

ISSN 0819-2642
ISBN 0 7340 2575 0



THE UNIVERSITY OF MELBOURNE
DEPARTMENT OF ECONOMICS

RESEARCH PAPER NUMBER 919

OCTOBER 2004

**PERSUASION: REFLECTIONS ON ECONOMICS,
DATA AND THE 'HOMOGENEITY ASSUMPTION'**

by

Adam FForde

Department of Economics
The University of Melbourne
Melbourne Victoria 3010
Australia.

Persuasion: Reflections on Economics, Data and the 'Homogeneity Assumption'

Adam Fforde

Forthcoming, *Journal of Economic Methodology*, 2004

Melbourne Institute of Asian Languages and Societies, The University of Melbourne

Email: Fforde@unimelb.edu.au

Words – 7875 plus Bib and tables @ Wednesday, 22 September 2004

This work draws upon a wide range of experience in policy advice and economic analysis, mainly relating to Vietnam. From this I owe much to many people too numerous to mention. For the work presented here, I thank for comments on earlier drafts: D. Gale Johnson, Adrian Wood, John King, Gus Edgren, Warwick McKibben, Doug J. Porter, Stefan de Vylder, Peter Robson and various anonymous referees, participants at an ADB seminar in Sept 2001, and the Experimental Economics "Global Club" for happily sharing materials with me. I thank Valerie Tan for Research Assistance. The mistakes of interpretation remain mine.

Abstract:

This paper discusses issues to do with the empirical basis of modern economics and points towards the need to look more closely at the 'homogeneity assumption' that underpins much economic theory. It argues that severe problems currently prevent economics from becoming more persuasive to both students of economics and those outside the discipline. The issue involves the management of disciplinary boundaries, and excessive use of the 'homogeneity assumption.' Three areas of concern are explored. First is the literature on causes of growth, and the role of policy. The paper documents reasons to doubt the existence of robust relationships between growth and policy variables. Second is the 'homogeneity assumption' that different countries are usefully viewed as members of a single population. Third is evidence suggesting that an assumption of 'normal' maximizing behaviour has to be justified, not just assumed, and that regular deviations from the usual maximizing assumptions occur with gender and culture.

The paper argues that a central issue in economic methodology and pedagogy should be, as North implicitly argues, the negotiation of disciplinary boundaries: what economics can versus cannot explain. It suggests more explicitly basing the choice of explanatory models on empirics identifying where the model applies.

Key words: homogeneity, neo-institutional economics, robustness testing, methodology, policy advice, experimental economics.

Persuasion: Reflections on Economics, Data and the 'Homogeneity Assumption'

Introduction

Preface

Economies are very important, obviously. It is important to know what can be understood about them, and economists are largely responsible for trying. Yet, economists disagree, often greatly, about many things, and their approach is much criticized by other disciplines. How persuasive, then, is what is said? How can we judge when we should allow ourselves to be persuaded? This paper attempts a contribution to such judgments.

I take a qualified if not agnostic position, believing, for example, that inference is above all something that is done (Winch 1992), meaning that a logic cannot 'take you by throat' and necessarily force agreement with its particular rules. To accept a proposition is to choose, or, in a more civilized - or multi-polar - world, to be persuaded. It follows, for me, that, rather than advancing clear yes or no answers to how persuasive or unpersuasive certain economic arguments are, it is wiser to present a discussion. I have found also that this approach has practical and concrete significance in the preparation and delivery of policy advice. It is unwise to give suggestions too readily, partly because people often react adversely to unwarranted certainty, and partly because they have learnt from bitter experience that confidently expressed advice has often turned out to be wrong. It is often useful to phrase policy advice in terms of the 'contingent' – for example, an openness policy may work in societies of type X, but not in societies of type Y.

In this paper I discuss two significant issues: the assumption of homogeneity, and an example of how economists have responded to contradictory evidence related to this assumption. My interest in the second is partly because this informs others' judgments of us as applied economists. I start, though, by discussing why this assumption is interesting.

Economics

Economics forms one of the central disciplines of modern social science. Examination of textbooks shows that economic theory makes strong assumptions about what does and what does not vary in the populations it studies: micro economics deals with consumers, firms and commodities; macro economics with nations, GDP and so on. For example, the behaviour of consumers, is, at least initially, discussed in abstraction from their gender, age and culture; the relationship between trade and growth is discussed in abstraction from issues such as historical context, culture and so forth. Of course, these issues can be and are brought into the analysis, but normally as a second step, and there is a strong tendency within economic theory to use models that are highly 'universal'.¹ There is a tendency to assume 'homogeneity', or, more precisely, 'ontological and epistemological universalism': I call this the 'homogeneity assumption'.² I discuss these terms below. There is a long history of arguments that criticize this approach a priori. These have not on the whole been found persuasive. In this paper, however, I discuss empirical work that implies a need to reconsider heterogeneity.

The 'homogeneity assumption', and the associated universalism, encourages lack of self-awareness, and so economics faces problems with negotiating disciplinary

boundaries.³ The assumption discourages dialogue since it must, if applied consistently, downplay the differences that other disciplines treat as central: cultural meanings, socially constructed institutions, gender, historical change, and so on. Economic judgments risk losing their persuasiveness, for, too often, the 'homogeneity assumption' is used without being justified empirically. Further, in statistical work, assuming in a technical sense that a sample is drawn from a homogeneous population, so as to apply a theory that assumes such homogeneity, may risk presenting spurious statistical results, when significance tests assume homogeneity to exist, yet it does not. I discuss this further below.

I argue, therefore, that poor attention to the negotiation of disciplinary boundaries is an indicator of deeper problems reflected in method and the treatment of data, and in a still rather underdeveloped ability to shift between universal and local explanatory frameworks. Until the issue of disciplinary boundaries is resolved, these weaknesses in generating persuasive arguments will likely therefore remain.

My examination of the literature suggests, albeit weakly, that while there are pressures within the discipline to move away from the 'homogeneity assumption', these remain fragile.

It seems unwise to teach theory when the teacher knows, and the literature shows, that it is badly supported by the data. This holds more forcibly for policy advice, especially that given to foreigners, who often will argue strongly for their being different from oneself. Persuasion seems often to require an overt acceptance of difference. Yet, consistent with the 'homogeneity assumption', economists usually assert the existence of cause-effect policy outcomes based upon a posited universal

'economic logic' applying across countries and time.

Basic issues and two core propositions

'Reversals' and the history of economics

Goldfarb 1997, in an unusual and valuable paper, examines shifts in established views amongst economists as to the nature of empirical reality, seeking to document 'emerging contrary results' and the causes (221). He contrasts the view that 'theory-mindsets determine beliefs' with the view that 'theories stand or fall by their ability to predict what the data reveal' (id.).⁴ His examination of papers in the *American Economic Review* reveals that around 10% of articles contain emerging contrary results (229). Further, he notes that the main explanation for such reversals (around 65% of his examples) is the availability of more data and different techniques. However, he asserts that the exact ways in which data influences beliefs is not clear, and that 'the relative fragility of empirical findings suggested by the existence of so many (emergent recalcitrant results) makes it more likely that theoretical preconceptions will be relatively impervious to empirical onslaughts' (238).

This view is both too pessimistic and too optimistic. The work on cross-country regressions discussed below presents profoundly negative results. One possibility is that these should lead economists to abandon the general 'homogeneity assumption' in their analyses and move to other analytical frameworks. One way of discussing what these might be is to propose that they should be thought of as 'contingent', by which I intend to suggest a far greater attention to the notion that the fundamental causal relations between the standard economic variables be, early on the analysis, taken to vary across contexts. I will come back to this later. Conversely, though, I

propose that we should only rarely and deliberately (when there is good empirical reason) adopt universalistic underlying assumptions.⁵ That is, the results are useful in telling us that the world is not one best modeled using uncontingent approaches. Goldfarb is thus too pessimistic in failing to see just how much of value the weak empirical foundation of much current economics is actually telling us. And he is too optimistic, in using a framework for analyzing economists' practice that does not reach very far outside economics.

However, he may be more correct in moving towards a 'dualistic' characterisation of economists' results. That is, reality can sometimes be shown empirically to be consistent with analysis based upon assumptions of homogeneity. At other times, it needs contingent approaches based on heterogeneity assumptions.⁶

A crucial point, already mentioned, is that positing universally valid relationships implies an assumption of homogeneity. There is, it would seem, likely to be a relationship between this and the statistical methods used in subsequent investigations of its statistical validity. Whilst the relationship between assuming homogeneity in analysis and the technical meaning of homogeneity in econometrics is not simple, it seems not unlikely that lack of clarity here may easily generate spurious posited relationships. Further, these may end up only tested through approaches, such as robustness tests, that in effect force the explanatory framework to confront its most basic assumptions. So long as that is not done, the weak empirical basis for theory may remain very hard to appreciate.

