sanity, humanity and science

post-autistic economics review

Issue no. 37, 28 April 2006 back issues at www.paecon.net

Subscribers: 8,571 from over 150 countries

Subscriptions are free. To subscribe, email "subscribe". To unsubscribe, email "unsubscribe". Send to: pae_news@btinternet.com

In this issue:

Why Research Assessment Exercises Are a Bad Thing

Donald Gillies (University College London*)

1. Introduction

In the UK a Research Assessment Exercise (henceforth abbreviated to RAE) was introduced in 1986 by Thatcher, and was continued by Blair. Now the idea seems to be catching on, and RAEs are being introduced in many countries. But are such RAEs really a good thing? In this paper I want to argue that they are not. The rationale for conducting an RAE is presumably that it will improve research output. However I will show that an RAE is likely to have the opposite effect, and make the quality of research produced worse than it was before.

An RAE usually involves a double use of peer review. A researcher has to submit publications, and these will in general have been peer reviewed. Then the review by the RAE panels is itself a peer review. This exclusive reliance on peer review is the first major defect of an RAE, for, as I will argue, it is likely to lead to a systematic failure to recognise ground-breaking research. Indeed the study of the history of science shows that peer review can give results which later turn out to have been quite erroneous. It often happens that researchers produce work which is judged at the time by their fellow researchers to be worthless, but which is later (sometimes much later) recognized to have been a major advance. In the next section I will give three examples of this phenomenon selected from different branches of science, namely (i) mathematics, (ii) medicine, and (iii) astronomy. Then in section 3 I will explain why this occurs, using Kuhn's philosophy of science. These results will be used in sections 4 and 5 to analyse the likely effects of an RAE on the quality of research and on wealth-generating technologies. The last section argues that the conclusions drawn can be extended to economics and the social sciences.

2. Examples of the Failure of Peer Review

Mathematics

My first example is taken from the field of mathematics and I want to consider an important advance in mathematical logic. This advance was made by Frege in a booklet published in 1879, and which is usually referred to by its German title of *Begriffsschrift*, which means literally: 'concept-writing'. Frege worked all his life in the mathematics department of Jena university.

In the *Begriffsschrift*, Frege presents for the first time an axiomatic-deductive development of the propositional calculus and of the predicate calculus (or quantification theory). These subjects are the core of modern mathematical logic, and are expounded in the opening chapters of most modern textbooks on the subject.

Frege's remarkable achievement has been fully recognised by experts in the field since the 1950s. William and Martha Kneale in their 1962 history of logic write: 'Frege's *Begriffsschrift* is the first really comprehensive system of formal logic. … Frege's work … contains all the essentials of modern logic, and it is not unfair either to his predecessors or to his successors to say that 1879 is the most important date in the history of the subject.'

However the significance of Frege's work was certainly not realised by his contemporaries working in the same field. There were 6 reviews of the *Begriffsschrift* – 4 by Germans, 1 by a Frenchman, and 1 by an Englishman. Of these 6, 1 was favourable, but the other 5 were not only hostile but even completely dismissive. Schröder, the leading German logician of the time, wrote: ' ... the present little book makes an advance which I should consider very creditable, if a large part of what it attempts had not already been accomplished by someone else, and indeed (as I shall prove) in a doubtlessly more adequate fashion.' Tannery in France wrote: 'In such circumstances, we should have a right to demand complete clarity or a great simplification of formulas or important results. But much to the contrary, the explanations are insufficient, the notations are excessively complex; and as far as applications are concerned, they remain only promises.' Venn in England entirely agreed with Schröder that Frege had made no advance in the subject, and had indeed taken a step backwards. He wrote: ' … it does not seem to me that Dr. Frege's scheme can for a moment compare with that of Boole. I should suppose, from his making no reference whatever to the latter, that he has not seen it, nor any of the modifications of it with which we are familiar here. Certainly the merits which he claims as novel for his own method are common to every symbolic method.' Venn concluded his review by saying: ' … Dr Frege's system … seems to me cumbrous and inconvenient.'

The importance of Frege's work only began to be recognised towards the end of the 19th century, twenty years after it has been published, and then only by a few avant-garde researchers such as Peano in Italy and Bertrand Russell in Britain.

Medicine

My second case-history (Semmelweis and antisepsis) comes from a completely different branch of science. Semmelweis's investigation was into the causes of puerperal fever, which was, at the time, the principal cause of death in childbirth.

Semmelweis was Hungarian, but studied medicine at the University of Vienna. In 1844 he qualified as a doctor, and, later in the same year obtained the degree of Master of Midwifery. From then until 1849, he held the posts of either aspirant to assistant or full assistant at the first maternity clinic in Vienna. It was during this period that he carried out his research.

The Vienna Maternity Hospital was divided into two clinics from 1833. Between 1833 and 1840, medical students, doctors and midwives attended both clinics, but, thereafter, although doctors went to both clinics, the first clinic only was used for the instruction of medical students who were all male in those days, and the second clinic was reserved for the instruction of midwives. When Semmelweis began working as a full assistant in 1846, the mortality statistics showed a strange phenomenon.

Between 1833 and 1840, the death rates in the two clinics had been comparable, but, in the period 1841-46, the death rate in the first clinic was 9.92% and in the second clinic 3.88%. The first figure is more than 2.5 times the second – a difference which is certainly statistically significant. Semmelweis was puzzled and set himself the task of finding the cause of the higher death rate in the first clinic.

After considering many different hypotheses, Semmelweis finally hit on the idea that some cases of puerperal fever might be caused by doctors transferring particles from corpses to the patients. In fact professors, assistants and students often went directly from dissecting corpses to examining patients in the first clinic. It is true that they washed their hands with soap and water, but perhaps some cadaverous particles still adhered to their hands. Indeed this seemed probable since their hands often retained a cadaverous odour after washing. The doctors and medical students might then infect some of the patients in the first clinic with these cadaverous particles, thereby giving them puerperal fever. This would explain why the death rate was lower in the second clinic, since the student midwives did not carry out post-mortems.

In order to test this hypothesis, Semmelweis, from some time in May 1847, required everyone to wash their hands in disinfectant before making examinations. At first he used *chlorina liquida*, but, as this was rather expensive, chlorinated lime was substituted. The result was dramatic. In 1848 the mortality rate in the first clinic fell to 1.27%, while that in the second clinic was 1.30%. This was the first time the mortality rate in the first clinic had been lower than that of the second clinic since the medical students had been divided from the student midwives in 1841.

Through a consideration of some further cases, Semmelweis extended his theory to the view that, not just cadaverous particles, but any decaying organic matter, could cause puerperal fever if it entered the bloodstream of a patient.

Let us next look at Semmelweis's theory from a modern point of view. Puerperal fever is now known as 'post-partum sepsis' and is considered to be a bacterial infection. The bacterium principally responsible is *streptococcus pyogenes*, but other *streptococci* and *staphylococci* may be involved. Thus, from a modern point of view, cadaverous particles and other decaying organic matter would not necessarily cause puerperal fever but only if they contain a large enough quantity of living *streptococci* and *staphylococci*. However as putrid matter derived from living organisms is a good source of such bacteria, Semmelweis was not far wrong.

As for the hand washing recommended by Semmelweis, that is of course absolutely standard in hospitals. Medical staff have to wash their hands in antiseptic soap (hibiscrub), and there is also a gelatinous substance (alcogel) which is squirted on to the hand. Naturally a doctor's hands must be sterilised in this way before examining any patient – exactly as Semmelweis recommended.

This then is the modern point of view, but how did Semmelweis's contemporaries react to his new theory of the cause of puerperal fever and the practical recommendations based on it? The short answer is that Semmelweis's reception by his contemporaries was almost exactly the same as Frege's. Semmelweis did manage to persuade one or two doctors of the truth of his findings, but the vast majority of the medical profession rejected his theory and ignored the

practical recommendations based upon it. This can be illustrated by one typical reaction. After Semmelweis had made his discoveries in 1848, he and some of his friends in Vienna wrote about them to the directors of several maternity hospitals. Simpson of Edinburgh replied somewhat rudely to this letter saying that its authors obviously had not studied the obstetrical literature in English. Simpson was of course a very important figure in the medical world of the time. He had introduced the use of chloroform for operations, and had recommended its use as a pain-killer in childbirth. His response to Semmelweis and his friends is very similar in character to Venn's review of Frege's *Begriffsschrift*.

The failure of the research community to recognise Semmelweis's work had of course much more serious consequences than the corresponding failure to appreciate Frege's innovations. In the twenty years after 1847 when Semmelweis made his basic discoveries, hospitals throughout the world were plagued with what were known as 'hospital diseases', that is to say, diseases which a patient entering a hospital was very likely to contract. These included not just puerperal fever, but a whole range of other unpleasant illnesses. There were wound sepsis, hospital gangrene, tetanus, and spreading gangrene, erysipelas (or 'St. Anthony's fire'), pyaemia and septicaemia which are two different forms of blood poisoning, and so on. Many of these diseases were fatal. From the modern point of view, they are all bacterial diseases which can be avoided by applying the kind of antiseptic precautions recommended by Semmelweis.

In 1871, over twenty years after his rather abrupt reply to Semmelweis and his friends, Simpson of Edinburgh wrote a series of articles on 'Hospitalism'. These contained his famous claim, well-supported by statistics, that 'the man laid on the operating-table in one of our surgical hospitals is exposed to more chances of death than the English soldier on the field of Waterloo'. Simpson thought that hospitals infected with pyaemia might have to be demolished completely. So serious was the crisis, that he even recommended replacing hospitals by villages of small iron huts to accommodate one or two patients, which were to be pulled down and re-erected periodically. Luckily the theory and practice of antisepsis were introduced in Britain by Lister in 1865, and were supported by the germ theory of disease developed by Pasteur in France and Koch in Germany. The new antiseptic methods had become general by the mid 1880s, so that the hospital crisis was averted. All the same, the failure to recognise Semmelweis's work must have cost the lives of many patients.

Astronomy

I now turn to my third example (Copernicus and astronomy). Copernicus (1473-1543) was born in which is now Poland and studied at universities in both Poland and Italy. Through the influence of his uncle, he obtained the post of Canon of Frauenberg Cathedral in 1497, and held this position until his death. Copernicus' duties as canon seem to have left him plenty of time for other activities, and he seems to have devoted much of this time to developing in detail his new theory of the universe. This was published as *De Revolutionibus Orbium Caelestium*, when Copernicus was on his death bed. In the preface Copernicus states that he had meditated on this work for more than 36 years.

There is little doubt that during Copernicus' lifetime and for more than 50 or 60 years after his death, his view that the Earth moved was regarded as absurd, not only by the vast majority of the general public, but also by the vast majority of those who were expert in astronomy.

Although the majority of expert astronomers of the period would have dismissed the Copernican view as absurd, a few such astronomers, notably Kepler and Galileo, did side with Copernicus and carried out researches developing his theory until, in due course, it won general acceptance by astronomers.

3. Kuhn's Distinction between Normal and Revolutionary Science

I have given three examples of the failure of peer review, and, of course, many others could be given. But why do such errors occur? How is it possible for experts in a field to judge as worthless what is later seen to be a major advance? This phenomenon is explained by Kuhn's theory of scientific development as set out in his *The Structure of Scientific Revolutions* (1962). Kuhn's view is that science develops through periods of *normal science* which are characterised by the dominance of a *paradigm*, but which are interrupted by occasional revolutions during which the old paradigm is replaced by a new one. During a period of normal science, the researchers in a given field all accept the dominant paradigm. So those who diverge from the paradigm are regarded as 'cranks' who 'don't know what they are talking about'. Usually such dissidents are indeed cranks who don't know what they are talking about, but every so often they turn out to be a Frege, a Semmelweis, or a Copernicus, and initiate a revolutionary advance in the subject. An important consequence of Kuhn's theory is that the mistaken judgements regarding Copernicus, Semmelweis and Frege are not features of science's past, but are likely to recur over and over again, *because they are features of the development of science in general*.

4. Analysis of the Likely Effects of an RAE

Let us begin by considering the effects of an RAE on normal science. In a period of normal science, those working in a branch of the subject will all accept the dominant paradigm, and no revolutionary alternative will have been suggested. It will then be an easier matter for the experts in the field to judge who is best according to the criteria of the dominant paradigm. Allocating research funding to these most successful 'puzzle solvers', as Kuhn calls them, will usually enable the normal science activity of puzzle solving to continue successfully.

Even in the relatively unproblematic case of normal science, however, an excessive reliance on peer review can lead to mistakes. Suppose research is required on some problem, and there are four different approaches to its solution which lead to four different research programmes. This situation is still possible, and indeed often occurs, in normal science, for the four different research programmes could all be compatible with the dominant paradigm. It may be almost impossible to say at the beginning of the research which of the four programmes is going to lead to success. Suppose it turns out that only research programme number 3 is successful. The researchers on programmes 1, 2 & 4 may be just as competent and hard-working as those on programme 3, but, because their efforts are being made in the wrong direction, they will lead nowhere. In a case like this, a thoughtless use of peer review as a tool could easily lead to wrong decisions. Suppose that programme 3, the one which eventually leads to the solution of the problem, is initially supported by only a few researchers. A peer review conducted by a committee chosen at random from those working on the problem might

well contain an overwhelming majority of researchers working on programmes 1, 2 & 4, and such a committee could easily recommend the cancellation of funding for research programme 3, a decision which would have disastrous long term results.

This point can be clarified and extended by introducing a distinction between two types of error (Type I error, and Type II error). A research assessment procedure commits a Type I error if it leads to funding being withdrawn from a research programme which would have obtained excellent results had the funding been continued. A research assessment procedure commits a Type II error if it leads to funding being continued for a research programme which obtains no good results however long it goes on. This distinction enables us to state a second major defect of an RAE. An RAE concentrates exclusively on eliminating Type II errors. The idea behind an RAE is to make research more cost effective by withdrawing funds from bad researchers and giving them to good researchers. No thought is devoted to the possibility of making a Type I error, the error that is of withdrawing funding from researchers who would have made important advances if their research had been supported. Yet the history of science shows that Type I errors are much more serious than Type II errors. The case of Semmelweis is a very striking example. The fact that his line of research was not recognised and supported by the medical community meant that, for twenty years after his investigation, thousands of patients lost their lives and there was a general crisis in the whole hospital system.

In comparison with Type I errors, Type II errors are much less serious. The worst that can happen is that some government money is spent with nothing to show for it. Moreover Type II errors are inevitable from the very nature of research. Suppose in our example of the 4 competing research programmes, programme 3 is cancelled in order to save money (Type I error), then all the money spent on research in the problem will lead nowhere. It will be a total loss. On the other hand if another unsuccessful programme (programme 5) is also funded, the costs will be a bit higher but a successful result will be obtained. This shows why Type I errors are much more serious than Type II errors, and why funding bodies should make sure that some funding at least is given to every research school and approach rather than concentrating on the hopeless task of trying to foresee which approach will in the long run prove successful.

So an RAE may well have a damaging effect even on normal science. Yet normal science tends to be routine in character and to produce small advances rather slowly. Surely, however, we want a research regime to encourage big advances in the subject, exciting innovations, breakthroughs, etc.

It is precisely here that an RAE is likely to fail in the most serious way. Any big advance is likely to have something revolutionary about it, something which challenges accepted ideas and paradigms. However it is precisely in these cases, as we have shown above, that an RAE with its excessive reliance on peer review is likely to have a very negative effect. Our conclusion then is that an RAE is likely to shift the research community in the direction of producing the routine research of normal science resulting in slow progress and small advances. At the same time it will have the effect of tending to stifle the really good research – the big advances, the exciting innovations, the major breakthroughs. Clearly then the overall effect of an RAE is likely to be very negative as regards research output.

5. The effects of an RAE on wealth-generating technologies

An RAE is also likely to impact very negatively on the production of wealth-generating science-based technologies. The reason for this is that the most striking technologies from the point of view of wealth-generation are often based on revolutionary scientific advances. This is well-illustrated by the three examples considered in this paper. Copernicus' new astronomy led to much better astronomical tables, and so to a much improved navigation. This greatly helped the profitable development of European sea-borne trade in the 17th and 18th centuries. The new mathematical logic introduced by Frege was essential for the development of the computer. It is significant here that Bertrand Russell was one of the first to recognise and develop Frege's work. Russell established an interest in mathematical logic in the UK, which passed on to two later researchers at Cambridge: Max Newman and his student Alan Turing. After the Second World War, Newman and Turing were part of the team at Manchester which produced the Manchester Automatic Digital Machine (MADM). This started running in 1948, and can be considered as the first computer in the modern sense. Thus Russell's early recognition of Frege's revolutionary innovations led indirectly to the UK taking an early lead in the computer field. This early lead was later lost, as we know, but this was owing to lack of sufficient investment by either the public or private sectors. There was no problem with the UK's research community in those pre-RAE days. Our third case was concerned with the revolutionary introduction of antisepsis in conjunction with revolutionary new theories about the causes of disease. We focussed on Semmelweis whose research work was rejected by the medical community of his time. As we remarked, however, Lister was more successful, and was able to persuade the medical community in the UK to accept antisepsis. This was obviously of great benefit to patients, but I would now like to add that it led to very successful business developments. For his new form of surgery Lister needed antiseptic dressings, and he devoted a lot of time and thought to working out the best design and composition of such dressings. As his ideas came to be accepted, the demand for these dressings increased and companies were formed to produce them. One of these was founded by a pharmacist Thomas James Smith. In 1896, he went into partnership with his nephew Horatio Nelson Smith to produce and sell antiseptic dressings. They called the firm Smith and Nephew. Today Smith and Nephew is a transnational company operating in 33 countries and generating sales of £1.25 billion. The company is still involved in wound care as one of its three main specialities, but it has expanded into orthopaedics and endoscopy. One of its well-known products is elastoplast which was developed in 1928. The general design of elastoplast is based on some of the antiseptic dressings developed by Lister. The commercial success of Smith and Nephew is a good illustration of the importance of having a satisfactory research regime in the UK. If Lister's research on antisepsis had met the same fate as that of Semmelweis only 17 years earlier, then the firm of Smith and Nephew would not be with us today.

