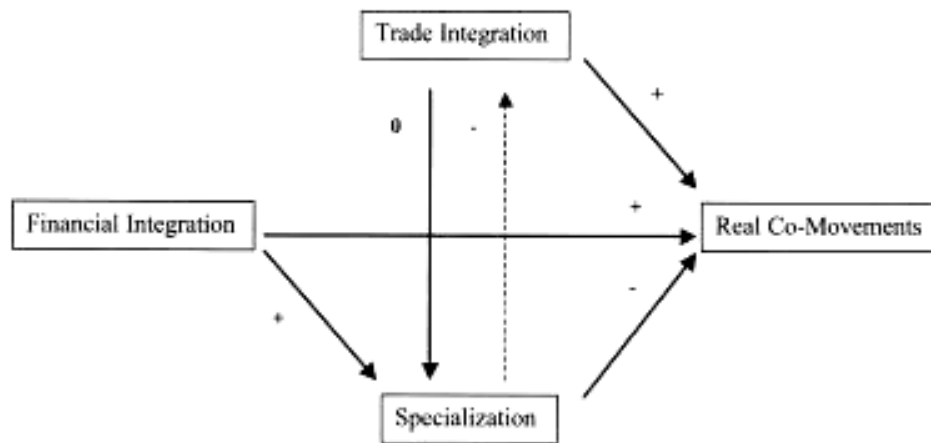
Source: Andrews and Kohler (2005)

**Fig. 1. Australian and US Business Cycles, Real GDP, year-ended growth rates**

## 2. Transmission Channels

Analysis of Australia's existing, and changing patterns of trade relationships, allows for crude postulation of the factors behind Australia's bilateral co-movement and interdependence in the international arena. It is argued by many that it is the linkages in these trade relationships that are the driving force behind business cycle co-movement in the international community. While many have been able to identify the extent of correlation between economy pairs of business cycles, there is still difficulty in identifying the transmission mechanisms which underpin the observed co-movement. Imbs (1999) sees limited possibility for countries to insulate themselves from foreign developments in an increasingly globalised environment, with most quick to point to *trade openness* as responsible for the cross-country correlations we observe.

Causation of business cycle co-movement can be demarcated into a number of chained starting points for further and careful consideration. The first initiative is to assess the effects that globalisation and integration within a globalised and continually 'globalising' economy, has had on Australia's interdependence with particular economic partners. The second task is to give value to the significance of observed linkages through international trade and to discern if it is indeed trade that is responsible for passing shocks between coupled economies. If trade is (at least in part) found to be a major driving force of interdependence and business cycle synchronisation, which channels of trade are involved? Trade as a broad transmission mechanism is composed into a further two broad channels: trade in goods and services, and trade in financial assets. These channels can then be further broken down into other pertinent aspects; goods and services trade can be decomposed to assess the direction and magnitude of fluctuations of the importing and exporting behaviours of both economies, and trade in financial assets can further lead to both increased financial liberalisations and greater financial integration. Imbs' (2004) investigation of transmission mechanisms centres on finance, specialisation and trade as the key determinants of synchronisation, pointing out that both relative openness to trade and financial liberalisations may induce specialisation – which he finds likely to have *direct* effects on co-movement between countries. See Figure 2 for a diagrammatic representation of the perceived direct and indirect effects. The interactions between trade openness, financial integration,



This figure summarises the empirically estimated interactions between goods trade, financial integration, specialisation, and real business cycle co-movements. Zero means the coefficient is estimated to be insignificant.

Source: Imbs (2004)

**Fig. 2. Direct and Indirect Channels for Transmission of International Influence**

specialisation, and real business cycle movements are, suggests Imbs (2004), “complex”, and policy makers require a thorough awareness of the significance of these international linkages. The fourth variable under consideration is the bearing of cross-country structural similarities, which the likes of Imbs (2004) and Otto, Voss and Willard (2001) give much consideration to. In addition to the broad channel of trade, the final predictor we bear in mind is cross-country policy similarity. The following section evaluates the theory and literary evidence on the importance of trade and other factors in determining business cycles synchronisation.

## 2.1 The Significance of Trade: Literary Evidence

There has been tremendous debate over the statistical importance of trade in shaping cross-country business cycle co-fluctuations. In this section we revisit the empirical literature. Frankel and Rose (1998) confirm the broadly accepted wisdom that closer bilateral trade linkages are a proponent for more closely synchronised business cycles, going on to suggest that taking steps towards enhancement and augmentation of the bilateral links in trade (advancing economic and trade integration) is likely to result in greater cross-country synchronisation. “Trade in goods and services is the most commonly identified mechanism by which fluctuations in the level of activity in one economy spill over into other economies,” [Otto, Voss and Willard (2001)]. Meyer (2001) states, in reference to the US, that trade is the ‘traditional’ channel for spillover effects, with the sharpest spillovers experienced in countries who demonstrate the highest ratios of exports relative to the US, relative to their income. After examining the correlation between trade and real GDP, Frankel and Romer (1999), are unable to discern direction of causation between the two suggesting instead the possible importance of the geographic components of trade. Empiric use of the gravity model [see Guttman and Richards (2004)] has found geography to be a powerful determinant of bilateral trade (at least in the models). They consider factors such as country size, geographic distance, whether two countries share a border, and whether the nations are landlocked. It is apparent that economic groupings such as the European Union would thus tend to trade more with their fellow EU partners relative to the rest of the world [Rose and Engel (2000)]. Similarly it is found that the same is true of the United States, with its large geographically proximal community – the many US states have relatively high levels of intranational trade, exhibiting greater synchronisation than the European economies [Clark and Van Wincoop (2001)]. It is these close linkages that seem to be a key ingredient in the formation of currency unions and other optimal currency areas. Concurrently international trade linkages can be expected to generate both demand and supply side spillovers across countries, for example, on the demand side, an investment or consumption boom in one country can generate increased demand for imports, boosting the economic outlook abroad (of course the converse, is also true). Via such spillover effects, greater linkages in international trade could result in higher business

cycles correlations observed across countries. However, the nature of trade flows could also induce increased specialisation of production.<sup>6</sup> The literature expresses mixed views as to the sign of the impacts of specialisation, but it is clear that it generates changes in the integrated relationship between economies and thus changes in the nature of business cycle synchronisation. Otto Voss and Willard (2001) suggest in addition to trade and exchange rate channels, common take-up of new technology and common language are highly important in explaining output correlations via their relation to trade openness.

Of growing interest is the degree of trade openness. Guttman and Richards (2004) assess the fit of the gravity model for aggregate country trade with fifteen variables under consideration including a trade liberalisation index, used as a measure of trade openness. Their findings are that trade openness is significant at the 2 per cent level, with the three most significant predictors including: product of total GDP (a measure of combined economic size of the trading nations), geographic distance and product of per capita GDP (a measure for the economic stage of development) – all significant at the 3 per cent level. However, after adjusting their model to remove scaling effects, adjusted R-squared decreases from 0.70 to 0.34. These authors go on to construct an ‘Openness Equation’ finding the most important measures of trade openness to be population, economic location and trade policies. While many economies have made efforts to open up to the international market by reducing barriers to trade, it may now be that transport costs (the costs of geographic distance) represent a larger effective rate of ‘protection’ from wider trade than any trade policy restrictions [Clark, Dollar and Micco (2004)]. The degree of economic and geographic distance seems to be important, at least as an indication of who trades with whom. The literature examines business cycle characteristics across broad samples i.e. industrial versus developed economies, English speaking and non-English speaking, OECD and other advanced economy groupings, down to the intrinsic nature observed between two countries.

Of interest is an additional finding that a historically common coloniser (i.e. colonising nation) provides some indication as to the strength of bilateral trade amongst economies, this variable was found to be significant at the 9 per cent level by Guttman and Richards (2004). This finding is partial confirmation of another stream of thought which examines the possibility of non-economic variable as determinants of co-movement. Imbs (1999, 2000) for example, suggests the explanatory significance of structural similarities, although this factor seems only able to explain co-movement amongst the OECD and industrialised economies, using structural shares of employment as a measuring indicator.

While it is generally agreed upon that a country’s economic development status provides some indication as to the group with which they are most likely to exhibit synchronisation of cycles, there is much less agreement on the possible reasons underlying this observed fact. It may either be interpreted that the considerable interdependencies observed between economies – characterised by their unique trade linkages – are transmitted via trading behaviour, resulting in spillovers through the international economy, or alternatively we may interpret the likelihood of synchronisation to be driven by common shocks. If it is indeed these common shocks that underlie correlated cyclical behaviour, then if two economies are also characterised by similar economic structures, they should respond similarly to these shocks [Otto, Voss and Willard (2001)].<sup>7</sup>

The previously-made point that trade as a mechanism for spreading influence has both direct and indirect spillover effects, is also likely to be characterised by differences in the magnitudes of these effects. The combined directional and magnitudinal strength of these spillovers from one economy to another can lead to the larger effects dominating the impact or influence of any others; the nature of the specific cross-country relationships may also mean that there is no one observed pattern. While the empirical evidence remains varied in discerning the importance of trade and its magnitude and direction(s) of influence, for business cycle synchronisation, there is wide-held agreement that trade is an obviously important transmission mechanism (at least in theory). Evidence that trade and financial integration enhance the spillover of macroeconomic fluctuations is strong, although this is only true for industrial countries [Imbs (2004)]. The direct effect of trade on synchronization as documented by Frankel and Rose (1998), estimates a strong and robust positive relationship between trade and cyclical synchronisation.

The three previously introduced transmission channels – trade in goods and services, trade in financial assets and the similarity of policy coordination – while not an all-inclusive framework for analysis, does capture what are likely to be the three most important, though broadly captured, channels. Roles are investigated for innumerable non-trade factors, as a starting point these include: corporate governance, structural economic policies, technology uptake (an indicator of economic development), ‘social distance’ as measured by geographic and economic distance and language. The following section highlights the specific

<sup>6</sup> See Imbs (2000, 2004) and Kose, Prasad and Terrones (2003)

<sup>7</sup> See also Imbs (1999, 2000, 2004)



experiences of Australia's cyclical co-movement with a group of its major economic partners, also introducing some of the potentially important non-trade predictors of synchronisation described above. We observe first correlations with the US, followed by New Zealand and a discussion of the increasing economic influence of China.

## 2.2 The Australian Experience: Correlations with the United States

Perceived as an 'engine of growth' in the international community [Chriszt (2001)], the United States does indeed seem to be an engineering force of business cycle correlation. The recent economic downturn of the US saw growth slow to 1.0 per cent in the fourth quarter of 2000, having peaked prior in the fourth quarter of 1999 at 8.3 per cent.<sup>8</sup> As the economy weakened, demand for US imports declined, as a result those economies which depend heavily on exporting their product to the US experienced a slowdown in growth which in turn fed back through the trade transmission mechanism to affect foreign demand for US goods. Gruen and Shuetrim (1994) estimate the contemporaneous impact of US GDP growth on the growth in Australian GDP to be between 0.4 and 0.6, consistent with other estimates of around 0.5. De Roos and Russell (1996) model US and Australian GDP finding that deviations between the long-run relationship between the two does not affect US GDP, only that of Australia, implying the relationship in output fluctuations may be one of direct causation and not correlation. A number of authors support this view finding that direct effects of US activity on Australia are quite small, although this has increased slightly since the early 1980's, coinciding with increased trade and financial market liberalization [De Roos and Russell (1996)]. Otto, Voss and Willard (2001) model Australian-US output growth correlation, finding that intensity of trade makes a relatively small contribution to the correlation; this is in opposition to naïve opinion that high levels of trade induce high levels of synchronicity. Table 3 below decomposes the predicted contribution of US output growth to Australian output growth.