Two propositions

I develop my discussion of the 'homogeneity assumption' by using two core

propositions, both empirically testable. The first, from macro economics, is that across countries and time there exist robust relationships between policy and outcomes. This is denied empirically by the now classic article by Levine and Zervos 1993 in the *American Economic Review*. They are supported in different ways by other work (for example, Rodriguez and Rodrik 1999, Brock and Durlauf 2000, Kenny and Williams 2001). They argue that, apart from very limited exceptions, there are no robust relations between economic policy variables and economic performance.⁷

Levine and Zervos, published in the world's leading economics journal, did not explicitly conclude that these results pointed to any fundamental methodological problem. Rather, they sought a refinement by testing for robustness.

The second proposition, related to micro economics, is that economics possesses persuasive assertions about economic behaviour that are reliable across different cultures, times and places. Econometric results inform judgment upon how persuasive these assertions are. Experimental economics also provides evidence about when this belief is and is not empirically justified (eg Cameron 1999). This may open the way to policy advice that is linked to particular contexts.

I start by discussing the first proposition.

An example from macro economics: the reported lack of robust observable relationships between policy and growth in econometric studies

Common positions in the development economics literature

A wide range of texts asserts that policy has a major effect upon economic performance. Two articles show this. Booth 1999 argues that there is 'widespread

agreement' on five explanations for fast growth: human and physical capital, the pattern of income and wealth distribution, the role of government and 'insulated bureaucracies', the high levels of physical investment and rapid export growth. Policy is an important part of this. Dowling and Summers 1998 with equal confidence posit 'general agreement' on three factors responsible for rapid growth: rapid growth in savings and investment, trading regimes which promoted exports, sound macroeconomic and broad sectoral policies, and a possible fourth factor, good initial conditions regarding education and skills.

Statements like these, which are common, argue that good policy leads to good performance. Now, the terms used to describe policy here are 'universals', that is, they do not imply that the relationship between policy and outcome is determined by context - in fact, quite the reverse: a rock is a rock. 'Policy' has the same meaning regardless of the where and the who. This amounts to an assumption of 'homogeneity'. An example is economic openness, the argument that economies where policies encourage international trade grow faster as a result, a position both these articles take.

What does the evidence argue?

Levine and Zervos 1993 argue an opposing view. Their data analyses suggest one cannot conclude that there are *robust* relationships between growth and variables that measure policy. The many published econometric results to the contrary have been based upon a particular selection of functional form, variables and data favoring the reporting of such relationships. The only robust relationship found was with the black market foreign exchange premium, and, perhaps, various financial sector variables:

... small changes in the right-hand side variables produce different conclusions regarding the relationship between individual policies and growth (426) ... a wide assortment of fiscal, monetary and trade-policy indicators have very fragile relationships with long-run growth. It is very difficult to find robust partial correlations between individual policy indicators and long-run growth (428) ...

Now, one might think that this *empirical* conclusion would have created considerable eddies in the profession.⁸ 'Homogeneity assumptions' make economists' analytical frameworks notorious amongst anthropologists and others for their a-historical lack of contingency. For example, 'monetary policy' effects are discussed with abstract models, typically with consequent limitations upon what causal links are envisaged. Levine and Zervos is a major challenge to this 'theory-mindset'.

As I discuss below, there were two basic reactions to Levine and Zervos in the economics literature. The first was to disregard the issue of whether the assumptions behind the approach are empirically reasonable. This could involve ignoring the basic issue of 'robustness', or it could instead entail addressing robustness by seeking data or formulations for a 'better fit'. Levine and Zervos themselves pointed in this latter direction. Second, was to address this assumption by seeking a change in data and theory that treats the data as selected from a heterogenous population. This second approach, pointing to a reversal of normally accepted views (Goldfarb 1997), was a minority position.

I discuss now an example of work in this field, reporting links between growth and trade policy. Later, I return to the reactions to Levine and Zervos.

Towards 'reversal'? An example: criticisms of accepted results in trade and growth

Rodriguez and Rodrik 1999 examine work by Dollar 1992, Sachs and Warner 1995, Ben-David 1993 and Edwards 1998. These widely-cited and influential studies argue

that there is a reliable relationship between indicators of trade policy and growth.

Rodriguez and Rodrik, however, conclude, similarly to Levine and Zervos, that when the econometric research is looked at in detail, it is far less persuasive than would appear at first sight:

Do countries with lower barriers to international trade experience faster economic progress? ... The prevailing view in policy circles in North America and Europe is that recent economic history provides a conclusive answer in the affirmative. (1) ... we are skeptical that there is a strong negative relationship in the data between trade barriers and economic growth, at least for levels of trade restrictions observed in practice. We view the search for such a relationship as futile. (38-39)

They convincingly criticize the proxies for 'openness' devised by Dollar 1992 and Sachs and Warner 1995. The Sachs and Warner index of 'openness' is based upon a dummy, which can take the value of zero or unity depending upon answers to a range of questions:

... The Sachs-Warner dummy's strength derives mainly from the combination of the black market premium (BMP) and the state monopoly of exports (MON) variables. Very little of the dummy's statistical power would be lost if it were constructed using only these two indicators. In particular, there is little action in the two variables that are the most direct measures of trade policy: tariff and non-tariff barriers (TAR and NTB). ... The extent to which (BMP and MON have) ... significance in explaining growth can be traced to their correlation with other determinants of growth: macroeconomic problems in the case of the black-market premium, and *location in Sub-Saharan Africa* in the case of the state monopoly variable ... (the dummy variable) serves as a proxy for a wide range of policy and institutional differences, and yields an upwardly-biased estimate of the effects of trade restrictions proper. (15 et seq, emphasis added)

Now, why do Rodriguez and Rodrik view the search as 'futile'? The answer, for them, lies in the methodological issues. They argue that classic tools for moving away from the universalistic generalisations lurking behind the 'homogeneity assumption' should be employed: exploration of the '*contingent* relationships between trade policy and growth' and local 'micro-econometric analysis of plant-level data sets' (39 – stress added). But they are not entirely explicit in addressing the issue of *homogeneity*. This

comes later, in the literature reacting to Levine and Zervos discussed below.

Two central points need to be made, and these relate directly to the persuasiveness of this literature. First, the scientific pretensions, for which economics is so often criticised, can here be seen operating in a healthy way. Levine and Zervos, it appears, are doing 'good econometrics'. The data, according to them, shows what it shows, *granted their statistical assumptions*.

The notion that Levine and Zervos are doing 'good econometrics' can be taken in a number of ways. Following the line epitomised by Winch, and from the perspective of a practitioner, an article published in the *American Economic Review* must certainly be taken to command respect. Further, whether there are reasons for thinking that there are grounds for criticising what they did, on technical grounds, at root they are looking at data, in ways that their reviewers accepted. Within econometrics viewed as a science, if their methods could be subject to criticism as the discussion proceeded, this does not necessarily imply that they were doing 'bad' econometrics. The central issue, which I believe was put squarely on the table by their work, is the extent to which economics as practised is aware of the implications of underlying assumptions; the response to their paper tells us something about this.

There has, indeed, been a range of comments on the methods used in their paper.⁹

These are useful, in that, like the arguments relating to 'ontological and epistemological universalism', energise the question of the risks that analysis faces if it does not engage with 'homogeneity'. Coming at the issue from the perspective of statistical theory and practice, these tend to argue that the issue is that tests for statistical homogeneity exist, *but tend to be ignored*. My own look at citations of

Levine and Zervos suggests that this is indeed the case. Thus, a deeper problem with their approach is that it was interpreted as offering techniques for checking 'robustness' that could produce better results *without* addressing the underlying homogeneity issue.¹⁰

But the point stands. Second, persuasive rational arguments about 'reality' need not just be based upon econometrics.

I turn now to look at insights from outside the discipline. This helps place the reactions to Levine and Zervos into context; more importantly, it shows the value of inter-disciplinary interaction.