6. General Conclusions

The examples I have given are taken from science, using this term in a broad sense to include mathematics and medicine, as well as the natural sciences such as physics and chemistry. But do the conclusions drawn apply also to economics and the social sciences? It seems to me clear that they do. If areas such as mathematics and experimental medicine, which

are normally thought of as unproblematic, raise severe problems as regards peer review, it can hardly be denied that such problems are going to worse in areas such as economics and the social sciences where political and ideological factors are much stronger. So for mathematics, the natural sciences, medicine, economics and the social sciences, we can draw the following conclusion.

An RAE is very expensive both in money and in the time which academics have to devote to it. Its likely effect is to shift the research community in the direction of producing the routine research of normal science resulting in slow progress and small advances, while tending to stifle the really good research – the big advances, the exciting innovations, the big breakthroughs. Thus a great deal of tax payers' money will be spent on an exercise whose likely effect is to make research output worse rather than better. Only one conclusion can be drawn from this, namely that RAEs should be abolished rather than introduced..

Note

This is a shortened version of a paper entitled: 'Lessons from the History and Philosophy of Science regarding the Research Assessment Exercise' which was read at the Royal Institute of Philosophy in London on 18 November 2005 in a series of talks on the Philosophy of Science. The series will be published by Cambridge University Press in 2006. For convenience of reading I have not included exact references and other academic apparatus in this shortened version, but they are to be found in the longer version, which is available on my website: www.ucl.ac.uk/sts/gillies.

Author contact: donald.gillies@ucl.ac.uk

SUGGESTED CITATION:
Donald Gillies," Why Research Assessment Exercises Are a Bad Thing " *post-autistic economics review*, issue no. 37, 28 April 2006, article 1, pp. 2-9, http://www.paecon.net/PAEReview/issue37GallagherZarsky.htm

Forum on Economic Reform

In recent decades the alliance of neoclassical economics and neoliberalism has hijacked the term "economic reform". By presenting political choices as market necessities, they have subverted public debate about what economic policy changes are possible and are or are not desirable. This venue promotes discussion of economic reform that is not limited to the one ideological point of view.

Rethinking Foreign Investment for Development

Kevin P.Gallagher and Lyuba Zarsky[1] (Boston University and Businesses for Social Responsibility, USA)

In the 1990s, foreign direct investment (FDI) came to be seen as a "miracle drug"—a jumpstart to economic growth and sustainable industrial development, especially in developing countries. Policies to attract FDI became the centerpiece of both national development strategies and supra-national investment agreements.

This paper examines case study and statistical evidence about the impacts of FDI in developing countries on economic growth, technology spillovers and environmental performance. Mirroring the heterogeneity of developing countries, we find that there is no consistent relationship: the impact of FDI on each variable has been found to be positive, neutral, or even negative. Key variables are domestic policies, capacities and institutions.

We conclude that the purported benefits of FDI are exaggerated and its centrality in development strategies misplaced. Rather than attract FDI *per se*, development policies should aim to promote endogenous local capacities for sustainable production. With the right national and global policy framework, FDI could help in that process.

I. Introduction

In the 1990s, foreign direct investment (FDI) came to be seen in much the same light as export-led growth in the 1970s—a jumpstart to sustained economic growth in developing countries. Optimism was fuelled by a surge in FDI inflows, outstripping other forms of public and private finance. Led by transnational corporations (TNCs), the hope was that FDI would transfer superior technology and management skills, stimulate domestic investment and growth, generate efficiency spillovers, and integrate developing country firms into global markets.

An added twist was that, directly and indirectly, FDI would also boost environmental performance. The direct benefits would be gained from the transfer of cleaner technology and the better environmental management practices of TNCs. The expectation was that TNCs would implement and diffuse in their developing country operations the high standards required in Europe, Japan, and the US. The indirect benefits would be gleaned largely from the impacts of FDI on economic growth. By increasing per capita income, growth would promote cleaner consumer goods and greater citizen demand for environmental protection.

[1] Kevin P. Gallagher is assistant professor of international relations at Boston University, Lyuba Zarsky is research director at Businesses for Social Responsibility. Both are senior researchers at the Global Development and Environment Institute, Tufts University. This paper draws on their earlier contribution to the volume, *International Investment for Sustainable Development* edited by Lyuba Zarsky (Earthscan, 2004).

Two standard prescriptions followed. First, since "more is better," policies to attract FDI should be at the heart of national development strategies. An understanding of what these policies should be has changed over the decades. In the 1980s, the emphasis was on "getting the prices right," that is, the elimination of domestic policies, such as energy and food subsidies, which create a cleavage between domestic and global prices. In the 1990s, the trend was toward "getting the macro-policies right", especially the deregulation of financial markets.

Currently, the focus is on fashioning the right "enabling environment" for FDI; that is, creating or strengthening legal, regulatory and political institutions which provide transparency, property protection, and financial stability to foreign investors. Even if desirable, the creation of such institutions is neither an overnight nor costless affair. If success in attracting FDI awaits their creation, many developing countries, especially the poorest, will be waiting a long time.

The second prescription is that investment agreements—global, regional, bilateral--should aim to "make the world safe" for FDI, including by expanding the protections for and the rights of foreign investors. Greater rights for investors have come at the expense of flexibility and diversity of national development policies. Moreover, TNCs have used new rights to challenge national environmental and health regulation.

This paper examines recent statistical and case study evidence about the impacts of FDI in generating efficiency spillovers, promoting growth and improving environmental performance in developing countries. The studies paint an ambiguous picture: FDI has been found to have positive, neutral, or even negative impacts on all three counts. The poorer the country, the more likely is the FDI impact negative. While further studies are needed, especially case studies of particular TNCs and sectors, the key variables appear to be the domestic institutional and policy context, on the one hand, and TNC practices on the other hand.

The central argument of the paper is that FDI is no "miracle drug" for economic development, environmentally sustainable or otherwise. Structuring development strategies and investment regimes around the assumption that it *is* a miracle drug may act, ironically, to undermine the positive contributions that FDI could potentially make to nurturing local capacities for sustained economic growth.

II. FDI in Developing Countries

FDI is a financial investment in a domestic enterprise by which a foreign investor gains a significant equity stake in the firm. In most national accounting systems, FDI is defined as an equity share of 10 percent or more. Besides selling equity, enterprises finance their operations via debt, including loans from banks and other financial institutions and corporate bonds.

The major players in FDI are transnational corporations (TNCs). The world's largest 100 TNCs held nearly $2 trillion in foreign assets at the end of the 1990s, and employed over six million people in their foreign affiliates. All of the top 10 and nearly 90 percent of the top100 TNCs are from the United States, Japan and the European Union (UNDTAD, 2002). Given the predominance of TNCs, a conventional definition of FDI is a "form of international inter-firm

cooperation that involves significant equity stake and effective management decision power in, or ownership control of, foreign companies" (De Mello, 1999, p. 135).

FDI, in short, is more than a flow of capital. It is a cross-border expansion of production undertaken primarily by large corporations. The internationalization of production is at the heart of the process of globalization. FDI takes place in two ways: "Mergers and Acquisitions" (M&As) that is, the purchase by TNCs of existing domestic companies, in whole or in part; and "Greenfield Investment," that is, additions to the capital stock and the creation of new productive capacity.

Why and Where? Determinants of FDI

Why does a TNC decide to expand production overseas rather than exporting products it makes at home? What determines which countries TNCs will invest in?

Mainstream trade theory suggests that FDI is driven by imperfections in markets for goods or factors of production, including labor and technology. To make FDI profitable, a firm must have some distinctive asset—technology, global marketing capacities, management skills—not possessed by domestic firms (Blomstrom and Kokko, 1996). The firm is thus able to earn a "rent" by producing in the host community.

Because the superior technology and/or management skills of foreign firms raise the efficiency and productivity of domestic firms, at least in theory, FDI is sometimes said to be"efficiency-seeking". In addition, FDI can be "market-seeking". If a government restricts imports, FDI may be prompted by efforts to get under import barriers or high tariffs, or to gain preferential treatment to third country markets. If transport costs are high, local production would promote competitiveness in local markets, especially if there are economies of scale.

Many developing countries seek to attract FDI on the basis of low labor costs. Lacking the technology and/or global marketing networks of the TNCs, developing countries conceive of FDI primarily in terms of providing capital to employ low or unskilled workers. Differences in labor costs between countries are kept large by immigration restrictions, that in, an "imperfection" in global labor markets.

An understanding both of what drives TNCs to invest globally and what determines the location of FDI is murky. Traditionally, studies have found that the most unambiguous "pull" factor drawing FDI to particular countries—both developed and developing--is the market size (per capita GDP) of the host economy (Chakrabarti, 2001). For the most part, it seems, TNCs invest where markets are large, either to expand markets or to earn rents, or both.

With globalization have come predictions that "non-traditional" factors would increasingly determine the location of FDI, opening up new possibilities for developing countries. These factors include cost differences between locations, the quality of infrastructure, the ease of doing business, and the availability of skills. However, a recent study found that globalization has not "changed the rules of the game" and confirmed that market size is still the dominant driver of the distribution of FDI in developing countries (Nunnenkamp and Spatz, 2002).

There is also close to a consensus that macroeconomic and political stability are needed to attract FDI. Countries with volatile exchange rates and high and growing trade deficits tend to be negatively correlated with FDI. Evidence on whether low (or high) labor costs attract FDI, on the other hand, is ambiguous. A review of sixteen studies found that, in six studies, low wages attracted FDI, while four studies found FDI to be correlated with higher wages and six found labor costs to be insignificant (Chakrabarti, 2001). Other studies emphasize the importance of human capital, in particular the level of education, in attracting FDI.

There has been much debate in recent years about the role of environmental factors, especially differences in enforced standards, in attracting or repelling FDI. No consistent statistical evidence has been found that differences in standards affect TNC location decisions, presumably because, in most industries, environmental costs are a small component of total costs (Zarsky, 1999). However, case studies have found that in certain "dirty" industries, such as leather tanning, more stringent standards in OECD countries propelled companies to shift production to countries with lower standards (Mabey and McNally, 1999).

Finally, there is the question of whether standards of treatment of foreign investors play a role in attracting FDI to particular countries. The surge in the number of Bilateral Investment Agreements (BITs) concluded in the 1990s was fuelled--at least on the part of developing countries--by the hope that expanding the rights and protections of foreign investors would attract FDI. However, there is no evidence that it has done so. Nonetheless, with over 2000 BITs in force, investor protections have become the norm. They may not provide marginal benefits, but their absence might have marginal costs in attracting FDI.

FDI Trends in the 1990s

Inflows of FDI soared to unprecedented levels during the 1990s. From 1970 to 1990, global FDI inflows averaged $58 billion a year, or less than one half of one percent of global GDP. Between 1990 and 1995, FDI inflows averaged $225 billion per year and surged to $828 billion per year between 1996 and 2000. With a total of $1.5 trillion in 2000 alone, FDI inflows were 4 percent of global GDP (Figures 1.1 and 1.2).

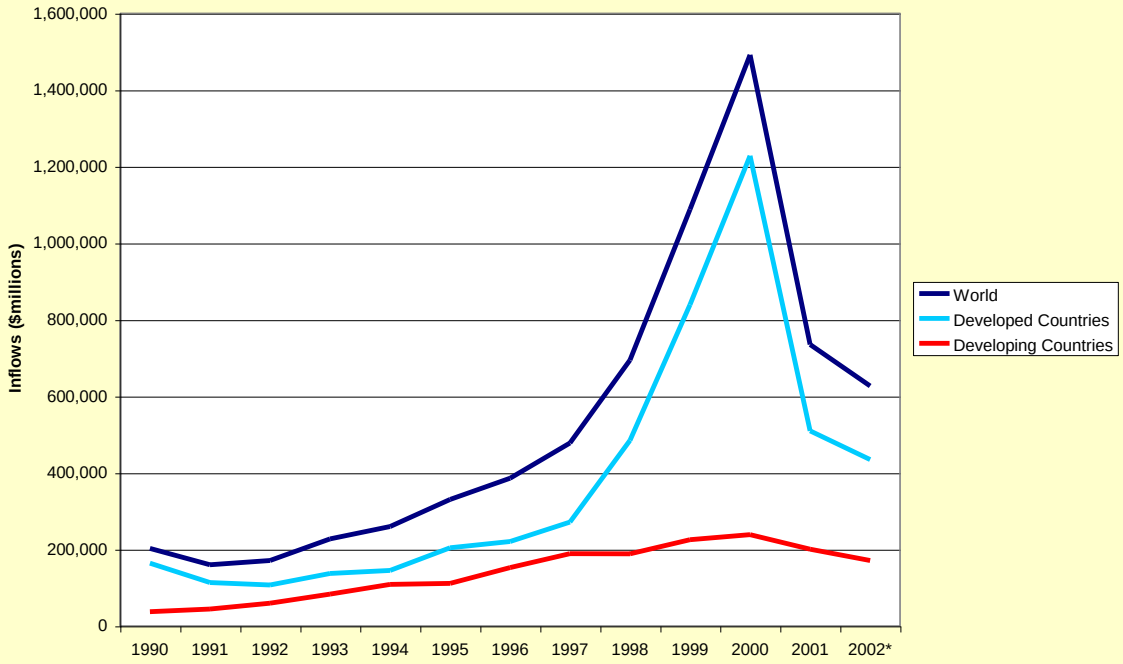
Originating from North America, Europe and Japan, the FDI boom affected both developing and developed countries. Indeed, the lion's share— nearly three quarters on average—of FDI inflows went from one OECD country to another. However, the boom was big news in developing countries because FDI inflows started from a low base. In 1990, FDI comprised only about a quarter of all capital flows into developing countries. By 2000, the share of FDI had climbed to 60 percent (UNCTAD, 2002a, p. 12).

Moreover, "official flows," that is, multi-lateral and bilateral development aid (ODA), remained stagnant during the 1990s, with an annual flow of $54 billion. Many analysts began to predict—or hope--that FDI would "dwarf" or replace ODA as the primary source of global development capital.

The promise of FDI as a replacement for ODA, however, largely remains to be fulfilled. For 55 of the world's 70 poorest countries, ODA flows outstripped FDI in the late 1990s. For 42 poor countries, ODA flows were twice the size of FDI. Indeed, FDI "dwarfed" ODA in only seven of the poorest countries.[1]

However, only a small part of global FDI inflows—about 30 percent on average between 1990 and 2001--went to developing countries. Indeed, the developing country share fell off sharply between 1997 and 2000, falling from 39 to 16 percent.

Figure 1.1 FDI Flows in the World Economy: 1990 to 2002



Source: UNCTAD, 2002.

FDI inflows are highly concentrated in ten, mostly large developing countries, led by China, Brazil, and Mexico. Between 1990 and 2000, the "top ten" garnered 76 percent of the total FDI flowing into developing countries. The trend towards concentration seems to be intensifying: in 2001, the top ten share rose to 81 percent (Table 1.1).

Though they are miniscule in the global picture, FDI inflows to poor, developing countries may comprise a significant part of total national investment and/or GDP. For example, between 1996 and 1999, FDI comprised about 10 percent of GDP in Bolivia, 26 percent in Lesotho, and 26 percent in Thailand (Overseas Development Institute, 2002, Figure 9). Overall, the ratio of FDI to GDP in developing countries rose from about 1 percent to 3.5 percent between 1990 and 2000.

Table 1.1 10 Largest Developing Country Recipients of FDI inflows
($US billions)

Top 10	***(1990-2000) ave***	***Top 10***	***2001***
China*	43.4	China*	69.7
Brazil	12.0	Mexico	24.7
Mexico	10.1	Brazil	22.5
Argentina	7.2	Bermuda	9.9
Singapore	7.1	Poland	8.8
Malaysia	4.7	Singapore	8.6
Bermuda	4.7	Chile	5.5
Poland	3.7	Czech Republic	4.9
Chile	3.3	Taiwan	4.1
South Korea	3.2	Thailand	3.8
Top 10 total:	99.5		162.5
Total For Developing Countries:	130.9		200.9
Top 10 share:	76%		81%
Top 3 share:	50%		58%

Source: UNCTAD, 2002.

The 1990s FDI boom will probably prove to be a bubble, though the long term trend is towards expansion. [2] Global FDI inflows declined sharply in 2001, falling by nearly 51 percent. Data for 2002 suggest that this trend will persist: FDI inflows into developing countries fell from about $170 billion in 2001 to an estimated $145 billion in 2002 (World Bank, 2003b). However, Latin America accounted for nearly all of the contraction, while FDI in Asia held steady.

Privatization Wave

In both developed and developing countries, the 1990s FDI boom was led by cross-border mergers and acquisitions. According to UNCTAD, the share of cross-border M&As in world FDI flows increased from 52 percent in 1987 to 83 percent in 1999. Although most cross-border M&As occur within OECD countries, M&As accounted for close to 70 percent of FDI inflows to developing countries in 1999 (UNCTAD, 2000a, p. 14 and Table I.5).

In developing countries, M&As were concentrated in newly privatized state-owned companies spanning water, energy, telecommunications and financial services. In Mexico, for example, the US company Citigroup acquired Banamex, Mexico's largest bank, for $12.5 billion in 2001. The single transaction accounted for half of all FDI inflows to Mexico in that year (*ibid*).