Another explanation for the observed Australia-US correlation is the impact of commonly experienced global shocks, which this author finds to hold some credence. The previously noted similarity in US and Australian economic structure seems a worthy argument for cross-country correlation given that the two economies may tend to respond similarly to commonly experienced shocks with the same economic tools in use. This is supported by empirical observations of higher cycle synchronisation during recessionary periods, the first of the major significant world-wide shocks being the oil price shock of the 1970's (refer to Figure 1). The subsequent recessions in the 1980's, early 1990's and most recently in 2000 are inherently characterised in the observed business cycle correlations via sharp increases in the correlation coefficient – the Australian US correlation relationship can be seen in Figure 3, also showing the trends in two of the three broad transmission mechanisms (structural similarity and trade) in addition to relative market flexibility.

Much consideration has also been given to the similarity of monetary policies, with Andrews and Kohler (2005) also suggesting non-negligible transmission channel: financial integration. Figure 4 and Figure 5 show economic policy correlations between Australia and the interest group, in addition to the UK and some others.

**Table 3**

### Contributions to Predicted Australia-US Output Growth Correlation

	Trade intensity	Exchange rate deviation	Country of legal origin	Accounting standards	technology take-up	– $y_i - y$
Contribution	0.04	-0.02	0.08	0.06	0.08	0.24
Percentage	17	-8	33	25	33	100

Source: Otto, Voss and Willard (2001)

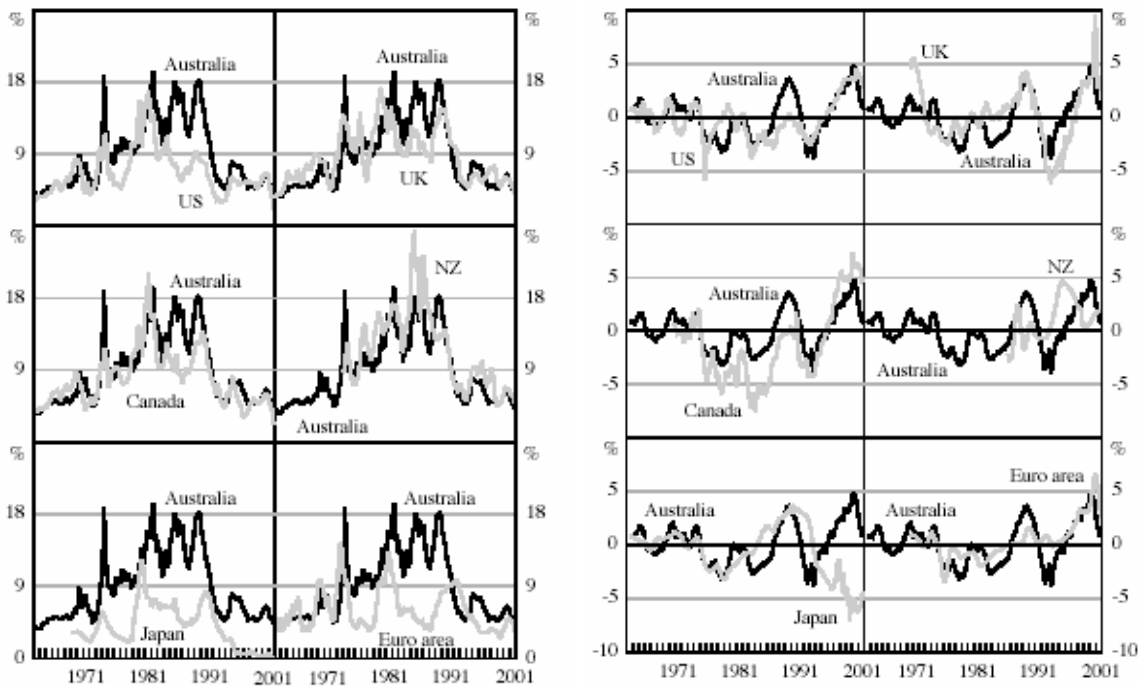
<sup>8</sup> See Chriszt (2001)



(a) The dates refer to end-point of 16-year moving window. Dashed lines denote Newey-West-corrected 95 per cent (asymptotic) confidence intervals.

Source: Andrew and Kohler (2005)

Fig. 3. Australian-US Correlation Trends through Time: 1980 - 2000



Source: Andrews and Kohler (2005)

Fig. 4. Monetary Policy Variables  
Nominal short-term interest rates (3-month maturity)

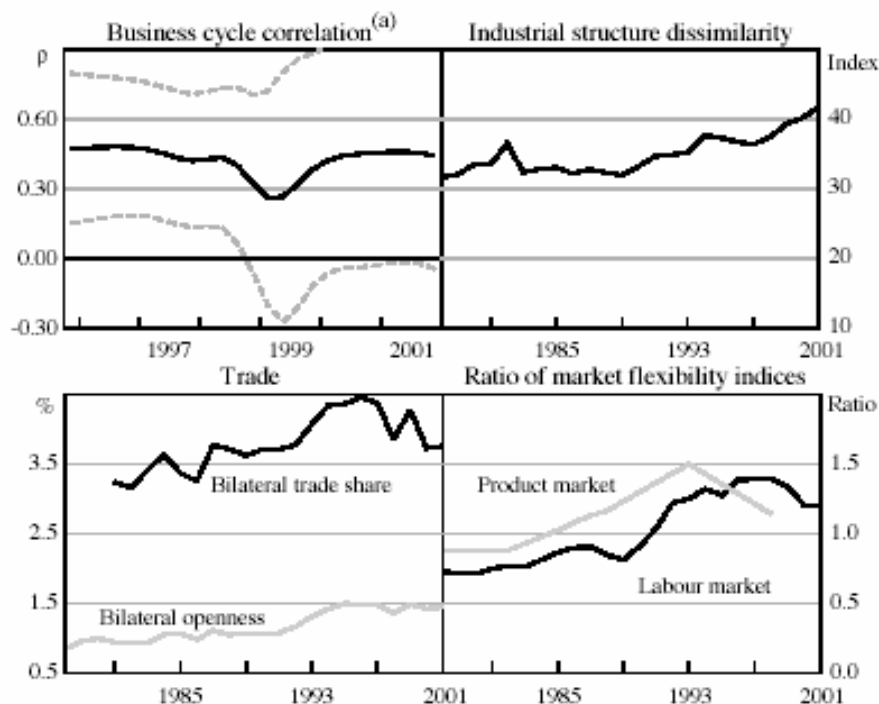
Fig. 5. Fiscal Policy Variables  
Primary budget deficit as per cent of GDP

Source: Andrews and Kohler (2005)

### 2.3 The Australian Experience: Correlations with New Zealand

“In a period when countries are becoming increasingly linked to one another through trade and capital flows, the management of the exchange rate regime is a critical factor in economic policy making,” [Drew, Hall, McDermott and St.Clair (2002)]. The topical issue as to whether New Zealand would do better to adopt the Australian currency has fair placement in the issues of cycle synchronisation. Notwithstanding the similarities of the two economies, whether the shocks are commonly experienced by the pair or their experiences are unique it remains arguable whether dollarisation in this case would benefit the New Zealand economy. Despite their structural similarities and geographic proximity, evidence of co-movement across this country pair is not as high as might be expected from their close economic and social ties. Figure 6 delineates the generally stable business cycle co-movement between the pair, as well as a growing measure of dissimilarity in economic structure and declining correlation between ratio market flexibilities (as is also apparent between Australia and the US – refer to Figure 3). Table 4 whilst not strictly a measure of Australia and New Zealand’s bi-lateral correlation, is perhaps implicative of a lack of co-movement indicating that the losses for New Zealand under the adoption of foreign policies (such as the proposed AUD adoption). This is contrary to the Euro area, whose success lies in the fact that there is quite a high correlation amongst its members. Figure 7 supports New Zealand’s low correlation tendencies with New Zealand appearing 5 times out of the 20 lowest output correlations.

If measures of bi-lateral trade are of any significance, it is worthwhile noting the relative shares of bi-lateral trade that New Zealand and Australia exhibit; Crosby (2003) calculates total trade (imports plus exports) between country pairs as a ratio of their GDP. As a ratio of Australia’s GDP, trade with New Zealand contributes 1.20 times GDP, on the flip side trade between the two contributes 8.33 times the GDP of New Zealand. Comparable to the Australian-US relationship, Australia seems to be in the position of power to exert its influence in the Australia-New Zealand correlations, just as the US is in the same position relative to Australia.



(a) The dates refer to end-point of 16-year moving window. Dashed lines denote Newey-West-corrected 95 per cent (asymptotic) confidence intervals.

Source: Andrew and Kohler (2005)

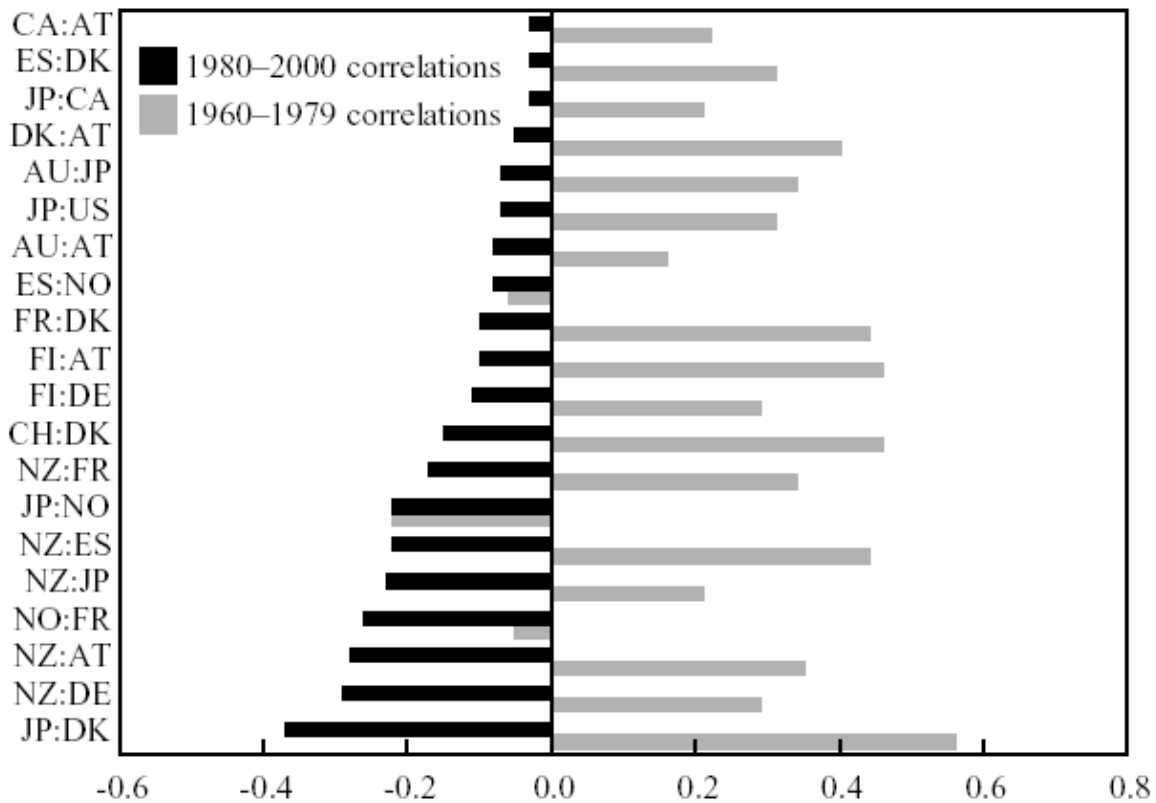
**Fig. 6. Australian-New Zealand Correlation Trends through Time: 1980 – 2000**

Table 4

Collation of Stochastic Simulation Properties into a Standard Loss Function

Simulation	Loss under Domestic Policy	Loss under Foreign Policy	Percentage difference in Loss
Domestic Shocks	1	3.8	273
World Shocks	1.3	2.9	120
Mixed Shocks	1.6	3.3	105
All Shocks	2.3	3.9	71

Source: Drew, Hall, McDermott and St.Clair (2002)



Source: Otto, Voss and Willard (2001)

Fig. 7. Twenty Lowest Output Growth Correlations

2.4 Analysis of the Emergence of China in the Global Economy – Implications for Australian-Chinese Correlations

While the inference is basic, changes in the patterns of trade (as delineated in Table 1) are a signal that the global economic environment is changing. Economists are proposing the emergence of a new economic force, with China’s recent reforms and integration into the world economy [see, for instance Crosby (2005)]. Bilateral trade with China as a fraction of Australia’s total GDP was 0.99 for the period 1980-1999, relative to Australia’s other major trading partners, Japan (6.18), the US (4.36) and New Zealand (1.20), does not exactly place China in a position to monopolise our other trade relationships as is of some current concern – yet – although it is higher than Australia’s trade-GDP ratio with China of 0.56, most certainly placing China in a *potential* position of power, if it’s position should improve relative to Australia’s other partners [Crosby (2003)].