The 'homogeneity assumption' from non-economic perspectives

The 'homogeneity assumption' is, by the current standards of other disciplines, somewhat extreme, and as such very interesting. In some areas of anthropology and post-development studies, for example, the common view is that knowledge is 'relative': that reported 'truths' are no such thing, rather reflections of the perspectives of those who articulate them, and indeed that positions that argue for reality as having some essential nature are best viewed as reflecting relations of power, domination etc. A classic retort, also made to Ayer's *Language, Truth and Logic*, is that such views are as absolute and certain as those which they attempt to criticise. Further, the empirical basis of such statements is frequently thin in the extreme. Thus, I have heard students regard much of the mainstream development policy literature as 'statist and realist', with pejorative connotations. Some argue, further, that policy, when actually observed, is inherently incoherent (Wright 1997). One danger here is that reasonable scepticism becomes 'post-Modernist' nihilism that dismisses almost any

reference to data.

Kenny and Williams 2001 argue from a methodological perspective that behind this econometrics is a combination of epistemological and ontological universalisms: the ideas that people can be understood the same way, and are everywhere the same: rocks are rocks are rocks. In other words, that the 'homogeneity assumption' holds - the sample is drawn from a single population.

Like Levine and Zervos, and Rodriguez and Rodrik, Kenny and Williams point out that:

Overall, attempts to divine the cause or causes of long-term economic growth, testing a wide range of possible determinants using statistical techniques, have produced results ... that are frequently contradictory to results reported elsewhere. That is, empirical evidence is hardly unanimous in support of a particular view of the growth process (1).

However, they are coming at the matter from different tacks. Levine and Zervos stress that there are very few – but not necessarily no - *robust* relationships observable between policy indicators and growth; for them, such relationships are possible.

Kenny and Williams, however, argue that there is no agreement, and that this implies, more profoundly, that basic assumptions are likely to be awry.

By the end of the decade expressions of the fundamental issues raised by Kenny and Williams were increasingly coming from within the econometrics literature itself. In 2000, Brock and Durlauf, in a piece intended for the *World Bank Economic Review*, echoed Kenny and Williams by arguing that the assumption of parameter heterogeneity is insufficiently discussed (9):

...economic theory does not imply that individual units ought to be characterized by the same behavioural functions. ... it is a matter of judgment when homogeneity assumptions are or are not to be made. Our contention is that the assumption of parameter homogeneity seems particularly inappropriate when one is studying

heterogeneous objects such as countries. ... reporting ... based on the assumption that all countries obey a common linear model may *understate* the uncertainty present when the data are generated by a family of models. (Emphasis added)

Now, it is all very well stating that assumptions are a matter of judgment, but how is one to judge whether these judgments result in persuasive positions, or not?

A useful natural science example comes from fluid dynamics. Gas flow can be modeled in different ways. A crucial distinction is made between laminar flow and turbulent flow.¹¹ Precise reference to empirical conditions is given to guide which model is to be used. Each 'homogeneity assumption' holds within certain empirically defined limits; there are theories that explain when those limits are reached, and when a different model is to be used. This example shows how different explanatory frameworks can be bounded by empirical statements. This suggests that one way of deciding whether to be persuaded or not is to ask whether the choice of model is empirically based, or not: one very different alternative, all too familiar from much econometric work, is algebraic tractability.

Back to economics: what can be learnt from this excursion beyond disciplinary boundaries?

On my reading, Levine and Zervos do not suggest that there is a fundamental methodological problem in the assumption of a homogenous sampled population. Rather, for them, the issue is the failure so far to establish robust relationships. One can read into this a concern that the basic homogeneity assumption may not hold, but that this is not their point.

Yet this is the point made, in different ways, by both Kenny and Williams and Brock and Durlauf 2000. To them there is an implicit fundamental methodological issue thrown up by the data analysis in Levine and Zervos that should be made explicit.

Baldly put, for them the data is telling us to move away from the current approach.

But how much recognition has there been in economics of this methodological issue?

What have economists made of the meaning of the 'homogeneity assumption': have the results of Levine and Zervos encouraged a shift towards assumptions of 'heterogeneity', or not? Empirically, one way of assessing the effect of Levine and Zervos upon the discipline is by tracking citations to their work, and to this I now turn.¹²

Reactions to Levine and Zervos: citations

Initially, many authors seem to have interpreted Levine and Zervos as follows. Since they did find some robust relationships (though there were extremely few, compared with what standard economic theory would have expected), the implication drawn is to 'try harder' by looking for other robust relationships. Later, however, the implicit methodological issue did start to come through (eg Bruton 1998, for whom the whole edifice of cross-country regressions has fundamental problems).¹³

We surveyed all references to Levine and Zervos up through September 2001. There were 40 citations, 33 available in English. We grouped the articles as follows.

- A. Those who disregard the methodological implications of Levine and Zervos.
- B. Those who largely disregard these methodological implications, but respond to some degree, usually by applying robustness tests after Levine and Zervos or Levine and Renelt.
- C. Those who accept the methodological implications, and work around them by

dealing with the empirics of contingency.

D. Those who accept and discuss these implications, and perhaps use fundamentally altered approaches; and

E. Those who do not enter into this discussion, explicitly or implicitly.

A. Those who disregard the methodological implications of Levine and Zervos. Total - 11

These eleven, one third of the citations, maintain the belief that cross-country regressions deliver plausible statements about relationships; they do not mention the possibility that the underlying 'homogeneity assumption' is implausible. They continue with cross-country regression analysis, and usually do not apply robustness tests (see Table 1).

B. Those who large disregard these methodological implications, but respond to some degree, usually by applying robustness tests after Levine and Zervos or Levine and Renelt. Total - 11.

This second group, again about one third of the citers, tend to apply robustness tests as a way of generating 'good results' (see Table 2). These researchers appear to view the robustness tests offered by Levine and Renelt as offering a way around the apparently negative results of Levine and Zervos.

C. Those who accept the methodological implications, and work around them by dealing with the empirics of contingency. Total 7.

This group, around one quarter of the researchers, points in various ways to 'reversal', and how to move away from the 'homogeneity assumption'. Wood has some particularly interesting remarks. See Table 3.¹⁴

D. Those who accept and discuss these implications, and perhaps use fundamentally altered approaches. Total – 1. See Table 4.

E. Citations of Levine and Zervos for other reasons not relevant for the implicit

fundamental methodological issue. Total – 3. See Table 4.

Conclusions from the data: reactions to Levine and Zervos in the discipline

This survey suggests that the fundamental implications of Levine and Zervos, after a short lag, have started to push themselves to the surface, and so a 'reversal' may be coming. But work that squarely addresses this issue, and would be capable of a dialogue with scholars outside or on the border of the discipline (such as Kenny and Williams) remains in a minority - perhaps one quarter to one third in total. This suggests that the challenges of abandoning the 'homogeneity assumption' are formidable.

Rather than confronting the evidence for major problems, which relate both to the 'sociology of knowledge' issue (Kenny and Williams) as well as to basic statistical method (Brock and Durlauf), the citations data shows that the majority of the profession has gone around them. The dominant trend has been to continue to seek results based upon the assumption that samples are drawn from a single population. A minority, though, has sought a more local and contextual approach, as the quote from Wood (fn. 14) shows. This points to an emergent 'reversal', if not yet strong, but whilst the direction is obvious the challenges are vast and far from fully taken on. The persuasiveness of economic judgments, therefore, remains weakened by the lack of agreement and reliable empirics.

I turn now to my second core proposition, relating to micro economics.

Some evidence related to micro economics: the issue of universals, whether epistemological or ontological, in comparative economic behaviour

Introduction

The cross-country regressions literature discussed above encourages reflection on the question of universals within macro economics. Other useful empirical evidence supports assumptions of heterogeneity and contingent explanatory frameworks in the context of micro economics. The evidence on gender and cultural differences is interesting.

Experimental economics is a remarkable source of empirical arguments. It has been around a long time (Roth 1995), but seems only recently to have examined issues of universality. Roth, like many others, points to a wide range of evidence that should suggest scepticism about the *empirical* support for the behavioural assumptions of much of modern microeconomics. This is certainly a telling point - Jehle and Reny 1998, a standard textbook, is strikingly data-free, particularly so in the area of the *empirical limits to the applicability of the standard behavioural models*.¹⁵

Camerer, in Hagel and Roth 1995, concludes from a wide ranging review of results:

The studies reviewed in this section suggest a variety of broad classes of anomalies of the standard utility theories under risk and uncertainty. ...These phenomena ... suggest that people use simple procedures to make choices, constructing their preferences from procedural rules rather than maximising over well-formed preferences. (673)¹⁶

To stress the point: in standard economics textbooks, argument based upon data is not developed by analogy to the natural sciences: 'this model works here, that model works there'. One way amongst others into this is to stress, as Simon 1986, the

assumption of 'instrumental rationality': that the subjectivity of the objects under study is *not* relevant to the method of analysis. The local meanings of economic phenomena are then not believed needed for the analysis to be persuasive. It follows from this that it is not so necessary for economists to report, unlike natural scientists, that 'this model works here, that model works there'. I examine this further below.

