World Economic Review
An Open Access Journal from the World Economics Association

Incorporating the *Rentier* Sectors into a Financial Model[1]

Michael Hudson
University of Missouri at Kansas City & Levy Institute, USA

Dirk Bezemer
University of Groningen, Netherlands

Abstract

Current macroeconomics ignores the roles that rent, debt and the financial sector play in shaping our economy. We discuss the Classical view on rents and policy responses to the *rentier* sector in the 19th century. The finance, insurance & real estate sector is today's incarnation of the *rentier* sector. This paper shows how financial flows can be conceptually and statistically studied separately from (but interacting with) the real sector. We discuss finance's interaction with government and with the international economy.

Key words: asset prices, debt, deflation, FIRE-sector, *rentier* sectors

1. Introduction

Now that the Bubble Economy has given way to debt deflation, the world is discovering the shortcoming of models that fail to explain how most credit creation today 1) inflates asset prices without raising commodity prices or wage levels, and 2) creates a reciprocal flow of debt service. This debt service tends to rise as a proportion of personal and business income, outgrowing the ability of debtors to pay – leading to 3) debt deflation. The only way to prevent this phenomenon from plunging economies into depression and keeping them there is 4) to write down the debts so as to free revenue for spending once again on goods and services.

By promoting a misleading view of how the economy works, the above omissions lead to a policy that fails to prevent debt bubbles or deal effectively with the ensuing depression. To avoid a replay of the recent financial crisis – and indeed, to extricate economies from their present debt strait-jacket that subordinates recovery to the overhang of creditor claims (that is, saving the banks from taking a loss on their bad loans and gambles) – it is necessary to explain how credit creation inflates housing and other asset prices, while interest and other financial charges deflate the 'real' economy, holding down commodity prices, shrinking markets and employment, and holding down wages in a downward economic spiral. We are dealing with two price trends that go in opposite directions: asset prices and commodity prices. It therefore is necessary to explain how credit expansion pushes asset prices up while simultaneously causing debt deflation.

Yet the typical MV=PT monetary and price model focused on commodity prices and wages, not on the asset prices inflated by debt leveraging. Similarly, today's Dynamic Stochastic General Equilibrium (or DSGE) models are characterized by the "absence of an appropriate way of modeling financial markets" (Tovar 2008, p. 29). While Schumpeter (1934, p. 95) already noted that "processes in terms of means of payment are not merely reflexes of processes in terms of goods. In every possible strain, with rare unanimity, even with impatience and moral and intellectual indignation, a very long line of theorists have assured us of the

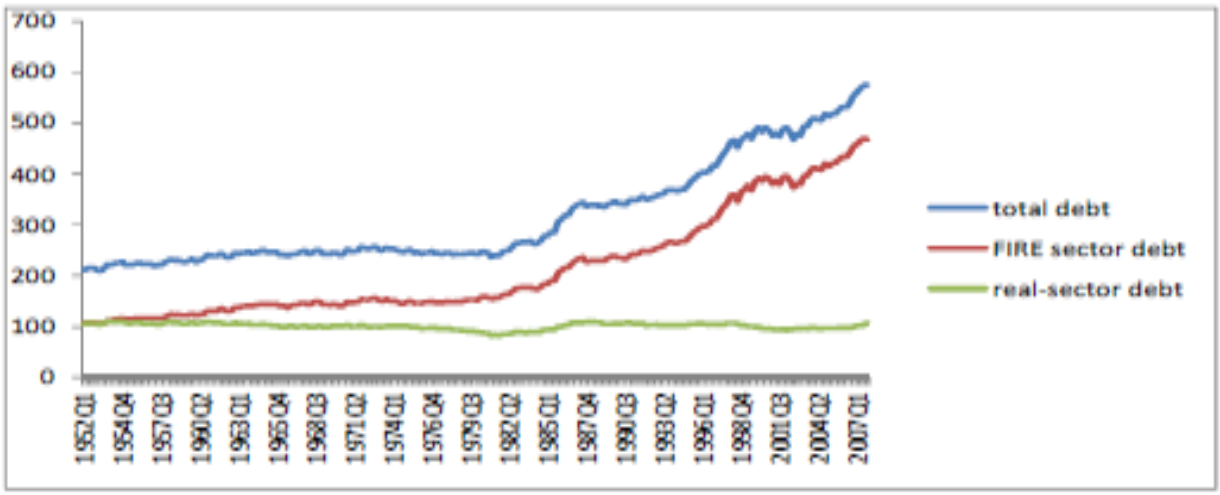
[1]An earlier version of this article will be in Hudson's forthcoming book, *The Bubble and Beyond*. The authors declare no conflict of interest.

opposite", this finds no place in DSGE models or, for that matter macroeconomics in the last decades. Cecchetti et al (2011, p. 2) describe current practice in writing that "for a macroeconomist working to construct a theoretical structure for understanding the economy as a whole, debt is… trivial… because (in a closed economy) it is net zero – the liabilities of all borrowers always exactly match the assets of all lenders. …With no active role for money, integrating credit in the mainstream framework has proven to be difficult". And yet credit and its counterpart, debt, have shaped our economic systems since prehistory (Hudson 2004). Understanding how credit is used is therefore a sine qua non for understanding our economy. That requires, in turn, to think about a fundamentally different model which can replace DGSE type models as the standard for analysis. Only then can the "naked emperor be dethroned" (Keen 2011).

2. Finance is not *the* economy

In the real world most credit today is spent to buy assets already in place, not to create new productive capacity. Some 80 percent of bank loans in the English-speaking world are real estate mortgages, and much of the balance is lent against stocks and bonds already issued. Banks lend to buyers of real estate, corporate raiders, ambitious financial empire-builders, and to management for debt-leveraged buyouts. A first approximation of this trend is to chart the share of bank lending that goes to the 'Fire, Insurance and Real Estate' sector, aka the nonbank financial sector. Graph 1 shows that its ratio to GDP has quadrupled since the 1950s. The contrast is with lending to the real sector, which has remained about constant relative to GDP. This is how our debt burden has grown.

Graph 1: Private debt growth is due to lending to the FIRE sector: the US, 1952-2007



Source: Bezemer (2012) based on US flow of fund data, BEA 'Z' tables

What is true for America is true for many other countries: mortgage lending and other household debt have been "the final stage in an artificially extended Ponzi Bubble" as Keen (2009, pp. 347–357) shows for Australia. Extending credit to purchase assets already in place bids up their price. Prospective homebuyers need to take on larger mortgages to obtain a home. The effect is to turn property rents into a flow of mortgage interest. These payments divert the revenue of consumers and businesses from being spent on consumption or new capital investment. The effect is *de*flationary for the economy's product markets, and hence consumer prices and employment, and therefore wages. This is why we had a long period of low cpi inflation but skyrocketing asset price inflation. The two trends are linked.

Debt-leveraged buyouts and commercial real estate purchases turn business cash flow (ebitda: earnings before interest, taxes, depreciation and amortization) into interest payments. Likewise, bank or bondholder financing of public debt (especially in the Eurozone, which lacks a central bank to monetize such

debt) has turned a rising share of tax revenue into interest payments. As creditors recycle their receipts of interest and amortization (and capital gains) into new lending to buyers of real estate, stocks and bonds, a rising share of employee income, real estate rent, business revenue and even government tax revenue is diverted to pay debt service. By leaving less to spend on goods and services, the effect is to reduce new investment and employment. Contemporary evidence for major OECD economies since the 1980s shows that rising capital gains may indeed divert finance away from the real sector's productivity growth (Stockhammer 2004, pp. 719-41) and more generally that 'financialization' (Epstein 2005) has hurt growth and incomes. Money created for capital gains has a small propensity to be spent by their *rentier* owners on goods and services, so that an increasing proportion of the economy's money flows are diverted to circulation in the financial sector. So wages do not increase, even as prices for property and financial securities rise – just the well-known trend that we have seen in the Western world since the 1970s, and which persists into the post-2001 Bubble Economy.

It is especially the case since 1991 in the post-Soviet economies, where neoliberal (that is, pro-financial) policy makers have had a free hand to shape tax and financial policy in favor of banks (mainly foreign bank branches). Latvia is cited as a neoliberal success story, but it would be hard to find an example where *rentier* income and prices have diverged more sharply from wages and the 'real' production economy.

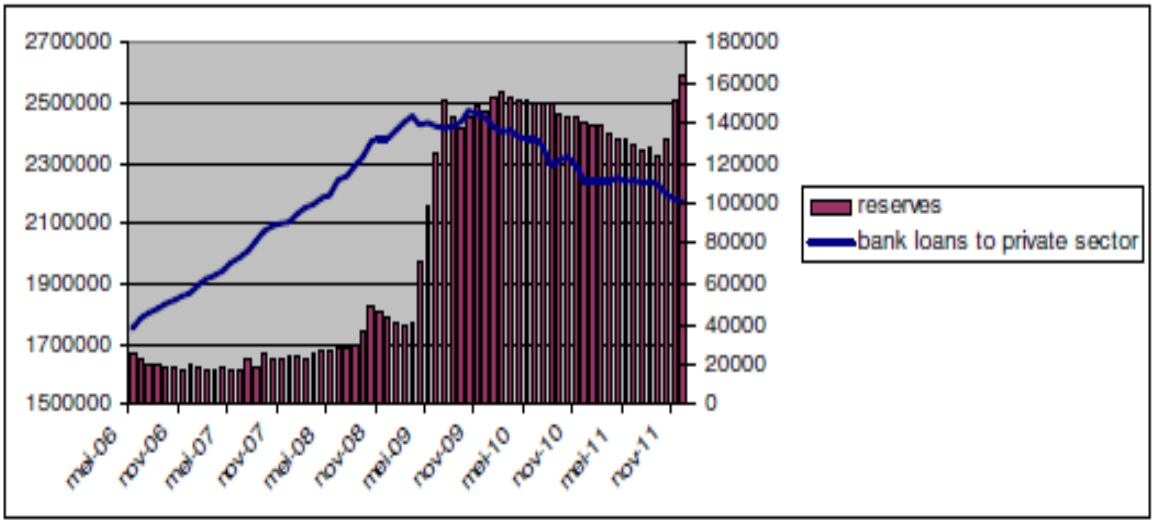
The more credit creation takes the form of inflating asset prices – rather than financing purchases of goods and services or direct investment employing labor – the more *de*flationary its effects are on the 'real' economy of production and consumption. Housing and other asset prices crash, causing negative equity. Yet homeowners and businesses still have to pay off their debts. The national income accounts classify this pay-down as 'saving', although the revenue is not available to the debtors doing the 'saving' by 'deleveraging'.

The moral is that using homes as what Alan Greenspan referred to as 'piggy banks' to take out home-equity loans was not really like drawing down a bank account at all. When a bank account is drawn down there is less money available, but no residual obligation to pay. New income can be spent at the discretion of its recipient. But borrowing against a home implies an obligation to set aside future income to pay the banker – and hence a loss of future discretionary spending.

3. Towards a model of financialized economies

Creating a more realistic model of today's financialized economies to trace this phenomenon requires a breakdown of the national income and product accounts (NIPA) to see the economy as a set of distinct sectors interacting with each other. These accounts juxtapose the private and public sectors as far as current spending, saving and taxation is concerned. But the implication is that government budget deficits inflate the private-sector economy as a whole. However, a budget deficit that takes the form of transfer payments to banks, as in the case of the post-September 2008 bank bailout, the Federal Reserve's $2 trillion in cash-for-trash financial swaps and the $700 billion QE2 credit creation by the Federal Reserve to lend to banks at 0.25% interest in 2011, has a different effect from deficits that reflect social spending programs, Social Security and Medicare, public infrastructure investment or the purchase of other goods and services. The effect of transfer payments to the financial sector – as well as the $5.3 trillion increase in US Treasury debt from taking Fannie Mae and Freddie Mac onto the public balance sheet – is to support asset prices (above all those of the banking system), not inflate commodity prices and wages. Similarly, the 2009 'quantitative easing' policy in Britain confused loans used in the real economy (which were stagnating or falling throughout the experiment) with boosting bank balances with the Bank of England which quadrupled over 2009 (Graph 3). Bezemer and Gardiner (2010) show that neither bank loans nor spending nor GDP increased noticeably during or after the exercise, but there was a curious stock market rally during 2009. A London Stock Exchange press release on 29 December 2009 reported that "a record £82.5 billion was raised through new and further issues of equity on the London Stock Exchange during the course of 2009… despite difficult market conditions". Finance is not the economy.

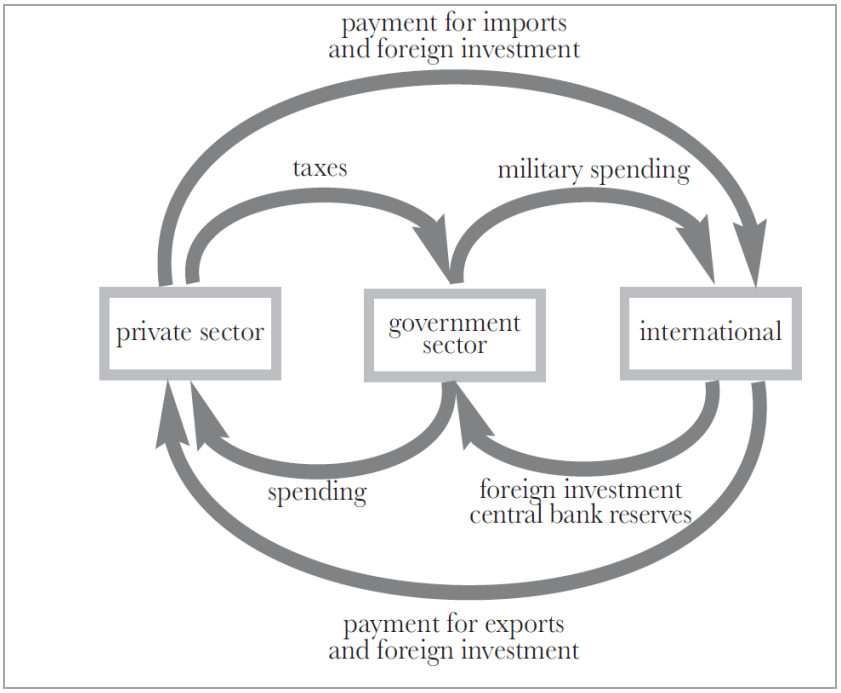
Graph 2: 'Quantitative Easing' in Britain increased bank reserves (right hand axis), but not lending to the real sector (left-hand axis) (billion Pound Sterling)



Source: Bezemer (2012) based on Bank of England data and author's calculations

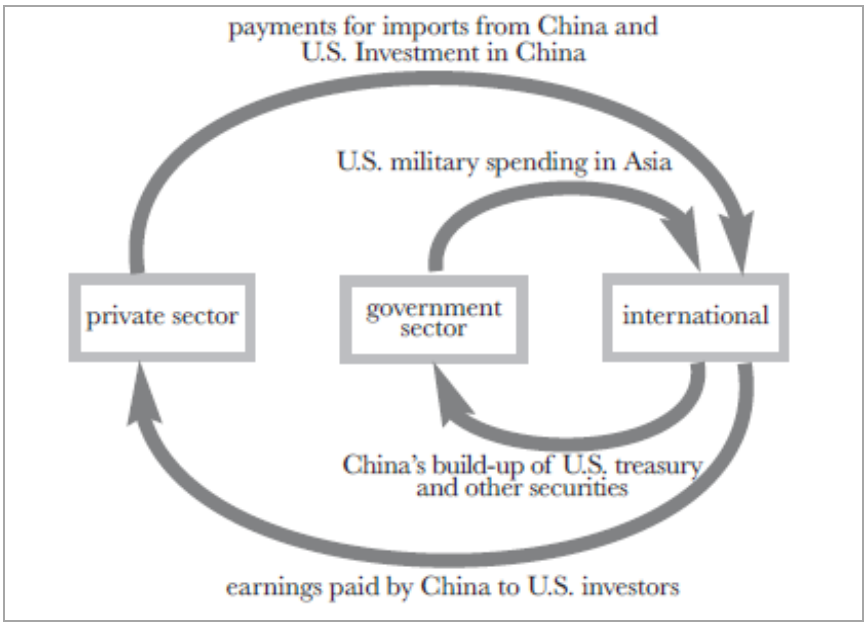
Most models treat the international sector either as a 'leakage' (as Keynes termed foreign trade and capital flows) or as a balancing item in the private/public sector surplus or shortfall (as in the Levy Institute model – see Zezza 2009, pp. 289-310 for an analytical description). But the international sector involves not only export and import trade and other current account items (emigrants' remittances, and above all, military spending) but also foreign investment and income – and foreign central bank reserves held in US Treasury and other securities, that is, loans to the US Government. Capital flows have swollen enormously since the turn of the millennium, and they have increasingly been matched by outflows of investments into dollar-denominated assets held both by private citizens and their governments. This was facilitated by new investment vehicles such as Sovereign Wealth Funds (SWFs). UNCTAD (2011, p. 119) reports 25 newly established SWFs since 2000 only. Thanks to capital inflows, the capital account is now moving independently from the current account. It is not as if the buildup of international savings requires current account surpluses. Even developing countries with current account *deficits* had accumulated foreign reserves as well as private investments in enormous quantities at the eve of the crisis, as Obstfeld (2009) reports. At the heart of this is the US economy and its financial markets. For instance, US consumers and businesses ran a trade deficit, and banks used the entire $700 billion QE2 supply of Fed credit for foreign currency arbitrage and other international speculation, not for lending to the domestic US economy. But the US Treasury received an inflow from foreign central banks building up their dollar reserves by buying Treasury securities and other US financial securities.

Figure 1: Private sector, government sector, international sector



This model can be used to trace US transactions with China. The economy runs a trade deficit with China, and also a private-sector investment outflow to China. There is some return of earnings from these investments to US companies. But on balance, there is a dollar outflow to China – which also receives dollars from its exports to third countries. China's central bank has recycled most of these dollar receipts to the US Treasury (and earlier, into Fannie Mae bonds and kindred investments), but was not permitted to buy US companies such as Unocal's refinery operations.

Fig. 1a: US transactions with China, private and government sectors



This public/private/international model may be made more realistic by treating the financial, insurance and real estate (FIRE) sector as distinct from the underlying production and consumption economy, as motivated in Graph 1.

4. The FIRE sector, rents, and the Progressive response

The FIRE sector deals with the economy's balance sheet of assets and debts, real estate, stocks and bonds, mortgages and other bank loans – and the payment of interest, money management commissions and other fees to the financial sector, as well as insurance payments and also rental payments for housing. The FIRE sector is today's form that the *rentier* class takes. *Rentiers* are those who benefit from control over assets that the economy needs to function, and who, therefore, grow disproportionately rich as the economy develops. These proceeds are rents – revenues from ownership "without working, risking, or economizing", as John Stuart Mill (1848) wrote of the landlords of his day, explaining that "they grow richer, as it were in their sleep". Classical economics from Adam Smith onwards analysed rents, its effects, and policies towards rents, but the very concept is lost on today's economics.

Just as landlords were the archetypal *rentiers* of their agricultural societies, so investors, financiers and bankers are in the largest *rentier* sector of today's financialized economies: finance controls the economy's engine of growth, which is credit in all its forms. Economies obviously need banking services, insurance services, and real estate development and so, of course, not all of finance is "without working, risking, or economizing". The problem today remains what it was in the 13th century: how to isolate what is socially necessary for 'retail' banking – processing payments by checks and credit cards, deciding how to relend savings and new credit under normal (non-speculative) conditions – from extortionate charges such as 29% interest on credit cards, penalty fees and other charges in excess of what is socially necessary cost-value.

In principle, all monopolies should be included in this *rentier* sector, as they represent a special privilege (control over markets, especially for necessities) whose return in the form of prices and income in excess of necessary costs of production is a form of economic rent, that is, a transfer payment rather than 'earned' income. But statistically there is no practical way to isolate monopoly rent in the NIPA, as this would include a large part of the information technology sector, pharmaceuticals, and much 'industry'. The ideal conceptual framework for statistics would be to separate economic rent from underlying cost value.