With Bodman and Crosby (2005) suggesting that one of the main reasons behind synchronisation with the US and other countries is through the synchronisation of monetary policy, the fact that Australia’s resources

sector and terms of trade is quite strongly integrated with that of China, may not hold validity in the light of extremely different economic policies. Having had a communist regime until only relatively recent, though comprehensive reforms, China is still faced with some significant obstacles. Crosby (2005) notes that while China has impressively integrated into the international economy – it still has a way to go before it can be definitively labelled one of the top economic powers. Despite issues of low quality investment – caused by the inefficient SOEs and dampening of private investment, China's GDP growth is expected to remain strong, although its growth experience may not be as phenomenal or rapid as that of the previous group of Asian Tigers. Crosby predicts, that with the gradual growth of the private sector and private investment, higher productivity growth will ensue, and it can be expected that China's income per capita could converge to current US levels within the next half century. While this rate may seem surprising, Crosby notes that economies such as Singapore have done so far more quickly. Crosby also note that the Australia-China bilateral trade relationship may mimic the earlier trend of Australia and Japan's trade pattern. In 1950, 3 per cent of Australia's exports were to Japan which grew to nearly 20 per cent over two decades [Crosby (2005)]. Crosby perceives that China's increasing demand for energy may become a point for increasing trade ties as China continues to penetrate the global economy. He also notes that there is strong evidence of vertical specialisation amongst Asian nations and that the issue of intellectual property rights may be a point of concern in the not too distant future.

China's rise, in conjunction with some general changes in the international economy (for instance the recent US downturn) may just be good timing on China's part; nonetheless the impact that its emergence will have and is already having is seen as major changing force in the global economic landscape. With the nature of the changing Australian-Chinese relationship still open to transformation.

### 3. Concluding Comments

By observing Australia's business cycle correlation with its major economic and trading partners, we have attempted to discern the causal factors responsible for driving those correlations. The empirical findings and analyses present a rather complex interaction of variables perhaps responsible for the complex international interactions economists and policymakers hope to explain through synchronisation. Three broad channels are distinguished by which transmission of shocks from one economic body may transfer influence through to another or others. The three channels: trade in goods and services, financial assets and policy similarity are capable of positive and negative, direct and indirect influences as the process of transmission takes place. Additionally these channels can be broken down into a number of integrated mechanisms, and combined with non-economic factors of influence; this entire gamut of variables is employed in modeling and theoretical analysis in order to decompose and explain possible causal and interactive forces driving business cycle synchronization across economic groups.

In particular we have examined Australia's correlated relationships with the US – perhaps the largest source of influence on Australian cycles, New Zealand – arguably Australia is in a position to exert its influence in this relationship, and China – the growing presence of China in the international economy is bound to have some significant impact on the future 'hierarchy' of economic power, and on the changing nature of international integration, and thus business cycle co-movement experienced amongst countries.

We find trade to be a definite link through which transmission occurs, but this would have to be the simplest role trade has in this dynamic process, fundamentally a process of continued globalisation trends expressed in the characteristics of the synchronicity in real GDP cycles. Factors such as language, legal structures, ICT uptake, geographic proximity, trade intensity, trade openness and exchange rate behaviour are all interactive variables which in some cases explain traits of other variables making it just another complex web to weave together in light of explaining what drives business cycle correlations. While cross-country synchronization is unlikely to be simply coincidental, it still remains to be better understood.

### References

- Andrews, D. and M. Kohler, 2005, International Business Cycle Co-movements through Time, in ed. Christopher Kent and David Norman, *The Changing Nature of the Business Cycle, Proceedings of the Reserve Bank of Australia 2005 Conference*, October.  
[http://www.rba.gov.au/PublicationsAndResearch/Conferences/2005/Andrews\\_Kohler.pdf](http://www.rba.gov.au/PublicationsAndResearch/Conferences/2005/Andrews_Kohler.pdf)
- Bodman, P. M. and M. Crosby, 2005, When the US Catches a Cold, Do We Need to Sneeze? Historical and Future Linkages Between the Australian and US Business Cycles, in ed. Christopher Kent and David

- Norman, *The Changing Nature of the Business Cycle, Proceedings of the Reserve Bank of Australia 2005 Conference*, October.  
[http://www.rba.gov.au/PublicationsAndResearch/Conferences/2005/Crosby\\_Bodman.pdf](http://www.rba.gov.au/PublicationsAndResearch/Conferences/2005/Crosby_Bodman.pdf)
- Canova, F. and H. Dellas, 1993, Trade, Interdependence and the International Business Cycle, *Journal of International Economics* 34, 23-47.
- Christz, 2001 US Sneezes, World catches Cold, Federal Reserve Bank of Atlanta, Publications  
<http://www.frbatlanta.org/invoke.cfm?objectid=87B67B00-6666-11D5-3390020352A7A95&method=display>
- Clark, X., D. Dollar, and A. Micco, 2004, Port Efficiency, Maritime Transport Costs and Bilateral Trade, NBER Working Paper no. 10353.
- Clark, T. E. and Van Wincoop, E., 2001, Borders and Business Cycles, *Journal of International Economics*, 55, 59-85.
- Cotis, J-P. and J. Coppel, 2005, Business Cycle Dynamics in OECD Countries: Evidence, Causes ad Policy Implications, in ed. Christopher Kent and David Norman, *The Changing Nature of the Business Cycle, Proceedings of the Reserve Bank of Australia 2005 Conference*, October.  
[http://www.rba.gov.au/PublicationsAndResearch/Conferences/2005/Cotis\\_Coppel.pdf](http://www.rba.gov.au/PublicationsAndResearch/Conferences/2005/Cotis_Coppel.pdf)
- Crosby, M. 2005, China's Increasing Economic Might, *The Melbourne Review* 1, no. 1, 49-53.
- Crosby, M., 2003, Business Cycle Correlations in Asia-Pacific, *Economic Letters* 80, no.1, 35-44.
- Dalgaard, T., J. Elmskov, and C-Y. Park, 2002, Ongoing Changes in the Business Cycle – Evidence and Causes, OECD Economics Department Working Papers, no. 315.
- De Brouwer, G. and J. Romalis, 1996, External Influences on Output: An Industry Analysis, Reserve Bank of Australia, Economic Research Department – Discussion Paper 9612, December.
- De Roos, N., and B. Russell, 1996, Towards and Understanding of Australia's Co-movement with Foreign Business Cycles, Reserve Bank of Australia, Economic Research Department – Discussion Paper 9607, November.
- Drew, A., V. Hall, J. McDermott and R. St.Clair, 2002, Would Adopting the Australian Dollar Provide Superior Monetary Policy In New Zealand, New Zealand Association of Economists Conference, June.
- Frankel, J. A. and D. Romer, 1999, Does Trade Cause Growth?, *American Economic Review* 89, no. 3, 379-399
- Gourinchas, P-O. and O. Jeanne, The Elusive Gains from International Financial Integration, IMF Working Paper WP/04/74.
- Gruen, D. and G. Shuetrim, 1994, Internationalisation and the Macroeconomy, in P. Lowe and J. Dwyer (eds), *International Integration of the Australian Economy, Proceedings of the Reserve Bank of Australia 1994 Conference*, 309-363.
- Heathcote, J., and F. Perri, 2003, Why Has the US Economy Become Less Correlated with the Rest of the World?, *American Economic Review* 93 no. 2, 57-62.
- Imbs, J., 1999, Co-fluctuations, Centre for Economic Policy Research Discussion Paper Series, no. 2267, October.
- Imbs, Jean, 2000, Sectors and the OECD Business Cycle, CEPR Discussion Papers, no. 2473
- Imbs, J., 2004, Trade, Finance, Specialization and Synchronization, *The Review of Economics and Statistics* 86 no. 3, 723-34.
- Kose, M. A., E. S. Prasad, and M.E. Terrones, 2003, How Does Globalization Affect the Synchronisation of Business Cycles?, IMF Working Paper WP/03/27, January.
- Kose, M. A. and K-M. Yi, Can the Standard International Business Cycle Model Explain the Relation Between Trade and Comovement?, Federal Reserve Bank of Philadelphia Working Papers, no. 05-3, November
- OECD, 2002, Ongoing Changes in the Business Cycle, draws on Dalgaard, T., J. Elmeskov and C-Y. Park, *Ongoing changes in the business cycle – evidence and causes, OECD Economics Department Working Papers, no. 315*.
- Rodriguez, F. and D. Rodrik, 2000, Trade Policy and Economic Growth: A Skeptic's Guide to the Cross-national Evidence, NBER Working Paper, no. 7081, April.
- Rose, A. K. and C. Engel, Currency Unions and International Integration, Centre for Economic Policy Research Discussion Paper Series, no. 2659.
- Thompson, H., 2006, *International Economics: Global Markets and Competition* 2<sup>nd</sup> edn, World Scientific, Singapore.
- Wacziarg, R. and K. H. Welch, 2003, Trade Liberalization and Growth: New Evidence, NBER Working Paper Series, no. 10152.

## THE WAGE GENDER GAP: ARE WOMEN CATCHING UP?

ANNIE NGUYEN\*

*This paper provides an overview of the gender wage differential, particularly in reference to Australia. It addresses the phenomenon of why women continue to earn lower wages than men, even after the enactment of equal pay decisions which changed the institutional framework of the country. Although there was a significant narrowing of the gap after the first wave of reform, stagnation in the following decades requires an examination of the underlying causes. In this regard, the legislative changes are deemed to have been most effective in the first round due to the institutional reform, yet less effective in redressing the labour market discrimination. Alternatively, while the human capital characteristics of women has increased in part as a result of legislation, and increased labour force participation they continue to be disadvantaged owing to the lower returns they receive for their productive characteristics. Thus, discrimination which is expected to be a relatively stable component continues to perpetuate a large proportion of the gap. The debate that surrounds occupational segregation is also discussed, due to its potential influence on the gap. It is concluded in this paper that occupational segregation does not seem to be a large determinant of the observed gap, thus the minimal influence of legislative reform on occupational segregation does not contribute much to understanding the remaining gap. Another finding of the paper is that while the gender wage gap is a widespread phenomenon, women in Australia are also found to be remunerated quite favourably on comparison with their counterparts, which reinforces the influence of country-specific institutional effects on the wage gap.*

### I. INTRODUCTION

The gender wage differential has been an issue that has continually dominated the study of labour markets in Australia. In particular, the wage differential between males and females that has persisted after the implementation of significant legislative reforms since the 1960's has been a subject of a number of academic papers over the past decade. This paper seeks to provide an introduction to the concept of the gender wage gap, with particular reference to the situation in Australia.