The experimental economics literature is rich and diverse. Indeed, it can appear to provide a support for much of the standard framework (eg Bergstrom and Miller 2000). This work examines gender and cultural differences in behaviour, and adds to the weight of other econometric studies.¹⁷

Gender

Econometricians often report that models of male wage determination simply 'do not work' when applied to women.¹⁸ Eckel and Grossman 1998 make similar points (see also Eckel and Grossman forthcoming 2004).

Research in every other social and behavioural science indicates substantial differences in the behaviour of men and women in noneconomic settings. The general conclusion drawn from this work is that women are more socially-oriented (selfless) and men are more individually-oriented (selfish). If these differences survive in economic decisions, when money is at stake, then theories that model agents as homogeneous, or drawn from a common distribution, may predict behaviour inaccurately. If instead the differences in behaviour are overwhelmed by monetary incentives, then economic decisions are fundamentally different from those examined in other social and behavioural sciences. (1998:726)

Their work sought to isolate gender differences. They attempted to isolate factors such as risk aversion and ways in which women react differently to the experimental conditions. This was done by retaining gender anonymity for both the partner and the experimenter. The results suggested that women were less selfish than men, which they term a 'baseline difference'.

In terms of the wider argument here, this suggests that the standard constrained-maximization model of behaviour is *more incorrect for women than it is for men*, since altruism is usually absent from this model.

Plott and Smith (forthcoming 2004) survey differences between the economic decisions of men and women. They conclude that risk levels are a core influence upon choice of behavioural model ('this model works here, that model works there').

Systematic differences are only revealed when subjects are not exposed to risk. Then, the choices women make are less individually-oriented and more socially-oriented.¹⁹

Note, though, that the concept of 'risk' is treated here as a universal, situated within Simon's sense of an 'instrumental rationality' (see above) within which local meaning is not significant.

Cultural differences

Apart from difference in behaviour based upon gender, what can be found about variation due to 'culture'? Experimental economics also throws light upon differences here.

Henrich *et al* 2001 looked at fifteen small-scale societies. Based upon standard economic experiments, they conclude that the 'Economic Man model is not supported in *any* society studied', that the main variation was between groups (rather than between individuals); and that behaviour in the experiments was generally consistent with economic patterns of everyday life in these societies.

In other unpublished work, Henrich and Smith (mimeo n/d) compare Latin American and US subjects' behaviour. They conclude that they find more evidence than earlier studies that inter-group characteristics are of considerable importance, and that

variation in individual attributes is of less significance to explaining differences. For them, as would be expected from such ethnographically-informed accounts, differences in economic life are mainly group-based.

Cameron 1999 reaches similar conclusions. She argues that increased stakes (thus responding to Eckel and Grossman's point about the effects of large monetary stakes on 'swamping' standard models of behaviour) move behaviour *further* from the standard game-theoretic norms and towards procedural solutions such as a 50:50 split.

A comment on cultural differences can also be found in Henning-Schmidt et al mimeo n/d. Here, based upon analysis of videotapes of experiments in negotiation, they conclude in a comparison of Chinese and German subjects that, in a stylised game, cultural differences were high. Confirming the discussion above, *individual* differences could be insulated from the data, and the analysis could thus bring out differences between the groups, categorised as 'cultural'.²⁰

The referential empirical base of experimental economics is interesting, but, unlike natural sciences, still does not point very clearly to where, empirically and in a measurable manner, boundaries between different sets of behavioural assumptions are to be found. It does, however, point to their existence. This existence may then be presented, in the context of policy advice, as a basis for developing the 'contingent' nature of policy advice: with specific cultural and gender contexts (e.g. the specific gender division of labour), the consequences of policies can be argued persuasively to differ across time and place.

Disciplinary boundaries – reflections on neo-institutional economics, instrumental rationality and universalistic assumptions

There is thus reason to conclude that much empirical work in economics is not very effective in persuading the acceptance of assumptions of universality: economic change varies in its essence too much across countries, and so the cross-country econometrics based upon homogeneity leads to disagreement rather than convergence. Evidence on gender and culture also suggests caution. What are the alternative ways forward?

The discussion so far, and especially the citations of Levine and Zervos, suggests that the homogeneity assumption in fact creates great difficulties for economics. How can economics negotiate the limits of homogeneity, or more precisely what Kenny and Williams refer to as the assumption of epistemological and ontological universalism? More fundamentally, how can it seek a basis for persuasive judgements as to where those limits are – that is, *when is that assumption reasonable?* Further, the discussion suggests that we can learn from the opportunities to get 'out of the box' offered by disciplines other than economics, which have therefore become more relevant to the search for persuasive economic arguments.²¹ Whether and how this happens, is in part an empirical question. This argues that we should watch closely the negotiation of disciplinary boundaries as an indicator of the evolving practice of economics. Here the sub-discipline of neo-institutional economics provides an illuminating example.

The rise of neo-institutional economics has been marked by the award of the Nobel Prize for Economics to one of its major practitioners, Douglass C. North.²² Neo-institutional economics has been viewed both negatively and positively.

Viewed negatively, as for example by Bates 1995, neo-institutional economics is a failed attempt to use economic theory to explain things it cannot explain. As a non-economist, Bates sees neo-institutional economics' core logic as asserting that '*when markets fail*' other organisational forms can and will arise to exploit "the opportunities for gain". Thus the idea of 'market failure' is used to explain the logic of non-market forms of social organisation. He points to what he sees as various problems with this view. Allegedly, it does not explain the politics or reality of change. It says what good government should be, but not how it might get there, and therefore fails to guide how to do so; neither does it help answer the question of who would pay for the costs of resourcing a new institution, for there are incentives to 'free ride'. This may, then, explain policy failure.

Here what is significant for my argument is the underlying negotiation over disciplinary boundaries: who can and should explain what. Bates' position is to criticise neo-institutional economics for, he says, trying to use economic logic to explain the non-economic. But is this what is happening? I argue that the allegations are overdone, for North 1995 takes a more positive tack.

A useful point here is to consider how certain North is that economics can explain change. He states that what is needed is the rejection of 'instrumental rationality', which rejection is, he argues, key to the success of neo-institutional economics as he seeks to practise it. By 'instrumental rationality' he means the idea that "it is not necessary to distinguish between the real world and the decision-maker's perceptions of it" and, "that it is possible to predict the choices that will be made by a rational decision-maker entirely from a knowledge of the real world and without a knowledge of the decision-makers perceptions or modes of calculation" (2).²³ The relevant point

here, given the discussions of econometrics and experimental economics above, is that perceptions may have both local and universal characteristics, and the distinction is there to be appreciated and negotiated.

This does not, as Bates would have it, simply seek to expand the realm to be explained by economics; rather, it seeks to establish where those boundaries may be, and then to negotiate. A classic study by economists of English state finance (North and Weingast 1989) must clearly confront the arguments of historians and other non-economists who have access to analytical frameworks and data unavailable to the authors. The reader must then decide among these competing approaches, perhaps taking pieces from each. In this way economics may find itself interesting and persuasive to non-economists, and vice versa.

North attributes his views to Simon 1986. Arrow 1986, in the same collection as Simon 1986, returns us to the issue of data and measurement. He argues that many of the results of neoclassical economists rely, not upon the underlying behavioural model (of constrained optimisation) but upon the particular assumptions made upon the functional forms involved (ie, of the shapes of indifference curves, their relationships one to another and so forth). This, though, is, or should be, as much an empirical as a theoretical issue.

I conclude that the issues raised by neo-institutional economics have a strong resonance with those brought up by my earlier discussions. Their increasing relevance within the discipline can be taken to show, especially through the discussion of 'instrumental rationality', that gathering importance is attached to the issue of contingent local meaning, the antithesis of which is the assumption of ontological and

epistemological universalism.

Conclusions

There are strong empirical arguments which suggest that economists should think far more about just what they are doing when they adopt the 'homogeneity assumption'. For macro economics, I argue that the failure of the cross-country regression work to produce clear and agreed results needs to be given serious thought, and that a significant minority of economists have done so. The evidence I have examined related to micro economics is far weaker, but points in the same direction, to the need to ponder on exactly what is being done when assumptions of 'epistemological and ontological universalism' are made.