In many countries, the acquisition of formerly public services by TNCs, especially water, has generated widespread concern and popular resistance. Despite promises of more efficient, reliable and equitable access to water, social advocates are concerned that privatized water systems under TNC management will price the poor out of the market. In Cochabamba, Bolivia, for example, a "water war" erupted after the city's water system came under the management of an international partnership involving Bechtel Corporation. When the price for water rose by

between 35 and 400 percent, city residents took to the streets and the municipal government canceled the contract (Dolinsky, 2001).[3]

The privatization wave changed the sectoral composition of FDI in developing countries, substantially increasing the share going to the service sector. In 1999, services accounted for 37.3 percent of FDI inflows, up from 20.7 percent in 1988 (UNCTAD, 2000). Though shrinking in relative terms in all regions except Africa, the manufacturing sector continues to account for the largest overall share, about 55 percent in 1999. In all regions except Latin America, the share of FDI in the "primary" sector, mostly agriculture and mining, decreased in the 1990s and accounted for only 5.4 percent in 1999 (*ibid).*

Despite its small share in aggregate flows, FDI in the primary sector forms the largest share of overall inflows in particular countries, including some of the poorest. Among the 49 "Least Developed Countries", four oil-exporting countries—Angola, Equatorial Guinea, and Sudan--receive nearly half of all FDI inflows. According to a recent report by UNCTAD, there is a strong link between dependence on primary commodities and poverty (UNCTAD, 2002b).

III. Does FDI Promote Economic Development?

The pursuit of FDI as an engine of growth is a formula prescribed by mainstream economic theory, as as well as the IMF and other global development organizations. "The overall benefits of FDI for developing country economies are well documented," claims a 2002 report undertaken for the OECD's Committee on International Investment and Multinational Enterprises (OECD, 2002b, p. 5). Based on consultations with OECD member governments and business, labor and NGO advisors, the report, titled *Foreign Direct Investment for Development, Maximising Benefits, Minimising Costs,* nicely sums up conventional wisdom about the promise of FDI:

> Given the appropriate host-country policies and a basic level of development, a preponderance of studies shows that FDI triggers technology spillovers, assists human capital formation, contributes to international trade integration, helps create a more competitive business environment and enhances enterprise development. All of these contribute to higher economic growth, which is the most potent tool for alleviating poverty in developing countries (*ibid*).

In addition, the report goes on, FDI "*may* help improve environmental and social conditions in the host country by, for example, transferring 'cleaner' technologies and leading to more socially responsible corporate policies" (*ibid;* emphasis added).

The more qualified endorsement of FDI's social and environmental benefits likely stems from the OECD's own commissioned work in this area, including by one of the authors of this paper, which shows that the environmental impacts of FDI may be positive, negative, or neutral, depending on the industrial and institutional context (Zarsky, 1999;OECD, 2002a). But caution is also warranted in assessing the relationship between FDI and economic development. An

examination of recent studies shows that, in this case too, the benefits of FDI are far from well-documented.

Efficiency Spillovers: Is Knowledge "Contagious"?

FDI generates rents to transnational corporations by virtue of their possession of superior technology, management and/or access to global markets. According to economic theory, host communities get "spillover" benefits of the superior asset(s). Indeed, "efficiency spillovers", which occur through the transfer of technologies and management practices, are increasingly seen as the primary way in which to gauge the contribution of FDI to economic growth.

Dubbed a "contagion" effect, knowledge is diffused to local firms and workers, raising the efficiency, productivity and marketing skills of domestic firms (Findlay, 1978). While knowledge diffusion is postulated for TNC investment in both developed and developing countries, it is the transfer from industrialized to developing countries that promises the greatest hope for global economic development.

Efficiency spillovers can occur through several routes, including the copying of TNC technology by local firms and the training of workers who then find employment in local firms or start their own. The most important conduit, however, is the linkage between TNC affiliates and their local suppliers (Lall, 1980). TNCs generate spillovers when they:

- Help prospective suppliers set up production facilities;
- Demand that suppliers meet high quality standards and develop capacities for product innovation—and provide training to enable them to do so;
- Provide training in business management;
- Help suppliers find additional markets, including in sister affiliates in other countries.

To what extent do TNCs actually undertake these activities in developing countries? Put another way, what is the empirical evidence that efficiency spillovers exist? Studies have been of two types: 1) statistical studies which examine trends in key macro-variables, such as domestic investment (gross fixed capital formation) and productivity; 2) case studies of particular industries, such as autos (Moran, 1998) and high-tech (Amsden and Chu, 2003).

In developed countries, limited evidence suggests that the productivity of domestic firms is positively correlated with the presence of foreign firms, though one study found no independent growth effect of FDI (Carkovic and Levine, 2002). However, some studies found the magnitude of spillovers to be very small (Lim, 2001). Moreover, tax policies and other incentives to attract FDI distort firm technology and investment choices and generate *negative* spillovers—a loss in the efficiency of local firms (Blonigen and Kolpin, 2002).

For developing countries, the evidence is mixed. Some studies have found clear evidence of spillover effects, while others have found limited or even negative effects (1.2). An IMF study found "overwhelming" evidence of productivity increases through technology transfer (Graham, 1995). However, a later literature review took a much more nuanced view, finding

that a host of country- and industry-specific variables determined whether FDI generated technology transfer and diffusion in developing countries (Kokko and Blomstrom, 1995).

Several studies suggest that the capture of spillovers depends on host country conditions. One found that the larger the productivity gap between foreign and domestic firms, the less likely were spillovers (Kokko, Tansini, and Zejan, 1996). Moran (1998) found that export performance requirements helped Mexico to capture spillovers from investment by US auto-makers. In their case study of Taiwan, Amsden and Chu (2003), found that spillovers were captured as a result of government policies, especially support for research and development for nationally-owned companies.

Some studies find that, rather than generate positive spillovers, FDI generates *negative* spillovers. Krugman (1998) argues that, generally, domestic investors are more efficient than foreign investors in running domestic firms—otherwise, foreign investors would have purchased them. However, in a financial crisis, such as the crisis which swept East Asia in the late 1990s, domestic firms may be cash-constrained and be available for purchase at "fire-sale" prices. Krugman concludes that a superior cash position, rather than efficiency-enhancing technology or management, drives FDI.

Another study argues that FDI is driven by the information advantage of foreign investors, who are able to gain—and leverage--inside information about the productivity of firms under their control. With their superior information, foreign firms can inflate the price of equities sold in domestic stock markets. The expectation of future stock market opportunities then leads to over-investment and efficiency (Razin, Sadka and Yuen, 1999).

Overall, of the eleven studies reviewed for this paper, only three found unequivocably found that FDI generates efficiency spillovers in developing countries. Two found the opposite, while six found that FDI may or may generate spillovers, depending on local productive, policy or financial conditions (Table 1.2). The evidence suggests that there is no automatic or consistent relationship between FDI and efficiency spillovers, either for developing countries as a whole or for all industries within a county. Realizing the promise of FDI to transfer technology and diffuse knowledge requires conducive policy, institutional and market environments.

Like the benefits of FDI itself, however, there is little consensus on what constitutes a "conducive" policy. Moran (1998) argues that a liberal trade and investment regime which allows TNCs maximum flexibility has the best chance of increasing the efficiency of local firms and integrating them into global supply chains. On the other hand, Moran also found that export requirements may work to speed up TNC host country investments which generate spillovers. Amsden and Chu (2003) conclude that the most important ingredient in capturing spillovers and indeed, in increasing productive capacity in "latecomer" states is a strong state acting to nurture domestic firms through effective, market-friendly and performance-related subsidies.

Table 1.2

Does FDI Generate Spillovers in Developing Countries?

Author/Study	**Year**	**Yes, No, Maybe**	**Explanation**
Aitken, Hanson & Harrison	1997	Yes	Foreign firms act as export catalysts for domestic firms (in Mexico)
Aitken & Harrison	1999	No	No evidence of spillovers (Venezuela)
Amsden & Chu	2003	Maybe	Government industry policies actively promote nationally owned firms (Taiwan)
Graham	1995	Yes	Increases productivity of domestic firms
Kokko & Blomstrom	1995	Maybe	Depends on industry and country characteristics, especially policy environment
Kokko	1994	Yes	Increases productivity of domestic exporting firms (Mexico)
Kokko, Tansini and Zejan	1996	Maybe	Productivity gap between foreign and domestic firms can't be too big (Mexico and Uruguay)
Krugman	1998	No	Domestic investors are more efficient but foreign investors have superior cash position.
Lensink & Morrissey	2001	Yes	Reduces costs of R&D and promotes innovation
Moran	1998	Maybe	Liberal investment climate must encourage integration of local firms into TNC global sourcing and production network
Razin, Sadka, and Yuen	1999	Maybe	Foreign investors can speculate on domestic stock prices, leading to over-investment and inefficiency.

Crowding In--or Crowding Out--Domestic Investment?

The central promise of FDI is that it promotes economic growth, not only through efficiency spillovers but by stimulating or "crowding in" domestic investment. By increasing the productivity and efficiency of local firms, spillovers themselves can help to stimulate domestic investment.

But the "crowding in" effect of FDI on investment may be gained whether or not there are technology spillovers or even if much value beyond labor is added in local production. Assembly operations, for example, where workers put together components made elsewhere, can still drive domestic investment and growth via increases in local consumer demand.[4]

On the other hand, TNCs may undermine local savings and "crowd out" domestic investment by competing in product, service and financial markets and displacing local firms. The loss of domestic firms can undermine market competition, leading to inflated prices and lower quality products.

TNCs can also crowd out domestic investment by borrowing in domestic capital markets, thus driving up interest rates and cost of capital to business. High domestic interest rates may also be the result of deliberate government policies to attract foreign capital. The higher-than-global-average interest rates will also cause the exchange rate to be overvalued, further crowding out domestic firms producing for export. While foreign firms will also suffer from loss of competitiveness, the impact is cushioned by their access to foreign sources of financing. While much is made of the potential for FDI to increase foreign exchange earnings, there is a risk that it will instead contribute to crises in the balance of payments by repatriating profits and by increasing the rate of imports faster than the rate of exports.[5]

Taken together, the risk is that FDI could lead to an overall contraction, rather than an increase, in domestic investment and economic growth. Indeed, in a study that generally argues for the potential benefits of FDI, Moran (1998) cautions that "the possibility that FDI might lead to fundamental economic distortion and pervasive damage to the development prospects of the country is ever present" (p. 2).

What is the more likely "face" of FDI? A host of studies over the past decade have examined the nature of economic benefits and the conditions under which they are—or are not—captured (Table 1.3). Moran (1998) reports on the findings of three separate "net assessments" of the impact of FDI covering 183 *projects* in some 30 countries over the past 15 years. Two studies found that FDI had a positive impact in 55 to 75 percent of the projects they studied. But one study found that FDI had "a clearly negative impact on the economic welfare of the host" in an astonishing 75 percent of the projects studied (p. 25).

Table 1.3

Does FDI Promote Economic Growth in Developing Countries?

Study Author(s)	Year	Yes, No, Maybe	Key Variables
Balasubrmayam, Salisu & Sapsford	1996 1999	Maybe	Requires open or neutral trade regime
Borensztein, Gregorio & Lee	1998	Maybe	Depends on education level of workforce
de Mello	1999	Maybe	Depends on degree of complementarity and substitution between FDI and domestic investment.
Graham and Wada	2001	Yes	Raised per capita GDP in Chinese provinces with FDI concentration
Graham	1995	Maybe	TNC's market power can generate negative impacts
Lensink & Morrissey	2001	Yes	Reduces costs of R&D and promotes innovation
Loungani & Razin	2001	Maybe	Risks
Lim	2001	Maybe	Depends on tax incentives, regulatory & legal impediments, macroeconomic instability
Marino	2000	Maybe	Requires open trade and investment policies
Mallampally & Sauvant	1999	Maybe	Requires human resource development; information and other infrastructure
Markussen & Venables	1999	Yes	Raises productivity and exports of domestic firms; generates spillovers
Rodrik	1999	No	Reverse causality: TNCs locate, rather than drive growth, in more productive and faster growing countries

Economy-wide studies generally have found both positive and negative impacts of FDI on domestic investment. For example, a study by the Brookings Institution covering 58 countries in Latin America, Asia, and Africa found that a dollar of FDI generates another dollar in domestic investment (Bosworth and Collins, 1999). On the other hand, many studies have found that the

investment and/or growth impacts of FDI could be positive *or* negative, depending on a variety of variables, mostly having to do with host country policies.

One study found that the impact of FDI is significantly positive in "open" economies, and significantly negative in "closed" economies (Marino, 2000). Others have found that positive impacts depend on the effectiveness of domestic industry policies; and on tax, financial or macroeconomic policies A World Bank study found that the impacts of FDI depend on the structure and dynamics of the industry, as well as host country policies (World Bank 2003a). In its recent report on the role of FDI in development, the OECD concluded that the overall benefits of FDI, while depend on "the appropriate host-country policies and a basic level of development" (OECD 2002b, p. 9).

Several studies suggest that, to capture the benefits of FDI, a country must *already have reached* some kind of "development threshold". One found that FDI raises growth only in countries where the labor force has achieved a minimum level of education (Borensztein. Gregorio and Lee, 1998). Another found "significant cross-country diversity" in terms of the catalytic role of FDI in developing countries and concluded that the key variables are "country-specific factors", including institutions and policies (de Mello, 1999, p. 148).

Overall, of the twelve studies reviewed for this paper, three found positive links between FDI and economic growth, while one found a negative link and eight studies found that "it depends" Like efficiency spillovers, the positive benefits of FDI on domestic investment and growth depend largely on domestic policies, capabilities, and institutions.

IV. FDI and the Environment

In the last decade, a surge of regional and bilateral investment agreements have promoted the liberalization of investment regimes. These agreements increase expand the rights of foreign investors but, with few exceptions, articulate no environmental or social responsibilities of either investors or governments. Many in the sustainable development community are concerned that, without an environmental framework, liberalization will accelerate environmental degradation (Mabey and McNally, 1999).

The impacts of FDI on the environment can be traced through three routes:

- *Environmental performance* of TNCS;
- *Impacts of economic growth on scale and composition of industry;*
- *Impacts on national and global environmental regulation.*

In this section, we sketch the key mechanisms at play in each of these routes and review some of the theoretical and empirical literature (for a deeper look at the FDI-environment relationship, see Zarsky, 1999).

Performance of TNCs

Two critical strategic and management decisions of TNCs affect their environmental performance. First is the choice of technology, *viz,* whether to invest in newer, cleaner "best available" or to "dump" older, dirtier technologies. In most industries, a range of technologies are in use. Efficiency and "clean-ness" may be a function as much of industry sector as of company choice: some industries are more technologically dynamic than others.

The second decision has to do with management practice, *viz,* whether the corporate parent has implemented an effective EMS (environment management system) and required the same of its overseas subsidiaries and suppliers. NGO advocacy campaigns have increasingly demanded that TNCs to adopt "voluntary initiatives" to go "beyond compliance" in global operations.

One of the promises of FDI for sustainable development is that TNCs, especially from the OECD, will help to drive up standards in developing countries by transferring both cleaner technology and/or better environmental management practices. The evidence for such a trend, however, is mixed. A case study of foreign investment in Chile's mining sector in the 1970s and 1980s found that foreign companies not only performed better than domestic companies but diffused better environmental management practices (Lagos and Valesco,1999). A volume of case studies of FDI in Latin America in the 1990s, including bananas in Costa Rica and soybeans in Brazil, concluded that FDI promotes better management practices (Gentry, 1998).

Other studies have failed to find a positive link between foreign firms and environmental performance in host countries. In statistical studies of Mexico (manufacturing) and Asia (pulp and paper), World Bank researchers found that foreign firms and plants performed no better than domestic companies. Instead, environmental performance was found to depend on 1) the scale of the plant (bigger is better); and 2) the strength of local regulation, both government and "informal" (Dasgupta *et al,* 1998; Hettige *et al,* 1996).

A recent study which examined the environmental performance of TNCs in India had nuanced findings (Ruud, 2002). On the one hand, TNCs were found generally to be transferring state-of–the-art production (though not necessarily pollution control) technologies, rather than dumping older and dirtier technologies. In addition, TNC affiliates were strongly influenced by corporate parents to improve environmental management.

On the other hand, there was no evidence that better environmental management practices were diffused by TNC affiliates to local partners, suppliers or consumers. The author concludes that there is no evidence that TNCs drive a "race to the top"—or the bottom. Rather, the chief insight is that local norms and institutions are of central importance in determining TNC practices and therefore, "FDI inflows do not automatically create a general improvement in environmental performance (Ruud, 2002, p. 116).

Many developing countries, especially the poorest, lack the capacity and/or political will to enforce environmental oversight of industry. In this context, TNCs are largely able to "self-regulate" and have one of three choices: 1) follow local practice and norms; 2) adopt company-

wide global standards based on home country standards; or 3) adopt international standards or "best practice" norms for corporate social responsibility.

Best practice may entail the embrace of broad and encompassing norms, not only for companies themselves but also for their suppliers, such as the OECD's Guidelines for Multinational Enterprises (MNEs) (Table 1.4). The Guidelines spell out eight good practices, including the implementation of an environmental management system and consultation with stakeholders. They also extend good practice to the supply-chains of MNEs. Other approaches to best practice include implementing standards developed by industry associations, such as the chemical industry's Responsible Care initiative.

Table 1.4

OECD's Environmental Guidelines for MNEs

1. Establish an environmental management, which includes
   a) information regarding environmental, health, and safety impacts
   b) measurable objectives and targets for improved environmental performance
   c) monitoring and verification of progress towards targets.

2. Provide adequate and timely information on the potential environment, health and safety impacts of the activities of the enterprise; and communicate and consult with stakeholder communities.

3. Consider full life cycle impacts of production processes, goods and services

4. Use precautionary principle

5. Maintain accident, emergency and contingency plans

6. Continually seek to improve environmental performance by
   a) adopting best practice.
   b) developing products with low environmental impact, including recyclable and energy-efficient
   c) promoting customer environmental
   d) research ways to improve environmental performance over the longer term.