The paper will begin by examining the declining trend of the differential since the 1960's, with particular attention paid to the institutional reform which resulted from Equal Pay decisions initiated at the end of the 1960's. Subsequently, examination of the stagnation in convergence of the gap in spite of reforms in the 1980's shall serve to assess the relative effectiveness of legislation changes

Examination of the status of Australian women through the papers of Daly, Meng, Kawaguchi and Mumford, (2006), and Blau and Kahn (2000a) shall be discussed to provide a cross-country comparison of differentials. This emphasizes Australian women's status as one of the highest earning groups relative to males when compared to fellow OECD members; and the importance of institutional frameworks and labour market characteristics upon the observed differential between countries.

The influences upon the wage differential of skills differences, labour market treatments and occupational segregation factors shall also be discussed. The final section will address the traditional measure of discrimination employed in the literature and discuss some of its shortcomings and the alternatives proposed.

### II. THE GENDER WAGE GAP IN AUSTRALIA

Prior to analysing the trends in the gender wage gap in Australia, it is important to make a distinction between the period before 1969 and from 1969 onwards. This enables an observation of the direct effect upon the gap due to prevailing institutional settings (Kidd and Meng, 1997) and understanding of the significant occupational segregation of the sexes.

For the greater part of the twentieth century, Australia implemented a centralised system of wage setting, a result of by Justice Higgins as described in Pocock (1999), that fixed a woman's wage at 54 per cent of her male counterpart. Thus as outlined by Borland<sup>1</sup> (1999) female workers were subjected to direct gender wage

\* Written for ECON2800, *Labour Economics*, Sem 2 2006.

<sup>1</sup> The decision was justified based on a needs criterion that determined that male's wages had to support a family while a female's wage need only support a single person described by Kidd and Shannon (1996) and as such, female workers were subjected to direct gender wage discrimination

discrimination. An exception as part of the ruling provided for the equal payment of women who worked alongside men as a means to prevent mass substitution of men for less expensive female equivalents (Pocock, 1999). Pocock (1999) asserts that this generated significant incentives for occupational segregation resulting in Australia becoming one of the most gender segregated nations. The consequence of this characteristic of the Australian labour market was observed through its impact upon the effectiveness of subsequent legislative reforms.

The Equal Pay decision of 1969 and subsequent decisions in 1972 and 1974 impacted upon the gender wage differential by facilitating the movement from a remuneration system that explicitly utilised sex as a means of wage discrimination. Kidd and Meng (1997) report that between 1969 and 1972, Australia implemented 'equal pay for equal work' policy. However, the strict interpretation of 'equal work'<sup>2</sup> in conjunction with the high degree of occupational segregation by sex resulted in only 18 per cent of women benefiting as reported in Pocock (1999) and Borland (1999). Alternatively, the effects were minimal due to limited coverage whereby 'female occupations' where 80 per cent of women at the time were concentrated did not satisfy the 'equal work' criteria (Jones, 1983).

Subsequent 'equal pay for work of equal value' comparable worth policies implemented between 1972 and 1975 attempted to rectify this limitation by expanding the definition of equal work according to Borland (1999). In response, the earnings ratios of women rose in the order of 22 per cent from 72 per cent to 94 per cent of the male weighted average hourly minimum, and a gain of 18 per cent from 59 per cent to 77 per cent of male full-time weekly earnings between 1968 to 1977 reported by Kidd and Meng (1997) due to the elimination of direct wage discrimination (Jones, 1983). The weekly earnings gap fell from 42 percent to 20 percent spanning the decade between 1969 and 1979, primarily underpinned by these reforms (Langford, 1995).

Chart 1 shows the significant narrowing of the gap in the 1970's was followed by relative stagnation in the 1980's. This result was contradictory to expectations as the strong effects of legislative reforms seemed evidentiary of similar advances accompanying further reforms. In particular, the Sex Discrimination Act introduced in 1984, and the Affirmative Action Act of 1986 emphasised equal treatment. The former sought to achieve this by outlawing gender discrimination in a number of areas<sup>3</sup>, whilst the latter<sup>4</sup> emphasised promotion, training access and redundancy decisions made absent of gender considerations (Kidd and Meng, 1997) and focusing on results according to Gunderson (1989). Statistics show that despite reforms, the ratio of female to male average weekly earnings remained fairly constant at 65 per cent for the duration of this period, whilst full-time female workers saw only a minor increase in their wage ratio from 76 per cent in 1982 to 79 per cent in 1990. Since this period, the wage gender ratio has remained fairly constant at 94.3 percent as reported by Borland (1999).

---

<sup>2</sup> The definition of 'equal work' as described in Kidd and (Meng, 1997) meant only jobs whereby women 'did the same work as men and were likely to displace them' would qualify, effectively excluding the majority of occupations where females were concentrated, thus the decision had minimal impact.

<sup>3</sup> The areas covered under the Act were quite substantial including employment; terms and conditions offered to new and existing employees; and training, educational or advancement opportunities (Kidd and Meng, 1997)

<sup>4</sup> Affirmative action is a more results based response which seeks to address and compensate for past inequalities and systemic discrimination. For more detail, refer to Gunderson (1989)



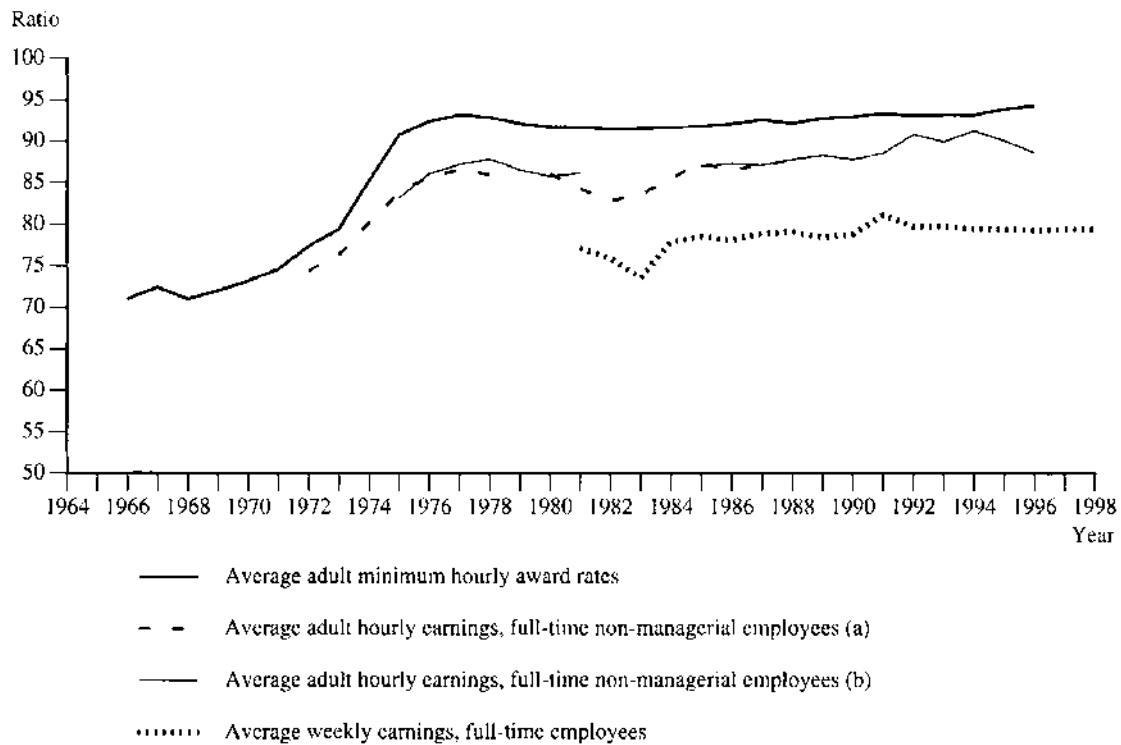


CHART 1  
Ration of Female to Male Earning, Australia 1964-1998

Source: Borland (1997)

#### THE PERSISTENCE OF THE GENDER WAGE GAP

The preceding section highlighted the dominant trends in the wage gender gap since the 1960's. Whilst successful in the 1970's, the slowing convergence of the 1980's and stagnation in the 1990's requires an examination of possible reasons for the gap's persistence.

Borland (1997) proposes that the implementation of the 'equal pay for work of equal value' in some female-dominated occupations has not been achieved, a view aligned with that of Short (as referenced in Langford, 1995), who argues that the Equal Pay decision is still to be fully implemented. Borland's (1997) article proposes that there is inherent difficulty associated with assessing work value where work in female-dominated occupations has no obvious counterparts in other occupation groups, a view supported by Strachan and Jamieson (1999). This study, centred on New South Wales documented the difficulty and confusion that has accompanied the legislation's implementation and interpretation as evidenced when occupations as diverse as nursing and coalmining were considered as counterparts for comparison.

Borland (1997) presents an alternative explanation that may help to explain the persistence of the wage differential. Existing studies utilise measures that solely consider the monetary compensation of employees. Thus the argument these measurements fail to account for differences, whether positive or negative, that workers accrue from non-monetary compensation. Because a worker's net utility is a function of all these various factors, Borland (1997) argues that it may be the case that women do not suffer any disadvantage, and that the advantage of employment for both genders is equal.

Another reason is afforded by Le and Miller (2001) who report that the low portion of the wage gap accounted for by productivity differences may indicate the persistent favourable treatment of males, and systematic unfavourable female treatments. Due to the stickiness of wages that accompany such treatments, females may find it increasingly difficult to move out of the disadvantaged state which is accentuated by the educational attainment and firm size attributes. Thus the persistence of the wage gap may refer to the same disadvantaged group of women, and importantly for policy implies labour market restructure as one means of redress.

### III. INTERNATIONAL COMPARISONS

Another area of interest for academics has been whether the observed wage differential in Australia is

consistent with evidence from other countries. For this purpose, a brief examination of the papers by Daly, Meng, Kawaguchi and Mumford, (2006), and Blau and Kahn (2000) shall be conducted. With all countries experiencing rising female participation, and increases in their ratios as shown in table 1, international differences indicate that the gap between the earnings of males and females are not influenced solely by gender-specific characteristics but are also sensitive to inter-country factors.

#### *Comparisons with Japan, France and Great Britain*

Daly et al (2006) performed a cross-country comparison of the gender wage gaps involving Australia, Japan, France and Great Britain. Chart 2 shows their results that despite significant change in institutions of wage determination and legislation, gender wage differentials still exist.