Following Winch's position that such ponderings are the 'doing' of something, are social activities and bearing in mind that economics is a discipline, it follows that such ponderings need to be made to happen: the assumptions must be questioned, whether in seminars, reviews of papers or wherever the self-policing of the discipline happens. Whether in terms of the universalism of analytical frameworks, or the statistical testing of the homogeneity of sampled populations, failure to identify these positions as assumptions, and to question them as such, risks much. Most importantly, in a world that is not only increasingly diverse, but in an academic universe increasingly permeated by antipathy to rigorous empirical investigation as economists understand it, failure to do so greatly weakens the wider persuasive power of results.

The econometrics and experimental work discussed shows that there is much to be learnt from this empirical literature. The assumption of homogeneity is not inherently unreasonable, yet articles such as Levine and Zervos, papers such as Rodriguez and

Rodrik 1999 and others provide a basis for reflecting on the wisdom of it. Two simple but tricky questions remain – where does this assumption come from and is it persuasive because *empirically* justified?

I believe it is unreasonable to retreat, as is common in many social sciences, to some a-factual post-Modernist world. Moreover, as the citations survey showed, many economists are aware of these issues and a minority is explicitly trying to do something about it.

'Reversal'? What does the data suggest should be done?

Growing awareness of the existence and nature of the 'homogeneity assumption' seems crucial to a 'reversal'. Yet, my review of citations of Levine and Zervos shows that only limited progress has been made in this direction: people are aware of the issue, and are trying to find ways of dealing with it, but they are a minority. The simplest reason is most likely the sheer scale of the task of creating economic theory that can handle heterogeneity.

This suggests, *pace* Goldfarb, a 'reversal' that can be sensed 'in the egg' even if it may in fact never arrive. In the cross-country research examined, a research program has generated results that challenge an assumption thought to be so fundamental and unquestioned *that the research projects of the program did not test it*. Yet it appears that the program can and did. Further, the work of Hendry and others argues that there exist techniques of econometrics that would have permitted testing of this assumption.

If this view is correct, then what has happened can be put simply. First, adopt the 'homogeneity assumption' for your theory and empirics - you do not question or test this directly; second, generate a wide range of conflicting but apparently

statistically correct results; third, draw from this the conclusion that some core assumption is awry, and look for it, using perspectives both inside and outside the discipline. In this case, the evidence points strongly towards the 'homogeneity assumption' as the culprit. What is interesting from this practice is that it seems that two necessary ingredients were, first, an empirical analysis that could address the stability of results (Levine and Zervos) combined with, second, inputs from outside the discipline (e.g. Kenny and Williams) that could critique the core aspects of the theory that were causing the problems. My review of the references to Levine and Zervos tends to show, however, that existing techniques within econometrics that could have tested for homogeneity were not used to engage with Levine and Zervos.

I find here the example from fluid dynamics discussed above both revealing and limiting. The normal practice in that discipline has been to rely upon the nature of the reality it has studied to establish well-negotiated boundaries between two different fields: laminar and turbulent flow. Levine and Zervos could have found many more robust relationships than they did, and then the negotiation of the relationships between global and local explanatory frameworks, between assumptions of homogeneity and their abandonment, could have been far easier. Yet they did not, and this is revealing. The comparison is limiting, though, in that the stuff of social science, and its inherent tolerance of long-lasting differences that would be quite unacceptable in natural sciences, means that the nature of methodological issues is far less straightforward.

The evidence presented above supports the view that there are moves in the direction of a 'reversal'. Abandonment of the 'homogeneity assumption' may be leading to a shift in analytical approach to one where the standard behavioural models are linked

to data defining the limits of their applicability; similarly, the cause-effect relations that underpin the standard notion of policy could become seen as, in any particular context, dualistic: some follow universal logics, some local ones. The results I report from experimental economics also support such local arguments, for culture and gender differences do discourage use of analytical approaches that ignore such differences. In both instances, I would argue that progress will likely be marked by a more intense and meaningful negotiation of economics' disciplinary boundaries: what it can explain, what it cannot. However, there is a long way to go. And the world is becoming more diverse, and so persuasion, rather than authority, more important.

Policy and policy advice?

As a trained economist who has sought to offer policy advice on development I have had to face many of the problems discussed above. In my own experience, what makes the difference is the presentation of persuasive arguments in ways that command respect, *and that this respect can usually only be won through a demonstrated familiarity with the institutional and other realities inhabited by one's interlocutors.* Further, this can only be done through prolonged interaction and discussion. This is made more difficult if policy-makers have had to previously deal with unwise policy advice given too readily. The result has been that people have learnt to cope with prescriptions articulated with unwarranted certainty, and confront the bitter experience that confidently expressed advice often turns out to be wrong. When they have learnt this, it may then be very hard to engage them in persuasive dialogue; they can, of course, like any human, smile and disregard what one says.

Granted that access, I have found that in practice it is actually rather easy to develop

the combination of universalistic and particular analytical frameworks to which I have argued the data is pointing us, and to base these in plausible statements about where the empirically-founded boundaries between them are. Moving 'beyond the core' in this way seems to require great attention to local detail, which is in any case needed for persuasive policy advice. This search benefits from contact with other disciplines, and does not at all require abandoning the basic assumption, of 'agnostic realism', that is common to so much empirically based research. The point is not to give up, but to base persuasive arguments on models that work; clearly, many do not.

References

- Ahmed, H. and Miller, S.M. (2000) 'Crowding-Out and Crowding-In Effects of the Components of Government Expenditure', *Contemporary Economic Policy*, 18 1 124-133 Jan.
- Alexander, W. R. J. (1997) 'Inflation and economic growth: Evidence from a growth equation', *Applied Economics*, 29, 233-238.
- Amoateng, K., and Amoako-Adu B. (1996) 'Economic growth, export and external debt causality: the case of African countries', *Applied Economics*, 28 21-27.
- Arrow, Kenneth J. (1986) 'Rationality of self and others in an economic system', *Journal of Business*, Vol 59 Issue 4 Part 2, Oct S385-S399.
- Bates, Robert H. (1995) 'Social Dilemmas and Rational Individuals: an Assessment of the New Institutionalism', in Ed Hariss et al *The New Institutional Economics and Third World Development*, London: Routledge.
- Ben-David, Dan, (1993) 'Equalizing exchange: trade liberalization and income convergence', *Quarterly Journal of Economics*, 108(3) 653-679.
- Benhabib, J. and Spiegel, M.M. (2000) 'The Role of Financial Development in Growth and Investment', *Journal of Economic Growth*, 5 341-360
- Bergstrom, Theodore C. and John H. Miller (2000) *Experiments with economics principles: microeconomics*, 2nd ed., Boston MA: Irwin McGraw-Hill.
- Berthelemy, J.C. and Varoudakis, A. (1996) 'Economic growth, convergence clubs and the role of financial development', *Oxford Economic Papers*, 48 300-328.
- Booth, Anne (1999) 'Initial Conditions and Miraculous Growth; Why is South East Asia Different from Taiwan and South Korea?', *World Development* 27 2 301-322.
- Bougheas, S., Demetriades, P.O., and Mamuneas, T.P. (2000) 'Infrastructure, Specialization and Economic Growth', *Canadian Journal of Economics*, 33 2 506-522.
- Brock, W.A., and Durlauf, Steven N. (2000), 'Growth economics and reality', NBER Working Paper 8041 December.
- Bruno, Michael and Easterly, William (1998) 'Inflation crises and long-run growth', *Journal of Monetary Economics*, 41 3-26.
- Bruton, H.J. (1998) 'A reconsideration of import substitution', *Journal of economic*