7. Provide training to employees on environment, health and safety

8. Contribute to the development of environmentally meaningful and economically efficient public policy

Source: OECD 2000

Companies make different choices, depending on company culture and the industry they are in. In the petroleum and mineral industries, case studies suggest that TNCs have tended to follow—or even to worsen--local practice (Rosenthal, 2002; Leighton et al, 2002). In many parts of the world, mining operations have generated severe environmental degradation and pollution, including the discharge of toxic substances into river systems, large volume waste disposal, the inadequate disposal of hazardous wastes, and the long run impacts of poorly planned mine closure (Sandbrooke and Mehta, 2002). Multinational oil companies have been the target of protest and criticism for widespread pollution and human rights violations in the Amazon region, Nigeria, Indonesia,. and, increasingly, the Caspian region.

In the high tech sector, on the other hand, American and European TNCs tend to adopt either company-wide standards or international "best practice" for environmental management and community consultation. Within the industry, however, there are "leaders" and "laggards" (Zarsky and Roht-Arriaza, 2002).

One mechanism by which TNCs affect environmental performance in developing countries is via requirements on suppliers and sub-contractors. An increasingly popular trend is for large TNCs to require that suppliers be certified to ISO 14,001, an international standard for environmental management systems.

Scale and Composition Effects

One of the potential, if not automatic, benefits of FDI is that it stimulates economic growth. Without adequate global and national regulation, however, economic growth is likely to accelerate environmental degradation—even if TNCs are good performers--through scale effects. The experience of East Asia, often described as an "economic success story," provides a tragic example. According to the Asian Development Bank, resource degradation and environmental pollution in both East and South Asia is so "pervasive, accelerating, and unabated" that it risks human health and livelihood (Asian Development Bank, 2001, p. 2).

The scale impacts of economic growth on the environment derive largely from unsustainable production and consumption patterns. If FDI was channeled into sustainably produced and sustainably transported goods and services, then the overall impact--even of rapid and high growth--on the environment would presumably be neutral or low. To date, however, rapid growth, in developing and developed countries alike, has tended to be associated with an increase in unsustainable production and consumption patterns (WWF, 2002).

Potential environmental benefits of FDI may flow from changes in the structure or "composition" of industry. In theory, FDI flows to more efficient companies and industries. Greater economic efficiency translates into greater environmental efficiency via reduction per unit output of productive inputs such as energy, water, and materials, as well as reductions in wastes. The "pollution intensity" of an economy overall, in other words, can be reduced by a change in the relative mix of industries, as well as by changes in the "eco-efficiency" of companies within industries.

While acknowledging that environmental impacts can worsen with an increase in the rate of growth, some economists argue that, over time, economic growth generates environmental improvements. The "Environmental Kuznets Curve" posits that environmental quality first worsens and then improves as per capita income (GDP) rises. [6] Reasons include the substitution of less polluting consumer goods; changes in the structure of industry; and greater political demands for environmental regulation. Early studies put the "turning point" at between US$3000 and US$5000 (Grossman and Kruger,

If true, the EKC suggests that, to a large extent, the pursuit of economic growth is *itself* a sustainable development strategy. One major concern, however, is that the environmental and resource degradation at lower levels of income often results in *irreversible* losses. Examples include loss of biological and genetic diversity and potable water due to degradation or destruction of "old growth" forests; depletion or destruction of fish stocks due to coastal degradation; and human deaths resulting from severe air pollution. Given the number of people on the planet living today at very low levels of per capita income, the potential environmental (and human) losses which must be endured before the global "turnaround" are staggering.

Another concern is that a positive relationship between income and environmental quality in one country or region might mask a relocation of dirty industry to another country or region, resulting in an overall neutral or even negative global environmental impact. Many East Asian studies in the 1980s and 1990s, for example, documented the correlation between improved environmental quality in Japan and the relocation of Japan's pollution-intensive industries to South East Asia (Mani and Wheeler, 1997).

Most important, a number of studies question the validity of the EKC hypothesis itself, especially for developing countries (Stern, 1998). The evidence in support of a "turning point" is limited to a small number of localized pollutants, primarily sulfur and particulate matter. For many other environmental problems such as water pollution, municipal waste, carbon dioxide, and energy use, no consistent evidence has been found that performance increases with higher levels of income.

Environmental Regulation: Stuck in the Mud?

Environmental and resource management is largely the preserve of nation-states. How does FDI affect national (and-sub-national) environmental regulation? There is evidence that TNCs themselves, wielding their substantial bargaining power, can help to drive local standards up—or down. In Chile in the 1970s and 1980s, foreign mining companies pressed for more coherent environmental regulation. In the Russian Far East, on the other hand, oil TNCs involved in obtaining leases for exploration and drilling off of Sakhalin Island in the 1990s flouted and undermined Russia's fledgling environmental laws Rosenthal., 2002).

The asymmetric bargaining power of TNCs vis-à-vis local governments is most troublesome in the context of the intense competition for FDI in both developed and developing countries. Given the absence of global environmental standards, would-be host governments seeking to attract FDI may be reluctant to make higher-than-average environmental demands on

individual TNCs. They may even be tempted to offer lower-than-average environmental demands to enhance the attractiveness of an overall package.

Dubbed by one of the authors of this paper as the "stuck in the mud" problem, the impact of intense global competition for FDI—absent common environmental norms--is thus likely to inhibit the rise of environmental standards (Zarsky, 2002; Oman, 2000). The problem afflicts both developed and developing countries: efforts in the 1990s to put a modest tax on carbon were roundly defeated in both the US and Australia by worries that investment would move offshore.

There is some evidence that, despite regulators' fears, high environmental standards do not, in fact, deter investors and in some cases, are even preferred by investors (Bradford and Gentry, 1998). Moreover, with the rise of the global corporate social responsibility movement, TNC and host-government expectations may be changing.

Overall, an examination of all three of the channels linking FDI and the environment suggests there is no determinate trend: FDI can improve, worsen or have no impact on environmental quality. Other factors—government regulation, the rate of economic growth, company culture, the particular industry in which the FDI takes place, the rules that govern FDI—are key variables.

The paper has also shown that there is a "disconnect" between the globally integrated and the domestic parts of the economy. As a result, hoped for employment and income benefits did not materialize. Most worrisome, FDI has not stimulated backward linkages or increased the endogenous capacity for innovation. In addition to its economic role, the capacity for innovation is crucial to reducing the environmental impacts generated by increasing the scale of economic activity.

V. Conclusions

The overwhelming evidence makes it easy to conclude that FDI is no "miracle drug". What needs fleshing out are the subtle and difficult questions that flow out of this analysis:

- What kinds of local policies would best capture the potential growth benefits and efficiency spillovers of FDI?
- Are the poorest, least developed countries better off without FDI?
- How much "room to move" do developing countries have to shape local policies towards FDI, given the constraints on national policy in regional and bilateral (and potentially global) investment agreements?
- Can voluntary TNC practices raise the probability that FDI will deliver environmental, social and economic benefits to host communities and countries?
- To best garner the potential development benefits of FDI, how should investment be governed at the supranational level?

Fully developed answers to these questions are beyond the scope of this paper. What our analysis does suggest, however, is that a first step is to eschew the search for a "miracle drug" and to embrace the need to develop more fulsome, domestically-oriented development strategies. Moreover, our analysis make clear that structuring investment regimes solely around the interests and concerns of foreign investors will not necessarily deliver economic, environmental or social benefits to host communities.

The centerpiece of a more fulsome, sustainable development strategy should be the nurturance of endogenous capacities for production and innovation. Rather than skew policies towards attracting *foreign* investment, macroeconomic policies should aim to enhance the overall *climate* for investment, both domestic and foreign. Rather than encouraging FDI to flow towards export platforms for the assembly of imported inputs, industry and technology policies should aim to develop local skills, local markets, and solid, world-class domestic firms. With the right set of local—and global—policies, FDI could potentially help in that process.

Endnotes

1. Vietnam, Angola, Lesotho, Ecuador, Turkmenistan, Azerbaijan and China.
2. For developed countries, the drop-off was 49 percent; for developing countries, it was 82 percent.
3. In response, Bechtel sued the city under a bilateral investment agreement between Bolivia and the Netherlands, where the partnership is incorporated.
4. Such operations, however, have "shallow roots" and are vulnerable to being relocated to other locales where labor and other production costs are cheaper.
5. This is precisely what happened in Mexico's manufacturing sector in the 1990s.
6. A landmark article by Simon Kuznets in 1955 posited that inequality first rises, then falls with increases in per capita income. Development policymakers evoked the theory for decades to argue that inequality could be ignored in the short term. More recently, empirical evidence has faded, leading economists to conclude "there is no empirical tendency whatsoever in the inequality-development relationship" (Fields, 1995).
7. This section draws from Kevin P Gallagher and Lyuba Zarsky, *Sustainable Industrial Development? FDI and the Integration Strategy,* paper presented to conference on "New Pathways for Mexico's Sustainable Development," Science, Technology and Development Program (PROCIENTEC), El Colegio de Mexico, Mexico City, October 20-21, 2003.
8. Criteria air pollutants are non-toxic air pollutants such as NOx, SOx, SO2, NO2,
VOC, HC, all particulates, and carbon monoxide.
9. Authors' calculations, based on Corbacho and Schwartz (2002).

References

Aitken, Brian and Ann Harrison (1999). "Do Domestic Firms Benefit from Foreign Direct Investment?" *American Economic Review,* 89 (3), pp. 605-18, June.

Aitken, B., G. H. Hanson, and A. Harrison (1997). "Spillovers, Foreign Investment and Export Behavior." Journal of International Economics, 43: 103-32.

Amsden, Alice and Wan-wen Chu (2003). *Beyond Late Development, Taiwan's Upgrading Policies,* Cambridge, Massachusetts: MIT Press.

Amsden, Alice (2001). The Rise of "The Rest," Challenges to the West From Late-*Industrializing Economies,* New York: Oxford University Press.

Asian Development Bank (2001) *Asian Environment Outlook,* Manila: ADB.

Balasubramanyam, V. N, M. Salisu, and D. Sapsford (1999). "Foreign Direct Investment As An Engine of Growth," Journal of International Trade and Economic *Development,* Vol. 8, No. 1, pp. 27-40.

Balasubramanyam, V. N, M. Salisu, and D. Sapsford (1996). "Foreign Direct Investment and Growth in EP and IS Countries," *Economic Jounral,* Vol. 106, pp. 92-105.

Banco de Mexico (2003). www.bancomex.mx

Barkin, David (2002). "The Greening of Business in Mexico," in Peter Utting, ed., *The* Greening of Business in Developing Countries, London: Zed Books.

Blomstrom, Magnus and Ari Kokko (1996). "The Impact of Foreign Investment on Host Countries: A Review of the Empirical Evidence," World Bank Policy Research Working Paper, 1745.

Blomstrom, Magnus, Ari Kokko and M. Zejan (1994). "Host Country Competition, Labour Skills, and Technology Transfer by Multinationals, " *Weltwirtschaftliches Archiv,* Vol 130, No. 3, pp 521-533.

Blonigen Bruce A. and Van Kolpin (2002). "Technology, Agglomeration, and Regional Competition for Investment," NBER Working Paper 8862, April.

Borensztein, E., J. de Gregorio, J-W. Lee (1998). "How does foreign direct investment affect economic growth? *Journal of International Economics* Vol 45. pp. 115-135.

Bosworth Barry P. and Susan M. Collins (1999). "Capital Flows to Developing Economies: Implications for Savings and Investment," Brookings Papers on Economic Activity: 1, Washington DC: Brookings Institution, pp. 143-69.

Carkovic, Maria and Ross Levine (2002). "Does FDI Accelerate Economic Growth?" University of Minnesota, May.

Chakrabati, Avik (2001). "The Determinants of FDI: Sensitivity Analyses and Cross Country Regressions," *Kyklos,* 54 (1), pp. 89-113.

CSIS (2003). *Economic Competitiveness in Mexico*. Washington, Center for Strategic International Studies.

Dasgupta, S., H. Hettige, and D. Wheeler (1998). "What Improves Environmental Performance? Evidence from Mexican Industry," Policy Research Working Paper 1877, Washington, D.C., Development Research Group, World Bank.

deMello, Luiz R. Jr. (1999). "Foreign direct investment-led growth: evidence from time series and panel data," *Oxford Economic Papers* 51, pp. 133-151

De Velde, Willem (2001). "Policies Towards Foreign Direct Investment in Developing Countries: Emerging Best Practices and Outstanding Issues," Overseas Development Institute, London, March.

Dolinsky, L. (2001). "Cochabamba's Water Rebellion—And Beyond," *San Francisco hronicle,* February 11.

Dussel, Enrique P. (1999). Reflexiones sobre conceptos y experiencias internacionales de industrializacion regional. Dinamica Regional y Competitividad Industrial. C. R. D. a. E. D. Peters. Mexico City, Editorial JUS.

Dussel, E. D. (2000). Polarizing Mexico: The Impact of Liberalization Strategy. Boulder, Lynne Rienner.

Dussel, Enrique P. (2003). "Industrial Policy, Regional Trends, and Structural Change," in Kevin J. Middlebrook and Eduardo Zepeda, eds., *Confronting Development: Assessing Mexico's Economic and Social Policy Changes*. Palo Alto, Stanford University Press, pp. 241-76.

Gallagher, Kevin (2003). "NACEC and Environmental Quality: Assessing the Mexican Experience," in John Knox and David Markel (eds.) *Greening NAFTA*, Stanford: Stanford University Press.

Gallagher, Kevin (2002). "Industrial Pollution in Mexico: Did NAFTA Matter?" in Carolyn L. Deere and Daniel C. Esty, eds., *Greening the Americas: NAFTA's Lessons for Hemispheric Trade,* Cambridge: MIT Press.

Gallagher, Kevin, and Lyuba Zarsky (2003*). Sustainable Industrial Development? The Performance of Mexico's FDI-led Strategy,* Global Development and Environment Institute, Tufts University.

Gentry, Brad ed. (1998). *Private Capital Flows and the Environment, Lessons from Latin America,* Cheltenham: Edward Elgar.

Gentry, B. and L. Fernandez (1998). "Mexican Steel," in Brad Gentry, ed. *Private Capital Flows and the Environment: Lessons form Latin America.* Cheltenham: Eward Elgar.

Graham, Edward H. (1995). "FDI in the World Economy," IMF Working Paper, wp/95/59, Washington D.C.

Graham, Edward M. and Wada, Erika (2001). "Foreign Direct Investment in China: Effects on Growth and Economic Performance," in Peter Drysdale, ed., *Achieving High Growth: Experience of Transitional Economies in East Asia,* New York: Oxford University Press.

Hanson, Gordon H. (2001). "Should Countries Promote Foreign Direct Investment?" G-24 Discussion Paper 9, UNCTAD, February.

Hettige, H, M. Huq, S. Pargal and D. Wheeler (1996). "Determinants of Pollution Abatement in Developing Countries: Evidence from South and Southeast Ais," in *World Development, Vol*. 24, No. 12, pp. 1891-1904.

INEGI (2003). Bureau of Economic Informations. Mexico City, nstituto Nacional de Estadística Geografía e Informática (INEGI): www.inegi.gob.mx.

INEGI (2000). Instituto Nacional de Estadistica Geografia e Informatica, "Sistema de Cuentas Economicas y Ecologicas de Mexico, 1993-1999" (Determining the Economic Costs of Ecological Degradation in Mexico), Mexico City: INEGI, www.inegi.gob.mx

Kokko, Ari (1994). "Technology, Market Characteristics, and Spillovers," *Journal of Development Economics,* Vol. 43, pp. 279-293.

Kokko, A. and M. Blomstrom (1995). "Policies to Encourage Inflows of Technology Through Foreign Multinationals," *World Development,* Vol. 23, No. 3, pp. 459-68.

Kokko, Ari, R. Tansini, and M. Zejan (1996). "Local Technological Capability and Spillovers from FDI in the Uruguayan Manufacturing Sector," *Journal of Development Studies,* Vol. 34. pp. 602-611.

Krugman, Paul (1998). "Firesale FDI," Working Paper, Massachusetts Institute of Technology.

Lagos, Gustavo and Patricio Velasco (1999). "Environmental Policies and Practices in Chilean Mining," in Alyson Warhurst (ed.), *Mining and the Environment, Case Studies from the Americas,* Ottawa: International Development Research Centre, Chapter 3. Available on-line at http://www.idrc.ca/books/focus/828/chapter3.html

Lall, S. (1980). "Vertical Interfirm Linkages in LDCs: An Empirical Study," *Oxford Bulletin of Economics and Statistics,* Vol. 42, pp. 203-226.

Leighton, Michelle, Naomi Roht-Arriza, and Lyuba Zarsky (2002). *Beyond Good Deeds: Case Studies and a New Policy Agenda for Corporate Accountability,* Berkeley, California: Natural Heritage Institute, available on http://www.nautilus.org/enviro/

Lensink, Robert and Oliver Morissey (2001). "Foreign Direct Investment: Flows, Volatility and Growth," paper presented at the Development Economic Study Group Conference, University of Nottingham, April 5-7.

Lim, Ewe-Ghee (2001). "Determinants of and the Relation Between Foreign Direct Investment and Growth: A Summary of the Recent Literature," Working Paper 01/75, IMF, November.

Lougani, Prakash and Assaf Razin (2001). "How Beneficial is Foreign Direct Investment for Developing Countries?" *Finance and Development,* 38 (2), International Monetary Fund, June.

Mabey, Nick and Richard McNally (1999). "Foreign Direct Investment and the Environment, From Pollution Havens to Sustainable Development, WWF-UK Report, August.

Malampally, Padma and Karl P. Sauvant (1999). "Foreign Direct Investment in Developing Countries," *Finance and Development* 36 (1), IMF, March.

Mani, Mathukumara and David Wheeler (1997). "In Search of Pollution Havens? Dirty Industry in the World Economy," PRDEI, Poverty, Environment and Growth, World Bank, April.