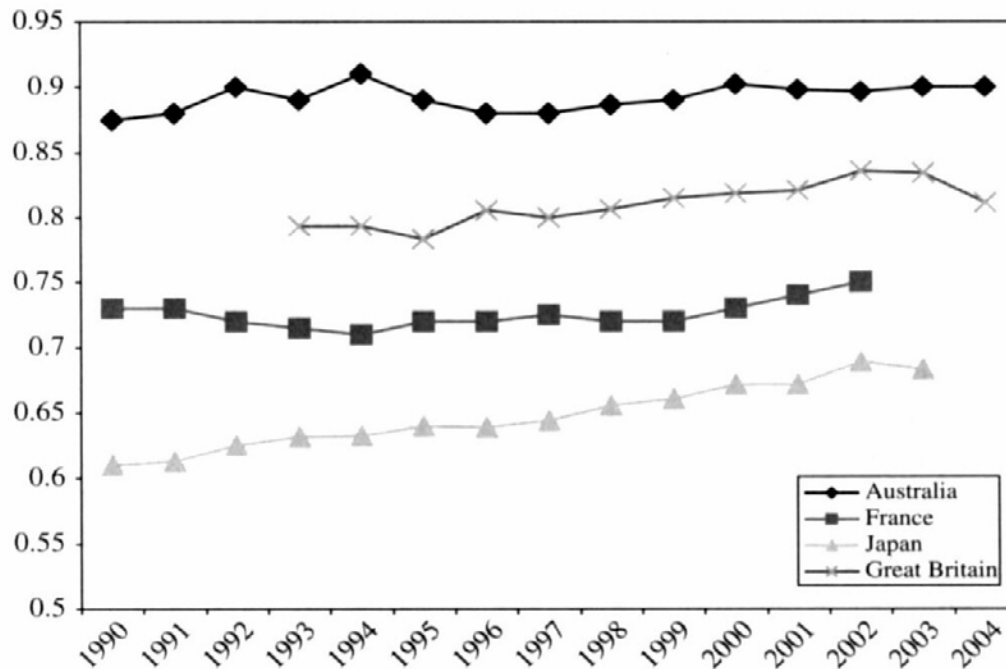


CHART 2  
Gender Wage Gap for Full-Time Employees

*Source:* Daly et al (2006)

Their findings indicate the magnitude of the difference between gaps across countries is highly sensitive to the country that is used as the basis for calculation. Differences in the institutional frameworks of these countries have been attributed as the main source of the differences in their gender wage gaps. In particular, Australia and Great Britain have moved towards a more decentralised wage determination systems; Japan has moved away from its traditional seniority system to a more results based system, however law does not prohibit inter-company gender wage discrimination<sup>5</sup>; whilst France's combined elements of centralised and enterprise bargaining.

Japan witnessed the most significant earnings ratio growth following reform, attributed to changes in endowments and rewards for experience with the current employer, in Daly et al (2006). While Australia, Britain and France all observed narrowing of the wage gap, the magnitude of the changes between 1990 and 2000 are much smaller. Decomposition showed that post-secondary qualifications generated the additional hourly earnings for men in France and Britain greatest while the returns for qualifications for men exceeded women in all cases. Firm size was important in Japan and Britain, impacting on hourly earnings for both genders, and the wages of those in large firms exceeded those in smaller firm counterparts, consistent with Le and Miller (2001) who report that intra-occupational wage effects may be partly explained by firm size, as larger firms are more able to implement classification systems that disadvantage women. The returns to potential experience as measured by tenure with current employer and labour market experience by were less than those to tenure for all countries except Britain which may be partly explained by labour market

<sup>5</sup> The perpetuation of inter-company discrimination arises through the two-rank system, where the first rank contains higher paying management roles requiring relocation, while the second is considered for lower-skilled individuals who are not expected to experience much career advancement and is where women are usually located. The decisions had impact due largely to the decline in importance of the seniority system which advantaged men.

operations. These returns to gender specific factors are significant for the conclusions drawn.

Australia had the smallest gap in the hourly wage, with a ratio of 93 percent. Britain and France's ratios were also high, at 82 percent and 83 percent respectively, however Japanese women earn less than two thirds the wage of their male counterpart. Comparatively, Australian women fare quite well to their counterparts in absolute proportion of male earnings. Cross-country calculations have also shown to be highly sensitive when the base country used for decomposition to separate the country effects. Alternatively the gender wage ratio is highly sensitive under the models whereby females are paid according to their counterparts in this study. The conclusion drawn is that rewards for endowments provide the largest source of the observed disparities between the countries.

### *The United States*

Blau and Kahn (2000a) conclude that the United States' male-female earnings ratio falls behind its OECD counterparts when compared in absolute levels. In 1979-1981, the United States' ratio was below the OECD average by 9 percentage points at 62.5 percent, while Australia exceeded the average by almost nine percentage points. Growth for the United States earnings ratio during the 1980's and 1990's outstripped the gains of fellow OECD members, with its ratio only marginally smaller than the OECD average at the end of 1998. Much of this convergence was attributed to decreases in the gap for labour market experience between the genders reported in Kidd and Shannon (1996). The growth of the United States earnings ratio, in percentage terms, was over twice that of Australia with Australia's female/male ratios for the weekly earnings of full-time workers increasing from 80 percent to 86.8 percent compared to the United States of 62.5 percent to 76.3 percent over the same period from the period spanning 1979 and 1998 (or a 0.068 and 0.138 percent change respectively).

Comparatively, the countries are similar along the dimensions of women's skill levels relative to men or the level of discrimination. However, the United States experiences higher commitment and/or attachment of women to the labour force<sup>6</sup>, Furthermore, the United States was at the forefront of equal pay initiatives implementing them considerably before other nations. This evidence of gender-specific factors and the gender-specific policies presented by Blau and Kahn (2000a) suggests that the disparity of the wage gaps between the United States and Australia should be smaller, however, they assert that institutional factors, namely the highly decentralised nature of the wage determination system in the United States is a large determinant of the disparity.

In particular, Australia's institutions of wage determination system including the ability of awards to cover multiple employers within industries or occupations, garner benefits for employees regardless of union status; and 'comparable wage justice'<sup>7</sup> stands in stark contrast to the United States. The United States' much higher levels of wage inequality, as evidenced by higher prices for skills and differentials in the wages paid by its sectors (Blau and Kahn 2000b) are also significant. In an earlier study, Blau and Kahn (1996) remarked upon the rising inequality<sup>8</sup> in the United States, which in the face of higher returns to measured and unmeasured skills would require women to witness improvements in both their skills and labour market treatment to maintain the constant gender wage gap. Furthermore, results showed that if the inequality prevailing in Australia existed in the United States, the gap between the wages of males and females would be similar to that in Australia (Blau and Kahn, 1996).

Differences in the gap between the two countries may be attributed to institutional frameworks and the characteristics of the labour market. In contrast to the United States, Australia has a high level of unionized employers and a highly decentralised system of wage determination, two institutional factors proven to be conducive towards reducing the wage gap due to resource effects which impact on productivity and employment as reported Blau and Kahn (1996). The high proportion of the wage gap within the United States attributed to inter-industry or inter-firm wage differentials demonstrates higher consequences for a decentralised system, while the system of minimum wage has aided Australia's to achieve a smaller wage gap (Blau and Kahn, 2000).

Compared to Japan, France and the United States, the earnings ratio of Australian women is very high. Although all these countries have implemented equal pay reforms, the decentralised nature of the wage

---

<sup>6</sup> This may be partly evidenced by its comparatively higher female participation rates and share of women employed in full-time work as reported in Blau and Kahn (2000a).

<sup>7</sup> The concept of 'comparable wage justice' as initiated in the policy changes tends to promote flow on effects as a means to maintain wage ratios across industries in Kidd and Shannon (1996), overall contributing to a smaller gap.

<sup>8</sup> This inequality was shown to arise both within and between both educational and experience groups, the impact of which is exacerbated when returns to skills will significantly change the observed gap.

determination system in Australia and high degree of unionisation has aided Australian women greatly in the pursuit of more equal pay (Blau and Kahn, 2000). After accounting for human capital differences between the genders across countries and wage structure effects, the returns to endowments, which is related to institutional structures, has been influential in the observed disparity in gender wage gaps between nations according to Daly et al (1996).

#### IV. INFLUENCES ON THE DIFFERENTIAL

The gender wage differential has been attributed to differences in skill levels and gender distribution across occupations reported in Borland (1999); and different labour market treatments in Le and Miller (2001).

##### *Gender-Specific Factors*

Differences in average qualification levels have been examined within the framework of the human capital model. This model relates the level of human capital investment to their expected working life according to Polachek (2004). Women overall have shorter expectations regarding both the length and continuity of their working lives (Blau and Kahn, 2000a) which in part relates to the gender division of labour and specialisation in the household sector as discussed in Neumark and McLennan (1995). For these reasons, the theory argues that women have less incentive to invest in both formal education and on-the-job training, the effect of which is exacerbated by the greater rate of depreciation women's skills undergo as a result of their absences from the labour market as a result of child-bearing (Vella, 1993). This has facilitated women's accumulation of less labour market experience as per Blau and Kahn (2000a) and tenure in Neumark and McLennan (1995).

Accordingly, the work decisions of women whereby they on average, work fewer hours signifies they invest less in human capital accumulation and are penalised with lower future earnings as discussed in Polachek (2004). Their discontinuous expectations of working life mean they are unwilling to make large investments and thus have a lower and flatter earnings profile relative to men according to Norris et al (2005).

The narrowing of the trend has therefore been affected by the simultaneous trend of rising female labour participation which has been occurring on a wide-scale base according to Polachek (2004). The implication in relation to the theory is that this has led to increased investments in human capital which in turn has reduced the gap between men and women's productive characteristics<sup>9</sup>.

The human-capital gap reported by Langford in 1995 comprises approximately 40-50 percent of the overall gap. With reference to specific characteristics, human capital factors account for between 23.91 to 32.66 percent. It has also been that since the implementation of equal pay policies, the gap in gender specific factors has fallen, perhaps as a consequence of the Sex Discrimination Act leading to increased investment in education and labour market experience as reported by Kidd and Meng (1997).

Overall, the human capital theory provides a valid explanation for the justified component of the wage gap whereby it validates women's increased investment and commitment to increasing their human capital. As a result of increased participation and subsequently working life expectation; and legislative reform have also served to women have been induced to increase human capital investments and thus reduce the gap in productive characteristics between the genders.

##### *Occupational Segregation*

In Australia, women have been most concentrated in nursing, reception and other clerical occupations Norris et al (2005). According to Fischel and Lazear (1986), this concentration may reflect either personal choice or 'crowding out'. The first as described in Fischel and Lazear (1986) proposes women choose jobs with attributes that will not conflict with household responsibilities, constituting a supply-side explanation according to Langford (1995).

Alternatively, 'crowding out' argues that women are segregated as a result of employer, or societal discrimination promoting some occupations as more acceptable as described in Borjas (2005). Thus women become concentrated in occupational niches according to Jones (1983). While it is difficult to distinguish between the two effects empirically, whether segregation arises from either source, does not alter the end

<sup>9</sup> This paper acknowledges that studies have been performed into the quality and nature of human capital investment, in particular the choice of higher education major. While the degree which is undertaken seems to matter and that women may be penalised for lacking skills in some areas (i.e. maths) in Langford (1995), they are not explicitly discussed here. This paper established only that the increased investment in women's human capital has reduced the gap.

result, which is increased concentration by women in some industries as attested to in Fischel and Lazear (1986).

Occupational segregation sorts groups by characteristics that become the basis for differential treatment (Reskin 1993). This leads to unfair treatment, in particular exclusion in some occupations and concentrations in others leading to male and female-dominated occupations. According to Norris, Kelly and Giles (2005), the traditional view between segregation and the wage gap was women's concentration in some industries characterised by their low wages was a large contributor to the perpetuation of the differential.