Classical political economists from the Physiocrats through Adam Smith, John Stuart Mill and their Progressive Era followers were reformers in the sense that they treated the *rentier* sectors as extracting transfer payments rather than earning a return for producing actual output ('services'). Their labor theory of value found its counterpart in the 'economic rent theory of prices' to distinguish the necessary costs of production and doing business (reduced ultimately to the value of labor) from 'unearned income' consisting mainly of land rent, monopoly rent, and financial interest and fees. The various categories of *rentier* income were depicted as the 'hollow' element of prices. Land rent, natural resource rent, monopoly rent and returns to privilege (including financial interest and fees) had no counterpart in necessary costs of production. They were historical and institutional products of privileges handed down largely from the medieval conquests that created Europe's landed aristocracy and banking practice that developed largely by insider dealing, legitimized by lending to kings to finance war debts in an epoch when money and credit were the sinews of war. So banking as well as military rivalries for land essentially involved the foreign sector. Mill (1848) asked "What claim have they, on the general principle of social justice, to this accession of riches? In what would they have been wronged if society had, from the beginning, reserved the right of taxing the spontaneous increase of rent, to the highest amount required by financial exigencies?".

The political aim of classical analysis, then, was to minimize the economy's cost structure by freeing industrial capitalism from these carry-overs from feudalism. The reformers' guiding idea was to minimize the role of *rentier* income (economic rent) by public investment, tax policy and regulation. We consider these in turn.

1. Direct public investment in basic infrastructure, including education, transportation systems, communication systems and other enterprises that were long kept in the public domain or publicly regulated from the late 19th century onward. The premier example of this is the French *Crédit Mobilier* bank founded by followers of the Count de Saint-Simon (1760–1825), who inspired key Classical economists including Karl Marx and John Stuart Mill. The *Crédit Mobilier* bank, founded in 1852, was named in contrast to the common mortgage bank (*Sociétés du Crédit Foncier*) or land banks, which lent money on the security of immovable property. The *Crédit Mobilier* aimed to loan to the owners of movable property and so to promote industrial enterprise, mining and the construction of railways and other infrastructure. Today, the bulk of bank lending is again to real estate and other property already in existence, not for the creation of new productive capacity and innovation of production processes. We need *Crédit Mobilier* –type financial institutions.

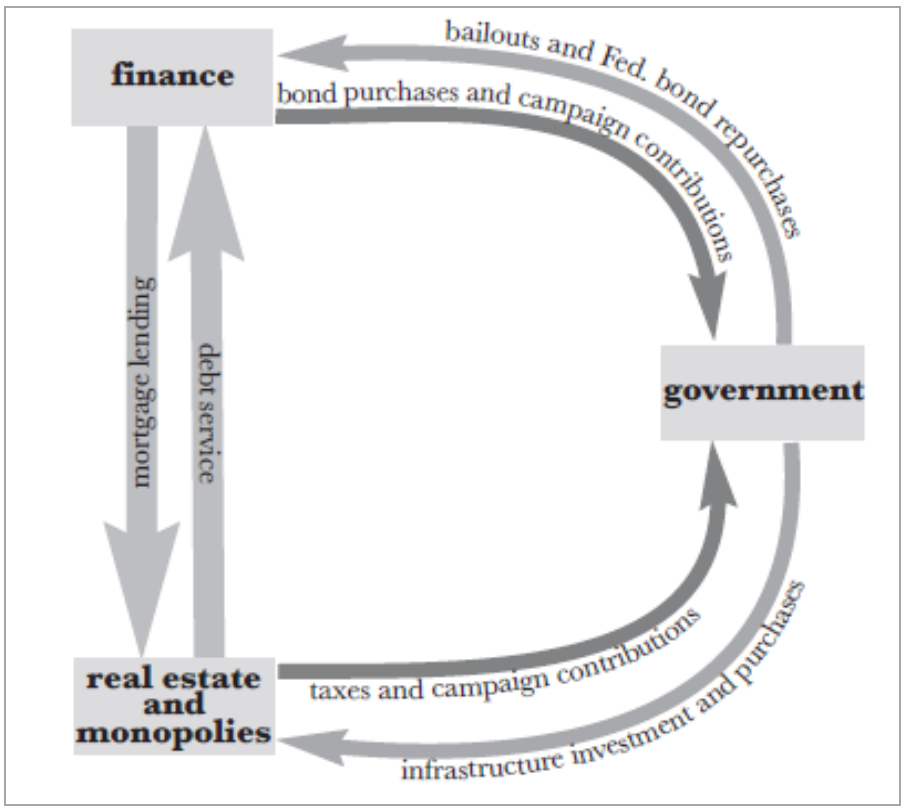
2. Tax policy (taxing land and natural resources). Here the foremost Classical-era name is Henry George (1839-1897). In his Progress and Poverty (1879) he observed that much of the wealth created by social and technological advances is captured by landowners and other monopolists via economic rents. This concentration of 'unearned' income – which strictly speaking is not income, though it is a revenue stream – in the hands of the few is, according to George, the cause of increasing poverty precisely in those areas which are more developed. The plight of the poor in the mature economy of New York struck him as much worse than the living standards of the poor in his native (then underdeveloped) California. Today, the impoverishing rent flows are (a) in payment for inflated assets prices and (b) in servicing loans against those assets. A large part of the economy's surplus flows to the property and finance sectors in payment of loans, interest and fees for the use of land and housing. And today just as in George's days, inequality has increased strongly as bank loans have been reoriented away from supporting the real sector and towards FIRE sector loans. This drives up asset prices and thus mortgages, increasing the drain from the real economy while enriching assets owners.

3. Regulatory policy to keep the prices charged by natural monopolies such a railroads, power and gas companies in line with actual production costs plus normal profit. The classical example of this is the US Sherman Antitrust Act (1890), enacted in response to the development of business conglomerates or 'trusts' in the last third of the 19th century, which often stifled competition and manipulated prices. Today again the global financial market place is dominated by a few giants; and in most economies three of four banks control 80% or more of domestic markets. The result is just the behavior that progressive Americans deplored in 19th century business, now played out in finance: artificial price increases for bank services and banker's remuneration, far above the level necessary to cover costs with a reasonable profit left; block buying and price fixing in the trading of financial products; and even fraud and intimidation of competitors. And after the crisis, small banks have been bankrupted in their hundreds while the large banks have been bailed out. Re-introduction of financial anti-trust policies will not be the end (in the first 10 years of existence of the Sherman Antitrust Act, many more actions were brought against unions than against big business). But it will be a start.

5. How the FIRE sector operates

The financial sector has become the leading *rentier* sector. Its 'product' is debt claims *on* the 'real' economy, underwriting, and money management on a fee basis. For this it receives interest and dividends from real estate and business borrowers, and from consumers. Over time, a real estate buyer typically pays more in interest to their mortgage lenders than the original purchase price paid to the property seller.

Fig. 3: Interaction between the FIRE and government sectors

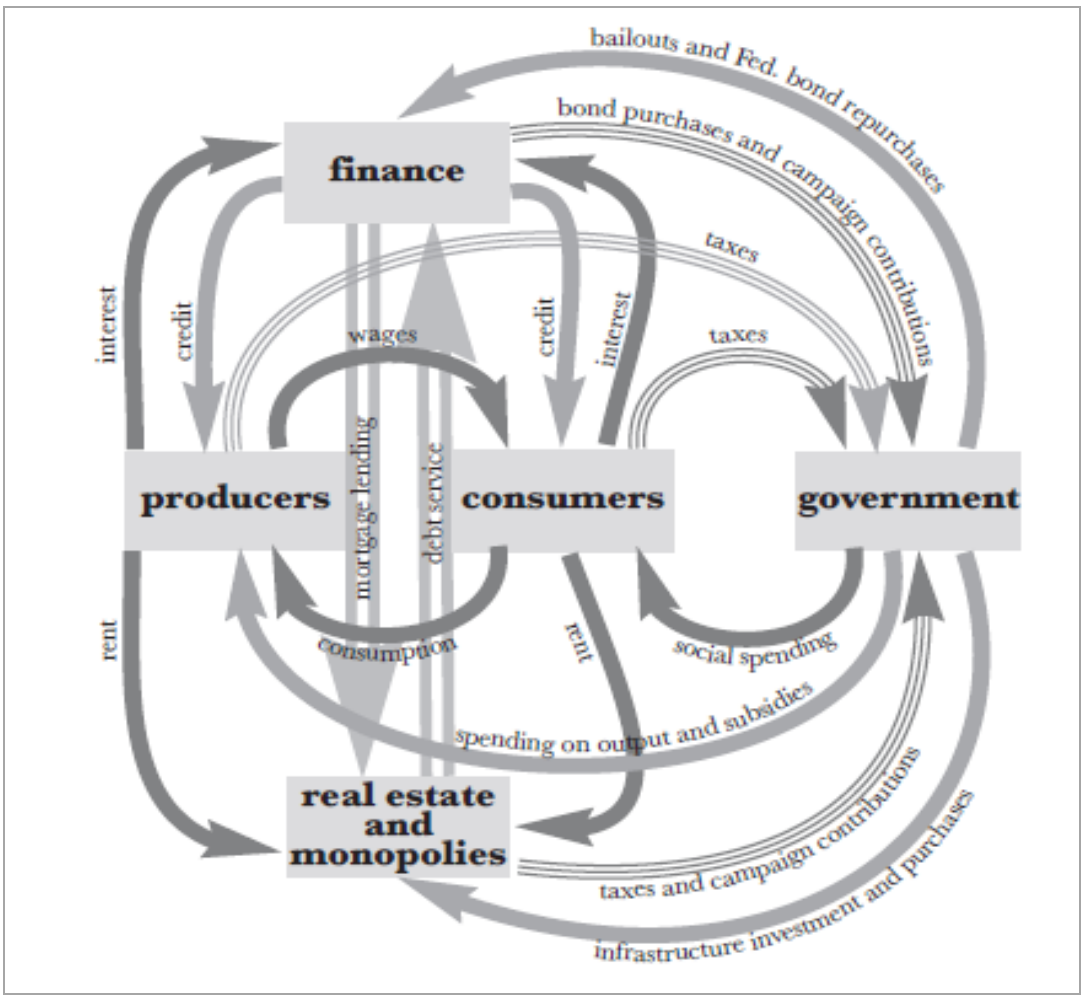


In its interactions with the government, the financial sector buys bonds (and also makes campaign contributions). The Federal Reserve pumps money into the banking system by purchasing bonds and, when the system breaks down, makes enormous bailout payments to cover the bad debts run up by banks and other institutions to mortgage borrowers, businesses and consumers. The government also enhances the real estate sector by providing transportation and other basic infrastructure that enhances the site value of property along the routes. Finally, the government acts as direct purchaser of monopoly services from health insurance providers, pharmaceutical companies and other monopolies. In the other direction, the US Government receives a modicum of taxes from real estate (mainly at the local level for property taxes), not much income tax but some capital gains tax in good years.

Hardly by surprise, the financial sector prefers to make itself invisible – not only to the tax collector and government regulators, but to voters. In fact, tax polices favor unearned income. The ordinary income tax rate in the US is twice the level of taxes on capital gains: for the 15 % income tax rate brackets, 5-year capital gains taxes are 8%; and for the 39.6% bracket, they are 18% (Kiplinger's, 2009). And yet, since capital gains are not income, higher capital gains tax is opposed on the grounds that this tax falls on (non-capital gains) incomes, which would therefore be unfairly taxed. Minarik (1992, p. 16) writes against capital gains taxation asserting that "the burden of proof should rest on those who would violate the basic principle of equal tax rates on incomes from whatever source". This conflates revenue streams with income.

Successful attempts to break out the *rentier* sector from the rest of the economy – and hence, balance sheet and debt transactions from the purchase of goods and services – have helped soften criticism of shifting the tax burden off land and monopoly rent, and off finance. Yet Epstein and Crotty report that "financial sector total financial assets grew from about a third of total US economy financial assets in the post-World War II decades to 45 percent of total financial assets. Their value was approximately equal to the US GDP in the early 1950s, whereas now it amounts to 4.5 times of the US GDP. Financial sector profit has grown from about 10 percent in the 1950-60s to 40 percent of total domestic profits in the early 2000s".

Fig. 4: Overall model of the FIRE sector: producers, consumers, government, world



The distinction between *rentier* and 'earned' income was not incorporated into the NIPA. It is as if all income was earned by playing a productive role, and in which money (and hence, credit and debt) were 'neutral', only a 'veil', not as affecting the distribution of income and wealth. Credit was spent only on goods and services, not on assets. And the financial sector's loans always took the form of productive credit, enabling businesses to pay back the loans out of future earnings while consumers paid out of rising future incomes. This is still the representation found in most textbooks today. For instance, Mishkin (2012: pp. 1 and 24) explains that "in our economy, nonbank finance also plays an important role in channeling funds from lender-savers to borrower-spenders… Finance companies raise funds by issuing commercial paper and stocks and bonds and use the proceeds to make loans that are particularly suited to consumer and business needs".

There thus was no explanation of how a credit bubble could inflate real estate prices and then collapse into a negative equity disaster. Finance seemed only to create wealth, not impoverish the underlying economy. Amazingly, this was claimed even for the exotic products whose proliferation preceded the 2008 crash. As late as 2006 academics asserted that "[f]inancial risks, particularly credit risks, are no longer borne by banks. They are increasingly moved off balance sheets. Assets are converted into tradable securities, which in turn eliminates credit risks. Derivative transactions like interest rate swaps also serve the same purpose [of eliminating credit risks, MH & DB]" (Das 2006).

Nor was there any way for mainstream models to distinguish government transfer payments to the financial sector (eg. the $13 trillion in post-2008 financial bailouts in the United States) from Keynesian-style deficit spending. Such transfer payments did not 'jumpstart' the economy. They turned a politically well-connected financial elite into new vested interests. All this is completely missed in conventional

macroeconomics, which cannot come to grips with the role of the financial sector in the economy. Eminent economists have described training in today's macro models as a useless, even socially wasteful activity (Buiter 2009; also Krugman 2009; Solow 2010).

One can understand why the financial sector has had so little interest in tracing the effect of rising money and credit on diverting income from the circular flow between producers and consumers, diverting business revenue from new capital formation, and stripping industrial assets and natural resources. Most model builders isolate these long-term structural, environmental and demographic feedbacks as 'externalities'. But they are part and parcel of reality. So one is tempted to say that the financial element of economic models is too important to be left to bankers and the think tanks they sponsor.

6. Effects on the environment, demography and the economy

Just as debt deflation diverts income to pay interest and other financial charges – often at the cost of paying so much corporate cash flow that assets must be sold off to pay creditors – so the phenomenon leads to stripping the natural environment. The so-called 'debt-resource-hypothesis' suggests that high indebtedness leads to increased natural resource exploitation as well as more unsustainable patterns of resource use (Neumayer 2012, pp. 127-141). This is what occurs, for instance, when the IMF and World Bank act on behalf of global banks to demand that Brazil pay its foreign debt by privatizing its Amazon forest so that loggers can earn enough foreign exchange to pay foreign bankers on the nation's foreign-currency debt. The analogy is for absentee landlords who pay their mortgages by not repairing their property but letting it deteriorate. In all these cases the effect of debt deflation extracting interest is not only on spending – and hence on current prices – but on the economy's long-term ability to produce, by eating into natural resources and the environment as well as society's manmade capital stock.

Demographically, the effect of debt deflation is emigration and other negative effects. For example, after Latvian property prices soared as Swedish bank branches fueled the real estate bubble, living standards plunged. Families had to take on a lifetime of debt in order to gain the housing that was bequeathed to the country debt-free when the Soviet Union broke up in 1991. When Latvia's government imposed neoliberal austerity policies in 2009-10, wage levels plunged by 30 percent in the public sector, and private-sector wages followed the decline (Sommers et al 2010). Emigration and capital flight accelerated: the *Economist* (2010) reported that an estimated 30,000 Latvians were leaving every year, on a 2.2m population. In debt-strapped Iceland, the census reported in 2011 that 8% of the population had emigrated (mainly to Norway).

In as much as investors today have come to aim more at 'total returns' (net income + capital gains) rather than simply income by itself, a realistic model should integrate capital gains and investment into the current production-consumption model. Producers not only pay wages and buy capital goods as in 'current economy' models; they also use their cash flow (and even borrow) to buy other companies, as well as their own stock. When they make acquisitions on credit, the resulting debt leveraging finds its counterpart in interest payments that absorb a rising share of corporate cash flow.

This has an effect on the government's fiscal position, because interest is a tax-deductible expense. By displacing taxable profits, the business revenue that hitherto was paid out as income taxes is now used to pay interest to creditors. The result in the early 1980s when debt-leveraged buyouts really gained momentum was that financial investors were able to obtain twice as high a return (at a 50% corporate income tax rate) by debt financing as they could get by equity financing. This tax incentive for debt leveraging rather than equity investment is the reverse of what Saint-Simon and his followers urged in the 19th century to become the wave of the future.

7. Conclusion

Only a portion of FIRE sector cash flow is spent on goods and services. The great bulk is recycled into the purchase of financial securities and other assets, or lent out as yet more interest-bearing debt – on easier and easier credit terms as the repertory of bankable direct investments is exhausted. So the pressing task today is

to trace how directing most credit into the asset markets affects asset prices much more than commodity prices. Loan standards deteriorate as debt/equity ratios increase and creditors 'race to the bottom' to find borrowers in markets further distanced from the 'real' economy. This increasingly unproductive character of credit explains why wealth is being concentrated in the hands of the population's wealthiest 10 percent. It is the dysfunctional result of economic parasitism.

Keynes recognized a 'leakage' in the form of saving (specifically, hoarding). But at the time he wrote in the midst of the Great Depression there was little motivation to focus on debt service, or on the distinction between direct capital investment (tangible capital formation) and financial securities speculation or real estate speculation (which had all but dried up as asset markets were shrinking to reflect the economy's shrinking). Saving took the form of non-spending, not of paying down debt. There was little lending under depression conditions.

Today's post-bubble attempts to incorporate balance-sheet analysis into NIPA statistics on current activity are too crude. Stock averages do not give an adequate quantitative measure distinguishing the flow of funds into land and capital improvements or industrial capital formation in contrast to speculation in financial securities. So monetary analysis needs to be reformulated along with a better structural breakdown of NIPA to distinguish between money and credit spent on goods and services from that spent on financial assets and debt service.

References

Buiter, W (2009), The unfortunate uselessness of most 'state of the art' academic monetary economics. *Financial Times,* March 3rd.

Cecchetti, S, M Mohanty and F Zampolli (2011), *The real effects of debt*. BIS Working Papers 52, Basel: Bank for International Settlements.

Das, D (2006), Globalization in the World of Finance: An Analytical History. *Global Economy Journal* 6 (1) article 2.

Economist (2010), Far from home: Migration and Latvia. *The Economist*, October 1st.

Epstein G and J Crotty (2012), *How big is too big? On the social efficiency of the financial sector in the United States*. Paper presented at the INET conference, Berlin. Available at www.ineteconomics.org/

Epstein, G (ed.) (2005), *Financialization and the world economy*. Aldershot: Edward Elgar.

George, H (1879), *Progress and Poverty: An Inquiry into the Cause of Industrial Depressions and of Increase of Want with Increase of Wealth, The Remedy*. Garden City, NY: Doubleday, Page, & Co.

Hudson, M (2004), 'The archeology of money: debt versus barter theories of money's origin. In: R Wray (2004) (ed.) *Credit and state theories of money: the contributions of Michael Innes*. Cheltenham UK: Edward Elgar.

Keen, S (2009), Household debt: The final stage in an artificially extended Ponzi bubble. *Australian Economic Review* 42, pp. 347-357.

Keen, S (2011), *Debunking economics: The naked emperor dethroned?* London: Zed Books.