- literature, Vol. XXXVI (June 1998) 903-936.
- Camerer, Colin (1995) 'Individual decision making', in Ed John H. Hagel and Alvin E. Roth, *The Handbook of Experimental Economics*, Princeton NJ: Princeton University Press.
- Cameron, Lisa (1999) 'Raising the stakes in the ultimatum game: experimental evidence from Indonesia', *Economic Inquiry*, Vol 37, No. 1, January, 47-59.
- Cheng, B.S., and Hsu R.C. (1997) 'Human capital and economic growth in Japan: an application of time series analysis', *Applied Economics Letters*, 4 393-395
- Cheshire, Paul and Stefano Magrini (2000) 'Endogenous processes in European regional growth: convergence and policy', *Growth and Change*, Vol 31. Fall, 455-479.
- Clarke, G.R.G. (1995) 'More Evidence on Income-Distribution and Growth', *Journal of development Economics*, 47 (2) 403-427 Aug.
- Comim F. (2000) 'On the Concept of Applied Economics: Lessons from Cambridge Economics and the history of growth theories', *History of Political Economy*, 32 147-176 Suppl. S.
- Demetriades, P.O., and Hussein, K.A. (1996) 'Does Financial Development Cause Economic Growth?' Time-Series Evidence from 16 Countries', *Journal of Development Economics*, 51 2 387-411 Dec.
- Dollar, David (1992) 'Outward-oriented developing countries really do grow more rapidly: evidence from 95 LDCs', *Economic Development and Cultural Change*, v40 n3 523-544.
- Dotsey, M. and Sarte, P.D. (2000), 'Inflation Uncertainty and Growth in a Cash-In-Advance Economy', *Journal of Monetary Economics*, 45 3 631-655 June.
- Dowling, Malcolm and Peter M. Summers (1998) 'Total factor productivity and economic growth: some issues for Asia' *The Economic Record* Vol 74 No 225 June 170-185.
- Easterly, W. and Levine, Ross, (1997) 'Africa's growth tragedy: policies and ethnic divisions', *The Quarterly Journal of Economics*, Nov 1203-1250.
- Easterly, W. (2001) 'The Lost Decade: Developing Countries' Stagnation in Spite of Policy Reform 1980-1998', *Journal of Economic Growth*, 6, 135-157.
- Easterly, W., Norman Loayza and Peter Montiel (1997) 'Has Latin America's post-reform growth been disappointing?' *Journal of International Economics* 43 287-311.
- Eckel, Catherine C., and Grossman, Philip J. (1998) 'Are women less selfish than men?: evidence from dictator experiments', *The Economic Journal*, 108(448): 726-735, May.
- Eckel, Catherine C., and Philip J. Grossman (forthcoming 2004) 'Differences in the economic decisions of men and women: experimental evidence', in Plott, Charles and Vernon L. Smith, eds., *Handbook of results in experimental economics*, New York: North-Holland.
- Edwards, Sebastian (1998) 'Openness, productivity and growth: what do we really know?' *The Economic Journal*, March 383-398.
- Frankel, Jeffrey A. and Romer, David (1999) 'Does trade cause growth?' *American Economic Review* June 379-399
- Goldfarb, R.S. (1997) 'Now You See It, Now You Don't: Emerging Contrary Results in Economics,' *Journal of Economic Methodology*, December 221-244.
- Goldfarb, R.S., and Stekler, Herman (2000) 'Why Do Empirical Results Change?: Forecasts as Tests of Rational Expectations,' *History of Political Economy annual supplement to Vol. 32* 95-116.

- Granger, C.W.G. Ed (1990) *Modelling economic time series: Readings in econometric methodology* Oxford: Clarendon Press.
- Hagel, John H. and Roth, Alvin E. (1995) *The Handbook of Experimental Economics*, Princeton NJ: Princeton University Press.
- Hanushek E.A., and Kimko D.D. (2000), 'Schooling, labour-force quality and the growth of nations', *The American Economic Review*, 90: (5) 1184-1208 Dec.
- Hausman, Daniel M. (1992) *The inexact and separate science of economics*, New York: Cambridge University Press.
- Hendry, D.F. and Mizon, G.E. (1990), 'Procrustean econometrics: Or stretching and squeezing data', in C.W.J. Granger Ed *Modelling economic times series: Readings in econometric methodology* Oxford: Clarendon Press.
- Henning-Schmidt, Heike, Li Zhu Yu and Yang Chao Liang (n/d) 'A cross-cultural study on negotiation behaviour. A video experiment run in Germany and the People's Republic of China, mimeo.
- Henrich, Joe and Smith, Natalie (n/d) 'Comparative experimental evidence from Peru, Chile and the US shows substantial variation among social groups', mimeo.
- Henrich, Joe, Robert Boyd, Samuel Bowles, Colin Camerer, Ernst Fehy, Herbert Gintis and Richard McElreath (2001) 'Cooperation, reciprocity and punishment in fifteen small-scale societies, mimeo, Jan.
- Hoover, Kevin and Perez, Stephen (forthcoming) 'Truth and robustness in cross-country growth regressions, *Oxford Bulletin of Economics and Statistics*.
- Jehle, Geoffrey A., and Reny, Philip J. (1998) *Advanced Microeconomic Theory*, Reading MA: Addison-Wesley.
- Kalaitzidakis, P., Mamuneas, T.P., and Stengos T. (2000) 'A Non-linear Sensitivity Analysis of Cross-Country Growth Regressions', *Canadian Journal of Economics*, 33 (3) 604-617 Aug.
- Kenny, Charles and Williams, David (2001) 'What do we know about economic growth? Or, why don't we know very much?' *World Development* Vol 29, No. 1 1-22.
- Kidd, Michael P. and Shannon, Michael (1997) 'Imputation of Female Labour Market Experience: Some Australian Evidence on the Zabalza and Arrufat Method', *The Economic Record*, Vol 73, No. 221, June 136-145.
- Kim, Sung and Willett, Thomas D. (2000) 'Is the negative correlation between inflation and growth real? An analysis of the effects of the oil supply shocks', *Applied Economics Letters*, 7 141-147.
- Kline, Morris (1980) *Mathematics: the loss of certainty*, New York: Oxford University Press.
- Leamer, Edward E. (1983) 'Let's take the con out of econometrics', *American Economic Review*, Vol 73 No.1 (Mar.), 31-43
- Lee, D.W., and Lee, T.H. (1995) 'Human-Capital and Economic-Growth - Tests Based on the International Evaluation of Educational-Achievement', *Economics Letters*, 47 (2) 219-225 Feb.
- Levine, Ross, and Renelt, David (1992) 'A sensitivity analysis of cross-country growth regressions', *American Economic Review*, Vol 82, (4) Sept 942-963.
- Levine, Ross and Zervos, Sara J. (1993) 'What have we learnt about policy and growth from cross-country regressions?', *The American Economic Review* Vol 82 Issue 2 *Papers and Proceedings ...* May 426-430
- Levine, Ross, and Zervos, Sara (1996) 'Stock market development and long-run growth', *World Bank Economic Review*, 10: (2) 323-339 May.
- Lindauer, David L. and Pritchett, Lant (2002) 'What's the big idea? The third

- generation of policies for economic growth', *Economia* 3:1 (Fall), pp.1-39.
- Lindh, T. and Lindstrom, T. (1997) 'Structural breaks in financial share, savings and growth in OECD: 1960-92', *Applied Economics* 29 1 33-41.
- McAleer, Michael, Pagan, Adrian R., and Volcker, Paul A. (1985) 'What will take the con out of econometrics?' *American Economic Review*, Vol. 75 No. 3 (Jun.), 293-307.
- Mehlkop, G. (2000) 'Analysis of Values and Economic Growth - Methodological Problems', *Zeitschrift fur Soziologie*, 29 (3) 217-
- North, Douglass C and Weingast, Barry R. (1989) 'Constitutions and Commitment: the Evolution of Institutions Governing Public Choice in Seventeenth Century England', *Journal of Economic History* Vol 49 # 4 Dec 803-832.
- North, Douglass C. (1995) 'The New Institutional Economics and Third World Development', in Ed Harriss et al *The New Institutional Economics and Third World Development*, London: Routledge.
- Panlilio, F. (1963) *Elementary theory of structural strength*, London: John Wiley and Sons.
- Pasinetti, Luigi and Solow, Robert M. (1994) *Economic growth and the structure of long-term development*, Proceedings of the IEA Conference held in Varenna, Italy, 1992 (IEA Conference Volume, No. 1), London: Macmillan.
- Plane, P. (1997) 'Privatization and economic growth: An empirical investigation from a sample of developing market economies', *Applied Economics*, 29, 161-178.
- Plott, Charles and Smith, Vernon L. Ed. (forthcoming 2004) *Handbook of results in experimental economics*, New York: North-Holland.
- Pugno, M. (1996) 'Structural stability in a cross-country neoclassical growth model', *Applied Economics*, 28, 1555-1566.
- Quinn, D. (1997) 'The Correlates of Change in International Financial Regulation', *American Political Science Review*, 91 3 531-551 Sept.
- Rodriguez, Francisco and Rodrik, Dani (1999) 'Trade policy and economic growth: a skeptics guide to the cross-national evidence', NBER Working Paper 7081 April.
- Roth, Alvin E. (1995) 'Introduction to Experimental Economics', in Ed John H. Kagel and Alvin E. Roth, *The Handbook of Experimental Economics*, Princeton NJ: Princeton University Press.
- Sachs, Jeffrey and Warner, Andrew (1995) 'Economic reform and the process of global integration', *Brookings Papers on Economic Activity* 1, 1-118.
- Seccareccia M., and Lavoie M. (1996) 'Central bank austerity policy, zero-inflation targets and productivity growth in Canada', *Journal of Economic Issues*, Vol XXX No. 2, June.
- Shi, Shouyong (1999) 'Search, Inflation and Capital Accumulation', *Journal of Monetary Economics*, 44 1 81-103 Aug.
- Simon, Herbert A. (1986) 'Rationality in Psychology and Economics', *Journal of Business*, Vol 59 Issue 4 Part 2, Oct S209-S224.
- Temple, J. (2000) 'Inflation and growth: stories short and tall', *Journal of Economic Surveys*, 14 (4) 395-426 Sept.
- Torstensson, J. (1994) 'Property rights and economic growth – an empirical study', *Kyklos*, Vol 47 Fasc. 2, 231-247.
- Widmalm, F. (2001) 'Tax Structure and Growth: Are Some Taxes Better than Others?' *Public Choice*, 107 199-219.
- Winch, Peter (1992) 'Persuasion', *Midwest Studies in Philosophy* XVII.
- Wood, Adrian and Ridao-Cano, Cristobal (1999) 'Skill, trade and international inequality', *Oxford Economic Papers* 5189-119