Indexes of segregation<sup>10</sup> (Reskin, 1993) allow calculation of the average earnings women would receive if distributed across occupations in the same manner as men. Trends signify that overall segregation has been on the decline<sup>11</sup>, however some authors have remarked that the fall in occupational segregation does not extend to job-level integration among the genders and that significant intra-occupation segregation has confined women to different and less desirable jobs (Reskin, 1993).

Previously, segregation was deemed to have raised average earnings of women (Norris et al, 2005, Kidd and Meng, 1997) with women's favourable distribution compensating for lack of human capital (Langford, 1995). However, Gunderson (1989) argues that studies have drawn inaccurate conclusion based upon broad aggregate classifications. Broad classifications underestimate inter-occupational effects upon the wage gap owing to the gender segregation of tasks within the aggregate that has been proposed as significant to the gap by Le and Miller (2001). Thus segregation remains a contentious issue, which has implication upon gap measurements, namely whether to model this variable as endogenous or exogenous.

### *Discrimination*

Wage discrimination pertains to consideration of non-productive characteristics when determining employee remuneration according to Langford (1995). Authors including Borland (1999) cite this as different returns for the same level of a characteristic. Predominantly, discrimination accounts for the largest part of the observed wage differential, but a significant fall from 44 to 13 percent between 1973 and 1989 supports the view that equal pay legislation eliminated a large proportion of institutional discrimination against women.

The model proposed by Becker in 1957 is the most widely cited theory of discrimination. It determined that there were three main sources from which discrimination could arise; employers, colleagues and customers as presented in Norris et al (2005). The distaste model specifies that in all cases, there is disutility incurred by interaction with the members of this group. In the case of employers, discrimination can present itself in many forms including promotion which may be in part influenced by the prominence of part-time employment in females. Connected to the human capital theory, Norris et al (2005) presents that part-employment satisfies gender considerations by offering flexibility, but hinders women as it introduces discrimination in the labour market due to unequal advancement opportunities in relation to limited career advancement opportunities, reduced human capital accumulation that results and perceptions of a commitment to career. The effect of co-worker distaste has not been examined, however, the presence of women in may retail and customer service occupations does not seem to indicate customer distaste as a significant source of discrimination.

A second model is that of 'statistical discrimination' which discriminates based upon the expected value of an individual's productivity which takes into account their membership of a certain group according to Norris et al (2005). Following on with the previous example, it may be that women are perceived by employers as more likely to quit based on historical trends. This can introduce statistical discrimination, as they perceive that any investment in skills upgrading may be futile due to women's traditional discontinuous and shorter working life. The traditional, intermittent characteristic of women to leave the workforce thus introduces a source of discrimination.

Traditional decompositions attribute discrimination as the largest component of the wage differential, accounting for approximately three quarters according to Le and Miller (2001). Specifically, education and experience are rewarded more highly for males and for females respectively as reported by Langford (1995). As such, decreasing gaps in labour market experience should benefit females more, through one avenue described in Blau and Kahn (2004) as signalling to employers a greater commitment or attachment to the

<sup>10</sup> The index of segregation most commonly used in the literature is the Duncan index as used in Reskin 1993, however others including Karmel/Machlachlan are also used to compare the distribution of genders across occupation

<sup>11</sup> For Australia and most industrialised nations, it appears segregation has declined slightly or remained fairly constant despite large increases in female participation rates as reported by Norris et al, (2005) and Reskin (1993)

labour force which reduces the basis for statistical discrimination against women<sup>12</sup>. The fall in discrimination can thus also be attributed to some extent on legislation in removing institutional discrimination and explicit gender discrimination, but the long term and relatively stable nature of discrimination means that legislation alone may not be sufficient to reduce it, and it may be through further signalling of women of their commitment that will produce further gains in their wages due to statistical discrimination reduction.

## V. DECOMPOSITION OF THE WAGE DIFFERENTIAL

The differential based upon raw earnings data has been attributed to two sources. The first relates to the differences in average skills or job types reported in Borland (1999), arising because women tend to possess comparatively fewer of the skills that have been theoretically proven to enhance future earnings potential in Le and Miller (2001). The second arises due to different treatment of the genders or difference in returns accruing to job and skill characteristics in Borland (1999). Economists have categorized these two components as justified and unjustified sources of the wage gap respectively, with the latter termed as wage discrimination in Borland, (1999).

### *Oaxaca Decomposition*

The traditional Oaxaca approach is used to decompose the gap into justified and unjustified effects. A brief outline of the manner in which the Oaxaca<sup>13</sup> decomposition is derived may be obtained by following Langford (1995).

The first step is to generate a model that includes gender discrimination. From Becker's model, discrimination against females results in a disparity between their wage and marginal productivity. Importantly, employers perceive that the wage paid to a female worker is greater than the true monetary cost of hire described in Borjas (2005). The relevant equations where women are discriminated against is described by the relationships between marginal productivity and wages for the genders<sup>14</sup>.

$$\begin{aligned} MP^F &= (1+d) w^F & (1a) \\ MP^M &= w^M & (1b) \end{aligned}$$

By taking the quotient of marginal product and taking logarithms, an equation separating discriminatory influences, and the genders' differences in marginal productivity is obtained:

$$\ln w^M - \ln w^F = \ln(1+d) + \ln MP^M - \ln MP^F \quad (2)$$

Human capital theory relates the log of an individual's earnings as a function of their 'productive characteristics' represented by  $X$  and  $b$  representing the returns to characteristics<sup>15</sup>. The wage rate for each gender is then modelled as:

$$\ln w_i^s = b^s X_i^s \quad (3)$$

where  $s =$  male (M), female (F) in the previous notation  
 $i = 1 \dots n$  individuals

<sup>12</sup> In the presence of feedback effects, it stands to reason that a positive review of employer's perceptions in regards to women's commitment will serve to increase their attitudes towards skills upgrading for women in the workplace as discussed in Blau and Kahn (2000a) as they perceive less risk of not receiving the full return on their investment. In accordance with human capital theory, this should help narrow the gap also.

<sup>13</sup> The Oaxaca decomposition assumes that both wage and occupational status are exogenously determined. Alternate decompositions such as that of Brown, Moon and Zoloth as used in Kidd and Meng (1997) endogenise these variables. Given the view of this paper in that occupational segregation is overall fairly insignificant to the wage gap, this approach is recognised but not discussed.

<sup>14</sup> These equations assume that women are discriminated against, but the hiring decision pertaining to male labour is unaffected by nepotism (whereby employers perceive the wage of a particular group is lower than the true monetary costs of hiring)

<sup>15</sup> The underlying intuition is that  $X$  which represents the individual's productive characteristics is an approximation of their marginal productivity (Langford, 1995)

Following this, Mincer's earnings function modelled the log of an individual's earnings as a regression function with education (ED), post-school labour market experience and Z which contains other explanatory variables with the potential to influence the wage gap. Education and experience have been associated with greater wages and thus are explicit explanatory variables even in the reduced form of Mincer's equation:

$$\ln w_i = b_0 + b_1ED + b_2EXP + b_3EXP^2 + b_4Z + \epsilon_i \quad (4)$$

Oaxaca combined Mincer's findings and the earlier notion of discrimination to partition the gender wage differential into the two components. The decomposition is performed by adding and subtracting the term  $(b^M \times X^F)$ <sup>16</sup>

$$\ln w^M - \ln w^F = (X^M - X^F)b^M + (b^M - M^F) X^F \quad (5)$$

This form is appropriate when the male wage structure is used as the basis for comparison and thus examines the amount a woman would earn if she were treated like a man. Similarly, the female wage structure could be used as the base wage:

$$\ln w^M - \ln w^F = (X^M - X^F)b^F + (b^M - M^F) X^M \quad (6)$$

### *Limitations and Alternative Measures*

A potential limitation to this approach remarked upon by Neumark (1988) is that the male wage structure as the basis may not be appropriate due to its assumption that women would receive equal returns for human capital investments when the market operates free of discrimination and thus the differential may differ markedly according to Gyimah-Brempong, Fichtenbaum and Willis (1992). Furthermore, when the wage paid is inconsistent with the value of a worker's marginal product, eliminating discrimination will lead to an equilibrium wage inconsistent with the wage paid to either gender (Gyimah-Brempong et al, 1992). This is especially relevant to (6), where if discrimination perpetuates premiums for males and lower wages for females<sup>17</sup>, elimination of discrimination will, ceteris paribus, lead to increased and reduced demand for female and male labour respectively. This results a discrimination-free wage higher than the female, but lower than the male wage (Langford 1995, Gyimah-Brempong et al (1992). Finally, it is sensitive to the variables used as determinants of productivity (Borjas, 2005) and omitting of potential explanatory variables will result in incorrect calculations. Further exacerbating this problem is the immeasurable nature of characteristics that could give rise to differences between males and females (Borjas, 2005).

In response, an alternative method that incorporates the no-discrimination wage structure is reported. Developed by Neumark (1988), the results are if the no-discrimination wage is equal to the male wage rate then the gender gap will equal that reported by equation (5) and vice versa if the no-discrimination wage is equal to the female wage rate<sup>18</sup>. The advantage of this over the traditional method is its ability to decompose the gap into differences in endowments; and also characteristic effects, comprised of the advantage of males in their treatment and the disadvantage female treatment effects.

## VI. CONCLUSION

The narrowing of the gender wage gap in Australia following the Equal Pay reforms is indisputable. However, the persistence of the wage gap into the 1990's has led studies into the underlying causes. Whilst female participation has been rising, and increases in human capital investment and experience by women have served to reduce the justified proportion of the wage gap, the substantial gap that remains points to discrimination inherent within the labour market. Discrimination which can manifest in many forms appears to have fallen since the introduction of equal pay reforms which signify the impact of institutional factors on the gap, yet a substantial amount remains in that there is disparity between the returns to characteristics and the perceptions of employers in particular.

<sup>16</sup> This is an algebraic manipulation which does not serve to change the overall value, but will allow the raw wage differential to be written as per equation (5). It follows the reasoning in Borjas (2005)

<sup>17</sup> In accordance with Langford (1995), if males command a premium, discrimination exists, there is higher demand for male hire, while lower demand for female labour accompanies their lower wage.

<sup>18</sup> For the derivation and assumptions behind this method, refer to Neumark (1988) or Langford (1995).

In light of legislative changes, the occupational segregation in Australia seems to have altered very slightly. In the Australian context, occupational segregation has been treated as either as slightly benefiting or have a neutral implications for the wage gap, however, recent studies may require a reassessment. Overall consensus seems to promote the negligible effects of occupational segregation, thus legislation has contributed fairly little to the narrowing of the gap through its effects on segregation.

While the gap in Australia remains a significant area of research, cross-country comparisons emphasise that the standing of Australian women relative to OECD counterparts is very high. Women in Australia have benefited from a highly centralised wage determination system that has seen their earnings ratio remain at a level higher than that observed in Great Britain, the United States, France or Japan. Thus country-specific factors, especially institutional characteristics are highly influential upon the observed wage gaps in countries.

In relation to measurement of the differential, the Oaxaca method traditionally used has been questioned upon its assumptions and treatment of the occupation segregation; however the findings of this paper are in agreement with the view that occupational segregation is not a highly influential variable and thus the Oaxaca decomposition method is a valid measurement.