Kiplinger (2009), Tax law changes for 2008–2017. *Kiplinger's*.

Krugman, P (2009), How did economists get it so wrong?, *New York Times*, September 6th.

Mill, J S (1848), Principles of political economy with some of their applications to social philosophy. Available on http://oll.libertyfund.org/index.php?option=com_staticxt&staticfile=show.php%3Ftitle=101&Itemid=27

Minarik, J (1992), 'Capital Gains Taxation, Growth, And Fairness', *Contemporary Economic Policy*, Volume 10, Issue 3, pp. 16-25, July 1992.

Mishkin F S (2012), *The economics of money, banking, and financial markets*. New York: Pearson.

Neumayer, E (2012), Does high indebtedness increase natural resource exploitation? *Environment and Development Economics,* 10, pp. 127-141.

Obstfeld, M (2009), *International finance and growth in developing countries: What have we learned?* IMF Staff Papers Volume 56, Number 1.

Schumpeter, J (1934), *The Theory of Economic Development.* New York: Harvard University Press.

Solow, R (2010), *Building a science of economics for the real world*. Prepared Statement of Robert Solow, Professor Emeritus, MIT, to the House Committee on Science and Technology, Subcommittee on Investigations and Oversight: 20 July.

Sommers, J D, D Bezemer and M Hudson (2010), The Human Costs of Financial Instability in Latvia. In: D Tavasci and J Toporowski (eds.) *Minsky, Financial Development and Crisis*. Basingstoke, Hampshire, UK: Palgrave MacMillan.

Stockhammer E (2004), 'Financialization and the slowdown of accumulation', *Cambridge Journal of Economics*, 28 (5), pp. 719-741.

Tovar, C (2008), *DSGE models and central banks*. BIS Monetary and Economic Department Working Paper No.258. Basel: Bank for International Settlements.

UNCTAD (2011), *Trade and development report 2012*. Geneva: United Nations Conference on Trade and Development.

Zezza, G (2009), 'Fiscal policy and the economics of financial balances', *European Journal of Economics and Economic Policies*, 6 (2), pp. 289-310.

Microfinance and the Illusion of Development: From Hubris to Nemesis in Thirty Years[1]

Milford Bateman
Department of Economics and Tourism, Juraj Dobrila Pula University, Pula, Croatia

Ha-Joon Chang
Faculty of Economics, University of Cambridge, Cambridge, UK

Abstract

The contemporary model of microfinance has its roots in a small local experiment in Bangladesh in the early 1970s undertaken by Dr Muhammad Yunus, the US-educated Bangladeshi economist and future 2006 Nobel Peace Prize co-recipient. Yunus's idea of supporting tiny informal microenterprises and self-employment as the solution to widespread poverty rapidly caught on, and by the 1990s the concept of microfinance was the international development community's highest-profile and most generously funded poverty reduction policy. Neoclassical economic theorists and neoliberal policy-makers both fully concurred with the microfinance model's celebration of self-help and the individual entrepreneur, and its implicit antipathy to any form of state intervention. The immense feel-good appeal of microfinance is essentially based on the widespread assumption that simply 'reaching the poor' with a tiny microcredit will automatically establish a sustainable economic and social development trajectory, a trajectory animated by the poor themselves acting as micro-entrepreneurs getting involving in tiny income-generating activities. We reject this view, however. We argue that while the microfinance model may well generate some narrow positive short run outcomes for a few lucky individuals, these positive outcomes are very limited in number and anyway swamped by much wider longer run downsides and opportunity costs at the community and national level. Our view is that microfinance actually constitutes a powerful institutional and political barrier to sustainable economic and social development, and so also to poverty reduction. Finally, we suggest that continued support for microfinance in international development policy circles cannot be divorced from its supreme serviceability to the neoliberal/globalisation agenda.

Key words: microfinance, microcredit, neoliberalism, impact, poverty, development.

1. Introduction

As originally conceived, microfinance (more accurately, microcredit[2]) involves the provision of a small loan, a microloan, that is used by a poor individual to support a tiny income-generating activity, thereby to generate an income sufficient to effect an exit from poverty. Since the early 1980s, the microfinance-supported proliferation of informal microenterprises and self-employment has been very widely promoted as the solution to poverty and under-development. By the 1990s, microfinance was the international development community's highest-profile and most generously funded poverty reduction policy (eg.

[1] The authors declare no potential conflicts of interest with respect to the research, authorship, and/or publication of this article. The authors received no financial support for the research, authorship, and/or publication of this article. This article is a substantially revised and updated version of a 2009 mimeo (Bateman and Chang 2009). We would like to thank all those who commented on our article on the World Economics Journal website, as well as those who in the past couple of years offered private comments on our original mimeo. The usual disclaimer applies.

[2] The term 'microfinance' is the most commonly used term today, so we use this term. Microfinance is actually the generic term covering all varieties of microfinancial interventions, such as microcredit, microsavings, microinsurance, micro-franchising, and so on.

Balkenhol 2007, p. 213). The expectation began to form that an historically unparalleled poverty reduction and 'bottom-up' economic and social development episode was in the making.

This article challenges the view that the microfinance model has a positive association with sustainable poverty reduction and local economic and social development. On the contrary, we find the microfinance model is most likely to lock people and communities in a 'poverty trap'. Moreover, in a growing number of 'microfinance-saturated' countries, regions and localities, the outcome of the microfinance model has been nothing short of catastrophic. Nonetheless, despite the growing evidence that it has failed in its original mission to reduce poverty, a fact that even long-standing proponents now concede, the microfinance model still largely retains its reputation and popularity within the international development community. To help explain why there is such a widespread misunderstanding of microfinance, we go on to argue that the microfinance model remains attractive to the international development community because of its huge political serviceability to the neoliberal worldview that centrally locates the main driver of economic development to be individual entrepreneurship.

The article is structured as follows. Section 2 briefly charts the rapid rise of the microfinance model after 1980 and its recent tribulations since mid-2007 that have contributed to people waking up to a completely new understanding as to its long-term impact. Section 3 then summarises the key areas where we feel the theory behind microfinance as opposed to its mere (possibly inefficient) execution, has proved to be most problematic. Section 4 explores the intimate links between the neoliberal globalisation project and the microfinance project. A brief conclusion summarises the argument.

2. Background

Broadly defined, microfinance has a long history and encompasses a diverse range of institutional formats, ranging from individual money-lenders through to more formal institutions, such as village banks, credit unions, friendly societies, financial cooperatives, building societies, state-owned banks for SMEs (Small and Medium-sized Enterprises), social venture capital funds, and specialised SME funds. The majority of these financial initiatives, especially those from the 18th and 19th century onwards, arose from a desire to transform the lives of the poor and the new industrial working classes**,** as they struggled to cope with the growing perils and exploitation associated with the rise of industrial capitalism. Noteworthy examples include the many Friendly Societies that were an outgrowth of the rapidly growing trade union movement (Thompson 1963) and the financial institutions established by the burgeoning Europe-wide cooperative movement that began in England and Scotland in the early 1800s (Birchall 1997). In short, the objective was not so much to help the poor to passively accept their poverty and exploitation under elite-dominated economic systems, but to challenge the emerging capitalist model and to genuinely empower the poor by enlarging the space of economic and social activity under their effective (and proto-democratic) ownership and control.

The recent explosion of interest in microfinance, and the foundation of a powerful 'microfinance movement', represents something quite different, however. At the forefront of this new microfinance movement was Dr Muhammad Yunus, the Bangladeshi-born and US-educated economist. Following a number of experiments in the mid-1970s with the provision of microcredit in and around the village of Jobra near Chittagong in Bangladesh, Yunus began to argue that the mere availability of a microloan would greatly benefit the poor everywhere, and especially women in poverty. The poor simply had to establish and operate an informal microenterprise in their local community and they would be well on the way to escaping their poverty. Yunus took to claiming that microfinance would "eradicate poverty in a generation" and he confidently predicted that very soon our children would have to go to a "poverty museum" to find out what all the fuss was about (eg. Yunus 1997).

The international donor community very much liked what Yunus was saying, and so agreed to underwrite his bold ideas for promoting self-help and individual entrepreneurship among Bangladesh's poor. This goal was to be achieved through a dedicated institution – the Grameen Bank. The Grameen

Bank was formed in 1983 and, largely based on Yunus's constant declarations that it was an enormous success, it was pretty soon being copied all over Bangladesh and then all over the world. Pretty soon, too, Yunus began to attract a dedicated band of followers, especially in the US, who all agreed (though often without any real analysis or evidence – see below) that microfinance would make massive inroads into global poverty. An efficient, private sector-led and market-driven model of poverty reduction and 'bottom-up' economic and social development appeared to have been found.

However, although neoliberal policymakers greatly appreciated the emphasis upon self-help and individual entrepreneurship, and thus also its implicit support for free market capitalism, they still had major reservations about the financing of the Grameen Bank microfinance model. This was because it soon became clear that Grameen Bank's operations, as with most microfinance institutions (hereafter MFIs) that had sprung up around the world at that time, actually depended upon a continuous inflow of subsidized capital. This funding was mostly provided by an MFI's own government and/or by the international development community. The neoliberal policymaking community began to feel increasingly awkward about using subsidies to keep the supposedly non-state, market-driven microfinance sector going. Spearheaded by the main Washington DC institutions – USAID and the World Bank – decisive action was therefore initiated to phase out the original Grameen Bank model of subsidised microfinance. The long-term solution to the 'problem' of subsidies in the microfinance sector was found in the idea to reconstitute microfinance as a privately-owned, profit-driven business model. Key advocates of commercialisation, notably Maria Otero (Otero and Rhyne 1994) and Marguerite Robinson (Robinson 2001) saw this new commercialised model, and the likely increase in the supply of microfinance, as being capable of generating huge benefits for the poor.

By the early 1990s a thoroughly 'neoliberalized' for-profit model of microfinance was being ushered in as the 'best practice' replacement for the original subsidized Grameen Bank model. This 'new wave' model (formally known as the 'financial systems' approach – Robinson 2001) quickly became the dominant template for microfinance programs. By the turn of the new millennium, the 'new wave' microfinance model was at the peak of its power and influence. Even the iconic Grameen Bank felt it had no other option but to finally agree to convert over to 'new wave' respectability, which it did in 2002 with the 'Grameen II' project. The UN declared 2005 to be the International Year of Microcredit. Numerous prestigious awards were also forthcoming for those involved in microfinance, famously including the 2006 Nobel Peace Prize jointly awarded to Muhammad Yunus and the Grameen Bank. And thanks to all these activities, the list of 'microfinance-saturated' countries (defined in terms of borrowers per capita) soon began to comprise not just the original pioneer Bangladesh, but also Bolivia, Bosnia, Mongolia, Cambodia, Nicaragua, Sri Lanka, Peru, Colombia, Mexico and India (Bateman 2011b, p. 4, Table 1.1). It seemed obvious to all involved that the world was undergoing an historically unparalleled episode of poverty reduction. But then the carefully constructed edifice of modern microfinance began to crumble.

Beginning in 2007, and in a most rapid, dramatic and unexpected fashion, hubris quickly turned to nemesis. It is widely recognised that the first spark was provided by the 2007 Initial Public offering (IPO) of the Mexican MFI, Compartamos. Rather than revealing commendable levels of poverty reduction among poor Mexican individuals – there still remains no evidence for this whatsoever – the IPO process revealed instead the Wall Street-style levels of private enrichment enjoyed by Compartamos's senior managers. These vast rewards were effectively made possible by quietly charging 195% interest rates on the microloans taken out by their poor – mainly female – clients[3]. The Compartamos IPO led to much public outrage against Compartamos and its senior staff, and then a tidal wave of criticism of the commercialised microfinance model in general. Even long-standing supporters of microfinance began to openly express their concerns at the way the microfinance concept was being destroyed in the hands of neoliberals and hard-nosed investors (notably Malcolm Harper – Harper 2011; Klas, 2011; Sinclair, 2012).

[3] See http://blogs.cgdev.org/open_book/2011/01/compartamos-and-the-meaning-of-interest-rates.php

Very soon the narrow criticism of the Compartamos IPO and commercialised microfinance was joined by a much more comprehensive critique of microfinance as an economic development model *per se* (Dichter and Harper 2007; Bateman and Chang 2009; Bateman 2010a, 2011a). Other researchers using new and supposedly more accurate Randomised Control Trial (RCT) methodologies found little to no impact arising from individual microfinance programs (Banerjee et al 2009: Karlan and Zinman 2009). Roodman and Morduch (2009) and Duvendack and Palmer-Jones (2011) mounted a serious challenge to the single most important study routinely cited as the best evidence that individual microfinance programs had a strong poverty reduction impact – a study undertaken in the 1990s by then World Bank economists Mark Pitt and Shahidur Khandker (Pitt and Khandker 1998). Re-examining the original dataset used by Pitt and Khandker, both sets of authors located serious mistakes in the original analysis and, as a result, declared that Pitt and Khandker's work did not confirm a positive impact from the microfinance programs studied[4].

Adding considerable impetus to the growing critique of the microfinance model were a number of hugely destructive sub-prime-style 'microfinance meltdowns' taking place around the globe. The first 'microfinance meltdown' had actually taken place in Bolivia in 1999-2000, but at the time microfinance supporters described it as a 'one-off' aberration caused by factors supposedly unrelated to the core of the microfinance model**,** such as unfair competition from a large MFI coming to Bolivia from Chile (Rhyne 2001). However, starting in 2008, a new round of even more destructive 'microfinance meltdowns' began in Morocco, Nicaragua and Pakistan, marked out by huge client over-indebtedness, rapidly growing client defaults, massive client withdrawal, and the key MFIs plunging into loss or forced to close or merge. These episodes were then followed in 2009 by the dramatic near-collapse of the hugely over-blown microfinance sector in Bosnia (Bateman, Sinković and Škare 2012).

By all accounts, the most devastating 'microfinance meltdown' to date started in late 2010 in the Indian state of Andhra Pradesh (Arunachalam 2011). With the poor increasingly taking out more and more microloans in order to repay earlier microloans that they had all too easily accessed, it was clear that the microfinance model in Andhra Pradesh had degenerated into nothing more than a vast Ponzi-like survival strategy for a very large number of the poor[5]. In late 2010, thanks to a deluge of personal over-indebtedness, defaults and MFI losses, Andhra Pradesh's microfinance industry effectively collapsed[6]. Further over-supply problems are also clearly emerging elsewhere, notably in Mexico, Lebanon, Peru, Azerbaijan and Kyrgyzstan[7].

In 2011 came a further quite devastating blow to the microfinance industry. This was a UK government-funded systematic review of virtually all of the impact evaluation evidence long said to confirm that microfinance has had a positive impact on the well-being of the poor (Duvendack et al 2011). The review found that the previous impact studies were almost all seriously biased, incomplete or else very poorly designed to the point of being quite unusable[8]. The Duvendack review reached an

[4] Notably, as Roodman and Morduch discussed in their revised paper published in 2011, Pitt and Khandker did not examine and rule out reverse causation, meaning that their reporting of a positive association between microcredit and household spending may indicate – as is the case in very many countries – that richer families simply borrow more.

[5] By late 2009 it was found that poor households in Andhra Pradesh were on average in possession of a total of 9.3 microloans, compared to between 2 to 4 microloans per poor household in the next most saturated states in India – Tamil Nadu, Orissa, Karnataka and West Bengal (see Srinivasan 2010)

[6] In mid-2010 the microfinance industry possessed a gross loan portfolio of nearly $3 billion (up from just $230 million in 2006), but it is predicted that it will almost entirely cease to exist by mid 2012. For example, with its once nearly £1 billion microloan portfolio in Andhra Pradesh almost entirely written off by the end of 2011, the largest MFI in the state – SKS - has announced it will move into new areas of operation as of early 2012, including rural insurance, rural payments and small business lending. See http://www.dnaindia.com/money/report_new-sks-head-talks-of-sea-change-in-business-model_1617016

[7] Private communications with MFI analysts: see also CGAP 2010.

[8] Most of these earlier studies were undertaken by, or contracted out by, the microfinance institutions themselves, as well as by the rapidly expanding raft of microfinance advocacy bodies. Genuine analysis of the microfinance model was overwhelmingly shunned in case it produced a negative result, an outcome that would have scuppered the chances of the external funding (from donors, governments, foundations, etc) that most MFIs and microfinance advocacy bodies desperately required. It is thus not too hard to locate the source and rationale for almost all of the massively exaggerated, and often openly false, claims relating to the power of

explosive conclusion, arguing that "[the] current enthusiasm [for microfinance] is built on [...] foundations of sand" (p. 75). Importantly (especially in the context of our comments below), the very final comment (p. 76) points to the case for microcredit having been made not so much on the basis of the *economics* (of poverty reduction and development), but to the *politics*, and the authors conclude that further research is required by political scientists in order to understand "[why] inappropriate optimism towards microfinance became so widespread".

One far-reaching result of all this bad news is that the microfinance industry has begun to drop the important claim to be facilitating poverty reduction, moving very quietly to redefine a new goal for itself in terms of facilitating the far more nebulous concept of 'financial inclusion'. However, in reality this new objective for microfinance appears to have even less substance to it than the failed poverty reduction objective it is designed to replace (Bateman 2012a).

We agree with the substance and direction of much of the growing criticism of microfinance. However, our own scepticism on this issue is not just rooted in our analysis of the faulty economic principles upon which the microfinance concept is based, as we will outline in the next section, but also in the important counterfactual that emerges from a careful examination of the economic history of the most successful national, regional and local economies. For if one looks at the advanced economies (US, Japan, Western Europe), as well as of the East Asian 'tiger' economies that burst on to the scene from the 1970s onwards (South Korea, Taiwan, Malaysia, China, Thailand and, most recently, Vietnam), one finds evidence of a successful national economic model that is almost the exact opposite of the market-driven microfinance model. As is now widely accepted (Amsden 2001, 2007; Chang 2002, 2006, 2007, 2011; Reinert 2007; Wade 1990), sustainable progress was forthcoming in all these countries largely thanks to a range of pro-active 'developmental state' interventions. In addition, a pivotal element underpinning the success achieved in many of these 'developmental state' countries also lies in what has been termed the 'local developmental state' (LDS) model – pro-active local development and growth strategies undertaken by local government level institutions (Friedman 1988; Weiss 1988; Oi 1995; Lall 1996; Bateman 2000; Thun 2006). This successful LDS model is very far removed indeed from the contemporary microfinance model**,** even though it may have some superficial similarities to it (for examples, Bateman 2010a, Chapter 7).

3. Why microfinance most often makes things worse, if not much worse

The above section has demonstrated that, after a seemingly auspicious beginning, in recent years the microfinance model has clearly run into a brick wall. In this section we identify the key factors that account for why it is that the microfinance model has had such an adverse impact at both the local community level and national economy level.

(a) The microfinance model ignores the crucial role of scale economies

By definition, *micro*finance produces *micro*enterprises – that is, enterprises and agricultural units that are very small and almost always operate below minimum efficient scale. However, it is widely accepted that for all enterprise sectors there remains an identifiable minimum efficient scale of production, and operating below this level makes it virtually impossible for any enterprise to survive and prosper in a competitive business environment.

In general, we may say that microfinance policymakers largely fail to register the crucial importance of minimum efficient scale. What matters above all, so their argument runs, is to construct a local financial system dominated by MFIs that can establish as many microenterprises as possible in the short term. Going further, microfinance supporters argue that a collection of the tiniest microenterprises

microfinance. The parallels with the adverse role that the three main ratings agencies played in creating the global financial breakdown starting in 2008 are obvious.

is actually the ideal foundation for sustainable development. As Dambisa Moyo (2009, p. 129) relates of her native Zambia, "Think of a woman selling tomatoes on a side street. …[T]his group – the real entrepreneurs, *the backbone of Zambia's economic future* – need capital just as much as the mining company" (italics added). The argument here is essentially that scale does not matter, and that many more of such tiny microenterprises will indeed provide the best possible (neoclassical textbook) foundation for sustainable development. It is an argument that has been extensively taken up by the microfinance industry as a whole: it is the *numbers* of microenterprises established that appear to matter the most, rather than their (initial) *size*. But is it an argument that holds water?