Marino, Andrea (2000). "The Impact on FDI on Developing Countries Growth: Trade Policy Matters," ISTAT-CEMAFI, France.

Markusen, James R. and Anthony J. Venables (1999). "Foreign Direct Investment as a Catalyst for Industrial Development," *European Economic Review,* 43, pp. 335-356.

Mattar, Jorse, Juan Carlos Moreno-Brid, and Wilson Peres (2003). "Foreign Investment in Mexico after Economic Reform," in Kevin J. Middlebrook and Eduardo Zepeda, *Confronting Development, Assessing Mexico's Economic and Social Policy Challenges,* Stanford: Stanford University Press, pp. 123-60.

Mercardo, Alfonso (2000). "Environmental Assessment of the Mexican Steel Industry," in Rhys Jenkins, ed., *Industry and the Environment in Latin America*, London: Routledge.

Moran, Theodore H. (1998). *Foreign Direct Investment and Development, The New Policy Agenda for Developing Countries and Economies in Transition,* Washington D.C.: Institute for International Economics.

Nadal, Alejandro (2003). "Macroeconomic Challenges for Mexico's Development Strategy," in Kevin Middlebrook and Eduardo Zepeda, eds., *Confronting Development, Assessing Mexico's Economic and Social Policy Challenges,* Palo Alto: Stanford University Press, pp. 55-88.

OECD (2002a). "The Environmental Benefits of FDI." Paris, OECD.

OECD (2002b). Foreign Direct Investment for Development, Maximising Benefits, Minimising Costs, Pais: OECD.

OECD (2000). *The OECD Guidelines for Multinational Enterprises,* Paris: OECD.

OECD (1998). Environmental Performance Review for Mexico, Paris: .OECD.

Oman, Charles (2000). *Policy Competition for Foreign Direct Investment,* Paris: OECD Development Centre.

PROFEPA (2000). www.profepa.gob.mx.

Overseas Development Institute (2002). *Optimising the Development Performance of Corporate Investment, The ODPCI Program,* Discussion Paper, August, available on-line at http://www.odi.org.uk/pppg/activities/country_level/odpci/index.html.

Razin, Assag, Efraim Sadka and Chi-wa Yuen (1999). "Excessive FDI Under Asymmetric Information," NBER Working Paper No. 7400.

RNIE (2003). Sistema del Registro Nacional de Inversiones Extranjeras. Mexico City, Secretary of Economy: www.economia.gob.mx.

Rodrik, Dani (1999). " The New Global Economy and Developing Countries: Making Openness Work," Policy Essay No. 24, Overseas Development Council, Washington D.C.

Romo Murillo, David (2002). "Derramas Tecnologicas de la Inversion Extranhera en la Industria Mexicana," *Comercio Exterior,* 53 n 3 (March): pp. 230-43.

Rosenthal, Erika (2002). "Conflicts Over Transnational Oil and Gas Developoment Off Sakhalin Island in the Russian Far East: A David and Goliath Tale," in Lyuba Zarsky, ed., *Human Rights and Environment: Conflicts and Norms in a Globalizing World,* London: Earthscan, pp. 96-122.

Ruud, Audun (2002). "Environmental Management of Transnational Corporations in India – Are TNCs Creating Islands of Environmental Excellence in a Sea of Dirt?" *Business Strategy and the Environment,* 11, pp. 103-118.

Sandbrooke, Richard and Pradeep Mehta (2002). "Rapporteurs' Report," in OECD, Foreign Direct Investment and the Environment, Lessons from the Mining Sector, pp. 9-16.

Schatan, C. (2002). "Mexico's Manufacturing Exports and the Environment Under NAFTA," in *The Environmental Effects of Free Trade*. Montreal, North American Commission for Environmental Cooperation.

Stern, David, (1998). "Progress on the Environmental Kuznets Curve?" *Journal of* Economic Development and the Environment, 3, 173-196.

Stromberg, P. (2002). The Mexican Maquila Industry and the Environment; An Overview *of the Issues*, Naciones Unidas, CEPAL/ECLAC**:** 1-56.

UNCTAD (2002a). *World Investment Report, Transnational Corporations and Export Competitiveness,* United Nations.

UNCTAD (2002b). Least Developed Countries Report 2002, Escaping the Poverty Trap.

UNCTAD (2000). World Investment Report, Crossborder Mergers and Acquisitions and Development, United Nations.

Urquidi, Victor L. (2003). "Mexico's Development Challenges," in Kevin J. Middlebrook and Eduardo Zepeda, Confronting Development, Assessing Mexico's Economic and Social Policy Challenges, Stanford: Stanford University Press, pp. 561-76.

Wackernagel, Mathis (2001). "Shortcomings of the Environmental Sustainability Index," Refining Progress, www.anti-lomborg.com/ESI%20critique.rtf.

Winn, P. (1992). *The Americas: The Changing Face of Latin America and the Caribbean*. Los Angeles, University of California Press.

Wisner, P. S. and M. J. Epstein (2003). "The NAFTA Impact on Environmental Responsiveness and Performance in Mexican Industry", pp 1-31.

World Bank (2003a). *Global Economic Prospects and the Developing Countries 2003,* Washington: World Bank, available on http://www.worldbank.org/prospects/gep2003/.

World Bank (2003b). *World Development Indicators,* Washington D.C.: World Bank.

WWF (2002). *Living Planet Report 2002,*
http://www.panda.org/news_facts/publications/general/livingplanet/index.cfm

Zarsky, Lyuba (2002). "Stuck in the Mud? Nation-States, Globalization, and the Environment," in Kevin P. Gallagher and Jacob Werksman (eds), *International Trade and Sustainable Development,* London: Earthscan, pp. 19-44.

Zarsky, Lyuba (1999). "Havens, Halos and Spaghetti: Untangling the Relationship Between FDI and the Environment," in *Foreign Direct Investment and the Environment,* Paris: OECD, pp. 47-73.

Author contact: kpg@bu.edu

SUGGESTED CITATION:
Kevin P.Gallagher and Lyuba Zarsky, "Rethinking Foreign Investment for Development" *post-autistic economics review*, issue no. 37, 28 April 2006, article 2, pp.10-32, http://www.paecon.net/PAEReview/issue37GallagherZarsky.htm

The Political Economy of Peer Production*

Michel Bauwens (Payap University and Chiang Mai University, Thailand)

Not since Marx identified the manufacturing plants of Manchester as the blueprint for the new capitalist society has there been a deeper transformation of the fundamentals of our social life. As political, economic, and social systems transform themselves into distributed networks, a new human dynamic is emerging: peer to peer (P2P). As P2P gives rise to the emergence of a third mode of production, a third mode of governance, and a third mode of property, it is poised to overhaul our political economy in unprecedented ways. This essay aims to develop a conceptual framework ('P2P theory') capable of explaining these new social processes.

Peer to Peer

P2P does not refer to all behavior or processes that takes place in distributed networks: P2P specifically designates those processes that aim to increase the most widespread participation by equipotential participants. We will define these terms when we examine the characteristics of P2P processes, but here are the most general and important characteristics.

P2P processes:

- produce use-value through the free cooperation of producers who have access to distributed capital: this is the P2P production mode, a 'third mode of production' different from for-profit or public production by state-owned enterprises. Its product is not exchange value for a market, but use-value for a community of users.
- are governed by the community of producers themselves, and not by market allocation or corporate hierarchy: this is the P2P governance mode, or 'third mode of governance.'
- make use-value freely accessible on a universal basis, through new common property regimes. This is its distribution or 'peer property mode': a 'third mode of ownership,' different from private property or public (state) property.

The Infrastructure of P2P

What has been needed to facilitate the emergence of peer to peer processes? The first requirement is the existence of a technological infrastructure that operates on peer to peer processes and enables distributed access to 'fixed' capital. Individual computers that enable a universal machine capable of executing any logical task are a form of distributed 'fixed capital,' available at low cost to many producers. The internet, as a point to point network, was specifically designed for participation by the edges (computer users) without the use of obligatory hubs. Although it is not fully in the hands of its participants, the internet is controlled through distributed governance, and outside the complete hegemony of particular private or state actors. The internet's hierarchical elements (such as the stacked IP protocols, the decentralized Domain Name System, etc...) do not deter participation. Viral communicators, or meshworks, are a logical extension of the internet. With this methodology, devices create their own networks through the use of excess capacity, bypassing the need for a pre-existing infrastructure. The 'Community Wi-

Fi' movement, Open Spectrum advocacy, file-serving television, and alternative meshwork-based telecommunication infrastructures are exemplary of this trend.

The second requirement is alternative information and communication systems which allow for autonomous communication between cooperating agents. The web (in particular the Writeable Web and the Web 2.0 that is in the process of being established) allows for the universal autonomous production, dissemination, and 'consumption' of written material while the associated podcasting and webcasting developments create an 'alternative information and communication infrastructure' for audio and audiovisual creation. The existence of such an infrastructure enables autonomous content production that may be distributed without the intermediary of the classic publishing and broadcasting media (though new forms of mediation may arise).

The third requirement is the existence of a 'software' infrastructure for autonomous global cooperation. A growing number of collaborative tools, such as blogs and wiki's, embedded in social networking software facilitate the creation of trust and social capital, making it possible to create global groups that can create use-value without the intermediary of manufacturing or distribution by for-profit enterprises.

The fourth requirement is a legal infrastructure that enables the creation of use-value and protects it from private appropriation. The General Public License (which prohibits the appropriation of software code), the related Open Source Initiative, and certain versions of the Creative Commons license fulfill this role. They enable the protection of common use-value and use viral characteristics to spread. GPL and related material can only be used in projects that in turn put their adapted source code in the public domain.

The fifth requirement is cultural. The diffusion of mass intellectuality, (i.e. the distribution of human intelligence) and associated changes in ways of feeling and being (ontology), ways of knowing (epistemology) and value constellations (axiology) have been instrumental in creating the type of cooperative individualism needed to sustain an ethos which can enable P2P projects.

The Characteristics of P2P

P2P processes occur in distributed networks. Distributed networks are networks in which autonomous agents can freely determine their behavior and linkages without the intermediary of obligatory hubs. As Alexander Galloway insists in his book on protocollary power, distributed networks are not the same as decentralized networks, for which hubs are obligatory. P2P is based on distributed power and distributed access to resources. In a decentralized network such as the U.S.-based airport system, planes have to go through determined hubs; however, in distributed systems such as the internet or highway systems, hubs may exist, but are not obligatory and agents may always route around them.

P2P projects are characterized by equipotentiality or 'anti-credentialism.' This means that there is no *a priori* selection to participation. The capacity to cooperate is verified in the process of cooperation itself. Thus, projects are open to all comers provided they have the necessary skills to contribute to a project. These skills are verified, and communally validated, in the process of

production itself. This is apparent in open publishing projects such as citizen journalism: anyone can post and anyone can verify the veracity of the articles. Reputation systems are used for communal validation. The filtering is *a posteriori*, not *a priori*. Anti-credentialism is therefore to be contrasted to traditional peer review, where credentials are an essential prerequisite to participate.

P2P projects are characterized by holoptism. Holoptism is the implied capacity and design of peer to peer processes that allows participants free access to all the information about the other participants; not in terms of privacy, but in terms of their existence and contributions (i.e. horizontal information) and access to the aims, metrics and documentation of the project as a whole (i.e. the vertical dimension). This can be contrasted to the panoptism which is characteristic of hierarchical projects: processes are designed to reserve 'total' knowledge for an elite, while participants only have access on a 'need to know' basis. However, with P2P projects, communication is not top-down and based on strictly defined reporting rules, but feedback is systemic, integrated in the protocol of the cooperative system.

The above does not exhaust the characteristics of peer production. Below, we will continue our investigation of these characteristics in the context of a comparison with other existing modes of production.

P2P and the Other Modes of Production

The framework of our comparison is the Relational Models theory of anthropologist Alan Page Fiske, discussed in his major work *The Structure of Social Life*. The fact that modes of production are embedded in inter-subjective relations -- that is, characterized by particular relational combinations -- provides the necessary framework to distinguish P2P. According to Fiske, there are four basic types of inter-subjective dynamics, valid across time and space, in his own words:

People use just four fundamental models for organizing most aspects of sociality most of the time in all cultures. These models are Communal Sharing, Authority Ranking, Equality Matching, and Market Pricing. Communal Sharing (CS) is a relationship in which people treat some dyad or group as equivalent and undifferentiated with respect to the social domain in question. Examples are people using a commons (CS with respect to utilization of the particular resource), people intensely in love (CS with respect to their social selves), people who "ask not for whom the bell tolls, for it tolls for thee" (CS with respect to shared suffering and common well-being), or people who kill any member of an enemy group indiscriminately in retaliation for an attack (CS with respect to collective responsibility). In Authority Ranking (AR) people have asymmetric positions in a linear hierarchy in which subordinates defer, respect, and (perhaps) obey, while superiors take precedence and take pastoral responsibility for subordinates. Examples are military hierarchies (AR in decisions, control, and many other matters), ancestor worship (AR in offerings of filial piety and expectations of protection and enforcement of norms), monotheistic religious moralities (AR for the definition of right and wrong by commandments or will of God), social status systems such as class or ethnic rankings (AR with respect to social value of identities), and rankings such as sports team standings (AR with respect to prestige). AR

relationships are based on perceptions of legitimate asymmetries, not coercive power; they are not inherently exploitative (although they may involve power or cause harm).

In Equality Matching relationships people keep track of the balance or difference among participants and know what would be required to restore balance. Common manifestations are turn-taking, one-person one-vote elections, equal share distributions, and vengeance based on an-eye-for-an-eye, a-tooth-for-a-tooth. Examples include sports and games (EM with respect to the rules, procedures, equipment and terrain), baby-sitting co-ops (EM with respect to the exchange of child care), and restitution in-kind (EM with respect to righting a wrong). Market Pricing relationships are oriented to socially meaningful ratios or rates such as prices, wages, interest, rents, tithes, or cost-benefit analyses. Money need not be the medium, and MP relationships need not be selfish, competitive, maximizing, or materialistic -- any of the four models may exhibit any of these features. MP relationships are not necessarily individualistic; a family may be the CS or AR unit running a business that operates in an MP mode with respect to other enterprises. Examples are property that can be bought, sold, or treated as investment capital (land or objects as MP), marriages organized contractually or implicitly in terms of costs and benefits to the partners, prostitution (sex as MP), bureaucratic cost-effectiveness standards (resource allocation as MP), utilitarian judgments about the greatest good for the greatest number, or standards of equity in judging entitlements in proportion to contributions (two forms of morality as MP), considerations of "spending time" efficiently, and estimates of expected kill ratios (aggression as MP).[1]

Every type of society or civilization is a mixture of these four modes, but it can plausibly be argued that one mode is always dominant and imprints the other subservient modes. Historically, the first dominant mode was kinship or lineage based reciprocity, the so-called tribal gift economies. The key relational aspect was 'belonging'. Gifts created obligations and relations beyond the next of kin, creating a wider field of exchange. Agricultural or feudal-type societies were dominated by authority ranking, that is, they were based on allegiance. Finally, it is clear that the capitalist economy is dominated by market pricing.

P2P and the Gift Economy

P2P is often described as a 'gift economy' (see Richard Barbrook for an example). However, it is our contention that this is somewhat misleading. The key reason is that peer to peer is not a form of equality matching; it is not based on reciprocity. P2P follows the adage: each contributes according to his capacities and willingness, and each takes according to his needs. There is no obligatory reciprocity involved. In the pure forms of peer production, producers are not paid. Thus, if there is 'gifting' it is entirely non-reciprocal gifting, the use of peer-produced use-value does not create a contrary obligation. The emergence of peer to peer is contemporaneous with new forms of the gift economy, such as the Local Exchange Trading Systems and the use of reciprocity-based complementary currencies; however, these do not qualify as peer production.

That is not to say that these forms are not complementary, since both equality matching and communal shareholding derive from the same spirit of gifting. Peer production can most easily operate in the sphere of immaterial goods, where the input is free time and the available surplus of computing resources. Equality matching, reciprocity-based schemes and cooperative

production are necessary in the material sphere where the cost of capital intervenes. At present, peer production offers no solution to the material survival of its participants. Therefore, many people inspired by the egalitarian ethos will resort to cooperative production, the social economy, and other schemes from which they can derive an income, while at the same time honoring their values. In this sense, these schemes are complementary.

P2P and Hierarchy

P2P is not hierarchy-less, not structure-less, but usually characterized by flexible hierarchies and structures based on merit that are used to enable participation. Leadership is also 'distributed.' Most often, P2P projects are led by a core of founders, who embody the original aims of the project, and who coordinate the vast number of individuals and microteams working on specific patches. Their authority and leadership derives from their input into the constitution of the project, and on their continued engagement. It is true that peer projects are sometimes said to be 'benevolent dictatorships'; however, one must not forget that since the cooperation is entirely voluntary, the continued existence of such projects is based on the consent of the community of producers, and on 'forking' (that is, the creation of a new independent project, is always possible).

The relation between authority and participation, and its historical evolution, has been most usefully outlined by John Heron:

There seem to be at least four degrees of cultural development, rooted in degrees of moral insight:

1. autocratic cultures which define rights in a limited and oppressive way and there are no rights of political participation;
2. narrow democratic cultures which practice political participation through representation, but have no or very limited participation of people in decision-making in all other realms, such as research, religion, education, industry etc.;
3. wider democratic cultures which practice both political participation and varying degree of wider kinds of participation;
4. commons p2p cultures in a libertarian and abundance-oriented global network with equipotential rights of participation of everyone in every field of human endeavor.

These four degrees could be stated in terms of the relations between hierarchy, co-operation and autonomy.