## References

- Blau, F.D. and Kahn, L.M. (1996). 'International Difference in Male Wage Inequality: Institutions versus Market Forces', *The Journal of Political Economy*, vol.104.
- Blau, F.D. and Kahn, L.M. (2000a). 'Gender Differences in Pay', *Journal of Economic Perspectives*, vol. 14.
- Blau, F.D. and Kahn, L.M. (2000b). 'Rising Wage Inequality and the U.S. Gender Gap', *American Economic Review*, vol. 84.
- Borjas, G. J. (2005), *Labour Economics* (Boston: McGraw-Hill/ Irwin)
- Borland, J. (1999). 'The Equal Pay Case – Thirty Years On', *The Australian Economic Review*, vol. 32.
- Daly, A., Meng, X, Kawaguchi, A. and Mumford, K. (2006). 'The Gender Wage Gap in Four Countries', *Economic Record*, vol. 82
- Fischel, D.R. and Lazear. (1986). 'Comparable Worth and Discrimination in Labour Markets', *The University of Chicago Law Review*, vol. 53
- Gunderson, M. (1989). 'Male-Female Wage Differentials and Policy Responses', *Journal of Economic Literature*, vol. 27.
- Gyimah-Brempong, K., Fichtenbaum, R., and Willis, G. (1992). 'The Effects of College Education on the Male-Female Wage Differential', *Southern Economic Journal*, vol. 58.
- Jones, F.L. (1983). 'Sources of Gender Inequality in Income: What the Australian Census Says', *Social Forces*, vol. 62.
- Kidd, M.P. and Shannon, M. (1996). 'Does the Level of Occupational Aggregation Affect Estimates of the Gender Wage Gap?' *Industrial and Labor Relations Review*, vol. 49.
- Kidd, M.P. and Meng, X. (1997). 'Trends in the Australian Gender Wage Differential over the 1980: Some Evidence on the Effectiveness of Legislative Reform', *The Australian Economic Review*, vol. 30.
- Langford, M.S. (1995). 'The Gender Wage Gap in the 1900's', *Australian Economic Papers*, vol. 34.
- Le, A.T. and Miller, P.W. (2001). 'The Persistence of the Female Wage Disadvantage', *The Australian Economic Review*, vol. 34.
- Polachek, Solomon W. (2004). 'How the Human Capital Model Explains Why the Gender Wage Gap Narrowed', Institute for the Study of Labour (IZA), Discussion Paper, no. 1102, Germany, April.
- Neumark, D. (1988). 'Employers' Discriminatory Behavior and the Estimation of Wage Discrimination', *The Journal of Human Resources*, vol. 23.
- Neumark, D. and McLennan, M. (1995) 'Sex Discrimination and Women's Labour Market Outcomes', *The Journal of Human Resources*, vol. 30.
- Norris, K, Kelly, R. and Giles, M. (2005), *The Economics of Australian Labour Markets* (Australia: Pearson Addison-Wesley Australia)
- Pocock, B. (1999). 'Equal Pay Thirty Years On: The Policy of Practice and Theory', *Australian Economic Review*, vol. 32.
- Reskin, B. (1993). 'Sex Segregation in the Workplace', *Annual Review of Sociology*. vol. 19
- Strachan, G. and Jamieson, S. (1999). 'Equal Opportunity in Australia in the 1990's', *New Zealand Journal of Industrial Relations*, vol. 24.
- Vella, F. (1994). 'Gender Roles and Human Capital Investment: The Relationship between Traditional Attitudes and Female Labour Market Performance', *Economica*, vol. 61



## JOURNAL ARTICLE REVIEW

AMY WAGNER\*

Vaithianathan, R (2004) 'A Critique of the Private Health Insurance Regulations' *Australian Economic Review*, **37**(3), pp. 257-270.

### 1. Background

Maintaining private health insurance coverage of Australians is seen as a critical strategy in preventing the overloading of the public health system. In 1997 the Private Health Insurance Incentives Scheme was implemented by the Australian government. It included three stages. The first involved a tax surcharge applied to high income earners who did not have private health insurance and a partial refund of premiums to low income earners. The second policy implemented was a subsidy of thirty percent on health insurance premiums regardless of income. The third policy change involved, "...the liberalisation of the regulations on policy benefits and community rating." (French, Hopkins and Macdonald, 2003, p. 59). Vaithianathan (2004) focuses primarily on the second two policies implemented.

### 2. Summary

Vaithianathan (2004) argues that due to the highly regulated nature of the Australian private health insurance market, the Private Health Insurance Incentives Scheme "...is unlikely to achieve the government's policy objective of a popular and profitable private insurance industry." (Vaithianathan, 2004, p. 258). She proposes that, in the presence of community rating, "...a subsidy in a competitive insurance market would result in a fall in the equilibrium (premium) price – further expanding coverage." (Vaithianathan, 2004, p. 259) and hence the thirty percent rebate should have been highly effective in increasing coverage. However, it was found that this policy was only slightly effective. Vaithianathan (2004) suggests that this was because "...community rating was ineffective." (Vaithianathan, 2004, p. 257). She also argues that the annual penalty for delaying insurance of two percent<sup>1</sup> is not high enough to be effective in encouraging people to join private health insurance when they are younger. Vaithianathan (2004) finds that insurance companies are targeting particular risk pools of consumers by offering a variety of plans with differences in coverage and premiums. Finally, she proposes two alternative policies; a risk-adjusted subsidy scheme and complete deregulation.

### 3. Underlying Assumptions

In her analysis, Vaithianathan (2004) assumes that health insurance industry is competitive. The validity of this assumption is questionable. The characteristics of a competitive market include; all firms being price takers, the products offered by firms being homogenous, firms being able to enter and exit the market freely and consumers being able to switch freely between suppliers (Pindyck and Rubinfeld, 2005). Whilst the market for private health insurance meets several of these criteria, it does not meet all of them. For example consumers should be able to move freely between insurers. However, health insurance companies often impose waiting periods during which time new customers are unable to claim for certain services. This could discourage consumers from moving between companies. The ability of firms to enter and exit the market freely could also be questioned. Once consumers have signed up with the insurance firm, the firm is required to fulfil its contractual obligations to the consumer until the end of the period for which they have paid to be insured.

Vaithianathan (2004) bases her models on the findings that the tendency of a person to want private health insurance depends on their age and income. However, she does not take into account other characteristics of people that influence their desire for health insurance. According to the Australian Bureau of Statistics (1998, p. 4) the proportion of coverage depends on the 'contributor unit type'<sup>2</sup>. Temple (2006, p. 63) finds that "... there is considerable variation in coverage by sex." The variation of coverage by sex seems unlikely to affect results, as gender numbers in Australia are approximately equal.

---

\* Written for ECON2030 *Microeconomic Policy*, Sem 2 2006.

<sup>1</sup> This was introduced as part of the lifetime community rating policy.

<sup>2</sup> Possible contributor unit types include; contributor only, contributor and dependent children, contributor and partner and contributor and partner with dependant children. Refer to Australian Bureau of Statistics (1998).

Vaithianathan (2004) draws conclusions regarding what she thinks are the primary motivators of the government's policy on private health insurance. She suggests that the two key principles are, "Everyone should have equal access to private health insurance." and "A sustainable and popular private insurance system is important for long-term viability of public hospitals." (Vaithianathan, 2004, p. 265). Whilst the report by the Parliament of Australia Senate (1998) confirmed the second of these principles, it does not seem to support (or reject) the first principle Vaithianathan (2004) proposes. However, it seems likely that they would support this first principle because if all the people in the high-risk categories were to use the public health system it would become overloaded.

In the models used, Vaithianathan (2004) assumes most people are risk adverse. It seems likely that people would consider insurance because they are risk adverse. However not all people would be risk adverse. This discrepancy could have some impact on the results, however it seems likely that it would not be large.

By suggesting that Australia deregulate its private health system like New Zealand, Vaithianathan (2004) is implying that Australia and New Zealand are similar with respect to variables that affect the private health insurance market such as demand for health care (both public and private), valuation of private health care by the public, costs of providing health care and competitiveness of the private health insurance market. To confirm this, further research would be required.

#### **4. Implications for Australia**

In Australia, health insurance companies are prohibited from including age in premium calculations. One of the conclusions Vaithianathan (2004) reaches is that health insurance companies are avoiding this restriction by offering a variety of different cover plans (with different premiums) which are designed to target different age groups. This implies that as people age, they will be obliged to change their insurance plan type to gain coverage for the areas they need and in doing so will be forced to pay higher premiums as they age. This defeats the purpose of lifetime community rating.

Various sources including French, H., Hopkins, S. and Macdonald, G. (2003), suggest that over time the percentage coverage of the Australian public by private health insurance will decline over time unless further action is taken. If research<sup>3</sup> finds that the variables that affect the private health insurance market in Australia and New Zealand are in fact very similar, then it may be that Australia would be better off with a deregulated market for private health insurance. However, this seems unlikely due to the problem of adverse selection in the private health insurance market. Complete deregulation could possibly cause more people to drop out of private health insurance, thus increasing pressure on the public health system. Vaithianathan (2004) acknowledges that deregulation of the private health insurance industry may cause premiums for some groups of consumers to increase dramatically. The unpopularity complete deregulation of the private health industry would be likely to face means that politicians are unlikely to propose this policy.

The risk adjusted subsidy scheme, involves paying a subsidy to insurers that makes the insurers indifferent between consumers with different amounts of risk. This would mean that insurers would not have to seek to balance their risk pool, thereby preventing premiums paid by individuals from increasing as they age and enabling private health insurance to maintain profitability. This seems like a more favourable scheme than complete deregulation of the private health insurance industry, as it will help to maintain the percentage of coverage of Australians by private health insurance. This will help to keep pressure off the public health system.

#### **5. Conclusion**

In 'A Critique of Private Health Insurance Regulations', Vaithianathan (2004) proposed that the thirty percent subsidy of private health insurance introduced as part of the Private Health Insurance Incentive Scheme should have been highly effective at increasing the percentage of people covered by private health insurance. She then suggested several reasons why this was not the case. These reasons included insurers using plan design to screen customers and the penalty for not obtaining insurance by the deadline for lifetime community rating not being high enough. She goes on suggest two alternative schemes; a risk adjusted subsidy scheme and complete deregulation of the private health insurance market. Whilst her analysis has some failings, it is obvious that the current public health coverage situation is unstable and other strategies will need to be considered to prevent adverse selection and other factors making the private health system unviable in the long term. Complete deregulation does not seem favourable due to the questionability of the

---

<sup>3</sup> Refer to the research suggested in section 3 of this paper.

competitiveness of the private health insurance market and the probable increase in pressure on the public health system. However, risk adjusted subsidies seems as though it might be a policy worth considering if the percentage private health insurance coverage in Australia continues to decline.

## References

- Australian Bureau of Statistics (1998) *Health Insurance Survey*, Cat. No. 4335.0, Australian Bureau of Statistics, Canberra.
- French, H., Hopkins, S. and Macdonald, G. (2003) 'The Australian Private Health Insurance Boom: Was It Subsidies or Liberalised Regulation?' *Economic Papers*, **22**(1), pp. 58-64.
- Parliament of Australia Senate (1998) *Report: Private Health Insurance Incentives Bill, Private Health Insurance Incentives Amendment Bill, Taxation Laws Amendment (Private Health Insurance) Bill*, Parliament of Australia Senate, Canberra, viewed 10<sup>th</sup> October, 2006, <[http://www.aph.gov.au/senate/committee/clac\\_ctte/completed\\_inquiries/1996-99/ptehealth/report/index.htm](http://www.aph.gov.au/senate/committee/clac_ctte/completed_inquiries/1996-99/ptehealth/report/index.htm)>.
- Pindyck, R and Rubinfeld, D (2005) *Microeconomics*, 6<sup>th</sup> edn, Pearson Education Inc., Upper Saddle River.
- Temple, J. (2006) 'Health Insurance Reform and Older Australians' *Australasian Journal on Aging*, **25**(2), pp. 63-68.
- Vaithianathan, R (2004) 'A Critique of the Private Health Insurance Regulations' *Australian Economic Review*, **37**(3), pp. 257-270.