First of all, we can say that Moyo's thesis holds no water in Africa. Africa already has more micro-entrepreneurs per capita than anywhere else on earth (African Development Bank and OECD 2005), and the rapidly expanding supply of microfinance is actually increasing this number year by year. For example, the share of the informal economy in GDP in Kenya is now as much as 72%, in Zambia around 58%, while even in more industrialised South Africa informal employment as a proportion of non-agricultural employment is likely to be above 70% (Rolfe et al 2010). However, Africa effectively remains trapped in its poverty *precisely* because the increasingly microfinance-dominant financial structure in Africa is suitable only to evolve an enterprise structure overwhelmingly composed of very tiny units operating way below minimum efficient scale. For a number of reasons, a careful study of economic development history (eg. Chang 2011, pp. 157-167) provides no evidence that might lead us to think that in Africa, or indeed anywhere else, entirely avoiding to reap economies of scale in productive activity will nevertheless still provide a suitable foundation upon which sustainable economic and social development can be achieved (see also below).

The situation in India is not dissimilar. Despite its rapid and well-publicised growth in recent years, India still has many huge development and poverty-related problems. One of the most pressing development problems is the need to fill the so-called 'missing middle' that exists between, on the one hand, the small number of large internationally well-known computing and manufacturing companies and, on the other hand, the hundreds of millions of 'survivalist' informal microenterprises. Put simply, India has so far failed to nurture an innovative and growth-oriented SME sector, one that would be capable not just of providing millions of desperately sought-after formal sector jobs, but also of acting as an efficient subcontracting and supplier base for the large firm sector. Meanwhile, the microfinance sector in India has been growing very rapidly indeed, especially in Andhra Pradesh state, as we noted above. As of 2006 microfinance constituted 15% of all commercial bank lending in the whole of India, while, as Arunachalam (2011) extensively documents, the non-bank microfinance sector has experienced a significant boom this last decade thanks to the entry of private entrepreneurs and other financial institutions and foreign investors. Crucially, the growth of funding for microfinance has arrived thanks to the diversion of funds away from other uses, particularly financial support for SMEs. Indeed, this substitution effect is one of the main features of the Indian banking sector this last decade, and it is at least partly driven forward by the Indian government's firm belief in the virtues of microfinance (commercial banks in India are required by law to allocate a certain percentage of their funds into the microfinance sector, usually via MFIs).

As Karnani (2007) points out, however, the growing focus on microfinance and the subsequent growth of tiny informal microenterprises in India, and the concomitant reduction in funding and support for SMEs, has quite dramatically *undermined* the productivity and overall efficiency of India's economy (Karnani, 2011)[9]. This is because the SME sector has seen what little hope it had of obtaining financial support recede even further into the distance. Providing finance to the SME sector is both risky and low margin work for India's banks, compared to investing in its large Indian and foreign companies, which is a secure and stable investment, and to investing into the country's booming microfinance sector, which (until recently at least) demonstrated very high returns. Moreover, India's Self-Help Group (SHG) movement, a movement that provides very poor women with a way of gradually accumulating a tiny

[9] See also 'Microcredit: why India is failing', *Forbes*, November 10th 2006.

amount of savings, by design does not lend to small or medium projects undertaken by members. Karnani's (2007, p. 39) view is that the millions of tiny survivalist microenterprises that have emerged in India in recent years do not provide anything approaching a solid foundation for India's growth and poverty reduction efforts. His conclusion is that it was wrong for Indian policy-makers to ignore the crucial importance of economies of scale in productive activity, because this has led to a seriously adverse economic structure where, "[t]he average firm size in India is less than one-tenth the size of comparable firms in other emerging economies. The emphasis on microcredit and the creation of microenterprises will only make this problem worse".

In neighboring Bangladesh - the spiritual home of modern microfinance – the situation in this anti-development respect is probably even worse than in India. With a high and growing share of the country's savings and commercial funds being recycled into highly profitable microloans, Bangladesh now has the highest microfinance penetration rate in the world (25% of the population are borrowers from MFIs – Bateman 2011b, p. 4). But the price that is being paid for this microcredit largesse is that Bangladesh's SME sector has effectively been displaced and starved of funding. Some of the international development agencies are now beginning to wake up to the damage being caused in Bangladesh as the far more productive SME sector is increasingly being left to wither on the vine. For example, research by DFID (Department for International Development), the UK government's aid arm, summarized the situation in Bangladesh (DFID 2008, pp. 2-3), as one where,

> "[t]he financial system – including banks, capital markets and the micro-finance sector - is inadequate to support long term investment financing for growth. Smaller firms, responsible for the lion's share of employment, have severely limited access to financial resources. *Rural areas, with the highest potential for lifting low income groups out of poverty, are cut off from most financing mechanisms*." (italics added)

If what the DFID study calls 'smaller firms' (that is, small firms that are not microenterprises) are finding it difficult to access financial support in the rural areas of Bangladesh, areas where the country's famed MFIs are increasingly in a desperate search for new microenterprise clients in order to keep themselves alive, then the 'smaller firm' funding situation is clearly very bad indeed. Informal microenterprises and poor individuals can very easily access – in fact, they are being *pushed* to access – far more funding than they can repay, while 'smaller firms' are increasingly being left without any finance to get established or to grow.

However, there is very little that the Bangladesh government appears capable of doing to stop the hugely unproductive informal sector from absorbing a large and growing part of the scarce funds available in that country (mainly savings and its vast remittance inflow). It certainly does not help that the 'big 4' MFIs in Bangladesh – Grameen Bank, ASA, BRAC and Proshika – are all very powerful political and economic institutions, and they have all tended to resist suggestions by the Bangladesh government and others that their lending programs should venture a little more into much less profitable, but perhaps more developmental, business areas, such as SME lending or housing mortgages[10]. In other words, just like in neighboring India, the massive microfinance industry in Bangladesh has turned out to be a major

[10] In 2011, a documentary by award-winning Danish filmmaker, Tom Heineman, famously exposed the Grameen Bank's reluctance to get involved in housing mortgages, even with donor grant funding explicitly offered for this purpose. Using previously secret documents held in the Norwegian state archives, Heinemann showed that in the mid-1990s Grameen Bank obtained a $100 million Norwegian government grant to be used to develop low-cost housing mortgages in Bangladesh. However, this grant was right away secretly transferred by Muhammad Yunus to a sister company (Grameen Kalyan) only for Yunus to then instantly transfer it right back to Grameen Bank as a loan to be used for far more profitable individual microloans. The exposure of this misappropriation of donor funds only came to the notice of the Norwegian government two years later, which immediately demanded that the $100 million be returned, which most of it was. Not unexpectedly, both parties to the transaction quietly agreed to keep the whole incident under wraps for fear of tarnishing the reputation of Yunus and the Grameen Bank, and that of the Norwegian aid authorities, as well as the reputation of microfinance in general. However, Heinemann's exposure of this misappropriation, as well as the huge publicity that ensued when his documentary went on to win a handful of major international documentary film-making awards, directly led on in 2011 to Muhammad Yunus being removed from his position as head of the Grameen Bank. See Sinclair 2012.

obstacle in terms of supporting the development of the enterprises operating at or above minimum efficient scale that Bangladesh very urgently needs in order to sustainably develop and reduce poverty.

The very same adverse dynamics have been identified as a major problem in Latin America too. In Mexico, for example, the manifest shift of resources into the hugely profitable microfinance sector has directly precipitated a booming sector of 'changarros' (informal microenterprises, or simply 'mom and pop stores'), but at the same time undermined the desperately required capitalization and expansion of the country's crucial SME sector. One result, as Levy (2007) argues, is that, "There are more resources to subsidize informal employment than formal employment" and so "Mexico is probably saving less and investing in less efficient projects". Mexico's biggest development problem today has become one of "Over-employment and over-investment in small informal firms that under-exploit advantages of size, [and so] invest little in technology adoption and worker training". Crucially, one of the reasons for this misallocation of capital scenario is the booming microfinance sector that has emerged in Mexico since the mid-1980s, and which has resulted in a growing percentage of the country's scarce capital resources being diverted into informal microenterprises and away from potentially higher value uses, such as formal SMEs[11].

Moreover, the IDB's far-reaching conclusion in a recent high-profile publication (IDB 2010) is that Mexico's adverse capital allocation and subsequent deindustrialisation problems have essentially been the main story throughout *all* of Latin America this last thirty years or so. As the IDB reports, Latin America has for too long remained trapped in poverty and under-development because it has channelled far too much of its scarce financial resources into low-productivity informal microenterprises and self-employment, and far too little into more productive formal small and medium enterprises. In other words, the massive microfinance-induced proliferation of informal microenterprises that has taken place in Latin America since the 1980s has not been its economic and social saviour, as analysts like the Peruvian economist Hernando De Soto have long propounded would be the case (De Soto 1989), but a factor that actually lies at the very *root* of that continent's recent economic and social malaise. As the IDB summed up (2010, p. 6), "the overwhelming presence of small companies and self-employed workers is a sign of *failure* (in Latin America), not of *success*" (our italics). Without perhaps having this objective in mind, the IDB has quite clearly blown out of the water the long-standing belief that the programmed expansion of microfinance in Latin America has been a positive development.

An equally dangerous 'primitivising' aspect of microfinance here is in relation to the agricultural sector, and against a background of food shortages and agricultural commodity prices rises that are (re)introducing food insecurity problems in many developing countries. It is well known that the microfinance sector has proved adept all around the globe at moving into the subsistence farming sector. Yet there is a wealth of evidence to show that tiny subsistence agricultural units are simply *not* the most appropriate agricultural units if a developing country wants to achieve sustainable rural jobs growth and local food security (eg. Sender and Johnston 2004). Inserting microfinance into supporting the expansion of such units is therefore counter-productive into the longer run in terms of rural sector development. Moreover, the proliferation of microfinance in the agricultural sector is likely to have undesirable political consequences in the form of a reduction in female empowerment, as micro-farms cannot survive without an increase in the exploitation of what is euphemistically known as 'non-contractable labour', that is, unpaid female labour (see Manji, 2006, for further discussion). But, at the other extreme, nor are the sort of large-scale plantation-style farms advocated by commentators such as Collier (2008) any better for the poor. In the main, such plantation farms employ few people on decent wages, may destroy the local ecology, and the often large profits go up to a tiny elite, which is often not even resident in the country concerned (and so valuable spending power is lost to the local economy)[12].

---

[11] For example, bank lending to formal enterprises (SMEs) fell in Mexico in the new millennium, going from 60% of total lending to just over 48% in only six years – see Dos Santos 2008, p.2.

[12] Obvious examples here include the commercially successful large-scale vineyards and wineries in parts of South Africa, which are also the location for the highest concentration of poverty in the country (see Du Toit 2004), and Kenya's horticultural export sector,

Instead, it is commercially viable, small (but not 'micro') family farms that are in many circumstances the most valuable in terms of contributing to efficient, sustainable and equitable agricultural sector development[13]. This is because family farms help to maximize the potential to adopt technologies that create rural employment opportunities, are big enough to make good use of irrigation schemes, raise agricultural productivity, re-localize the consumption of food, address food security issues, and all without unduly damaging nature's goods and services (Norberg-Hodge et al 2002; Pretty 2005). Notwithstanding, the microfinance sector today continues to recycle a country's valuable financial resources into the tiniest of subsistence farms, which are the *least* efficient forms of farming, while ignoring the family farming units that are likely to bring about *most long-term* benefits to the local community overall. It is difficult to conceive of a more damaging local financial structure in terms of facilitating the programmed long-term destruction of the agricultural sector.

An obvious illustration of the structural damage to agriculture brought about thanks to microfinance is in the Indian state of Andhra Pradesh – a global pioneer in increasing the supply of microfinance, as noted above. By all accounts, from the 1990s onwards the profit-driven channelling of large quantities of microfinance towards tiny subsistence farming units has precipitated a human and economic disaster. With evidence of a growing over-indebtedness to a new breed of commercial MFI, offering immediate access to a microloan but all too often at a deceptively high rate of interest (for example, thanks to a lot of hidden charges), the Andhra Pradesh rural economy began to implode. In 2003 the state authorities commissioned a major report to look into the problems (see Commission on Farmers Welfare, 2004). The report centrally noted that "Agriculture in Andhra Pradesh is in an advanced stage of crisis[.]… The heavy burden of debt is perhaps the most acute proximate cause of agrarian distress. The decline of the share of institutional credit, and the lack of access to timely and adequate formal credit, in the state have been a big blow to farmers, particularly small and marginal farmers" (*ibid*). Notwithstanding these findings, nothing was done to stop rural over-indebtedness to the new highly commercial MFIs, which rose even *more* dramatically than before[14]. A serious microcredit bubble was created, which in 2006 collapsed in the shape of the 'Krishna Crisis' (named after the Krishna District in which the over-indebtedness problem first became apparent – see Arunachalam, 2011).

The core problem here was that the *least* productive subsistence farms (generally less than two hectares) were all too easily able to access a microloan, when it should have been clear that they could really do almost nothing with it. Any marginal increase in output was simply not enough to cover the high interest rate charges on the microloan that gave rise to it. Of course, many subsistence farmers were desperate, and so it was easy for the local MFIs to persuade those already in deep debt to accept more of virtually *any* form of credit at *any* rate of interest in order to try vainly to resolve their long-standing problems. But the result of the subsistence farming community accessing microfinance in Andhra Pradesh was the gradual entrapment of several hundreds of thousands of its tiniest and least productive subsistence farms in a vicious downward cycle of dependency and growing microdebt (see the illuminating discussion in Taylor 2011). Just under 82% of farmers in Andhra Pradesh were in debt by the mid-2000s, the highest figure in all of India (Patel 2007).

Crucially, precisely because of their very small size and low productivity, very little additional agricultural output was actually secured by accessing so much microcredit: in fact, most subsistence farms in serious debt ground to a virtual halt. One reason for this was that high interest rate payments on

which is very successful for its mainly European owners, yet the local workforce receives poverty-level wages (see Pollin, Githinji and Heintz 2008). But see also the discussion in Cramer, Oya and Sender 2008, which shows some plantations operating in a somewhat better light.

[13] The definition of a 'family farm' is not an easy one to provide and it will vary from country to country. However, we may say it lies somewhere in the space above the inefficient subsistence farm variant described by Sender and Johnston, in that there is a significant marketable surplus, but well below the plantation-type farm promoted by Collier, which is almost entirely geared up to producing for often distant markets.

[14] See 'The Makings of a Debt Trap in Andhra Pradesh', *The Hindu Times*, April 20th 2006.

microloans effectively pushed many of the tiniest farms into financial loss-making territory. These farms then chose to slow down, or even stop farming completely, rather than rack up even *more* losses trying to fund the next agricultural cycle (Commission on Farmers Welfare 2004). Tragically, this reduction of output also arose because of the rising number of rural suicides in Andhra Pradesh[15]. At any rate, thanks to so many tiny subsistence farms languishing and failing outright under the burden of microdebt, while more commercially-oriented small family farms were increasingly unable to access capital on affordable terms and maturities, rural incomes fell by 20 per cent in Andhra Pradesh in the decade after 1993 (*ibid*). Even worse in retrospect, it was largely the commercial failures in the rural sector that encouraged Andhra Pradesh's MFIs subsequently to move into its urban areas in search of a completely new raft of poor clients to service, to quite devastating effect, as we saw above.

All told, the most obvious result of focusing upon expanding the numbers of the very tiniest informal microenterprises and farming units is the de facto shift of resources away from the far more productive above-minimum efficient scale enterprises and farms. This has resulted in what one astute critic of the microfinance model has denoted as 'the microcredit paradox': a situation where "the poorest people can do little productive with the credit, and the ones who can do the most with it are those who don't really need microcredit, but larger amounts with different (often longer) credit terms" (Dichter 2006, p. 4). More broadly, such a shift has led to the proliferation of 'infantilizing' development trajectories. Almost everywhere where the microfinance model has entered into the enterprise and agricultural sectors we find little real sustainable progress, while major opportunity costs are manifestly evident.

(b) *The microfinance model ignores the 'fallacy of composition'*

As the late Alice Amsden (2010) argued, it has been a major mistake when dealing with poverty in developing countries to assume that there is no local demand constraint, and that every local economy therefore has the elastic ability to productively absorb an unlimited number of the unemployed through the expansion of the local enterprise sector. Amsden noted that this form of Say's Law – "supply creates its own demand" – is a seductive lure for policy-makers seeking to help the unemployed through supply-side measures (such as enterprise development and training) but, as she demonstrated (see also Galbraith 2008; pp. 151-163), it has no basis in reality.

Other things being equal, new and expanded microfinance-induced microenterprises do not raise the *total* volume of business/demand so much as redistribute or subdivide amongst market participants the prevailing volume of business/demand (on this important point, see also Davis 2006). This point is, of course, the 'fallacy of composition' and it has quite serious implications for the presumed efficacy of microfinance. This fallacy is most vividly manifested in the statement by Muhammad Yunus that "[a] Grameen-type credit program opens up the door for limitless self-employment, and it can effectively do it in a pocket of poverty amidst prosperity, or in a massive poverty situation" (Yunus 1989, p. 156).

The reality in virtually all developing countries is that local economies have been saturated with simple informal microenterprises for many years: indeed, an informal microenterprise has long been the default activity for those without any type of formal employment or income – the vast majority in some countries (ILO 1972; Breman 2003). The scale and scope of the local informal sector was and is mainly determined by local demand. With the arrival of microfinance in the 1980s, however, an artificial supply-side MFI-driven increase in the numbers of informal microenterprises was stimulated without any compensating intervention on the demand side. This inevitably created hyper-competition at the local level, which in turn precipitated reduced turnover in existing individual microenterprise units and downward pressure on local prices and incomes in general (thus negatively affecting both new and incumbent microenterprises). As a result, we find, not surprisingly, that from the 1990s onwards,

[15]The cause and actual numbers of rural suicides, including those directly and indirectly caused by over-indebtedness to local microcredit institutions, remains a matter of hot dispute. See 'Death by microcredit', *Times of India*, September 16th 2006.

incomes, wages, profits and work-life conditions for those struggling in the informal microenterprise sector began to deteriorate quite markedly across the globe[16].

Two negative but largely unregistered outcomes are uppermost as a result of microfinance programs in this specific context: first, significant job and income displacement effects across the community and, second, significantly higher levels of exit by incumbent producers.

Consider first the issue of displacement. In Mexico, the typical local economy has for some time been bursting at the seams with informal microenterprises. Few market gaps remain. The result of new entry and expansion thanks to microfinance is that prices on most of the very simple products and services have been falling. In addition, lower turnover in individual microenterprises, as local market demand is shared out among a growing population of microenterprises, has been precipitating lower margins and incomes. In many sectors and in many regions of Mexico, poor individuals are hugely angry at the declining margins and wages, as well as longer working hours, brought about by the unremitting inflow of 'poverty-push' microenterprises supported with microfinance[17].

Noticeably in the wake of NAFTA[18] in 1994, which quickly closed many industries in Mexico, and so stimulated an extensive wave of new informal microenterprises composed mainly of the newly redundant, the end results were quite adverse. Popli (2008) reported that poverty levels in the (newly enlarged) informal microenterprise sector very rapidly increased after NAFTA. Even as some economic growth reappeared in the Mexican economy in the mid-1990s, poverty levels in the informal sector continued to rise. The simple dynamic here involved existing local market demand (and in many areas, *declining* demand, because very many small farmers after NAFTA lost their local market and incomes due to cheaper imported US corn) being shared out within the now enlarged informal microenterprise sector. Very little, if any, *net* employment or additional income was actually generated through the recession-driven surge in new microenterprise entry.