1. Hierarchy defines, controls and constrains co-operation and autonomy;
2. Hierarchy empowers a measure of co-operation and autonomy in the political sphere only;
3. Hierarchy empowers a measure of co-operation and autonomy in the political sphere and in varying degrees in other spheres;
4. The sole role of hierarchy is in its spontaneous emergence in the initiation and continuous flowering of autonomy-in-co-operation in all spheres of human endeavor.[2]

P2P and Communal Shareholding

With P2P, people voluntarily and cooperatively construct a commons according to the communist principle: "from each according to his abilities, to each according to his needs." The use-value created by P2P projects is generated through free cooperation, without coercion toward the producers, and users have free access to the resulting use value. The legal infrastructure that we have described above creates an 'Information Commons.' The new Commons is related to the older form of the commons (most notably the communal lands of the peasantry in the Middle Ages and of the original mutualities of the workers in the industrial age), but it also differs mostly through its largely immaterial characteristics. The older Commons were localized, used, and sometimes regulated by specific communities; the new Commons are universally available and regulated by global cyber-collectives, usually affinity groups. While the new Commons is centered around non-rival goods (that is, in a context of abundance) the older forms of physical Commons (air, water, etc.) increasingly function in the context of scarcity, thus becoming more regulated.

P2P and the Market: The Immanence vs. Transcendence of P2P

P2P and the Market

P2P exchange can be considered in market terms only in the sense that individuals are free to contribute, or take what they need, following their individual inclinations, with a invisible hand bringing it all together, but without any monetary mechanism. They are not true markets in any real sense: neither market pricing nor managerial command are required to make decisions regarding the allocation of resources. There are further differences:

- Markets do not function according to the criteria of collective intelligence and holoptism, but rather, in the form of insect-like swarming intelligence. Yes, there are autonomous agents in a distributed environment, but each individual only sees his own immediate benefit.
- Markets are based on 'neutral' cooperation, and not on synergistic cooperation: no reciprocity is created.
- Markets operate for the exchange value and profit, not directly for use value.
- Whereas P2P aims at full participation, markets only fulfill the needs of those with purchasing power.

The disadvantages of markets include:

- They do not function well for common needs that do not involve direct payment (national defense, general policing, education and public health). In addition, they fail to take into account negative externalities (the environment, social costs, future generations).
- Since open markets tend to lower profit and wages, they always give rise to anti-markets, where oligopolies and monopolies use their privileged position to have the state 'rig' the market to their benefit.

P2P and Capitalism

Despite significant differences, P2P and the capitalist market are highly interconnected. P2P is dependent on the market and the market is dependent on P2P.

Peer production is highly dependent on the market because peer production produces use-value through mostly immaterial production, without directly providing an income for its producers. Participants cannot live from peer production, though they derive meaning and value from it, and though it may out-compete, in efficiency and productivity terms, the market-based for-profit alternatives. Thus peer production covers only a section of production, while the market provides for nearly all sections; peer producers are dependent on the income provided by the market. So far, peer production has been created through the interstices of the market.

But the market and capitalism are also dependent on P2P. Capitalism has become a system relying on distributed networks, in particular on the P2P infrastructure in computing and communication. Productivity is highly reliant on cooperative teamwork, most often organized in ways that are derivative of peer production's governance. The support given by major IT companies to open-source development is a testimony to the use derived from even the new common property regimes. The general business model seems to be that business 'surfs' on the P2P infrastructure, and creates a surplus value through services, which can be packaged for exchange value. However, the support of free software and open sources by business poses an interesting problem. Is corporate-sponsored, and eventually corporate managed, FS/OS software still 'P2P': only partially. If it uses the GPL/OSI legal structures, it does result in common property regimes. If peer producers are made dependent on the income, and even more so, if the production becomes beholden to the corporate hierarchy, then it would no longer qualify as peer production. Thus, capitalist forces mostly use partial implementations of P2P. The tactical and instrumental use of P2P infrastructure, (collaborative practices) is only part of the story. In fact, contemporary capitalism's dependence on P2P is systemic. As the whole underlying infrastructure of capitalism becomes distributed, it generates P2P practices and becomes dependent on them. The French-Italian school of 'cognitive capitalism' stresses that value creation today is no longer confined to the enterprise, but beholden to the mass intellectuality of knowledge workers, who through their lifelong learning/experiencing and systemic connectivity, constantly innovate within and without the enterprise. This is an important argument, since it would justify what we see as the only solution for the expansion of the P2P sphere into society at large: the universal basic income. Only the independence of work and the salary structure can guarantee that peer producers can continue to create this sphere of highly productive use value.

Does all this mean that peer production is only immanent to the system, productive of capitalism, and not in any way transcendent to capitalism?

P2P and the Netarchists

More important than the generic relationship that we just described, is the fact that peer to peer processes also contribute to more specific forms of distributed capitalism. The massive use of open source software in business, enthusiastically supported by venture capital and large

IT companies such as IBM, is creating a distributed software platform that will drastically undercut the monopolistic rents enjoyed by companies such as Microsoft and Oracle, while Skype and VoIP will drastically redistribute the telecom infrastructure. In addition, it also points to a new business model that is 'beyond' products, focusing instead on services associated with the nominally free FS/OS software model. Industries are gradually transforming themselves to incorporate user-generated innovation, and a new intermediation may occur around user-generated media. Many knowledge workers are choosing non-corporate paths and becoming mini-entrepreneurs, relying on an increasingly sophisticated participatory infrastructure, a kind of digital corporate commons.

The for-profit forces that are building and enabling these new platforms of participation represent a new subclass, which I call the netarchical class. If cognitive capitalism is to be defined by the primacy of intellectual assets over fixed capital industrial assets, and thus on the reliance of an extension of IP rights to establish monopolistic rents, (as the vectoral capitalists described by Mackenzie Wark derive their power from the control of the media vectors) then these new netarchical capitalists prosper from the enablement and exploitation of the participatory networks. It is significant that Amazon built itself around user reviews, eBay lives on a platform of worldwide distributed auctions, and Google is constituted by user-generated content. However, although these companies may rely on IP rights for the occasional extra buck, it is not in any sense the core of their power. Their power relies on their ownership of the platform.

More broadly, netarchical capitalism is a brand of capital that embraces the peer to peer revolution, all those ideological forces for whom capitalism is the ultimate horizon of human possibility. It is the force behind the immanence of peer to peer. Opposed to it, though linked to it in a temporary alliance, are the forces of Common-ism, those that put their faith in the transcendence of peer to peer, in a reform of the political economy beyond the domination of the market.

Transcendent Aspects of P2P

Indeed, our review of the immanent aspects of peer to peer, on how it is both dependent on and productive of capitalism, does not exhaust the subject. P2P has important transcendent aspects which go beyond the limitations set by the for-profit economy:

- peer production effectively enables the free cooperation of producers, who have access to their own means of production, and the resulting use-value of the projects supercedes for-profit alternatives.

Historically, though forces of higher productivity may be temporarily embedded in the old productive system, they ultimately lead to deep upheavals and reconstitutions of the political economy. The emergence of capitalist modes within the feudal system is a case in point. This is particularly significant because leading sectors of the for-profit economy are deliberately slowing down productive growth (in music; through patents) and trying to outlaw P2P production and sharing practices:

- peer governance transcends both the authority of the market and the state

- the new forms of universal common property, transcend the limitations of both private and public property models and are reconstituting a dynamic field of the Commons.

At a time when the very success of the capitalist mode of production endangers the biosphere and causes increasing psychic (and physical) damage to the population, the emergence of such an alternative is particularly appealing, and corresponds to the new cultural needs of large numbers of the population. The emergence and growth of P2P is therefore accompanied by a new work ethic (Pekka Himanen's *Hacker Ethic*), by new cultural practices such as peer circles in spiritual research (John Heron's cooperative inquiry), but most of all, by a new political and social movement which is intent on promoting its expansion. This still nascent P2P movement, (which includes the Free Software and Open Source movement, the open access movement, the free culture movement and others) which echoes the means of organization and aims of the alter-globalization movement, is fast becoming the equivalent of the socialist movement in the industrial age. It stands as a permanent alternative to the status quo, and the expression of the growth of a new social force: the knowledge workers.

In fact, the aim of peer to peer theory is to give a theoretical underpinning to the transformative practices of these movements. It is an attempt to create a radical understanding that a new kind of society, based on the centrality of the Commons, and within a reformed market and state, is in the realm of human possibility. Such a theory would have to explain not only the dynamic of peer to peer processes proper, but also their fit with other inter-subjective dynamics. For example, how P2P molds reciprocity modes, market modes and hierarchy modes; on what ontological, epistemological and axiological transformations this evolution is resting; and what a possible positive P2P ethos can be. A crucial element of such a peer to peer theory would be the development of tactics and strategy for such transformative practice. The key question is: can peer to peer be expanded beyond the immaterial sphere in which it was born?

The Expansion of the P2P mode of production

Given the dependence of P2P on the existing market mode, what are its chances to expand beyond the existing sphere of non-rival immaterial goods?

Here are a number of theses about this potential:

- P2P can arise not only in the immaterial sphere of intellectual and software production, but wherever there is access to distributed technology: spare computing cycles, distributed telecommunications and any kind of viral communicator meshwork.
- P2P can arise wherever other forms of distributed fixed capital are available: such is the case for carpooling, which is the second most used mode of transportation in the U.S.
- P2P can arise wherever the process of design may be separated from the process of physical production. Huge capital outlines for production can co-exist with a reliance on P2P processes for design and conception.
- P2P can arise wherever financial capital can be distributed. Initiatives such as the ZOPA bank point in that direction. Cooperative purchase and use of large capital goods are a possibility. State support and funding of open source development is another example.
- P2P could be expanded and sustained through the introduction of universal basic income.

The latter, which creates an income independent of salaried work, has the potential to sustain a further development of P2P-generated use-value. Through the 'full activity' ethos (rather than full employment) of P2P, the basic income receives a powerful new argument: not only as efficacious in terms of poverty and unemployment, but as creating important new use-value for the human community.

However, as it is difficult to see how use-value production and exchange could be the only form of production, it is more realistic to see peer to peer as part of a process of change. In such a scenario, peer to peer would both co-exist with and profoundly transform other intersubjective modes.

A Commons-based political economy would be centered around peer to peer, but it would co-exist with:

- A powerful and re-invigorated sphere of reciprocity (gift-economy) centered around the introduction of time-based complementary currencies.
- A reformed sphere for market exchange, the kind of 'natural capitalism' described by Paul Hawken, David Korten and Hazel Henderson, where the costs for natural and social reproduction are no longer externalized, and which abandons the growth imperative for a throughput economy as described by Herman Daly.
- A reformed state that operates within a context of multistakeholdership and which is no longer subsumed to corporate interests, but act as a fair arbiter between the Commons, the market and the gift economy.

Such a goal could be the inspiration for a powerful alternative to neoliberal dominance, and create a kaleidoscope of 'Common-ist' movements broadly inspired by such goals.

Resources

Pluralities/Integration monitors P2P developments and is archived at:
http://integralvisioning.org/index.php?topic=p2p

A longer manuscript and book-in-progress on the subject is available at:
http://integralvisioning.org/article.php?story=p2ptheory1

The Foundation for P2P Alternatives has a website under construction at:
http://p2pfoundation.net/index.php/Manifesto

Notes

* This paper appeared on www.ctheory.com in December 2005.

1. Fiske website. http://www.sscnet.ucla.edu/anthro/faculty/fiske/relmodov.htm

2. Personal communication with the author

Bibliography

Barbrook, Richard. *Media Freedom*. London: Pluto, 1995.

Ferrer, Jorge N. *Revisioning Transpersonal Theory: A Participatory Vision of Human Spirituality*. Albany: SUNY, 2001.

Fiske, Alan Page. *Structures of Social Life*. New York: Free Press, 1993.

Gunderson, Lance H. and C.S. Holling. *Panarchy: Understanding Transformations in Systems of Humans and Nature*. Washington, D.C.: Island Press, 2001.

Heron, John. *Sacred Science*. Llangarron, Ross-on-Wye, UK: PCCS Books, 1998.

Galloway, Alexander. *Protocol: How Control Exists After Decentralization* Cambridge, Mass.: MIT Press, 2004.

Himanen, Pekka. *The Hacker Ethic and the Spirit of the Information Age*. New York: Random House, 2002.

Inglehart, Ronald. *Culture Shift in Advanced Industrial Society*. Princeton, N.J.: Princeton University Press, 1989.

Kane, Pat*. The Play Ethic: A Manifesto for a Different Way of Living*. London: Macmillan, 2003.

Lazzarato, Maurizio. *Les révolutions du capitalisme*. Paris: Les Empêcheurs de penser en rond, 2004.

Lessig, Lawrence. *The Future of Ideas*. New York: Vintage, 2002.

Lessig, Lawrence. *Free Culture*. New York: Penguin U.S.A., 2004.

Raymond, Eric. *The Cathedral and the Bazaar*. Sebastopol, CA: O'Reilly, 2001.

Sagot-Duvauroux, Jean-Louis. *Pour la gratuité*. Paris: Desclée-De Brouwer, 1995.

Sahlins, Marshall D. *Stone Age Economics*. Chicago: Aldine, 1972.

Skolimowski, Henryk. *The Participatory Mind*. New York: Penguin USA, 1995.

Skrbina, David. *Panpsychism in the West*. Cambridge, MA: MIT Press, 2005.

Stallman, Richard. *Free Software, Free Society*. Boston, MA: GNU Press, 2002.

Stewart, John. *Evolution's Arrow*. Canberra: Chapman Press, 2000.

Surowiecki, James. *The Wisdom of Crowds*. New York: Anchor, 2005.

Tuomi, Ilkka. *Networks of Innovation*. Oxford: Oxford University Press, 2003.

Vercelonne, Carlo, dir. *Sommes-nous sorti du capitalisme industriel*? Paris: La Dispute, 2003.

von Hippel, Eric. *The Democratization of Innovation*. Cambridge, MA: MIT Press, 2004.

Wark, McKenzie. *A Hacker Manifesto*. Cambridge, MA: Harvard University Press, 2004.

Weber, Steve. *The Success of Open Source*. Cambridge, MA: Harvard University Press, 2004.

Author contact: michelsub2004@gmail.com

SUGGESTED CITATION:
Michel Bauwens, "The Political Economy of Peer Production" *post-autistic economics review*, issue no. 37, 28 April 2006, article 3, pp. 33-44, http://www.paecon.net/PAEReview/issue37Bauwens37.htm

Can a Heterodox Economist Use Cross-country Growth Regressions?

Matthew McCartney (School of Oriental and African Studies, UK)

1. Introduction

This paper is concerned with how economic growth is analysed by economists. Over the last fifteen years an extremely common method has been through cross-country growth regressions. Section two shows how economic theory demonstrates that there should be a strong link between changes in economic policy and economic growth. Despite the implications of such theory *empirical* results using cross-country growth regressions remain disappointing. Section two demonstrates this using relevant empirical results in both a general manner and specifically those between economic growth and fiscal policy, investment, education and R+D.

The third section shows that the long-term averages typically used in cross-country growth regressions hide an important empirical reality of growth in developing countries. The medium-term growth averages used by cross-country regressions conceal the periods of stagnation, growth spurts, structural breaks, volatility and instability that actually characterise growth in developing countries. When confronted with these 'empirical' problems researchers typically stop and try to confront what they perceive as an empirical challenge. Researchers seek for better proxies for variables such as 'education', refine measures of 'trade openness' more precisely, and perhaps most commonly seek out longer and better data sets. Other methods have included the use of panel data and techniques such as 'trimmed least squares'. Some researchers venture further realising that there may be more than an empirical problem at work, in particular that the theory relating growth and economic policy may be more complex than allowed for by simple cross-country regressions. If this is the case then the emphasis on improved data and technical refinements to the econometrics may be fruitless.

Section four explores a number of *theoretical* reasons why cross-country regressions may be an intrinsically poor method to isolate the link between changes in policy and changes in economic growth rates. Those analysed here are complementarity among policy variables, the relation between different theories of growth, hysteresis effects and dynamics. As demonstrated in this section, responses by researchers to these theoretical problems have been much more ad hoc.

Section five demonstrates the final problem of cross-country growth regressions that has rarely been faced by orthodox researchers. Far from being a positivist statistical exercise, cross-country growth regressions are bound to an underlying neo-classical assumption – that the growth process is universal. Each individual country in cross-section according to this view will provide evidence that can be used to elucidate the one underlying universal economic relation. An increase in openness for example is hypothesised to have the same effect on growth in all countries. There is a limited amount of evidence that can be teased out of existing cross-country

growth regressions that suggests each growth experience should be treated as potentially unique, i.e. as a case study.

The last section concludes by suggesting that heterodox or post-autistic economists should open up the assumption of universalism to greater scrutiny and ask why the growth process may differ across time and space. In practical terms this would question the neo-classical assumption of universalism and with it the 'one-size-fits-all' programmes of liberalisation emanating from the World Bank, IMF and other institutions. A case study approach to economic growth would be justified on the assumption that growth processes are not universal. Comparative and historically informed case studies allow researchers to question the assumption of universality rather than be forced to assume it true a priori.

2. The Analysis of Growth and Empirical Results

This section shows there is a strong theoretical link between policy change and economic growth but also that the empirical evidence for this link is very weak. This is shown here in both a general sense and specifically in relation to fiscal policy, investment, education and R+D.

Policy provides the most straightforward explanation for changes in economic growth. A typical example is the World Bank[1], which purports to show that 'strong adjusters' (policy) in Sub-Saharan Africa during the 1980s experienced increased rates of economic growth. The apparent link between policy and growth has been enhanced by theoretical developments. The original Solow growth model predicted that policy (investment) would only impact on the level of growth not the long-run growth rate. Endogenous growth models by contrast were motivated by the lack of convergence to steady-state among developing countries and the inability of traditional models to account for cross-country differences in income and growth rates.

Romer developed an equilibrium model of technological change in which optimising agents drove long-run growth through the accumulation of knowledge. The creation of knowledge by one firm has a positive external effect on the production possibilities of other firms[2]. Adding to capital and labour a third input (usually education) generates externalities. Due to the externality, these models yield a sub-optimal equilibrium/ market solution and in turn generate a potential role for the state. Policy is shown to effect growth through its impact on incentives to accumulate capital and knowledge and so generate technological change.