## JOURNAL ARTICLE REVIEW

MADELINE VEENSTRA\*

Vaithianathan, R. (2002) 'Will Subsidising Private Health Insurance Help the Public Health System?' *The Economic Record*, **78**(242), pp.27-283.

### 1.0 Introduction

The article under discussion puts forward the argument that by increasing subsidies for private health care this will not guarantee a decrease in pressure on the public health system.

### 2.0 Main Message

The author uses a simple model to deduce that although these subsidies on private health insurance may 'expand' the number of consumers using private insurance coverage, this may not necessarily mean a reduction of demand for public health services. It is stated in the article that many governments have come to rely on this method in order to decrease the increasing pressure on the public health system. However the question of whether or not this process is 'helpful or harmful' is an ever pressing and most important issue that must be addressed. The author suggests that it is not the subsidies that will have the desired effect on the public health system, but the change in regulations that are currently governing the private health system.

The decline in the percentage of Australians who had private health insurance, fell by 20 points between 1985 and 1999. It was this occurrence that caused the government's concern over the ever-decreasing stability of demand for public health services, which subsequently led to the introduction of the Private Health Insurance Incentives Act (1998). It was argued by the government that the universal rebate of 30 per cent on health insurance premiums would cause an increase in the usage of the private health system. This thinking arose due to idea that consumers who originally were unable to purchase private health insurance due to their lack of purchasing power were now able to do so. The government reasoned that *ceteris paribus* the pressure on the public system would decrease solely because of this policy introduction. However the effects of this policy have been argued to have little effect. There are numerous reasons for this, which have been discussed throughout the article. It has been suggested that public hospitals are more efficient than private, hence the lack of desire from the population to purchase private health insurance. A welfare loss will be the result if the government encourages a switch from public to private services, without first ensuring that the consumers will benefit from suitable services.

As with any market, the market for private health insurance is hindered due to externalities. One that is under discussion in the article is the idea of consumers having imperfect information. Not all consumers will purchase private health insurance, as they do not know whether or not that they have the need for it. They do not anticipate illness so therefore they do not purchase any insurance, regardless of government rebates. Another significant problem in the private health care market is adverse selection. This causes the issue of 'unfair' premiums on private health insurance. Because of community ratings all consumers must be charged one price, essentially pricing many consumers who may add little cost to the insurance company out of the market. The author raises these issues of externalities to highlight the fact that the utilisation of private health insurance and the private system is a complicated issue and is not one that can easily be fixed under one simple policy.

As stated earlier the author suggested that although there may be an expansion within the market for private health insurance, this does not necessarily mean a reduction in demand for public health services. This statement was deduced due to the fact that price is not the only contributing factor to the consumption of private health care; there are other qualitative factors that also contribute substantially. Some of these factors include that the public system offers a more comprehensive service. Another reason for the pressure on the public system is the fact that many individuals have ideological reasons for using public health care and as a result are reluctant to use the private system.

The author suggests that there is a heterogenous behaviour of consumers when purchasing health care, and that they may be broken into two groups. These groups include publics and privates. The publics will use the public system in the absence of insurance. The presence of insurance however may encourage them to use the private system. The privates would use private care in the absence of insurance and would thus be more

---

\* Written for ECON2030 *Microeconomic Policy*, Sem 2 2006.

willing to purchase health insurance if it were at a reasonable price. This group in particular is hindered from purchasing private health insurance and thus they opt to self-insure instead. It would therefore be more beneficial to liberalise the premium policies, and allow these consumers to purchase policies for prices they deem more acceptable.

The model used by the author has deduced that a considerably higher subsidy would have the desired effect of expanding private coverage of the public, which would then relieve the burden on the public system. It would seem more reasonable if the government were to subsidise private health care and not insurance. This would have the desired effect of easing the burden on the public health system.

One concerning issue that is raised is the possible transfer effect that the current subsidy may have between the public and private system. It is stated that a subsidy may cause consumers due to the subsidy incentive to over consume private health care. This in turn may offset the gains of the decrease in pressure on the public system, by causing an increasing instability of the private market.

Due to the community-rated insurance system, the subsidy will have little effect on drawing the public group away from the public health system. They may only purchase insurance at the actuarially fair price, which is not currently offered due to the universal premium prices.

The overall meaning from this article is that a price subsidy or reduction on private health care would be more effective than an insurance subsidy. This option may prove to alleviate the burden on the public system, rather than current insurance subsidy policies have done. Vaithianathan even goes as far as to suggest that the health insurance subsidies will have the opposite effect of pushing more consumers into the public health care system in order to avoid the expected rise in private insurance premiums. This is one issue that the government must take careful steps to rectify. A balance must be found between the two systems, and a health care subsidy as opposed to an insurance subsidy offered to attract more public consumers into the private health system.

### **3.0 Evaluation of the Underlying Message**

In Frech et al (2003, pp. 58.), it is found that the price subsidy introduced in the period of July 1997 to July 2000 did not increase the overall number of the population that had private coverage. The effect of this subsidy was intended to decrease the pressure of the public system. This was a part of three-policy package over three years. This included a two-part process of price subsidies as discussed in the article under review and a liberalised regulation of private health coverage. However it was found in Frech et al that the increase in coverage was due to the liberalisation of regulation not the price subsidies. It was further stated, "The increase was commensurate with the existing estimates of the price elasticity of demand for health insurance" (2003, pp. 60). The response to the means-tested price subsidy, intended to increase the coverage of low-income earners was disappointing. It was found that the impact of price subsidies was off set by the increase in insurance premiums and the expectation of further price premium increases. During the time of introduction (June 1997- December 1998), the number of people who had private health insurance *declined* from 31.9 to 30.1 percent. The following price subsidy that was not means tested also proved to have dismal effects on increasing the number of Australians with private health insurance (Frech et al, 2003, p.61). During this policy introduction the number of people whom had private health insurance increased by only 1.2 per cent. It was the introduction of the lifetime community rating that had impressive effects. This caused an increase of the population who had private health insurance by 42.8 per cent in 2000.

As illustrated in Hall's article (2001, pp.320) the market for private health care inevitably as will most other markets be plagued by market externalities. The problems that she discusses are a good indicator and strong argument for the article under review's main message. As stated in the article under review, consumers suffer from imperfect information regarding their health, causing them to under consume private health care. In Gans (2004, pp.473) it was suggested that the market for private health insurance is affected by adverse selection due to the ever increasing number of people who consume private insurance being of a higher risk in terms of health, thus pushing premiums higher and causing other people who are risk averse to go without private insurance. Hall's article (2001) also raises the idea that health and health care are two very different concepts and do not necessarily come hand-in-hand (Hall, 2001, pp. 322). Hall further states that the choice between public and private systems will come down to consumer preferences and their search for 'effective' health care as opposed to health care in general (Hall, 2001, pp.323).

The event of purchasing health insurance may be a distressing issue for some consumers and can be classified as a bad consumption, one that will involve 'apprehension or pain'. Thus many consumers may avoid any thought to health care insurance let alone purchase of an effective policy for their individual needs, this is described by Hall as a process utility (2001, pp.324). This issue further highlights the fact that

consumers are ill equipped to consume adequate amounts of health care insurance and thus must rely on the public system if suddenly becoming ill (Brown et al, 005, pp.281).

#### 4.0 Implications for Australia and on a global level

In Frech et al a most important statement was made, that more liberalisation of insurance products may be necessary to avoid the downward trend that occurred in the mid 1980s to 1997. In another article by Vaithianathan (2004, pp.257), two options for reform to the private health insurance system were put forward. These include a complete deregulation of the private insurance market. This however may be a more radical approach to the problem. This system has currently been introduced in New Zealand and Canada (Gwen Gray, 1998, pp.36) the latter to not so great an effect as New Zealand's deregulation. It has been shown that as long as the market is priced competitively that it will work to its best advantage. The second option is to introduce a partial reform that would allow the government to be able to target the subsidy towards those that would otherwise cause too much pressure on the public system.

Several sources have suggested that the subsidy may cause an undesired transfer effect, hence causing too much pressure on the private health care system and pushing private insurance premiums to unstable levels. In an attempt to adjust the problems that are occurring within the Australian health system, Gans and King (2003, pp. 473) suggest that the private insurance system be offered as a purely supplementary product. It is argued that this would remove the current "anti-insurance" behaviour amongst the two systems and allow more consumer choice between the public and private systems (Gans and King, 2003, pp.490).

The USA is another country that is facing difficulty with over priced premiums as a result of age-based adverse selection (Brown and Connelly, 2005, p.281). It has been thought that a policy called the Guaranteed Renewability Scheme may be the answer to this problem. The scheme is thought to allow consumers to insure against future increases in health insurance premiums and insure against expected medical expenditures as a result of illnesses incurred during previous periods.

However it has been found from a collection of sources, that the Australian and other OECD countries' health insurance systems cannot be stabilised by one policy. Nor can that policy create a perfect symmetry between the public and private systems. Further research must be conducted to find a more effective policy reform.

#### 5.0 Conclusion

The main message discussed throughout the article, is the ideal that the current government subsidy for private health insurance, may not have the desired effects on alleviating the public health system. It is further discussed, that the policy may in fact have a transfer affect and cause unwanted pressure on the private system causing private insurance premiums to rise, thus pricing the low-income earners out of the private health market. Throughout the supplementary articles it has been suggested that a liberalisation of the private insurance premiums be introduced, as to charge an actuarially fair premium price to consumers. Another residing policy is to offer rebates for private health *care* as opposed to health insurance. The most important factor among this issue is to remember that not one policy can work to alleviate pressures from the public system and that consumers have preferences which are not always dictated to by price.

#### References

- Brown, H.S., Connelly, L.B. (2005) Market Failure in long-term private health insurance markets: a proposed solution, *Applied Economic Letters*, **12**, pp.281-284.
- Frech III, H.E.; Hopkins, S. and Macdonald, G. (2003) The Australian Private Health Insurance Boom: Was it Subsidies or Liberalised Regulation?, *Economic Papers*, **22**(1), pp.58-64.
- Gans, J. (2004) Does Australia's Health Insurance System Really Provide Insurance?, *Policy*, **20**(3), pp.1-6.
- Gans, J.S., King, S.P. (2003) Anti-Insurance: Analysing the Health Insurance System in Australia, *The Economic Record*, **79**(247), pp.473-486.
- Gray, G. (1998) Access to Medical Care under Strain: New Pressures in Canada and Australia, *Journal of Health Politics, Policy and Law*, **23**(6), pp.1-44.
- Hall, J. (2001) Health, Health Care and Social Welfare, *Australian Economic Review*, **34**(3), pp.320-31.
- Palangkaraya, A., Yong, J. (2005) Effects of Recent Carrot-and-Stick Policy Initiatives on Private Health Insurance Coverage in Australia, *The Economic Record*, **81**(254) pp.262-272.
- Vaitianathan, R. (2004) A Critique of the Private Health Insurance Regulations, *The Australian Economic Review*, **37**(3), pp.257-70.



## Careers and Cocktails



**What:** ESA Students and Employers Networking

**When:** Tuesday, March 20 @ 7pm.

**Where:** Bribie/Moreton Room, Holiday Inn, Roma St.

**Cost:** \$15 members  
\$20 non-members

<http://www.uqesa.com>