Thus seen, the proliferation of MFI-financed microenterprises simply *redistributes* poverty within the poorest communities, if indeed it does not *exacerbate* it: it certainly does not *resolve* it. More importantly, the poor do not always meekly accept to pay this social cost on behalf of society. Violent reaction against their fellow micro-entrepreneurs and local government officials (who mistakenly think that stimulating new entry is always and everywhere a good thing) has all too often emerged as incumbent wages and working conditions have declined, as was the case a few years ago in Mexico's several million strong community of mobile street vendors[19].

Turning to the related issue of an MFI's clients failing, we find, first of all, that such failure is even more pronounced in relation to informal microenterprises than in formal small enterprises, because the former are generally much more likely to be established on the basis of 'poverty-push' factors rather than 'opportunity/profit pull' factors. Failure rates of informal microenterprises are often very high indeed in developing countries (for an example from India, see George 2006). The core problem with client failure, however, is that this event very often plunges the hapless individual into much deeper, and possibly irreversible, poverty. This is because a failed microenterprise often means the poor lose not just their already minimal income flow, but also any additional assets, savings and land they might have invested into their microenterprise, or else are forced to sell off (often at 'fire-sale prices') in order to repay the microloan. Social networks and reputational capital are also lost.

An all too real illustration of what we mean here is to be found in Bosnia. As elsewhere, Bosnia's microenterprise sector is defined by its high failure rate, with up to 50% of microenterprises failing within just one year of their establishment (Demirgüç-Kunt et al 2007). Behind this dry statistic,

[16] For example, see ILO 2009.

[17] See International Press Service (IPS), Mexico City, September 2nd 2003.

[18] North American Free Trade Agreement.

[19] International Press Service (IPS), Mexico City, September 2nd 2003.

however, lies the fact that a very significant number of Bosnia's poor individuals failing in their microenterprise project have ended up in much deeper poverty, vulnerability and insecurity.

Bateman, Sinković and Škare (2012) find that here are several reasons for this adverse outcome. First, those failing in a microenterprise but who chose (or were effectively forced) to continue to repay their microloan ended up drawing down family assets (especially family savings) and selling off other family assets – family land, housing, private vehicles, machinery, and so on. Second, many in Bosnia were forced to divert other important family income flows into microloan repayment, such as remittance income and pensions. Third, very many individuals in Bosnia got hooked into taking out multiple microloans, using each new microloan to repay existing microloans, but in the process building up a mountain of personal debt that at some point needed to be repaid. As a result, the interest payments required to service these individual debt mountains constituted a growing proportion of household income, thus reducing the amount of income available for other important household items. Fourth, even those quite unconnected to a failing microenterprise, such as the estimated 100,000 individuals who guaranteed a microloan for friends and family, as is the common procedure in Bosnia, ended up severely disadvantaged by being forced to repay a microloan on someone else's behalf.

All told, there is no shortage of evidence from the field that routine displacement and client exit factors have often completely frustrated the poverty reduction goals of microfinance. However, partly because of the familiar neoliberal position that the 'opportunity' and 'freedom' to establish a new enterprise is all that really counts, and not other conditions, such as the capabilities of the entrepreneurs involved or if there is real demand for their simple outputs, these adverse features of the microfinance model have long been completely ignored. Today, the view that displacement and client failure are important factors is coming to be accepted by many individual analysts and institutions, though certainly not by all[20]. One example of this new realism is the ILO's recent response to the global financial crisis and rising unemployment, which was to argue *against* further stimulation of the informal microenterprise sector, since "[a]s was the case in previous crises, this could generate substantial downward pressure on informal-economy wages, which before the current crisis were already declining" (ILO 2009, p. 8).

(c) *The microfinance model helps to de-industrialise and infantilise the local economy*

Entrepreneurship theory and studies in institutional economics show that it is new, creative, technically innovative ideas and institutions that are the key engine in economic development (Schumpeter 1987/1942; North 1990; Baumol et al 2007). To develop in a sustainable fashion, and thus to reduce poverty, developing countries need to master key technologies, better understand 'state of the art' industrial products and processes, develop at least some innovative capabilities in domestic enterprises, and establish a tissue of pro-active development-focused institutions and organizations (UNCTAD 2003; Amsden 2007; Chang 2007).

However, given the high interest rates and short maturities demanded by most MFIs, it is generally only the most simple and unsophisticated microenterprises that can service a microloan. Typically, these microenterprises are very simple trading, retail and service operations, with perhaps some very small production-based operations that can add value very quickly (such as food preparation). We also know that very few growth-oriented microenterprises or SMEs using more sophisticated technologies can effectively get started or expand with the assistance of microfinance, as their returns are of longer term-nature. Within the 'new wave' microfinance paradigm, moreover, there is an in-built bias against longer term projects which are likely to be of much more value to the local

[20] In 2010 the EU launched the European Progress Microfinance Facility, a major €100 million program designed to support the unemployed in recession-hit Western Europe. It was built on an implicit assumption that there is sufficient local demand to un-problematically underpin a new wave of microenterprises set up by the unemployed. However, the evidence the EU has used to underpin this assumption is derived from evaluations of microenterprise growth and survival undertaken in the early 2000s, which showed that there was no shortage of local demand for microenterprises. That today's local demand situation is so radically different to the pre-global financial crash period appears to have been ignored. See 'Creating Jobs in recession-hit Communities in Europe: Why Microcredit will not help', *Social Europe Journal Blog*, May 15th 2012. Go to: http://www.social-europe.eu/2012/05/creating-jobs-in-recession-hit-communities-in-europe-why-microcredit-will-not-help/

community, but which would struggle to repay high interest rates in their initial period of operations. Nor does it help that many high-profile commercial banks are increasingly 'downscaling' out of traditional SME lending into higher profit microfinance.

Overall, then, to the extent that the financial sector shifts in favour of microfinance – as we are indeed seeing right around the globe – the more an economy's scarce financial resources are effectively directed *towards* the very simplest 'no-tech/no-capital' **–** mainly petty-trade-based – microenterprise projects, and so channelled *away* from more sophisticated and technology/innovation-based projects that offer far more to the economy and society in the medium to longer term. As Baumol (1990) among others have shown, we find many developing countries have, thanks to microfinance, evolved an enterprise structure that is structurally (in addition to the scale economies problem noted earlier) incapable of giving rise to sustainable productivity growth, and so also poverty reduction.

Consider once more the case of Sub-Saharan Africa (see also Chang 2011, pp. 157-167). With the microfinance sector rapidly expanding this last decade, local savings and remittance incomes are increasingly being recycled (and very profitably so) into the very simplest of trade-based operations and inefficient subsistence farms. This is helping to expand Africa's already giant informal microenterprise sector. At the same time, however, this emphasis upon microfinance has effectively reduced the financial backing required for the 'bottom-up' industrial transformation of Africa, particularly through reducing support for innovative and growth-oriented SMEs. In short, with the help of microfinance Sub-Saharan Africa's economic structure is increasingly becoming characterised by the 'missing middle' phenomenon – it is a continent of hundreds of millions of simple traders coexisting uneasily with a handful of large companies (eg. oil companies and copper and diamond mines), but very little else. Even in those countries where a natural resource bounty has made the availability of finance much less of a problem than elsewhere, such as in oil-rich Nigeria, the informal microfinance sector has ignored the obvious oil-sector related opportunities (subcontracting, servicing, etc) and demonstrated the usual overwhelming predilection to work with only the very simplest microenterprises – nearly 80% of microfinance in Nigeria (and the sector is growing rapidly at the expense of more traditional uses (ie. SMEs)), is channelled into simple cross-border petty trade-based microenterprises (Anyanwu 2004).

Africa's escape from poverty and under-development simply will not be facilitated upon the microfinance-induced entry of more of the simplest 'buy cheap, sell dear' trade-based microenterprises. Africa's growth requires instead the gradual construction of a robust light industrial and agro-processing foundation that will enable its entry into at least some mainstream production and manufacturing-based enterprises capable of productivity-growth. This in turn means that Africa urgently requires not even *more* microfinance than at present, but a raft of robust and far-sighted private and public financial institutions willing to socialise risk, carefully build productive capabilities where appropriate, and hold steady to a longer-term development and industrialisation vision. This need is not being addressed, however. In fact, (no) thanks to Dambisa Moyo's internationally well-received book setting out her own solutions to the continued poverty and underdevelopment in her native Africa – especially her belief that very much more microfinance is needed (Moyo 2009, Chapter 8) – we would argue that the real solution to Africa's problems has become more elusive than ever.

In short, microfinance greatly reduces the ability of developing countries to promote their industrial upgrading as one of the keys to eventual economic success and poverty reduction. This is not only because the microfinance sector misdirects scarce resources into the wrong type of enterprise (ie. mainly into simple trade-based microenterprises), but also because it draws scarce development funds away from financial institutions that *are* perhaps up to the required task (eg, Korean/Brazilian-style development banks, SME technology funds). Meanwhile, in the formerly industrially sophisticated and institutionally quite rich countries of Eastern Europe, an obvious and valuable industrial inheritance – an inheritance that most developing countries are desperately wishing to possess – has been largely abandoned despite being the potential starting point for a new generation of relatively technology-intensive enterprises.

(d) Microfinance fails to connect with the rest of the enterprise sector

Another important factor that we now know lies behind successful local economic and enterprise development is 'connectability' between enterprises of all sizes. It is now very well understood that the tissue of horizontal (clustering, networks) and vertical (subcontracting) connections within the local enterprise sector is a crucial determinant of a local economy's ultimate sustainability through industrial development (Pyke 1992). Indeed, as Weiss (1988, p. 210) concludes in reflecting on the successes of both the Italian and Japanese microenterprise and SME sectors since 1945, "the core of modern micro-capitalism is not competitive individualism but collective endeavour".

Wherever the microfinance model has been in the ascendancy, however, such beneficial grassroots dynamics have largely been undermined. While succeeding in terms of producing some new (albeit largely temporary) informal sector microenterprises, the overwhelming majority of these new entrants have no need, wish or ability to meaningfully cooperate in order to begin to forge the required productivity-enhancing horizontal ('proto-industrial districts') and vertical (sub-contracting) connections. The result in many developing and transition countries has been little movement towards a more 'connected' local economy. This gives rise to some significant handicaps. For example, large firms are unable to expand their operations by tapping into a local structure of quality suppliers, but must import instead. A lack of potential sub-contracting partners also typically dissuades investments in large-scale operations, especially 'greenfield' FDI. Important cluster building programmes simply cannot function when there are few, if any, local enterprises that can meet the technology, market and scale requirements to benefit from cooperating with their counterparts.

In short, the microfinance model pays no heed to the important requirement that enterprises be of the right type (size, quality, use of technology, innovative products and processes, etc.) that might both facilitate and benefit from local 'connectability'. The microfinance model therefore operates like a football academy that exists solely in order to turn out players with individual skills, but all of whom have no ability to engage in the vital organisational cooperation – the teamwork – required to actually win the match.

(e) *The microfinance model is pre-programmed to precipitate a sub-prime-style over-supply of microfinance*

Hyman Minsky (1986) predicted that neoliberal policies were likely to be especially destructive when played out through the financial sector, with an inevitable tendency towards Ponzi-style booms and busts in the supply of finance. It has become increasingly apparent through a series of financial crises, culminating in the 2008 global financial crisis, that Minsky was correct. Minsky's predictions also very much pertain to the local financial sector. For example, Black (2005) extensively documents a Minskyian-style adverse trajectory in the shape of the boom and then spectacular bust of the US Savings and Loans (S&Ls) institutions in the 1980s.

As the growing number of 'microfinance meltdowns' indicates, the microfinance sector has proved very receptive to Minskyian dynamics. In fact, the massive sub-prime-style over-supply of microfinance and various Ponzi-style dynamics are now intrinsic features of the microfinance model.

Two important sub-prime-style drivers are important here. First, as in any private business, pushing out a continuously increasing volume of microcredit is the most important way that an MFI can both justify and physically provide the financial space that allow for the generous salaries, bonuses and other perks that are increasingly the norm in the microfinance sector. All that matters is that, somehow, an MFI's clients are able to absorb whatever output of microcredit is forthcoming, even if only to repay microloans already taken out (as very much in Andhra Pradesh state in India). Second, the larger an MFI becomes, the more likely it is that its senior managers will be able to benefit when the time comes for the expected transition to publicly owned company status via the IPO route. The primary mechanism

that can provide for this private enrichment is found in the fact that an MFI's senior managers typically accumulate shares in their own MFI, almost always using interest free loans from their own MFI to do so. These shares are then offloaded at the time of the IPO. In the two most notorious microfinance IPOs to date – Compartamos in Mexico, and SKS in India – senior managers were able to garner several tens of millions of dollars of personal gain from the sale of shareholdings they had built up over previous years using the interest free loan route[21].

In a very real sense, then, the microfinance model contains the seeds of its own destruction as a development intervention. Microfinance today is about making large sums of money for the *providers* of microfinance, and not about resolving the poverty situation of the poor *recipients* of microfinance (Klas 2011; Sinclair 2012). MFIs become super-charged into selling as much microfinance as they can, and, unlike in other product markets (furniture, food, clothes, etc), it is not difficult to convince the poor that there is no upper limit to how much microcredit they can 'consume'. Both providers and recipients within microfinance are thus automatically stimulated into excessive supply and demand respectively, thereby providing the fuel for the inevitable 'microfinance bubble'.

(f) *The microfinance model ignores the crucial importance of solidarity and local community ownership and control*

It has long been recognised that community solidarity, trust, volunteerism, equality, cooperation and goodwill are intimately and positively linked to the wider issue of 'community liveability' (eg. Zamagni and Zamagni 2010). But as many have argued (eg. Leys 2001), whenever community development and poverty reduction activities are constituted as commercial operations, this quite dramatically increases the likelihood that such important outcomes for society are undermined.

In many ways the microfinance model undermines these important 'community liveability-building' processes. Perhaps most important of all, the local hyper-competition that follows in the wake of microfinance is a patently unsuitable foundation upon which to build 'community liveability'. As Davis (2006) reports, it is *precisely* the unrelenting growth of informal microenterprises that accounts for the destruction of the sense of local community and solidarity in many developing countries. As Davis argues,

> "[t]hose engaged in informal sector competition under conditions of infinite labour supply usually stop short of a total war of all against all: conflict, instead, is usually transmuted into ethnoreligious or racial violence… the informal sector, in the absence of enforced labour rights, is a semi-feudal realm of kickbacks, bribes, tribal loyalties, and ethnic exclusion… the rise of the unprotected informal sector has too frequently gone hand in hand with exacerbated ethnoreligious differentiation and sectarian violence" (p. 185).

Put very simply, the informal microenterprise sector simply does not possess the sort of 'transformational power' and solidarity-building capability widely claimed for it by the microfinance industry and its ideological supporters. On the contrary, the inevitable local hyper-competition and the resulting brutalization of poor individuals and intensification of their day-to-day workload and suffering are an unlikely precursor to 'community liveability', or for any other desirable economic and social development outcomes. Local solidarity is inevitably destroyed as the distorted business ethics and morals that inevitably emerge under such Hobbesian conditions gradually percolate into other enterprise structures (ie. SMEs), other institutions (ie. government) and across all levels of society.

[21] For example, see 'SKS and Compartamos – Catalysts for Catastrophe', *India Microfinance*, October 19th 2011. Available from: http://indiamicrofinance.com/sks-compartamos-milford-bateman.html

3. Microfinance is used as a vehicle for neoliberalism

One of the major assumptions about microfinance is that it is ideology-free and simply about 'helping the poor'. However, microfinance in its commercialised form is actually almost perfectly in tune with the core doctrines of neoliberalism, the reigning ideology of our time: that is, the need to vector all economic activity through private individual initiative; the need to avoid any aspect of planning or conscious guidance of the market mechanism; the need for all MFIs to attempt to 'earn their keep on the market'; and, the need to ensure that all economic organizations are also as much as possible owned and controlled by the private sector (Harvey 2006). So might one of the reasons for the almost unlimited well of support for microfinance be related to the political economy of neoliberalism? After all, at least since the time of Marx, and more recently re-emphasised by the conservative institutional theorist Douglass North (North 1990), 'bad' organisations are allowed to survive, and may even be encouraged to flourish, simply because it is in the interests of the powerful for this to happen. In this section, we briefly adumbrate the intimate association that clearly exists between microfinance and neoliberalism.

(a) Microfinance provides a model for poverty alleviation that is politically acceptable to the neoliberal establishment

A pervasive and continuing fear among neoliberals is that the poor will opt to use the democratic process or popular pressure to demand the establishment or strengthening of state and collective institutions capable of remedying their plight. As Bromley (1978) pointed out, neoliberals were very quick to see the informal sector in general, and, we would argue here, the microfinance sector in particular, as a way to pre-empt more radical alternatives that might upset the prevailing economic system and distribution of power and wealth.

Microfinance offers to neoliberals the hope that informal sector activities backed up by microfinance will become universally embedded as the only legitimate exit route out of poverty for both the individual and the community, thereby also removing from the political and policy agenda a wide range of progressive policies. These include demands for constructive state intervention, robust social welfare programmes, quality public services accessible to all, income and wealth redistribution (including land reform), and all forms of state, collective and cooperative ownership. The microfinance narrative helps to legitimise not only the entrepreneurial process as the core foundation of any society, but also the vastly unequal rewards (wealth and power) that inevitably arise in the process. After all, an opportunity to be successful in entrepreneurship in Dhaka, Abuja or Quito (thanks to obtaining a microcredit), or else as an entrepreneur in London, New York or Paris, essentially requires all parties to adhere to the same rules, regulations and processes: only the final rewards are different. And because such rewards (supposedly) depend on the amount of individual talent and effort put into the venture, there should also be no complaint as to any unequal outcome.

In this context, microfinance can thus be deployed to delegitimize and dismantle all possible 'bottom-up' attempts to propose alternative development policies which might primarily and directly benefit the majority but which would circumscribe the power and freedom of established elites. Put simply, microfinance offers to neoliberals a highly visible way of being seen to be addressing the issue of poverty, but in a way that offers no real challenge to the existing structures of wealth and power. Those who fail to put in sufficient effort to establish a successful microenterprise, or, worse, do not even attempt to establish a microenterprise, can very effectively now be blamed for their own poverty situation.

(b) *Microfinance can be used to undermine the concept of basic state service provision and to support privatisation and private sector provision*

In a very real sense, microfinance has been consciously positioned as the substitute for social welfare spending (and international donor support). Once the poor can be made to accept that they are now in control of their individual and family destiny by using microfinance, it becomes much easier for the government to fully absolve itself of continued responsibility towards them. Governments can also, if they so wish, even begin to dismantle social welfare systems constructed after years of social mobilization and collective struggle.

For example, microfinance has been deployed as part of the goal to promote private local service provision, rather than collective service provision through the state or local community. This has been a long-time goal of neoliberal policymakers everywhere. A major aspect of neoliberal Structural Adjustment Programs (SAPs) has involved the dismantling of important public services and utilities serving the common good, and their gradual replacement with private provision based upon user fees. However, not surprisingly, almost all of these programs meet determined resistance from the poor, and this is where microfinance comes in. Because it can spread the cost of access to private provision over a longer period of time, microfinance can dampen down the initial anger inevitably felt when important services are privatised and put on a 'full cost recovery' basis. Shiva (2002) reports that microfinance programmes have been successfully used to ensure a less precipitous, and thus less politically damaging, decline in water demand after privatisation.

Once the negative effects of the introduction of user fees are softened by the provision of microfinance, it is hoped that the poor will begin to accept that they must permanently pay for service provision, and so begin to see microfinance as the only way to find the larger sums of cash required to gain regular access to private provision. Even though collective provision of social services by the local state is usually the most efficient, including when directly compared to microfinance (eg. Mader 2011), in this way it might still be possible to encourage the poor to begin to rely upon much less efficient private sector provision. In some cases, unconcerned government officials and politicians hope that the poor can be fobbed off with microfinance rather than state activity. In India, for example, Harper (2007, p. 258) reports that government officials are increasingly deflecting community demands to support better basic education and health services on the grounds that poor people 'now have microfinance' and should individually seek to purchase such services (albeit at high prices) from the private sector rather than through taxpayer-funded public provision.