- Wood, Adrian (1994) 'Comment', in Ed Pasinetti, Luigi and Robert M. Solow, 1994, Economic growth and the structure of long-term development, Proceedings of the IEA Conference held in Varenna, Italy, 1992 (IEA Conference Volume, No. 1), London: Macmillan.
- Wood, Adrian (1997) 'Openness and wage inequality in developing countries: the Latin American challenge to East Asian conventional wisdom', The World Bank Economic Review Vol 11 No 1 33-57.
- Wright, Susan (1997) 'Policy: A new field of anthropology', in Ed Chris Shore and Susan Wright, Anthropology of Policy: Critical Perspectives on Governance and Power London: Routledge.
- Yonay, Yuval P. (1998) The struggle over the soul of economics, Princeton: Princeton University Press.

TABLE 1 - Citers of L&Z – Group A Those who disregard the methodological implications of L&Z.

Alexander 1997:233 “... any positive effect of inflation on growth is more than outweighed by the negative effects”.

Benhabib and Spiegel 2000:341 “... researchers have found that the correlation between financial development and economic growth is *uniquely* robust” – referring to L&Z, emphasis added.

Easterly 2001:136 quotes L&Z's *positive* results.

Easterly and Levine 1997 argue that “... the goal is not to establish ‘robustness’ as defined by Levine and Renelt (1992)” (1208).

Easterly et al 1997 simply quotes L&Z's *positive* results, relating to the “black market premium (which) has proven to be robustly correlated with growth performance in previous studies” (292) and then run a range of regressions to make their point, which is that the response of Latin American countries to reforms has “not been disappointing” (287).

Kim and Willett 2000:141: “Critics ... point to findings that the estimated relationships were quite unstable ... Such instability should not be surprising ... (for) the reasons ... are not well understood, and the relationship is far from being a tight mechanical one”.

Lee and Lee 1995:219 quote L&Z as an example of empirical research that uses measures of human capital that “are not directly linked to the notion of human capital” and then continue with their econometrics, without applying robustness tests.

Lindh and Lindstrom 1997:33 report L&Z's *positive* results (i.e. that “In cross-country regressions of growth, the positive correlation with financial development is one of the few findings that are robust to alternative sets of control variables”).

Plane 1997:171 reports L&Z's *positive* results (i.e. that there is a “robust correlation between the ratio of investment to GDP and economic growth”).

Sub-group - 'Fence-sitters': cannot ignore L&Z, but apparently want to.

Dotsey and Sarte 2000:632 straddle the fence. Thus “In studying the effects of monetary policy on real growth, we desire a model that is consistent with a number of stylised facts ... First, the model should display a negative relationship between average inflation and average growth. However, despite numerous studies on the link between ...inflation ... and growth, there still exists some controversy over the robustness of this relationship”.

Shi 1999:99 reports the L&Z finding of the absence of any robust relationship between growth and inflation, positive or negative, and suggests that this may imply a non-monotonic relationship between aggregate capital/output and inflation (i.e., positive with low inflation and negative with high).

TABLE 2 - Citers of L&Z – Group B - Those who large disregard these methodological implications, and respond to some degree, usually by applying robustness tests after L&Z or Levine and Renelt.

For example, Bougheas, Demetriades and Mamuneas 2000 apply the robustness tests given in Levine and Renelt, whilst Bruno and Easterly 1998:19-20 report "... in contrast to what is said by the existing literature, pooled cross-country datasets are not informative about what happens at lower ranges of inflation. In contrast, the data on discrete high inflation episodes speaks clearly – there is a strong and *robust* relationship between high inflation and growth" (emphasis added).

Cheshire and Magrini 2000:464-465 are an intriguing example of 'double think'. They conclude that the problem of instability of results as model specification changes is so great as to suggest that these methods be avoided; convergence estimates are dependent, not upon data, but upon model specification. Yet their approach to growth analysis is, like others, to 'add-in' novel explanatory variables, with universal character – "Magrini's model originated from the view that technological knowledge has a very important tacit component that has been neglected in formal theories of endogenous growth (455).

Clarke 1995 uses the robustness tests applied in Levine and Renelt to argue that his result, that inequality is negatively correlated with growth, "is robust across different inequality measures, and to many different specifications of the growth regression" (403).

Kalaitzidakis, Mamuneas and Stengos, 2000, develop the Levine and Renelt robustness check by assessing the robustness of models that include a common non-linear specification (L&Z only covered linear models).

Levine and Zervos 1996:332 simply refer to L&Z (and Levine and Renelt 1992) as sources of techniques for assessing the robustness of relationships.

Mehlkop 2000 finds no robust effect of posited explanatory variables in his work, which thus negatively critiques another work cited.

Quinn 1997 uses the robustness tests applied in Levine and Renelt "to increase reader confidence in the results" (531). He concludes that some policy measures are robustly associated with growth, in this sense with capital account liberalisation.

Seccareccia and Lavoie 1996 call for robustness testing. "... as L&Z (1993:430) affirm, "unless researchers study the sensitivity of their results to small variations in the sample of countries and changes in the conditioning information set, the results should be regarded with scepticism... Given the ravages of the zero inflation policy ... how can the Bank of Canada base its continued high real interest-rate policy on such shady theory and evidence?"

Torstensson 1994: 235 treats L&Z as a source of techniques for assessing the robustness of results, which are then presented as being sounder. "We have not been able to empirically relate state ownership to growth. ... (But) the analysis in this paper indicates quite clearly that arbitrary seizures of property decelerate economic growth".

Widmalm 2001:200 and 213 treats L&Z as implying a great need for caution, and for applying their robustness tests (from Levine and Renelt 1992).

TABLE 3 - Citers of L&Z – Group C - Those who accept the methodological implications, and work around them by dealing with the empirics of contingency.

Ahmed and Miller 2000:127. “Such estimation, however, (OLS – AF) can create problems of interpretation if country-specific characteristics are not considered (see L&Z 1993:420) ... The use of panel data, however, provides an approach to address that problem”. Panel data, of course, simply assumes homogeneity over time.

Amoateng and Amoako-Adu 1996:21 accept the strong warning implicit in L&Z, and shift the analytical approach to a focus upon *causality*. They conclude (26) that “(t)he evidence indicates that, generally, there is a joint feedback effect between export revenue, external debt service and economic growth”.

Cheng and Hsu 1997:39 quote L&Z to confound their own analysis – “Another issue concerning regression specification is that these studies used cross-sectional data, which implicitly assumed the existence of an invariant cross-section relationship between human capital and economic growth ... However, differences in the importance of effectiveness of education across countries defy this assumption. Thus, the results of causality tests depend very much on model specification”.

Demetriades and Hussein 1996:387 present essentially negative results: “Our findings also clearly demonstrate that causality patterns vary across countries and, therefore, highlight the dangers of statistical inference based on cross section country studies which implicitly treat different economies as homogenous entities”. They also suggest that the main explicit message of L&Z is the need to apply sensitivity analyses (390 fn 6), i.e. that the implicit methodological issue is hidden.