While the link between episodes of growth and stagnation and changes in policy seems intuitively reasonable and is supported by economic theory there is very little empirical evidence. Levine and Renelt took a number of variables commonly used in econometric growth analysis and ran them in thousands of regressions with different conditioning sets of other variables – judging them robust if they remained significantly related to growth. Their tests excluded variables that are only correlated with another factor that has a causal relationship with growth, i.e. those factors with an indirect impact on growth. They found only investment was robustly related to economic growth[3]. Even those factors many would accept as self-evidently related to economic growth, fiscal policy, investment, education and R+D have an ambiguous empirical relation to economic growth as revealed by conventional cross-country regression analysis. The relevant theory and empirical results concerning these four policy variables are analysed in turn.

2.1. Fiscal Policy and Growth.

Theory linking fiscal policy to economic growth is very clear. The link principally revolves around how increases in tax rates lower the return to private investment and hence permanently lowers the rate of investment. Barro measured government intervention as the ratio of real government consumption less spending on education and defence to real GDP. He found a significant negative association between this variable averaged over 1970-85 and real growth 1960-85[4]. More generally there is no robust relation between growth and the ratio of total government expenditure to GDP, government consumption expenditure, government capital formation, or government educational expenditure[5]. The coefficient in Barro becomes insignificant when the ratio of exports to GDP is included in the conditioning set[6]. There are good theoretical reasons for the relation between government expenditure and growth to become more complex once trade openness is considered. Openness may increase the cost of government intervention by raising the elasticity of taxed factors[7]. Rodrik finds a positive correlation between a country's exposure to international trade and the size of its government. A possible explanation he suggests is that government plays a risk-reducing role in economies exposed to a significant amount of external risk/ openness[8].

There are severe empirical problems with any attempt to quantify the role of the state through regressing the rate of growth on the level of government expenditure. Keyensian demand management and automatic stabilisers imply that government expenditure will increase with poor economic performance. This will generate a spurious negative relation between the 'size' of government and economic growth. Governments also influence the economy in many ways that do not involve expenditure, such as regulation. Tax exemptions and fiscal transfers may have identical effects but have different implications for the measured size of government. The demand elasticity for government services is typically greater than one (Wagner's Law). The level of government expenditure would then be determined endogenously.

2.2. Investment and Economic Growth

Investment was the one factor found robustly related to economic growth[9]. The average investment rate is frequently used as an independent variable in growth regressions, though there remain severe theoretical problems in identifying causality, and the traditional use of instruments as a 'solution' is problematic. "So many variables could be used to explain growth that it is difficult to find variables that are not only highly correlated with the endogenous variable but can also be plausibly excluded from the regression."[10]

There is some closer empirical work on this question. There is a positive correlation between investment in specifically machinery and equipment and productivity growth. The relationship holds for countries with 1960 levels of GDP per worker greater than 25% of the US level, between 1960 and 1985 and the result is causal, robust, strong and statistically significant[11]. There is also a positive and significant relation between the ratio of imported to domestically produced capital goods for a large cross-country regression between 1960 and 1985[12]. Ultimately such a causal link remains ambiguous. Between 1950 and 1988 the composition of investment in the OECD shifted sharply, the share of investment in producer durables, from 3 or 4% to more than 7% of GDP in France, Germany, the US and the UK and in

Japan from 3.5% to 9%, growth showed no upward trend. Others have found growth induced subsequent capital formation in a large cross-section of countries between 1965 and 1985[13].

2.3. Education and Economic Growth

Intuitively education has an evident link with economic growth, but again there is no clear empirical link. Pritchett finds a robust and *negative* correlation between higher school enrolment and educational attainment and total factor productivity growth in developing countries. Between 1960 and 1985 educational capital grew faster in sub-Saharan Africa and South Asia than in East Asia even though the latter region grew more rapidly[14]. Bils and Klenow find only a *weak* relation between initial schooling and subsequent economic growth, even allowing for the indirect effects of schooling in permitting greater technology absorption. They find the relation either spurious, the expected return and incentive to acquire education increases in an expanding economy when the skilled wage is growing rapidly, or reflective of omitted variables related both to initial schooling rates and subsequent economic growth rates[15].

A particular problem for regression analysis is finding a satisfactory measurement of human capital[16]. A large part of investment in education takes the form of forgone earnings by students. In addition, explicit spending on education takes place by the individual, family and state. Not all education expenditure is intended to generate productive human capital (for example the teaching of philosophy verses literacy). The typical proxy used in many cross-country regression equations is the share of the working-age population in secondary school. This fails to measure the quality of education, and the learning-on-the-job that takes place in the workforce. Good explanations of missing empirical support are possible external effects and the endogenous relation between education and economic growth.

2.4. R+D and Economic Growth

Theory and intuition suggest there is a clear link between R+D and economic growth. Again, this link has not been satisfactorily uncovered by empirical analysis. Between 1950 and 1988 the total number of scientists engaged in R+D in the US increased from 200,000 to over 1,000,000. A similar pattern was evident in Germany, France and Japan. Measured by R+D expenditure the results are similar. Despite this extra R+D there has been no permanent increase in growth in these countries[17].

3. An Empirical Problem: Episodes of Growth and Stagnation in Least Developed Countries

This section shows that using long-run averages typical of cross-country growth regressions hide an important empirical reality of growth in contemporary developing countries. Growth averages over the medium-term (25-30 years) conceal the periods of stagnation, growth spurts, structural breaks, volatility and instability that actually characterise growth in developing countries.

3.1. The Historical Experience of Developing Countries

Theoretical and empirical research on growth has focused on averages over the medium-term (25-30 years)[18]. A decade of ten-percent growth followed by another of zero-percent drops into a Barro-type regression with the same average as two decades of five-percent growth. This problem has real implications for the analysis of growth in developing countries. Brazil enjoyed rapid growth between 1965 and 1980, and stagnated during the 1980s. A medium-term average doesn't distinguish between the average of 3.1% between 1960-92 and the importance of the structural break. Per capita GDP in Cote D'Ivoire increased by 3.1% p.a. between 1960 and 1980 and declined by an average of 4.1% p.a. between 1980 and 1992. Ignoring the structural break average growth was 0.22%, almost the same as Senegal (0.18%) which stagnated throughout the whole period[19].

Growth averages over the medium-term (25-30 years) conceal the periods of stagnation, growth spurts, structural breaks, volatility and instability that actually characterise growth experiences in developing countries. The overall average is *not* a good summary indicator of growth performance. Countries show shifts in growth rates, often in clear episodes, such as the slowdown in Latin America in the 1980s. GDP growth is not well characterised by a single exponential trend. For forty percent of least developed countries the R^2 on such a trend is less than 0.5, suggesting that shifts and fluctuations are the dominant feature of GDP growth. There are instead six distinct patterns of growth, before and after statistically chosen structural breaks, steep hills, hills, plateaus, mountains, plains and accelerations[20]. Growth is very unstable across time periods. The correlation of per capita growth between 1977-92 and 1960-76 across 135 countries is only 0.08[21].

There are very striking instances of growth accelerations and collapses among developing countries. There are 14 episodes of growth in Africa between 1960 and 1996 including South Africa between 1960 and 1974 (5.1%), Cote D-Ivoire 1960 to 1978 (9.5%), Gabon 1965 to 1976 (13.1%), and Namibia 1961 to 1979 (6.4%)[22]. Ten countries in Africa between 1967 and 1980 had growth of more than 6% p.a., including Gabon, Botswana, Congo, Nigeria, Kenya and Cote D'Ivoire all which were outperforming both Malaysia and Indonesia[23]. Hausmann et al conducted a very broad empirical test to locate episodes of growth by finding the year that maximises the F-statistic of a spline regression with a break at the relevant year[24]. Countries can have more than one acceleration. This filter yields 83 growth accelerations, capturing most well-known episodes such as China 1978, Argentina 1990, Mauritius 1971, Korea 1962, Indonesia 1967, Brazil 1967, Chile 1986, and Uganda 1989. The magnitude of accelerations is striking. Their definition is conditional on a growth acceleration of at least 2% p.a.; the average acceleration though was 4.7% p.a. There are many episodes with acceleration of 7% or more such as Ghana 1965 (8.4%), Pakistan 1962 (7.1%), and Argentina 1990 (9.2%). The occurrence of an episode is quite common, of 110 countries in their sample between 1957 and 1992 54.5% had at least one episode of growth and 20.9% two. The occurrence is also common across space: 21 episodes occurred in Asia, 18 in Africa, 17 in Latin America, 12 in Europe and 10 in the Middle East and North Africa.

4. Cross-country Growth Regressions: Theoretical Problems

Recent theorising on endogenous growth models is clear that there should be a strong link between policy and growth. This section shows that any empirical link between changes in public policy and changes in growth rates will be difficult to isolate using traditional cross-country regression analysis for theoretical not just empirical reasons. These theoretical problems include complementarity among policy variables, the relation between different theories of growth, hysteresis effects, and dynamics.

4.1. Complementarity among Policies

Policy variables typically enter the right hand side of regressions separately without diagnostic tests allowing for any but very limited interaction among them. Theory *does* suggest complementarity is important. For example, investment may be causally related to growth only in the presence of strong property rights, reforms causally linked to growth only if considered credible or if correctly sequenced. There is some limited empirical support for the importance of complementarity between policies. Mosley finds that complementarity between inflation, openness and the government share to be minimal but when corrected for sequencing the coefficient(s) increases and becomes significant[25].

Econometrics has coped with this problem in an ad-hoc manner, splitting country samples by region or income level to look for changes in the strength and direction of causal relations or including occasional ad-hoc interaction effects between two variables. It would in theory be possible to add all possible interaction effects by adding multiplicative relations in a regression between all combinations of variables and adding a welter of dummy variables for all possible structural breaks and geographical regions. The resulting loss of degrees of freedom would then render the regression all but meaningless.

4.2. The Relation between Different Theories of Growth

There are numerous cross-country econometric studies finding some indicator of national policy to be linked to economic growth. There is however no clear consensus on what policy variables to include in cross-country regression analysis. Economic theory rarely generates a complete listing of variables to be held constant when trying to gauge the impact on the relation between the dependent and independent variable. Mauro for example adds measures for corruption and Knack and Keefer likewise add proxies for trust to standard Barro-type regressions[26]. There is no means to compare the merits of these two approaches and the relationship between these and other theories remains confusing. A causal relation between two variables (e.g. trade and growth) does not imply the falsity of another (e.g. democracy and growth). Levine and Renelt find “statistical relationships between long-run average growth rates and almost every particular policy indicator considered by the profession are fragile: small alterations in the ‘other’ explanatory variables overturn past results”[27].

4.3. Growth and Hysteresis Effects

Hysteresis effects are likely to exist in the process of economic growth. Hysteresis implies that a temporary economic shock can have a permanent impact on future growth. The implication being that growth is not a linear process and regression analysis will have trouble capturing this effect. There may be virtuous and vicious circles at work in growth connected with threshold effects[28]. If a country has a critical initial mass of human and physical capital, growth will be virtuous, capital accumulation attracting yet more capital. Green revolution technology for example depends on the availability of both seeds and fertilisers through access to adequate credit (and hence collateral). Households with enough collateral can invest in the necessary skills and technology to get the virtuous circle going. The option is not open for poor households without collateral. With an exogenous shock there is the potential for hysteresis effects. A disaster can wipe out the liquid assets of a household and leave it in a poverty trap, unable thereafter to invest in green revolution technology. Potential poverty traps make households and an entire economy more vulnerable to shocks. If a country were near the critical mass level of capital, then a terms-of-trade shock that rendered part of that country's capital stock useless might shift that country from strong growth to decline. The same shock may have little effect on a country far from the threshold. This implies there may be no primary causal factor but an interlocking circular process with feedback. Neat econometric models with fixed coefficients will by definition then be impossible to find.

4.5. Cross-Country Growth Regressions and Dynamics

Theories of cyclical and adjustment dynamics of output are not well developed within growth theories. Reliable data sets for many traditional growth determinants (inflation, government expenditure, tariffs, inequality etc) have typically run for twenty-five plus years. Averages over this sample length are too short for history and too long to model macroeconomic policy changes and short-run dynamics. In cross-sectional regression analysis it is not clear whether variables affect long-term growth, the steady-state, or both. Some growth effects are contemporaneous (macroeconomic and cyclical factors), some take several years (transitional dynamics due to changed investment incentives), others even decades (incentives effecting the rate of technical change). Some right-hand-side variables may have output/ growth effects at all three horizons - cyclical, transitional and steady-state. There is no reason to assume these are of the same magnitude or even the same sign[29]. What little ad hoc empirical work has been carried out finds regression parameters commonly are unstable over time. Knack and Keefer for example look at the relation between social characteristics and growth and find that social variables have different signs on growth before and after 1980[30]. With such findings common attempted explanations are notable only by their absence.

5. Universalism: A Problem of Neo-classical Economics?

In order to run large cross-country regressions researchers are tightly constrained to the assumption of universalism. Conventional growth analysis assumes growth parameters are identical across countries. Far from being a positivist statistical exercise, cross-country growth regressions are bound to an underlying neo-classical assumption – that the growth process is

universal. Each individual country in cross-section according to this view will provide evidence that can be used to elucidate the one underlying universal economic relation. An increase in openness for example is hypothesised to have the same effect on growth in all countries. There are a small number of exceptions that for example allow the constant term to differ across countries (controls for fixed effects) using panel data or on occasional a dummy variable is added for regions and notable events. The over-riding assumption behind cross-country growth regressions is a that of a universalist growth process.

Many studies explain Africa's slower growth as a function of different levels of explanatory variables[31]. They seek to explain growth as the result of a *common* growth process beginning from different *levels* of the same explanatory variables. Significant regional dummies remain common in much of this literature, and especially so for Sub-Saharan Africa. The usual assumption is that significant dummy variables are capturing the influence of missing variables, which must then be unearthed. This has led researchers to propose ever more variables in the hope that the dummy variable will be rendered insignificant and growth will finally be 'explained'[32]. The *alternative* route is to relax the assumption that only the levels of explanatory variables are different and explore the idea that the growth process in Africa works differently. There are a limited number of studies that suggest this latter notion may be true. The implication being that cross-country growth regressions are an intrinsically poor mechanism to analyse growth and each growth experience should be treated as potentially unique i.e. as a case study.

Block conducts a more flexible analysis and allows for the slope coefficients to differ and finds openness in Sub-Saharan Africa has a much stronger effect on growth and that growth is less responsive to fiscal policy than his sample average[33]. Brock and Durlauf find "the operation of ethnic heterogeneity on growth is different in Africa, not just the levels of ethnic heterogeneity. A comparison of other regressor coefficients for Africa with those of the rest of the world makes clear the growth observations for African countries should not be treated as partially exchangeable with the growth rates of the rest of the world."[34]. Asiedu finds that FDI is less responsive to openness in Africa than in other regions. For a given level of trade openness, infrastructure and return on capital, Sub-Saharan Africa receives less FDI[35]. Mosley finds that inequality only has a negative impact on growth in regions other than sub-Saharan Africa[36].

5. Summary

Cross-country growth regressions assume that economic growth operates according to universal laws across all economies through time and space. This is one of the key ideological foundations of neo-classical economics. There are only a few exceptions and the occasional dummy variable for regions and notable events. Discussion in this paper has demonstrated that there is evidence the growth process differs significantly between different regions and countries and over time. Finding fragility and heterogeneity of regression coefficients by region and country is only a beginning. Opening up the assumption of universalism to greater scrutiny leaves us asking why the growth process may differ. In doing so we would throw doubt on the neo-classical assumption of universalism and with it the 'one-size-fits-all' programmes of liberalisation emanating from the World Bank, IMF and other institutions. A case study approach to economic growth would be justified on the assumption that growth processes are not universal. The recent collection of country case studies together with an introduction drawing general lessons edited by

Rodrik (2003) is a useful step in this direction[37]. Comparative and historically informed case studies allow researchers to question the assumption of universality rather than be forced to assume it true a priori. A heterodox or post-autistic economist should begin with case studies and only then proceed to cross-country growth regressions with all due caution.

Notes

1. World Bank '*Adjustment in Africa: Reforms, Results and the Road Ahead*', (Oxford University Press, New York 1994).

2. P.M.Romer, 'Increasing Returns and Long-Run Growth', *Journal of Political Economy*, 94:5 (1986) p1003.

3. R.Levine and D.Renelt, 'A Sensitivity of Cross-Country Growth Regressions'*, Amercian Economic Review*, 82:4 (1992), p942-963.

4. R.J.Barro, 'Economic Growth in a Gross Section of Countries', *Quarterly Journal of Economics,* 106, (1991), p407-43l.

5. Levine and Renelt (1992).

6. Barro (1991) and Levine and Renelt (1992:951).

7. J.Slemrod, 'What do Cross-Country Studies Teach About Government Involvement, Prosperity and Growth', *Brookings Papers on Economic Activity*, 2 (1995), p405.

8. D.Rodrik, 'Why Do More Open Economies Have Bigger Governments?', *Journal of Political Economy*, 106:5 (1998), p997-1032.

9. Levine and Renelt (1992).

10. J.Temple 1999:128) 'The New Growth Evidence', *The Journal of Economic Literature*, 38, (1999), p128.

11. J.B.De Long and L.H.Summers, 'Equipment Investment and Economic Growth', *Quarterly Journal of Economics*, 106, (1991), p445-502.

12. J-W.Lee, 'Capital Goods Imports and Long-run Growth', *Journal of Development Economics*, 48, (1995), p91-110.

13. M.Blomstrom, R.E.Lipsey and M.Zejan, 'Is Fixed Investment the Key to Economic Growth', *Quarterly Journal of Economics*, 111 (1996), p269-276.