(c) *Microfinance underpins the drive towards financial sector liberalisation and commercialisation*

Microfinance has played an important role in the promotion of global financial liberalisation and commercialisation. As Weber (2002) shows, MFIs that have achieved financial self-sufficiency provide working examples to developing country governments of 'efficient', subsidy-free, financial institutions. It is thus expected that all other financial institutions will have to follow suit. If a financial institution serving the poorest people can be profitable, the reasoning goes, all other financial institutions with a better clientele profile should aim for profitability as well.

Most recently, commercial funding of microfinance programmes, including the outright purchase of established MFIs by the commercial banking industry, has increasingly separated the microfinance industry from its roots in the NGO sector. As increasingly a part of the global financial complex, microfinance can be portrayed as a good example of how Wall Street and the global financial sector in general 'cares' and how it directly addresses core societal problems. At least until the global financial crash of 2008, it was hoped by many in the international development community that this 'public service function' provided by MFIs, and very publicly supported by many of the largest banks on Wall Street

(eg. CitiGroup), would contribute to obtaining continued government and public support for the ongoing liberalisation of the financial sector in general.

(d) *Microfinance acts as an important 'safety valve' within the globalisation project*

Perhaps the most important factor of all, however, is the 'containment' role that microfinance has been allocated within the neoliberal globalisation project. It is widely argued by neoliberals that globalisation has the potential to provide a major reduction in poverty. Yet it is hardly a coincidence that globalisation has been determinedly driven by a handful of the wealthiest of the developed countries – by the US most of all (Gowan 1999) – which (or rather whose elites) are expected to be, and have by far been, its major beneficiaries (Stiglitz 2002; Chang 2007). But as globalisation increasingly concentrates wealth and power into the hands of a small number of countries, regions and corporate elites, the flipside, as Faux and Mishel (2000) explain, is a growing worldwide population of the unemployed, powerless, marginalised, hyper-exploited and insecure. And the rub here is that these 'losers' are beginning to reject both the outcome assigned to them and, most dangerous of all for neoliberals, the globalisation process itself. Symptomatic of this rejection is the rising social unrest, increased social and gang violence, explosion in substance abuse, increasing crime and illegal business activity, huge rise in pseudo-religions and cults, collapsing levels of social capital in the community, and associated violent conflict (Davis 2006).

In the potentially explosive situation emerging in many developing and transition countries today, dramatically made worse by the Wall Street-precipitated global financial crisis, and particularly acute in the growing number of 'mega-cities', microfinance provides a crucial 'safety valve'. The logic is well known. Universal social welfare systems are being dismantled under the guidance of the main international financial institutions, secure public employment opportunities are rapidly disappearing, and formal sector employment are an increasing rarity too. Nevertheless, the hope (not always a realistic one[22]) is that the microenterprise sector can engage the most articulate and vocal of the poor, who might otherwise be thinking about resisting, or proposing realistic alternatives to, neoliberalism and the globalisation project.

4. Conclusion

This article has raised issues of serious concern relating to the contemporary microfinance model. While accepting that there are some minor and largely temporary short-run benefits for a small minority of 'winners', just as in any casino a few lucky punters will end up on a winning streak, we argue that the microfinance model has very serious limitations as development policy. In many respects, in fact, microfinance constitutes a very powerful 'poverty trap'. We outlined the main flaws in the microfinance model. We then provided at least part of the answer as to why the international development community continues to lavishly support the microfinance model in spite of these fatal, and now increasingly well-publicised and accepted, flaws. The microfinance concept is linked to neoliberalism and the globalisation project. It is therefore supported so strongly and uncritically because it is in agreement with the international development community's preferred economic and societal model based on self-help and individual entrepreneurship.

Finally, a word on what might be better local and national alternatives. We very much believe that there are better financial institution alternatives, such as financial cooperatives, credit unions, building societies and local and national development banks. Fully adumbrating the advantages of these generally community-owned and controlled alternatives would, of course, require another article (however, see Bateman 2010b, 2012b: Chang 2007).

[22] The young and well qualified people who led the recent Arab Spring uprisings were not just risking their lives to bring down dictators, but also very centrally arguing against being stuck in petty retail and service jobs (ie, in informal microenterprises).

References

African Development Bank and OECD (2005), *African Economic Outlook 2005*. Paris: OECD

Amsden, A (2001), *The Rise of 'The Rest': Challenges to the West from Late-Industrializing Economies*, Oxford: Oxford University Press.

Amsden, A (2007), *Escape from Empire: The Developing World's journey through heaven and hell*. Cambridge, Massachusetts: MIT Press

Amsden, A (2010), 'Say's Law, poverty persistence, and employment neglect', *Journal of Human Development and Capabilities*, Vol 1 (1) pp. 57-66.

Anyanwu, C M (2004), '*Microfinance institutions in Nigeria: policy, practice and potentials*', paper presented at the G24 Workshop on 'Constraints to Growth in Sub Saharan Africa,' Pretoria, South Africa, November 29-30.

Arunachalam, R (2011), *The Journey of Indian Micro-Finance: Lessons for the Future*. Chennai: Aapti Publications.

Balkenhol, B (2007), Policy implications. In: B Balkenhol (ed.) *Microfinance and Public Policy: Outreach, Performance and Efficiency*. London: Palgrave Macmillan, p. 213

Banerjee, A, E Duflo, R Glennerster and C Kinnan (2009), '*The miracle of microfinance? Evidence from a randomised evaluation*.' Available at: http://econ-www.mit.edu/files/4162.

Bateman, M (2000), 'Neo-liberalism, SME development and the role of Business Support Centres in the transition economies of Central and Eastern Europe', *Small Business Economics*, 14 (4), pp. 275-298.

Bateman, M (2003), ''New Wave' micro-finance institutions in South-East Europe: towards a more realistic assessment of impact', *Small Enterprise Development*, 14 (3), pp. 56-65.

Bateman, M (2006), '*The informalisation of the BiH economy and the role of the microfinance model*', paper presented at the Economic Policy Research Unit (EPRU) Seminar 'Balkan economies: Studies on Non-Standard Phenomena', September 26th 2006, Sarajevo, Bosnia.

Bateman, M (2010a), *Why Doesn't Microfinance Work? The Destructive Rise of Local Neoliberalism*. London: Zed Books.

Bateman, M (2010b), 'The power of the community', *Making It*, Vol 4 (10), pp. 22-3. Geneva: UNIDO.

Bateman, M (2011a) (ed.), *Confronting Microfinance: Undermining Sustainable Development*. Sterling, Virginia, US: Kumarian Press.

Bateman, M (2011b), Introduction: looking beyond the hype and entrenched myths. In: Bateman, M (2011b) (*op cit*).

Bateman, M (2012a), 'Let's not kid ourselves that financial inclusion will help the poor', Guardian PovertyMatters Blog, May 8th 2012. Available from: http://www.guardian.co.uk/global-development/poverty-matters/2012/may/08/financial-inclusion-poor-microfinance

Bateman, M (2012b), 'A New Local Financial System for Sustainable Communities' in: E Campiglio and F Campagna (2012 eds) *What Are We Fighting For? A Radical Collective Manifesto*. London: Pluto Press.

Bateman, M and H-J Chang (2009), *The Microfinance Illusion*, Working Paper. University of Juraj Dobrila Pula, Croatia, and University of Cambridge, UK.

Bateman, M, J P Duran Ortĩz and D Sinković, (2011) 'Microfinance in Latin America: the case of Medellín in Colombia'. In: Bateman (2011), (*op cit*).

Bateman, M, D Sinković and M Škare (2012). 'Bosnia's microfinance meltdown', paper presented at the American Economics Association annual conference, Chicago, January 7th.

Baumol, W (1990), 'Entrepreneurship: productive, unproductive, and destructive', *Journal of Political Economy*, 98 (5): pp. 893–921.

Baumol, W, R Litan and K Schramm (2007), *Good Capitalism, Bad Capitalism, and the Economics of Growth and Prosperity.* New Haven, Connecticut: Yale University Press.

Birchall, J (1997), *The International Co-operative Movement*. Manchester: Manchester University Press.

Black, William K (2005), *The Best Way to Rob a Bank is to Own One: How Corporate Executives and Politicians looted the S&L Industry*. Austin, Texas: University of Texas Press.

Breman, J (2003), *The Labouring Poor: Patterns of Exploitation, Subordination and Exclusion.* Oxford: Oxford University Press.

Bromley, R (1978), 'Introduction – the urban informal sector: why is it worth discussing?' *World Development,* Vol. 6 (9/10), pp. 1033-1039.

CGAP *(*Consultative Group to Assist the Poor*)* (2010), *Growth and Vulnerabilities in Microfinance*, *Focus Note no 61*, February.

Chang, H-J (2002), *Kicking Away the Ladder – Development Strategy in Historical Perspective.* London: Anthem Press.

Chang, H-J (2006), *The East Asian Development Experience: The Miracle, the Crisis and the Future*. London: Zed Books and Third World Network.

Chang, H-J (2007), *Bad Samaritans: Rich Nations, Poor Policies and the Threat to the Developing World.* London: Random House.

Chang, H-J (2011), *23 Things they don't tell you about Capitalism.* London: Allen Lane.

Collier, P (2008), 'The politics of hunger: how illusion and greed fan the food crisis, *Foreign Affairs*, November/December.

Commission on Farmers' Welfare (2004), *Report of the Commission on Farmers' Welfare, Government of Andhra Pradesh*. Hyderabad: Government of Andhra Pradesh.

Cramer, C, C Oya and J Sender (2008), 'Rural Labour Markets in Sub-Saharan Africa: A New View of Poverty, Power and Policy', *Policy Brief*, No 1, November. London: Centre for Development Policy and Research, SOAS.

Davis, M (2006), *Planet of Slums*, London: Verso.

Demirgüç-Kunt, A, L Klapper and G Panos (2007), *The Origins of Self-Employment,* Development Research Group (February). Washington DC: World Bank.

DFID (Department for International Development) (2008), *The Road to Prosperity through Growth, Jobs and Skills*, (Discussion Paper). Dhaka, Bangladesh: DFID Bangladesh:.

De Soto, H (1989), *The Other Path: The Invisible Revolution in the Third World*. London: I B Taurus.

Dichter, T (2006), *Hype and Hope: The Worrisome State of the Microcredit Movement.* Available from http://www.microfinancegateway.org/p/site/m//template.rc/1.26.9051

Dichter, T and M Harper (2007) (eds.), *What's wrong with Microfinance?* London: Practical Action Publishers.

Dos Santos, P (2008), *The World Bank, the IFC and the Antecedents of the Financial Crisis*, Bretton Woods Project Update, November 27th.

Du Toit, A (2004), ''Social exclusion' discourse and chronic poverty: a South African case study', *Development and Change*, 35 (5).

Duvendack, M and R Palmer-Jones (2011), '*High noon for microfinance impact evaluations: re-investigating the evidence from Bangladesh*.' (Working Paper 27, DEV Working Paper Series). East Anglia, UK: The School of International Development, University of East Anglia.

Duvendack, M, R Palmer-Jones, J G Copestake, L Hooper, Y Loke and N Rao (2011), '*What is the evidence of the impact of microfinance on the well-being of poor people?'.* London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.

Faux, J and L Mishel (2000), 'Inequality and the Global Economy'. In Hutton, W and A Giddens (eds.), *On the Edge: Living with Global Capitalism.* London: Jonathan Cape.

Friedman, D (1988), *The Misunderstood Miracle: Industrial Development and Political Change in Japan*. Ithaca, NY: Cornell University Press.

Galbraith, J (2008), *The Predator State: How Conservatives Abandoned the Free Market and Why Liberals Should Too.* New York: Free Press.

George, A (2006), Response to 'Death by microcredit', September 25th. Available from http://mifin.wordpress.com/

Gowan, P (1999), *The Global Gamble: Washington's Faustian bid for World Dominance.* London: Verso.

Harper, M (2007), 'Some final thoughts'. In: Dichter, T and M Harper (eds.) (*op cit*).

Harper, M (2011), 'The commercialisation of microfinance: resolution or extension of poverty?'. In: Bateman, M (2011), (*op cit*).

Harvey, D (2006), *A Brief History of Neoliberalism*. Oxford: Oxford University Press.

IADB (Inter-American Development Bank) (2010), *The Age of Productivity: Transforming Economies from the Bottom Up.* Washington DC: IADB.

ILO (International Labour Organisation) (1972), *Employment, Incomes and Equality: A Strategy for Increasing Productive Employment in Kenya*. Geneva: ILO.

ILO (International Labour Organisation) (2009), *The Financial and Economic Crisis: A Decent Work Response.* Geneva: ILO.

Karlan, D and J Zinman (2009), *Expanding Microenterprise Credit Access: Using Randomized Supply Decisions to Estimate the Impacts in Manila (*mimeo).

Karnani, A (2007), 'Microfinance misses its mark'. *Stanford Social Innovation Review* (Summer).

Karnani, A (2011), *Fighting Poverty Together: Rethinking Strategies for Business, Governments, and Civil Society to Reduce Poverty*. New York: Palgrave MacMillan.

Klas, G (2011), *Die Mikrofinanz-Industrie: Die große Illusion oder das Geschäft mit der Armut* (The Microfinance Industry: The Great Illusion or the Business of Poverty), Hamburg: Association A.

Lall, S (1996), *Learning from the Asian Tigers: Studies in Technology and Industrial policy.* London: MacMillan Press.

Levy, S (2007), '*Informality, productivity and growth in Mexico*' presentation at the World Bank ABCDE Conference 'Private Sector and Development' Bled, Slovenia, May 17-18th.

Leys, C (2001), *Market-driven politics: Neo-liberal democracy and the public interest*. London: Verso.

Mader, P (2011), '*Attempting the production of public goods through microfinance: the case of water and sanitation'*. Paper presented at the University of Pula 5th International Scientific Conference 'Entrepreneurship and Macroeconomic Management: Reflections on the World in Turmoil', Croatia, March 2011.

Manji, A (2006), *The Politics of Land Reform in Africa: From Communal Tenure to Free Markets*. London: Zed Books.

Minsky, H (1986), *Stabilising an Unstable Economy*. New York: McGraw Hill Professional.

Moyo, D (2009), *Dead Aid: Why Aid Is Not Working and How There Is a Better Way for Africa*. New York: Farrar, Straus and Giroux.

Norberg-Hodge, H, T Merrifield, and S Gorelick (2002), *Bringing the Food Economy Home: Local Alternatives to Global Agribusiness*. London: Zed Press.

North, D (1990), *Institutions, Institutional Change and Economic Performance.* Cambridge: Cambridge University Press.

Oi, J (1995), 'The role of the local state in China's transitional economy', *The China Quarterly*, No 144.

Otero, M, and E Rhyne (1994), *The New World of Microenterprise Finance: Building Healthy Institutions for the Poor*. London: Intermediate Technology Publications.

Pitt, M and S Khandker (1998), 'The impact of group-based credit programs on poor households in Bangladesh: does the gender of participants matter?', *Journal of Political Economy* 106 (5), pp. 958-996.

Pollin, R, M W Githinji and J Heintz (2008), *An Employment targeted Economic Programme for Kenya.* Cheltenham: Edward Elgar.

Popli, G (2008), *Trade liberalization and the self-employed in Mexico*. UNU-WIDER Research Paper No. 2008/05.

Pretty, J (2005) (ed.), *The Earthscan Reader in Sustainable Agriculture.* London: Earthscan Publications.

Pyke, F (1992), *Industrial Development through Small-firm Cooperation*. Geneva: ILO.

Reinert, E (2007), *How Rich Countries Became Rich, and Why Poor Countries Stay Poor*. London: Constable.

Rhyne, E (2001), *Mainstreaming Microfinance: How lending to the poor began, grew, and came of age in Bolivia.* West Hartford, Connecticut: Kumarian Press.

Robinson, M (2001), *The Microfinance Revolution: Sustainable Finance for the Poor.* Washington DC: World Bank.

Rolfe, R, D Woodward, A Ligthelm and P Guimaraes (2010), 'The viability of micro-enterprise in South Africa.' Paper presented at the Conference on 'Entrepreneurship in Africa', Whitman School of Management, Syracuse University, Syracuse, New York, April 1-3.

Roodman, D and J Morduch (2009), 'The impact of microcredit on the poor in Bangladesh: revisiting the evidence.' *Working Paper Number 174* (June). Washington DC: Center for Global Development (Revised version 2011).

Schumpeter, J (1987) (1942), *Capitalism, Socialism, and Democracy*, 6th edition. London: Unwin.

Sender, J and D Johnston (2004), 'Searching for a weapon of mass production in rural Africa: unconvincing arguments for land reform.' *Journal of Agrarian Change,* Vol. 4 Nos. 1 and 2, January and April, pp. 142–164.

Shiva, V (2002), *Water Wars; Privatization, Pollution, and Profit*. Cambridge, Massachusetts: South End Press.

Sinclair, H (2012), *Confessions of a Microfinance Heretic: How Microlending Lost Its Way and Betrayed the Poor*. San Francisco: Berrett-Koehler

Srinivasan, N (2010), *Microfinance India: State of the Sector Report 2010.* London: Sage Publications.

Sriram, M S (2010), *Commercialisation of microfinance in India: a discussion on the emperor's apparel*. (Working paper, March). Ahmedabad: Indian Institute of Management.

Stiglitz, J (2002), *Globalisation and its Discontents*. London: Allen Lane.

Taylor, M (2011), 'Freedom from poverty is not for free': rural development and the microfinance crisis in Andhra Pradesh, India.' *Journal of Agrarian Change*, Vol 11 (4), October, pp. 484-504.

Thompson, E P (1963), *The Making of the English Working Class*. London: Victor Gollancz.

Thun, E (2006), *Changing lanes in China: Foreign direct investment, local government, and auto sector development*. Cambridge: Cambridge University Press.

UNCTAD (2003), *Trade and Development Report 2003*. Geneva: UNCTAD

Wade, R (1990), *Governing the Market*. Princeton, New Jersey: Princeton University Press.

Weber, H (2002), The imposition of a global development architecture: the case of microcredit, *Review of International Studies,* Vol. 28, No.3, pp. 537–55.

Weiss, L (1988), *Creating Capitalism: The state and small business since 1945*. Oxford: Blackwell.

Yunus, M (1989), 'Grameen Bank: organization and operation'. In: Levitsky J (ed.) *Microenterprises in Developing Countries*. London: Intermediate Technology Publications.

Yunus, M (1997), Remarks from Dr Muhammad Yunus's acceptance speech given on the occasion of his receiving the Help for Self-help Prize of the Stromme Foundation, September 26, 1997 in Olso, Norway. See Vol 1 (2), November/December, *Newsletter of the Microcredit Summit Campaign*.

Zamagni, S, and V Zamagni (2010), *Cooperative Enterprise: Facing the Challenge of Globalization*, Cheltenham: Edward Elgar.

External Fragility or Deindustrialization: What is the Main Threat to Latin American Countries in the 2010s?[1]

Roberto Frenkel
Principal Research Associate at CEDES and Professor at University of Buenos Aires, Argentina

Martín Rapetti
Associate Researcher at CEDES and researcher at IIEP, University of Buenos Aires, Argentina

Abstract

In this paper we evaluate whether the surge of capital inflows to Latin American countries after the 2007-08 global financial crisis poses a threat for these economies. Recent IMF's documents have warned that capital inflows could lead to boom-and-bust cycles ending up in external and financial crises as in the past. We provide evidence that the external conditions of these economies are far more robust than in periods prior to crises. The evidence that Latin American countries are not showing signs of external fragility does not imply, however, that the current flow of capital does not pose a threat for them. In our view, capital inflows could harm economic development in the region by weakening the expansion of modern tradable activities. We show that capital inflows have induced an appreciation of real exchange rates and a deterioration of tradable sector profitability. Signs of deceleration of growth in manufactures and tradable services have started to emerge.