Hanushek and Kimko 2001:1185 “This paper addresses the measurement problem of labor-force quality directly ... we construct new measures of quality based on ... various international tests of academic achievement ... quality differences measured in this way prove to have extremely strong effects on growth rates”.

Temple 2000:413. “We also know from Levine and Renult how sensitive results can be to the choice of specification. *There is not much to be said here*, except to note that inflation tends to emerge as particularly sensitive to the choice of control variables ...”. Emphasis added.

Wood 1994:67-68: “The value of this message is clear ... None of us ... will ever rely so casually and so heavily upon cross-country regressions. ... (yet) I have a basic reservation about the methodology used, which is powerful but blunt.”

TABLE 4 - Citers of L&Z – Groups D and E –

D - Those who accept and discuss these implications, and perhaps use fundamentally altered approaches.

Bruton 1998:932: “L&Z is valuable because it is a good clear statement of the dangers of using cross-country regressions as if they met the necessary assumptions of formal statistical and econometric theory”.

E – Citations of L&Z for other reasons not relevant for the implicit fundamental methodological issue.

Berthelemy and Varoudakis 1996:310, refers to L&Z, sense unclear.

Comim 2000:162 looks at the activities of the cross-country growth research within a survey of the concept of applied economics, and points out that this field, reinforced by L&Z, has tended to examine issues such as robustness precisely because it “follows its own research agenda with much less influence from theoretical quarters than the influence generated by the neoclassical theories on applied studies”.

¹ The disputes about whether sub-disciplines such as agricultural economics and development economics have anything theoretically specific to bring to the discipline support my position here.

² I take as a working definition of the 'homogeneity assumption', the assumptions of ontological and epistemological universalism discussed by Kenny and Williams 2001: that is, that both the things economics studies, and how they should be understood, are the same through time and space. This is the same as the assumption that mass, as a characteristic of things, is the same here as on Mars, and should be analysed the same way in both places. See below for the relationship between these issues and the assumption of 'instrumental rationality'.

³ Many economics departments, at least in Australia, are now situated within Faculties of Commerce rather than with the other social sciences. I do not wish to imply that the methodologies of other disciplines are, essentially, better guides to persuasion than those of economics: acceptance of 'relativism' does not really help people cope with intentional and rationalised choice, which is how policy remains understood and articulated, as practice.

⁴ Explanations of where these 'theory-mindsets' come from is not something adequately explained by economics alone.

⁵ I would propose, thus, that typical questions asked of the presenter of an attempt to model, for example, the determinants of growth in a particular context, should include those that require the researcher to have considered whether variables like 'inflation' are indeed, conceptually and empirically, the same in his or her context as elsewhere, and to what extent they stand in causal relationships with other variables that are the same as those elsewhere: that is, that the 'ontological and epistemological' assumptions be explored – see below.

⁶ Goldfarb and Stekler 2000 seek to develop an empirical methodology for assessing changes in economists' views. They conclude, however, that *further* arguments are needed to provide a convincing explanation of the conflicting results in the literature (110). In their case study, they observe relatively

little empirical relationship between empirical evidence and the acceptability of the rational expectations hypothesis (112). Compare this methodology with the more discourse-based approach of Yonay 1998, in discussing the debates between neoclassical and institutional economists before WWII.

⁷ In my own teaching experience exposure to this work has a somewhat electric effect upon students both within and outside economics. This result is interesting.

⁸ This article was in fact preceded by another, Levine and Renelt 1992, which reached similar but less provocative conclusions ('Almost all results are fragile ... to small changes in the conditioning information set' (942)). Later, in Easterly and Levine 1997, Levine used cross-country regressions techniques to explore African growth processes, 'adding-in' a 'newly discovered variable' (in this case 'ethnicity') to find that '... ethnic diversity helps explain cross-country differences in public policies and other economic indicators' (1203).

⁹ I am referred to McAleer, Pagan and Volcker 1985 questioning the original paper by Leamer 1983. Also Hendry and Mizon 1990. More recently, Hoover and Perez forthcoming examine what happens if the robustness criteria are relaxed. Perhaps not surprisingly, if the allowable range of variation is increased, more robust relationships can be found. Yet, what basis is there for judging what an allowable degree of significance is acceptable? We could chose to come back to Winch's position, and regard research as something that is human business, which may find the Truth, but can never be certain of it. To quote Hoover and Perez forthcoming: "We can never guarantee that the specifications selected by the general-to-specific approach are true. But the approach is part of a critical, indeed dialectical, methodology" (19). Interestingly, this work concludes that "What is surprising is how few of the variables matter in the end and how much is left unexplained. The preferred regression explains only 42 percent of the variability of countries' growth experiences ..." (19). This could be taken to suggest that the rest of the story is 'contingent'; as I argue below, this conclusion, like others, points to a possible consensus that articulates agreement over cause-effect relations in terms of an amalgam of universalist and local relations: *yes*, we agree that these relations hold for all of us; *yes*, we agree that these other relations hold for the here and now, and not necessarily for others. This would seem a wise formula.

¹⁰ It seems to me that a proper awareness of the homogeneity assumption, refined and expressed in terms of statistical analysis - i.e. that the sample is drawn from a single population - is needed to make meaningful the statistical tests for parameter estimates. If it is unwarranted, and unless measures are taken to deal with this, significance results are literal non-sense. Thus the risk would be that a technique, if one could be found, for securing robustness would remain founded upon sand: the robust results would reflect random order ... leading to a range of conflicting but significant results, as in the case of the cross-country regressions. Whether viewed as an issue in statistical analysis, or as a central assumption in the selection of analytical frameworks, the question of homogeneity is a valuable but (according to my own review of responses to Levine and Zervos under-used) check upon the degree to which positions that purport to be based upon empirical work should be taken as persuasive.

¹¹ The significance of the transition from laminar to turbulent flow is that wings often stall when this happens, which can kill people.

¹² The work surveyed covers a wide range, and I have had to be selective in the ways we have used it for this review. I apologise for any apparent misrepresentation, which is unintentional.

¹³ See Kline 1980 for a discussion of alternative responses to situations where plausible assumptions lead to contradictory results.

¹⁴ Wood here is arguing for investigative methods that focus more upon context and locality. Thus "... what is important for growth is not mainly *how much* you trade, but *what* you trade and *how you manage* your trade" (68).

¹⁵ Compare eg Panlilio 1963 for a discussion in a natural science.

¹⁶ See also Hausman 1992.

¹⁷ Here a caveat is in order. Such 'differences in behaviour' can be nested or non-nested. That is, they can suggest different underlying models, along the lines of the comment from Camerer quoted above, **or** they can suggest that the parameters of the model are different in the different cases. In the simpler

language of the engineer, Mars rocks are either not rocks at all, or they are, simply redder.

¹⁸ Yvonne Dunlop, personal communication. See Kidd and Shannon 1997 for an example. Standard methods assume lifelong employment, which is not what most women experience. In this case the difference is one of parameters, but if one reads Kidd and Shannon differently, especially their closing remarks about wage determination, they point to more fundamental issues.

¹⁹ They discuss three types of experiments: Public Goods, Ultimatum and Dictator. On reflection, and granted that in many human societies women bear contingent liability in the case of local 'system breakdown' (ie for the children in the case of family break-up), it might be argued that their apparent selfishness when risks appear high is, rather, a qualitative shift in strategy reflecting a need to look after themselves so they can cope with the contingent liability.

²⁰ Ethnographically, an interesting point here was the Chinese willingness to spend time exploring opponents' activities for signs of pre-known strategies, referenced to classic stories and literature on war and its standard stratagems. This the Germans did not seem to do. The interactions were videotaped. In terms of the discussion of instrumental vs procedural rationality, this would seem to be a clear example of search for procedures, and selection of responses in consequence, where knowledge of the local meaning was essential to any persuasive analysis.

²¹ See for two stimulating studies, Lindauer and Pritchett 2002 and Yonay 1998.

²² For a fascinating study of the nature and relationship between interwar neo-classical and institutional economics, see Yonay 1998. He argues that both were swept away by the mathematisation of economics as Samuelson's work became influential in the late 1950s. He also points to a far more widespread influence of institutional economic thinking than some other authors, and also to their concern to find new models to explain the economics data that was coming available (an endeavour that, he argues, failed).

²³ Thus "institutions are unnecessary; ideas and ideologies do not matter; and efficient markets – both economic and political – characterise economics" (2).