14. L.Pritchett, '*Where has all the Education Gone?*', World Bank (1999), p3.

15. M.Bils and P.J.Klenow, 'Does Schooling Cause Growth?', *American Economic Review*, 90:5 (2000), p1160-1183.

16. N.G.Mankiw, D.Romer and D.N.Weil, 'A Contribution to the Empirics of Economic Growth', *Quarterly Journal of Economics*, 107 (1992), p418-9.

17. C.I.Jones, C.I, 'R+D Based Models of Economic Growth', *Journal of Political Economy*, 103:4 (1995), p759-784.

18. Beginning with Barro (1991).

19. L.Pritchett, 'Understanding Patterns of Economic Growth: Searching for Hills among Plateaus, Mountains, and Plains', *World Bank Economic Review*, 14:2 (2000), p221-50.

20. Pritchett (2000).

21. W.Easterly and R.Levine (2001:195) 'It's Not Factor Accumulation: Stylised Facts and Growth Models', *World Bank Economic Review*, 15:2 (2001) p195, see also Temple (1999:116).

22. Defined as a ten-year plus period in which the five-year moving average of annual GDP growth exceeds 3.5%, J-C.Berthelemy and L.Soderling, 'The Role of Capital Accumulation, Adjustment and Structural Change for Economic Take-Off: Empirical Evidence from African Growth Episodes', *World Development*, 29:2 (2001), p323-343.

23. T.Mkanawire, 'Thinking about Developmental States in Africa', *Cambridge Journal of Economics*, 25 (2001) p2989-313).

24. R.Hausmann, L.Pritchett and D.Rodrik, '*Growth Accelerations*', JFK School of Govt, Harvard Universitry (2004).

25. Adding a premium for example when reforms were conducted in the correct sequence, deflation before devaluation, liberalisation of the current before capital account etc, P.Mosley, 'Globalisation, Economic Policy and Convergence', *World Economy*, 23:5 (2000) p613-634.

26. P.Mauro, 'Corruption and Economic Growth', Quarterly Journal of Economics, 110 (1995) p681-712). S.Knack and P.Keefer, 'Does Social Capital Have an Economic Payoff: A Cross-Country Investigation', *Quarterly Journal of Economics*, 112 (1997), p1251-1288.

27. Levine and Renelt (1992:943).

28. W.Easterly, '*The Elusive Quest for Growth: Economists' Adventures and Misadvantures in the Tropics*', (MIT Press, Cambridge, 2001), Chapter 10.

29. Temple (1999:124).

30. Knack and Keefer (1997).

31. W.Easterly and R.Levine, 'Africa's Growth Tragedy: Policies and Ethnic Divisions', *Quarterly Journal of Economics*, 112 (1997), J.D.Sachs and A.M.Warner, 'Fundamental Sources of Long-Run Growth', *American Economic Review*, 87:2, (1997) p184-8 among many others.

32. A good example of an exhaustive attempt to remove the significant dummy variable is P.Collier and J.Gunning 'Explaining African Economic Performance', *Journal of Economic Literature*, 37:1 (1999), p64-111.

33. S.A.Block, 'Does Africa Grow Differently?', *Journal of Development Economics*, 65, (2001), p443-467.

34. W.A.Brock and S.N.Durlauf, 'Growth Empirics and Reality: What Have We Learned from a Decade of Empirical Research on Growth?', *World Bank Economic Review* 15:2, (2001), p264.

35. E.Asiedu, 'On the Determinants of Foreign Direct Investment in Developing Countries: Is Africa Different?', *World Development*, 30:1 (2002), p107-119.

36. Mosley (2000).

37. D.Rodrik, '*In Search of Prosperity: Analytic Narratives on Economic Growth*', (Princeton University Press, Princeton, 2003).

Author contact: mm80@soas.ac.uk

SUGGESTED CITATION:
Matthew McCartney, "Can a Heterodox Economist Use Cross-country Growth Regressions?" *post-autistic economics review*, issue no. 37, 28 April 2006, article 4, pp. 45-54, http://www.paecon.net/PAEReview/issue37McCartney.htm

Prying Open American Political 'Science'

Bruce Cumings and Kurt Jacobsen (University of Chicago, USA)

Numbers are seductive. We love what we are good at so it is hardly surprising when a math wiz imagines that numbers contain the whole truth and nothing but the truth. A mathematically talented teen recently told one of us of an alleged encounter centuries ago of a famous French philosopher with a Russian mathematician who proceeded to spout an algebraic equation and to claim, because the equation made sense, that he had proven there was a god. The philosopher, according to the story, was dumbstruck, which our young friend interpreted as abject surrender to a superior mind. We replied that the philosopher doubtless was flabbergasted that a bright fellow could be so gullible as to believe that a perfectly enclosed and self-referential system like mathematics necessarily had anything reliable to say about the wider and wilder world around it.

Perhaps we were interpreting too, but the point stands. Run numbers through a complicated enough set of procedures and they enchant especially managerial mentalities who like to conjure a tidy abstract universe where there's no need to use careful judgment based on extensive research and hard-won experience about the way societies operate. Fill in the blank spaces to a formula and, presto, you've solved the problem. Skip all the steps in between and forget there was any processing as to what the numbers mean. A great deal gets lost when numbers are used without humility or reflection. Lousy policies are one result. Critics argue, for example, that environmental costs cannot be expressed adequately in money terms. What figure captures all the harm of polluted air, soil or water? The use of GDP to calculate prosperity is misleading since it counts disasters positively - the costs of clean-up raise GDP. Numbers may get you from here to the Moon or Baghdad but won't tell you if the trip is worth it. Instead of regarding numbers as a necessary evil we need to beware of, econometricians typically treat them with adoration. Economists, laden with glittering faux Nobel prizes, have led a strong trend toward quantification in all the social science by deploying econometric models - models, moreover, that tend to favour neoliberal market schemes. (After all, nothing commodifies you like a number does.)

Accordingly, in American political science today any research paper lacking phalanxes of jitterbugging numbers seems suspect, not science at all. Political science is hardly alone these days in privileging method and numbers. Economics, of course, is fully colonized by this mentality, to the great detriment of economic history. Sociology neglects its political and historical dimensions, and thus rarely produces figures like C. Wright Mills who command both a disciplinary and general reading public; and even in history, owing to different theoretical proclivities, political history is considered boring and passe. The result is that many students in top universities have trouble finding courses on actually-existing politics. A revolt was brewing among some of their disgruntled teachers.

So in October 2000 an anonymous American political scientist (or perhaps several of them) under the name "Mr. perestroika' dispatched a scorching email to a number of noted senior figures in the discipline. The email excoriated the domination of political science by enthusiasts of

formal theory and of quantitative methods, who tend to make common cause. The problem was not so much that certain factions within political science were ascendant but that formal theorists did not believe anything except their own brand of theory mattered and that many exponents of quantitative methods did not believe anything but their own manipulation of mathematical symbols deserved the label 'method.' Hence, they were disinclined, as they gained control of departments, to heed or hire any but their own. "Mr Perestroika" deplored these "poor game-theorists who cannot for the life of me compete with a third-grade economics student" yet are allowed to crush "diversity of methodologies and areas of the world that [American Political Science Association] 'purports' to represent."

'Perestroika' lamented that the cost to knowledge in the study of politics stood to be enormous in the sense of the fabled man with a hammer as his only tool treating everything as a nail, or seeing only hammers and nail-like items as worth knowing anything about. Disgruntlement with this dogma had been growing all along among other scholars, the sort who want to know their subjects well before playing reductionist games with them, and they were only awaiting a spark. So the email – likely inspired by the post-autistic economics movement in France - ignited a rousing scholarly movement in America.[1] One could not hope to assemble a more unlikely band of insurrectionists, ranging from apprehensive grad students to greying professors ensconced in named chairs.

Several hundred tenured scholars signed a petition charging that a dangerous fad for formal models and number-crunching was squeezing out valuable forms of research. At the 2001 American Political Science Association (APSA) meetings attendance at several perestroika-themed panels spilled into the hallways. Well-known panelists included Penn's Rogers Smith, Harvard's Steve Walt, Johns Hopkins University's Margaret Keck, Colorado's Sven Steinmo, Indiana's Gregory Kasza, Chicago's Susanne Hoeber Rudolph and Lloyd Rudolph and Yale's Ian Shapiro and James Scott, the lattermost becoming perestroika's first representative on the APSA Council. Political science has "been taken over by methodological parochialists who believe that the only worthwhile scholarship in political science speaks the language of mathematics," said Chicago Professor John Mearsheimer of foes whom, he warned, were formidable. For the latter, only counting matters because mathematics conveys a sense of precision, as if numbers never lie. As the belief that quantitative data are not themselves a form of interpretation becomes institutionalised, this naïve orthodoxy excludes important sources of insight. One consequence is that economists and political scientists seem to have less and less to say about anything that we recognize as the world we move in. Most have nothing to say about expanding social inequalities, neo-imperialistic crusades and ecological woes.

The perestroika movement – approximately 900 out of 15 thousand American Political Science Association members - is the latest round in a recurrent battle between different notions of legitimate research in the social sphere. Anthropologist Clifford Geertz once joked that in social science old ideas tend less to fade away than to go into second editions. The struggle today harks back to the 1960s when the Caucus for a New Political Science arose to combat pretensions of behavioral 'hard science,' to battles in sociology departments in the 1950s such as Columbia where Paul Lazersfeld's statisticians and C. Wright Mills more plain spoken disciples squared off, and even stretch centuries back to the piquant (and self-serving) Latin dictum that only knowledge expressed in numbers matters.[2]

Obviously there are scholars well-versed in both quantitative and qualitative methods who deftly integrate them. These few are not the problem. The field today is again enticed by the ambition that establishing a dominant method, as in 19th century physics, is the sine qua non of a first-rate science. It is just as reasonable to argue that the pursuit of a single paradigm is really the pursuit of the right not to have to think or raise discomfiting questions. Plug in pre-digested data and let the paradigm do the work for you. How very handy. One can see why this quest appeals to certain sorts of intellects. Numeracy is a wonderful thing - so long as it does not entail illiteracy in other fields.

As Thomas Kuhn, and other historians and philosophers of science, attest, every paradigm is a selective device, making a particular kind of informational demand on the multi-layered nature of reality, and deliberately excluding other aspects of it.[3] There is no way, within the boundaries of a single dominant paradigm, to discover if one is mistaken about the importance of those excluded aspects, and of those shunned perspectives. The paradigm you adopt pre-determines your answers. Further, because the profession seeks out, and rewards, generalizable propositions derived from studies of many ('large-N") cases, then those propositions simply are assumed to have counterparts in reality, no matter how Procrustean the methods used to slice reality down to such convenient size. Any savvy philosopher – and anyone who knows the cases on the ground intimately - can spot how flawed, indeed foolish, these notions are.

Rational choice theory derives from neoclassical economics and deploys simplifying assumptions about human behaviour to boil down complex experiences to prioritised "rational" choices that we presumably make in order to maximise our "utility" in any situation. Rational choice, and mathematical models that accompany it, have merits when used with humility, especially in studies of collective action. But its critics point out that the results are often trivial or else remote from reality.[4] Formal theory, unchecked, gives carte blanche to cram all manner of oblong pegs into little square holes. Examples of silliness abound. "Is it our task to understand politics," Perestroikan Greg Kasza of Indiana University therefore asks, "or to grapple with the logic of imaginary games?" What he and other perestroikans propose instead is 'the idea of an ecumenical science. It is based on three principles: problem-driven research, methodological pluralism, and interdisciplinary inquiry.' Apart from methodological dogmatism, Perestroikans are concerned about the stacked deck of APSA elections (where an unelected committee appoints the president for one year terms), reform of professional journals, and, most difficult of all, hiring practices.

The movement was heartened in February 2001 when the APSA nominating committee selected Theda Skocpol of Harvard, an academic with diverse methodological skills, as president elect. Both Skocpol and predecessor Robert Putnam began including Perestroikans on decision bodies of the Association. The American Political Science Review, flagship journal of the APSR, and the journals of regional associations, came under scrutiny because they are often used by departments as a short-cut certifying process for faculty recruitment and promotion. If you don't see print there, you often are in trouble. An initiative set in motion earlier by in-house critics of the APSR, to launch a journal, Perspectives on Politics, as an alternative to the parochialism of the APSR, was accelerated by Perestroika's presence. A new editor of the APSR recognized the grievances concerning the absence of diversity, and promised change. All was not sweetness and light internally either. Skocpol, in a show of presidential impartiality, chided Perestroika itself

as unrepresentative while APSA nominating committee member Joan Vecchiarelli Scott said – taking the role of Simon Schama regarding the French revolution – that, ho hum, reforms were in the pipeline anyway.

Still, the movement gained a major success in 2003 when the APSA presidency went to perestroikan Susanne Hoeber Rudolph of the University of Chicago. 'One of the effects of 9/11 is a renewed awareness that Americans need enhanced capacity to understand and interpret the 'other," remarked Rudolph at the time. "Corridor talk in the Association is that my special focus in political science, comparative politics and India, which would ordinarily be a disadvantage in an association whose membership is concentrated on America and the West, [probably was] an advantage."At first glance, the elegant and erudite Rudolph was not Hollywood central casting's idea of a rabble rouser, but academic trends drove many prominent people to the metaphorical barricades. Rudolph received a batch of fretful letters warning darkly of thermidore or of cooptation by wily formalists. Some colleagues worry that the new journal Perspectives on Political Science launched in order to broaden the association's appeal is fated to be a second class ghetto. A self-nominated committee on reform of Association governance formulated proposals for competitive elections, which ultimately were stymied during Rudolph's term. Rudolph's successor Margaret Levi, the 3rd woman President in a row, was not a perestroikan but current President Ira Katznelson of Columbia University is regarded as sympathetic.

So is perestroika an internet forum for the exchange of views, or a movement, or both? "Perestroika is more an attitude, a set of concerns, an adherence to certain values as a scholar, that lead each of us to question the dominant paradigm in Political Science, but obviously for many different reasons," perestroikan Michael Bosia sums up. "Perestroika is a movement of critique, disorganized and uncentered, a forum for discussion, but never an organization seeking power for itself. Perestroikans are all of us who choose to identify as critics of established orthodoxies in Political Science."

Notes

1.On the post-autistic economics movement see Edward Fullbrook, ed, The Crisis in Economics (London: Routledge, 2004).

.

2. On the Caucus, see Philip Green and Sanford Levinson, eds, Power and Community: Dissenting Essays in Political Science (New York: Pantheon, 1969).

3. See Thomas Kuhn, The Structure of Scientific Revolutions (Chicago: University of Chicago, 1962).

4. See, for example, Donald Green and Ian Shapiro , Pathologies of Rational Choice Theory (New Haven: Yale University Press, 1996) and Steve Walt, "Rational Choice or Rigor Mortis?: Rational Choice and Security Studies" International Security 23, 4 (Spring 1999), and the exchange in the next issue.

Author contact: jkjacobs@uchicago.edu

SUGGESTED CITATION:
Bruce Cumings and Kurt Jacobsen, "Prying Open American Political 'Science'" *post-autistic economics review*, issue no. 37, 28 April 2006, article 5, pp. 55-58, http://www.paecon.net/PAEReview/issue37CumingsJacabsen37.htm

Comment on Kaye-Blake

We are writing to comment on "Economics is a Structured Like a Language" (William Kaye-Blake, post-autistic economics review, issue no. 36, 24 February 2006, pp. 25-33) http://www.paecon.net/PAEReview/issue36/KayeBlake36.htm.

Kaye-Blake argues that Lawson's critical realism and Ruccio's postmodern ontologisms are closed, though they claim to be open. Their appearance of openness is dismissed as complexity: "There is a problem with both of these formulations of openness. They reflect complexity: multiple forces acting either on the surface of society or surging up from its depths; they are still deterministic." In order to offer indeterminacy, Kaye-Blake presents an ontological formulation stylized by the linguistic ideas of Saussure.

According to Saussure, meaning influences meaning. Aside from this influence, the relationship between the signifier and signified is arbitrary – leaving only synchronicity, i.e. the mere act of understanding – to encapsulate meaning. Otherwise meaning is to be achieved at the end of time, or at the end of the chain of meaning. Approaching economic questions with a similar ontology permits indeterminacy because causality can only be explained after the event, or at the end of time. Kaye-Blake explains, "the present is thus open not because of the past but because of the future."

Complexity science handles indeterminacy by means of bifurcation. Complexity agrees that cumulative causation shapes a system trajectory, supporting Kaye-Blake's critique quoted above, but if the parameters of a system are changed, however, attractor basins (or the initial conditions that lead to a particular final state) may appear or disappear. The system itself lacks the vitality to jump from one dependent path to another, but a system reconfiguration may cause bifurcation, thus permitting openness (Allen 2001).

Reference

Allen, Peter M. (2001) Knowledge, ignorance, and the evolution of complex systems. In John Foster and J. Stanley Metcalfe (Eds.), *Frontiers of Evolutionary Economics: Competition, Self-Organization and Innovation Policy*. (pp. 313-351). Cheltenham, UK: Edward Elgar.

Concetta Balestra
Frank Fagan
University of Bologna

Author contact: fjfagan@gmail.com

SUGGESTED CITATION:
Concetta Balestra and Frank Fagan, "Comment on Kaye-Blake" *post-autistic economics review*, issue no. 37, 28 April 2006, p. 59, http://www.paecon.net/PAEReview/issue37BalestraFagan37.htm

---

Articles, comments on and proposals for should be sent to the editor at
pae_news@btinternet.com

Subscriptions to this email journal are free.

Back issues of this journal and other material related to the PAE movement are available at www.paecon.net.

To subscribe to this journal, send an email with the message "subscribe" to
pae_news@btinternet.com

To unsubscribe to this journal, send an email with the message "unsubscribe" to
pae_news@btinternet.com