Key words: Real exchange rate, Latin America, Dutch disease, economic development

1. Introduction

This paper analyzes the challenges posed by persistent capital inflows to Latin America that started in late 2009. Several countries in the region experienced boom-and-bust cycles in the past, all of them associated with capital inflows. Based on these experiences, some analysts have recently begun to warn about the threats related to current flows of capital to the region. Although this is a valid concern, we believe that the main threat to Latin America lies not so much on the possibility of crises in some future, but on the effect of capital inflows on the real exchange rate (RER)[2]. More concretely, our concern is that capital inflows may lead to excessive RER appreciation, which could damage the profitability of manufacturing activities, reduce employment and productivity and ultimately hurt the development prospects of the region.

The paper is organized as follows. After this introduction, we analyze the external context that most Latin American countries are facing today and argue why it is likely to persist in the foreseeable future. In section 3, we review the evolution of RERs in Latin America during the last two decades and suggest that current levels are overvalued. In section 4, we show that the appreciation

[1] An earlier version of this paper was prepared for the Economic Development Division of ECLAC (Frenkel and Rapetti 2011). The authors thank ECLAC for the authorization to translate the article into English and to publish it in this journal. The authors also want to thank Osvaldo Kacef, Jaime Ros, Jayati Ghosh, John Weeks and, especially, Norbert Häring for their thoughtful comments, and Eleonora Tubio and Emiliano Libman for their research assistance. The authors declare that they have no personal or professional link to any industry, company or institution mentioned in this paper and that there is no material conflict of interest.

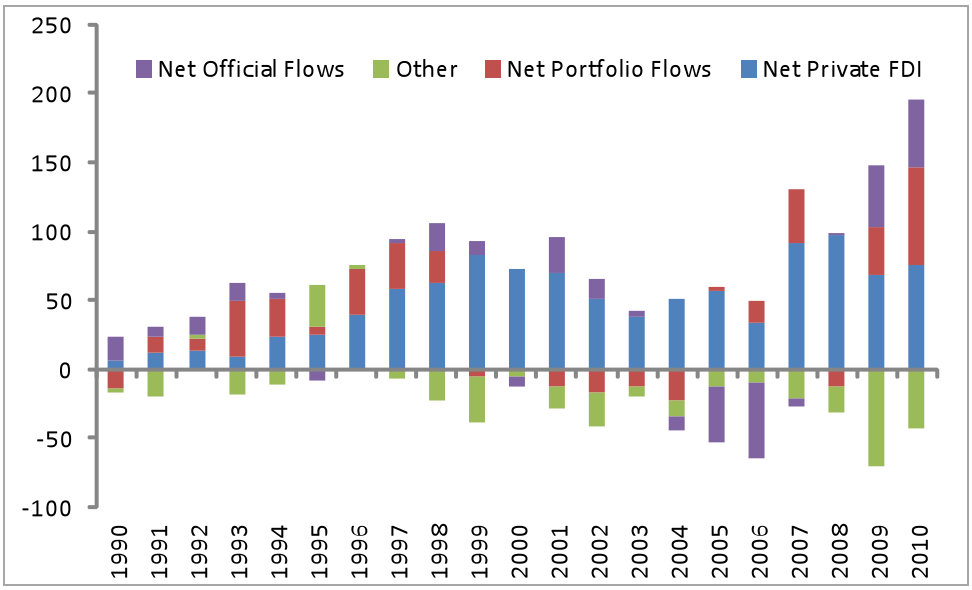
[2] We follow the definition of nominal exchange rate as the domestic price of a foreign currency. A rise (fall) in the exchange rate implies a depreciation (appreciation) of domestic currency. Similarly, a higher (lower) RER implies a real depreciation (appreciation).

of RERs have resulted in competitiveness loss in manufacture activities. We also show some evidence suggesting that the loss of competiveness is affecting the performance of manufacture and services activities. Finally, in section 5, we present a proposal for the conduct of macroeconomic policy to avoid excessive RER appreciation and its negative effects on employment and productivity in tradable activities.

2. This is just the beginning

It seems clear that the wave of capital inflows to emerging markets starting circa 2010 is influenced by the high returns that assets from these countries offer in comparison with those from advanced countries. Certainly current low GDP growth and interest rates in advanced countries are not permanent phenomena. Their real and financial yields will both probably rise in some future. However, we believe that the high rates of growth that emerging markets have been experiencing since the early 2000s will continue. This seems to us a more persistent phenomenon. Although growth rates in emerging markets and advanced countries had shown a high correlation since the 1980s, they started to diverge in the 2000s for the first time in the period of financial globalization (IMF 2010). This trend has persisted during and after the global financial crisis of 2007-08.

Figure 1: Net capital inflows to Latin America and the Caribbean (in billions dollars)



(FDI – foreign direct investment)

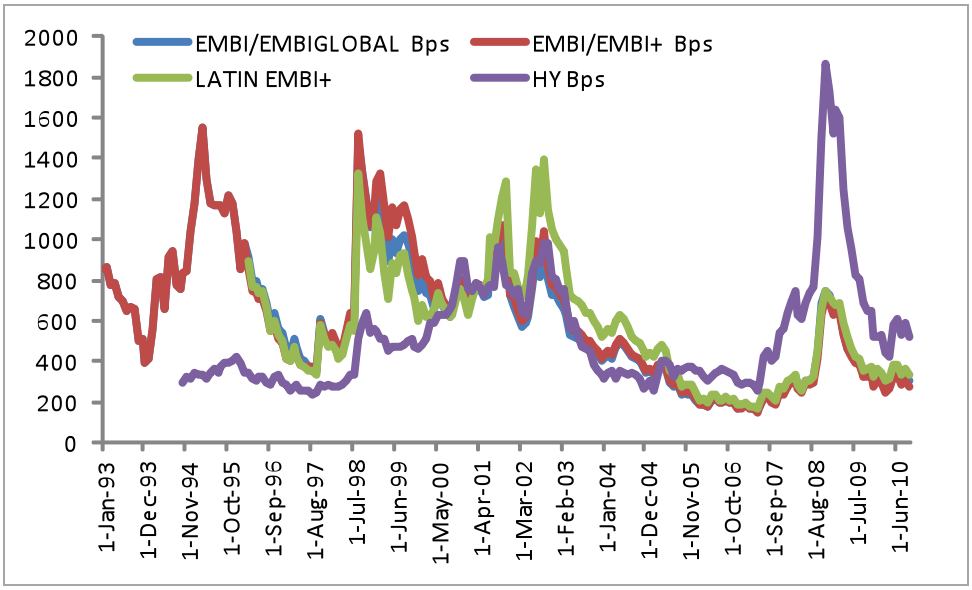
Source: *World Economic Outlook*, April 2011, International Monetary Fund

Besides the yield differentials, current capital inflows are determined by the reduction in the perceived risks in emerging markets. Regarding this factor, important changes have been observed in the way these economies participate in international financial markets since the Asian and Russian crises in 1997-98 (Frenkel and Rapetti 2010a). One key change was the switch from current account deficits to surpluses in their balance of payments, which also involved a change in the direction of net capital flows between advanced countries and emerging markets. Other relevant changes that reduced perceived risks are the substantial accumulation of foreign exchange reserves and the implementation of more flexible exchange rate regimes. These changes helped reduce the segmentation of emerging market assets and also the risks of contagion and herd behavior within this class of assets. As a result, the reduction in the perceived risks also spread to those emerging

market economies that kept running current account deficits or did not move towards more flexible exchange rate regimes.

Due to these factors, risk premia in emerging markets followed a falling trend since late 2002. By mid-2005 they went below the minimum levels reached before the Asian crises in 1997-98 and by early 2007 they reached historical minimum levels. This trend was reversed in mid-2007 once concerns about the US housing and financial markets became apparent. The resulting jump in emerging market risk premia during the subprime crises was, however, short-lived and since early 2009 they began to fall again. Figure 2 shows that the above description applies to sovereign risk premia in Latin America. The more accentuated reduction in Latin American countries risk premia compared to emerging markets average since the early 2000s is attributable to Argentina and Brazil's sovereign risk premia, which began the decade from very high levels.

Figure 2: Sovereign risk premia in emerging markets and Latin America, and US high-yield bonds spread (in basic points)



(EMBI – Emerging Markets Bond Index; EMBI+ – Emerging Markets Bond Index Plus; Bps – basis points; HY – High Yield; EMBIGLOBAL – Emerging Markets Bond Index Global)

Source: Bloomberg

The global financial crisis was a stress test for emerging markets. With the exception of a few European countries, none of them suffered external or financial crises and there was no sovereign debt default. Moreover, the same pattern of international financial integration persisted after the crisis. The increase in the IMF's financial resources and the flexibility of its assistance programs also played an important role in the prevention of crises in emerging markets. Overall, the results of the stress test and the changes in the IMF reinforced the previous perception about the profitable opportunities in emerging markets. Thus, we expect that the low risks and capital inflows to emerging markets will continue in the foreseeable future.

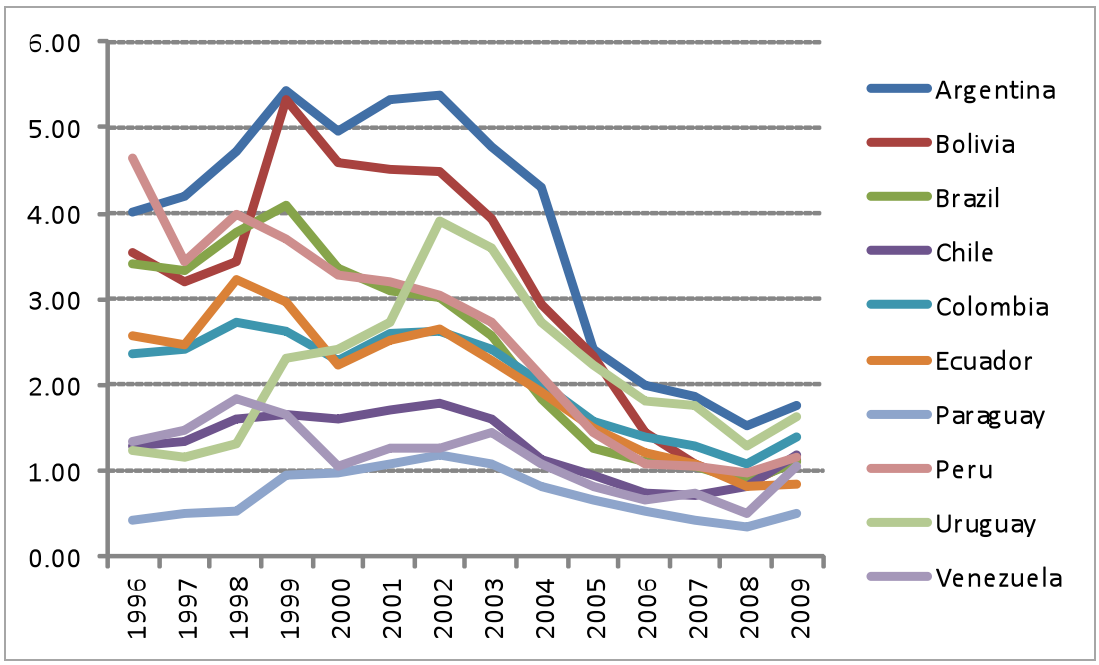
We now turn our analysis to Latin America. Between 2003 and 2007, the region as a whole ran a current account surplus. In 2008, it turned into a deficit that widened up until 2010, when it reached a local maximum. In fact, Mexico, Colombia and most of Central American and Caribbean countries had already been running current account deficits all along the 2000s. Thus, the dynamics described above resulted from the behavior of most South American countries. Assuming no major changes in current economic policies, forecasts – including those of the IMF (IMF 2011a) – indicate that current account deficits in these economies will tend to widen.

Do increasing current account deficits represent a threat in terms of external and financial crises as they did in the past? The experiences of capital inflow booms that ended up in crises in Latin America resulted from sustained current account deficits that led to excessive foreign debt accumulation. Rapid foreign debt accumulation and rising current account deficits in these experiences occurred in contexts of fixed exchange rate regimes and appreciated RERs (Frenkel and Rapetti 2009). The observed rise in sovereign risk premia was precisely associated with the perception that countries in those contexts would have a hard time meeting debt services. None of these features are currently observed in Latin American countries.

First, most countries in the region have adopted flexible exchange rate regimes – mostly, managed floating regimes – and have been accumulating large stocks of foreign exchange reserves. These elements give monetary authorities greater flexibility to absorb negative external shocks and to avoid sharp exchange rate corrections in contexts of low liquidity of foreign exchange.

Second, foreign debts in Latin America shrank substantially during the 2000s and reached historically low levels, as shown in Figures 3 and 4. The emergence of current account deficits in this situation is novel for the countries in the region. Since their reincorporation to the international capital markets in the late 1980s, Latin American countries have been dealing with heavy debt burdens inherited from the debt crises in the early 1980s. The new configuration suggests that most countries have substantial margins to accumulate foreign debt before reaching high debt-to-GDP ratios.

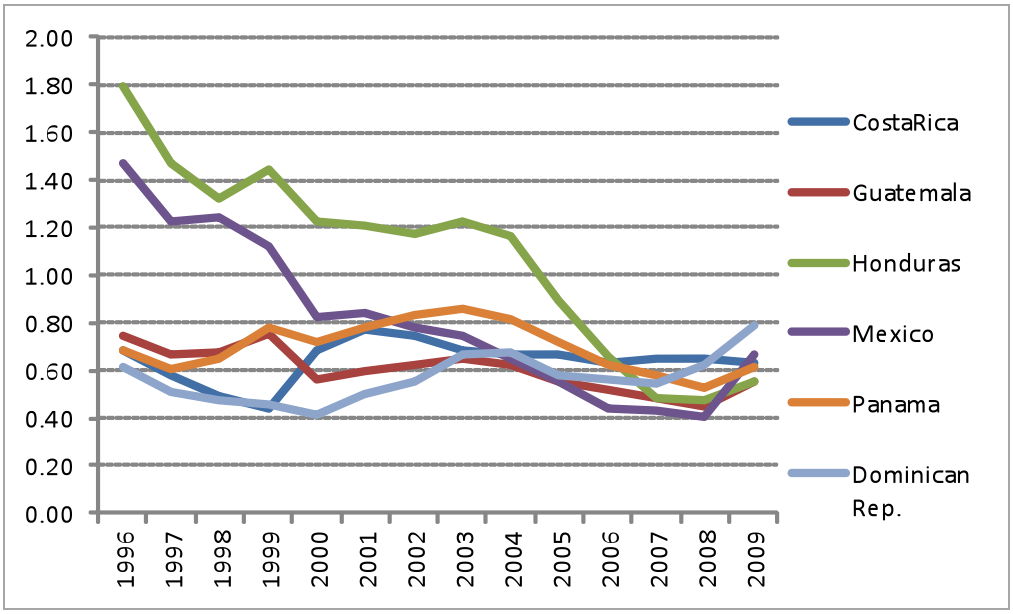
Figure 3: Foreign debt/total exports, South America



Source: CEPALSTAT, ECLAC[3]

[3] In the case of Argentina, all calculations including the CPI were re-calculated using the IPC-7 series elaborated by CENDA. The data on Brazilian exports were obtained from IPEADATA. (See Appendix.)

Figure 4: Foreign debt/total exports, Mexico and Central America



Source: CEPALSTAT, ECLAC

There is another important element making the threat of crises even less likely. Because of the reduction of foreign debts during the 2000s, the weight of interest payments in the factor income account of the balance of payments has reduced significantly. In contrast to the previous 30 years of financial globalization, current account deficits in most Latin American countries are now largely influenced by dividend payments of foreign direct investment (FDI). This represents an important change. Interest payments have to be paid in foreign currency – typically US dollars – and since they are contractual obligations, they constitute a source of foreign currency outflow that is delinked from the business cycle. On the contrary, FDI dividends are largely obtained in domestic currency – making their value in foreign currency depend on the exchange rate – and are highly correlated to the business cycle. This implies that in the case of a capital inflow deceleration or reversal, the magnitude of FDI dividend payments tends to contract due to both the depreciation of the domestic currency and the deceleration or contraction of domestic economic activity. Furthermore, a significant portion of FDI dividends are normally re-invested in the recipient economy – being registered in the balance of payments as a new inflow of capital – without even going through the foreign exchange market. This implies that part of factor income account deficits has a relatively automatic source of funding. Finally, in cases of severe scarcity of foreign exchange, authorities can impose transitory restrictions on the remittance of FDI dividends to alleviate the excess demand for foreign exchange.

Table 1 illustrates the change in the composition of current accounts in some Latin American countries. The current account deficit of Brazil was 49.5% of total exports in 1999. This deficit was almost equivalent to the gross factor income remittances (41%). 39.7% of these remittances corresponded to interest payments and the other 60.3% to FDI dividend payments. The current account deficit represented only 20.3% of total exports in 2010. Once again, gross factor income remittances were virtually of the same magnitude as the current account deficit (20%). The difference is that in 2010 virtually all of these remittances (88.6%) corresponded to FDI dividend payments.

Table 1: Current account of the balance of payments and foreign direct investment (as a share of total exports)

		Current Account	Trade Balance	Transfers	Factor income (debit)	Interests (debit)	Utilities (debit)	Foreign Direct Investment
Brazil	1999	-45.9%	-14.8%	3.1%	-41.0%	-16.3%	-24.7%	51.8%
					100%	*39.7%*	*60.3%*	
	2010	-20.3%	-4.6%	1.2%	-20.0%	-2.3%	-17.8%	20.7%
					100%	*11.4%*	*88.6%*	
Chile	1999	0.5%	8.0%	3.1%	-14.7%	-6.0%	-8.7%	41.7%
					100%	*40.8%*	*59.2%*	
	2010	4.7%	18.1%	5.4%	-25.0%	-1.9%	-23.2%	18.5%
					100%	*7.4%*	*92.6%*	
Colombia	1999	4.8%	4.1%	10.4%	-16.2%	-13.4%	-2.8%	10.8%
					100%	*82.8%*	*17.2%*	
	2010	-19.4%	2.0%	10.0%	-30.7%	-8.1%	-22.7%	23.5%
					100%	*26.3%*	*73.7%*	
Mexico	1999	-9.5%	-5.6%	4.2%	-11.1%	-8.4%	-2.7%	9.4%
					100%	*75.7%*	*24.3%*	
	2010	-1.9%	-4.1%	6.9%	-6.3%	-4.0%	-2.3%	5.7%
					100%	*63.8%*	*36.0%*	
Peru	1999	-17.9%	-15.7%	12.2%	-22.9%	-21.5%	-1.4%	25.2%
					100%	*93.7%*	*6.3%*	
	2010	-5.9%	11.9%	7.7%	-28.4%	-2.7%	-25.8%	18.5%
					100%	*9.3%*	*90.7%*	

Source: CEPALSTAT, ECLAC[4]

In Chile, similar to developments in Brazil, between 1999 and 2010 the proportion of interest payments went from 40.8% to 7.4%, It went from 82.8% to 26.3% in Colombia and from 93.7% to 15.6% in Peru. The exception has been Mexico, where the proportion only shrank from 75.7% in 1999 to 63.8% in 2010. In all the countries with current account deficits in 2010 (Brazil, Colombia, Peru and Mexico), the deficits were entirely financed with FDI, a large proportion of which was re-investment of dividends.

Because of the reasons discussed above, we do not see signs of excessive external fragility in Latin America and therefore do not think that the current wave of capital inflows represents a threat in terms of immediate crises. This assessment may look more 'optimistic' than a recent evaluation carried out by the IMF (IMF 2011b), in which the institution warns about the increase in current account deficits in the region and the potential risks of capital flow reversals. This concern was a reason behind the recent change in the IMF's view about the benefits of capital inflows to developing countries. The institution now promotes a more cautious approach and it even suggests that countries should consider the possibility of adopting capital controls transitorily (IMF 2011b).

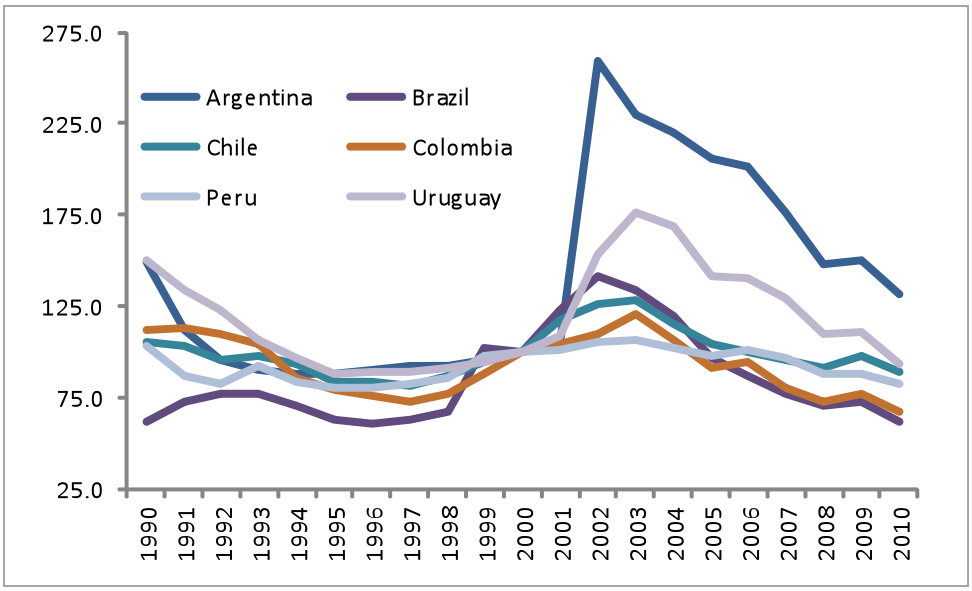
Our assessment that sudden stops and crises are not highly likely in the immediate future should not be understood as a statement that the current flow of capital is harmless and does not pose any threats to Latin American countries. Quite on the contrary, we believe that national authorities should worry about them and adopt measures to discourage them and to mitigate their effects. Our concern, however, is not so much the possibility of crises but the effects that capital inflows have on the real economy via their effects on real exchange rates (RER). More concretely, our concern is that massive capital inflows to Latin America may have pernicious effects via an excessive appreciation of the RER, which could lead to a contraction in output and employment in tradable activities with negative effects on long-run growth. Our concern, in other words, is the possibility of Dutch Disease.

[4] Data on balance of payments, external debt, bilateral RER with the US, effective RER, wages and value added in constant prices are all from CEPALSTAT, ECLAC. The data on Brazilian exports were obtained from IPEADATA. (See Appendix)

3. The evolution of real exchange rates in Latin America

Between 1990 and 2010, the behavior of RERs in South American has been different from those in Mexico, Central America and the Caribbean. In the first group, RERs tended to appreciate between the early 1990s until the eruption of the Asian and Russian crises in 1997-98. To deal with these shocks, Brazil, Chile and Colombia adopted floating and inflation targeting regimes in 1999. Peru had been using a managed floating regime since the early 1990s and only formally adopted inflation targeting in 2002. Argentina and Uruguay maintained fixed exchange rates and overvalued RERs until the 2001-02 crises and then opted for managed floating arrangements. In all these countries, RERs reached maximum levels in 2002-03 and then followed a persistent appreciation trend, transitorily interrupted by the effect of the subprime crisis and its global contagion. Within this group, Argentina's RER has been the most volatile and Peru's the least. Figures 5 the evolution of the bilateral RERs with the US between 1990 and 2010 for South American countries.

Figure 5: Bilateral real exchange rates with the US, South America (100=2000)



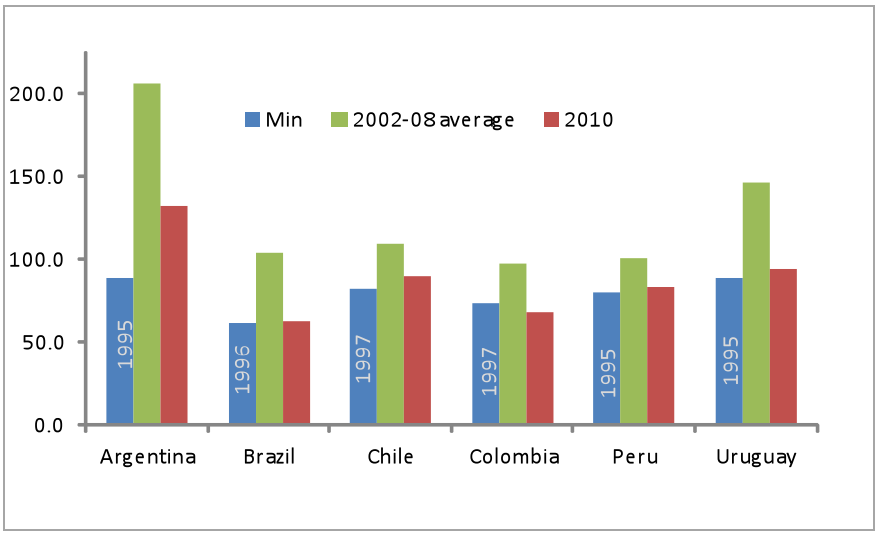
Source: CEPALSTAT, ECLAC[5]

There are some aspects worth highlighting. First, within each country, 2002-03 is the period in which RERs reached their highest levels since countries regained access to the international financial markets in 1990. Second, the rise of the RERs in the early 2000s improved current account balances before the commodities prices boom started circa 2004-05. Third, because of the high levels at which they started, during 2002-2008 RERs remained on average relatively high compared to the 1990s, despite their persistent downward trend. Fourth, the rise of RERs of 2008-09 represented only a mild and transitory detour from their downward trend.

Figure 6 helps to give a neater view of the fall of the RERs experienced in South America during the 2000s. With the exception of Argentina's, RER levels in 2010 were similar to the lowest levels in the 1990s.

[5] In the case of Argentina, all calculations including the CPI were re-calculated using the IPC-7 series elaborated by CENDA. (See Appendix.)

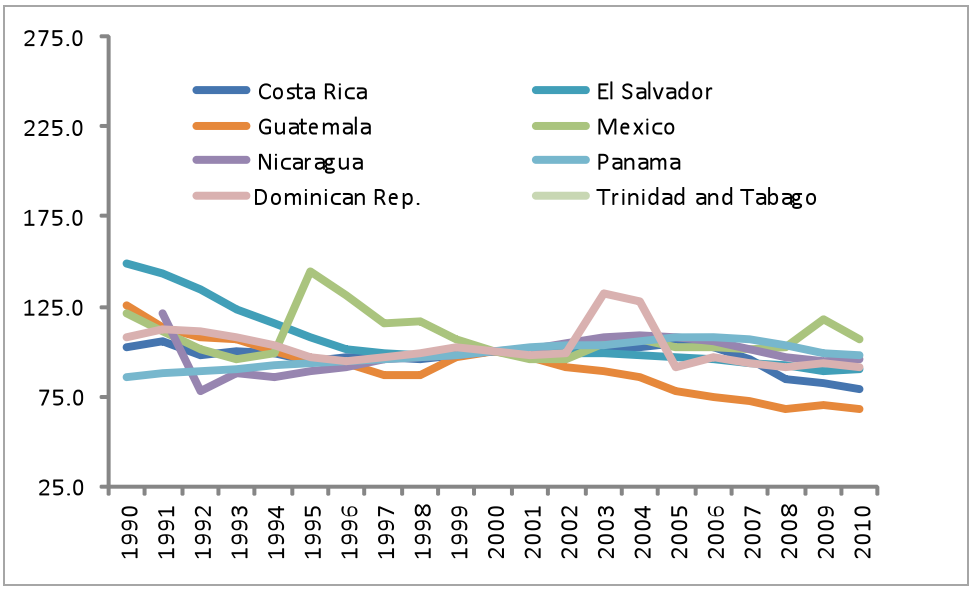
Figure 6: Bilateral real exchange rates with the US, South America. Minimum levels in the 1990s, 2002-08 average and 2010 (100=2000)



Source: CEPALSTAT, ECLAC[6]

The behavior of bilateral RERs with the US in Mexico, Central America and the Caribbean has been different. Figure 7 shows that they were substantially less volatile and that most of them followed a persistent downward trend all through the two decades.

Figure 7: Bilateral real exchange rates with the US, Mexico, Central America and the Caribbean (100=2000)



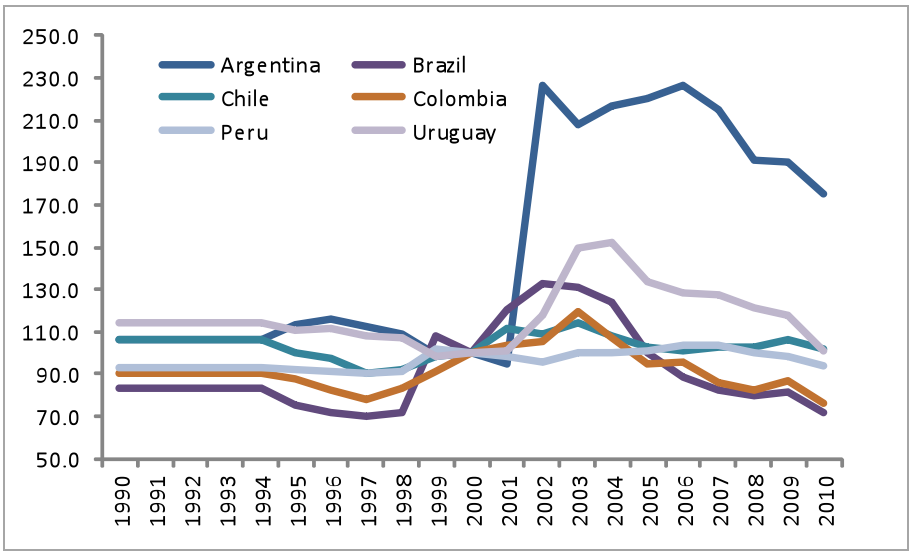
Source: CEPALSTAT, ECLAC

Figures 8 and 9 show the evolution of effective RERs in Latin American countries. In the case of Mexico, Central America and the Caribbean bilateral and effective RERs are very similar due to the

[6] In the case of Argentina, all calculations including the CPI were re-calculated using the IPC-7 series elaborated by CENDA. (See Appendix.)

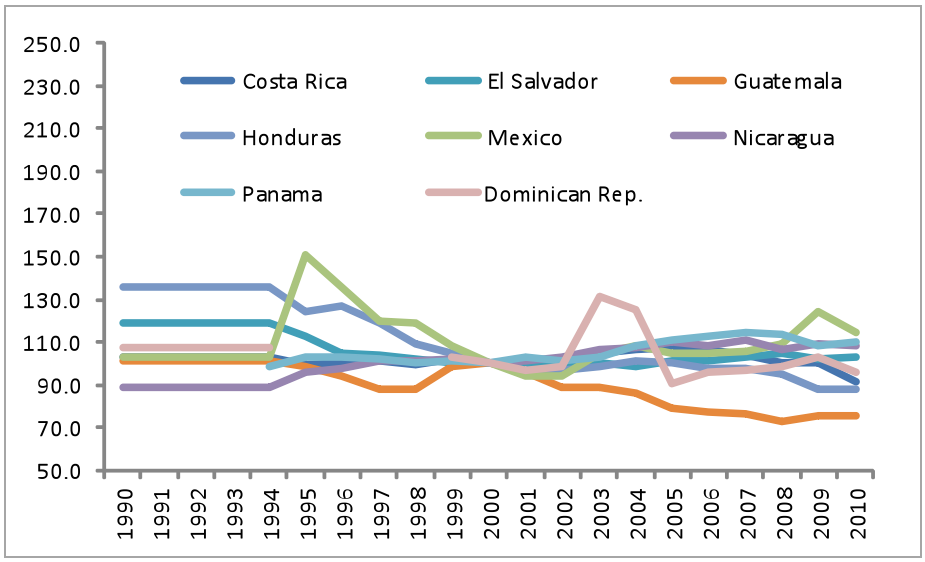
high weight that the US has on these countries' international trade. Effective RERs in South American countries are less volatile than bilateral RERs. This is because trade between these countries is high and therefore the correlation of bilateral RERs observed in Figure 5 reduces the volatility of effective RERs. Despite the fact that competitiveness gains and losses against all trade partners are less pronounced than against the US exclusively, the trajectories followed by effective RERs have been similar to those of bilateral RERs. The 2010 levels of effective RERs in South America were also similar to the minimum values of the 1990s as Figure 10 illustrates.

Figure 8: Effective real exchange rates in South America (100=2000)



Source: CEPALSTAT, ECLAC[7]

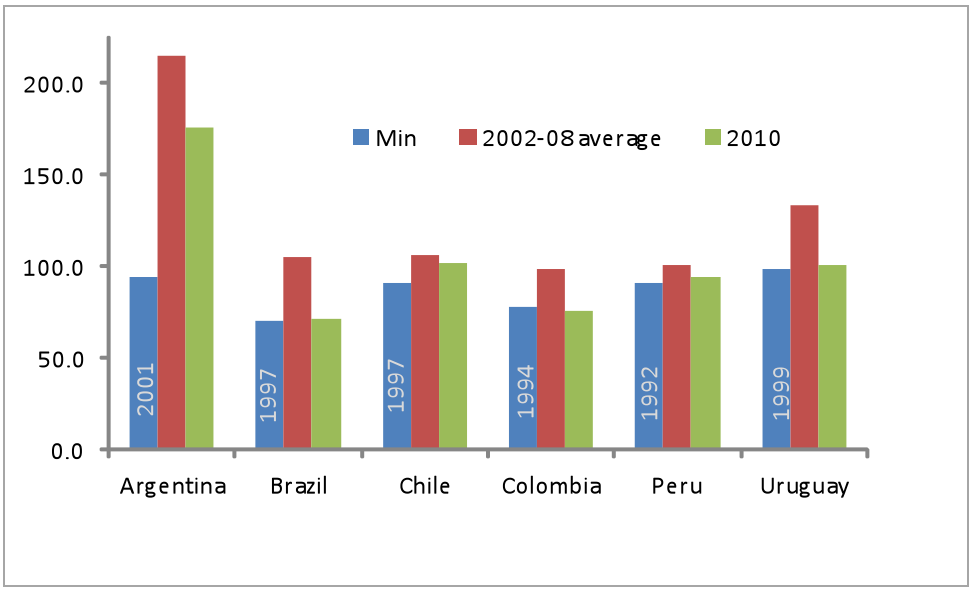
Figure 9: Effective real exchange rates, Mexico, Central America and the Caribbean (100=2000)



Source: CEPALSTAT, ECLAC

[7] In the case of Argentina, all calculations including the CPI were re-calculated using the IPC-7 series elaborated by CENDA. (See Appendix.)

Figure 10: Effective real exchange rates, South America.
Minimum levels in the 1990s, 2002-08 average and 2010 (100=2000)



Source: CEPALSTAT, ECLAC

4. Capital inflows, Dutch disease and economic development

The case in which massive capital inflows appreciate the RER leading to a contraction in output and employment in the manufacturing sector is commonly indicated as a variant of the Dutch Disease phenomenon. Some authors conceive Dutch Disease as an equilibrium outcome with no relevant effect on long-run economic growth.[8] Under this view, a positive shock – like the discovery of an oil field, a permanent increase in the price of an agricultural or mineral commodity or even a sustained flow of FDI – would represent an increase in national wealth. The consequent rise in the actual and expected flow of foreign currency income would lead to an equilibrium appreciation of the RER. With a more appreciated (ie. lower) RER, some other tradable activities – manufactures and services – would become uncompetitive and would perish against international competition. This outcome would not be problematic because the labor freed by these activities would be absorbed by the expanding sectors. This type of de-industrialization would be an equilibrium outcome and *a priori* would not affect long-run economic growth. Under this perspective, current capital inflows to Latin America – unless they represent a transitory phenomenon – should not be a source of concern for national authorities.

We find this view problematic for several reasons. First, it is not clear that capital inflows represent an increase of the recipient country's wealth in foreign currency, as in the standard case of Dutch Disease. Capital inflows can be a source of finance for a current account deficit – which would imply an increase in net foreign debt – or an exchange of foreign for domestic assets – without altering the net international investment position. A green field foreign investment certainly represents an increase in the capital stock of the recipient country, but it is typically made with the expectation that the discount value of future dividends will be higher than the original investment. Second, it is impossible from the viewpoint of a policy-maker to know *ex-ante* whether a wave of capital inflows represents a transitory or permanent phenomenon. Third, it is equally uncertain whether the labor displaced from the industrial and services sectors resulting from the appreciation of

[8] This appears to be the view, for instance, of Magud and Sosa (2010) in their assessment of the empirical literature on Dutch Disease and on the RER-growth relationship.

the RER will be absorbed by other sectors. There is, on the other hand, a much higher degree of certainty regarding the effects of a transitory but sustained RER appreciation on industrial employment and output. Transitory but lasting RER appreciations have typically led to the destruction of firms and employment, human and organizational capital, vertical and horizontal linkages and access to foreign markets. These outcomes have been formalized (Krugman 1987 and, Ros and Skott 1998) and documented empirically (Sachs and Werner 2001). Moreover, there are several examples in Latin American economic history of sustained RER appreciation leading to de-industrialization. These are, for instance, the experiences of Argentina and Chile between late 1970s and early 1980s and that of Argentina during the 1990s[9].

We also find problematic the view that Dutch Disease and RER appreciations do not have effects on long-run growth. Economic development is associated with the expansion of modern tradable activities (ie. manufactures and services intensive in knowledge). The expansion of these activities generates a variety of positive externalities – learning-by-doing, network externalities and technological spillovers – that tend to accelerate economic growth. They also increase the net supply of foreign currency and thus reduce the possibility of stop-and-go dynamics or excessive foreign debt accumulation and crises that hamper long-run growth. For these reasons, a competitive RER provides an environment conductive to economic development by stimulating investment in tradable activities. A recent body of econometric research has found a robust association between growth acceleration and competitive RERs[10]. Moreover, this relationship has been observed in several episodes in the economic history of Latin America: the most successful cases of sustained growth accelerations have occurred when governments oriented their macroeconomic policy to sustain competitive and stable RERs that protected industrial activities and promoted non-traditional exports Frenkel and Rapetti 2012).

Consequently, current capital inflows to Latin America – even if they do not represent a threat in terms of external vulnerability and crisis – could excessively appreciate the RER, harm the development of the industrial sector and its employment level and negatively affect long-run growth. Given current conditions and our expectation about their continuation, we believe that RER levels that guarantee external sustainability in Latin American countries are more appreciated (ie. lower) than those required to promote economic development. Macroeconomic policy should care about not only the RER level that guarantees external sustainability, but also the one that promotes the expansion of modern tradable activities, employment and economic development[11].

It is therefore very important to evaluate whether manufacturing activities in Latin American countries are experiencing profitability problems that could constrain their long-run development. A simple way to do this is calculate the unit labor cost in foreign currency (ULC$), which measures domestic wage rates relative to foreign wage rates in common currency (US dollar), adjusted by relative productivity[12]. Changes in ULC$ over time indicate the evolution of the profitability of tradable activities intensive in labor, as manufactures and modern services. A rise (fall) in ULC$ suggests a fall (rise) in the profitability of these sectors. For its calculation, we use the simple average of the rate of variation of GDP per capita of the US, Germany, China and Brazil as a proxy of the rate of variation of foreign productivity. These countries influence the productivity trends in the dollar and

[9] See Ffrench-Davis (2004) for the Chilean experience and Damill (2002) for that of Argentina.

[10] See, among others, Aguirre and Calderon (2005), Gala (2007), Rodrik (2008) and Rapetti et al (2012).

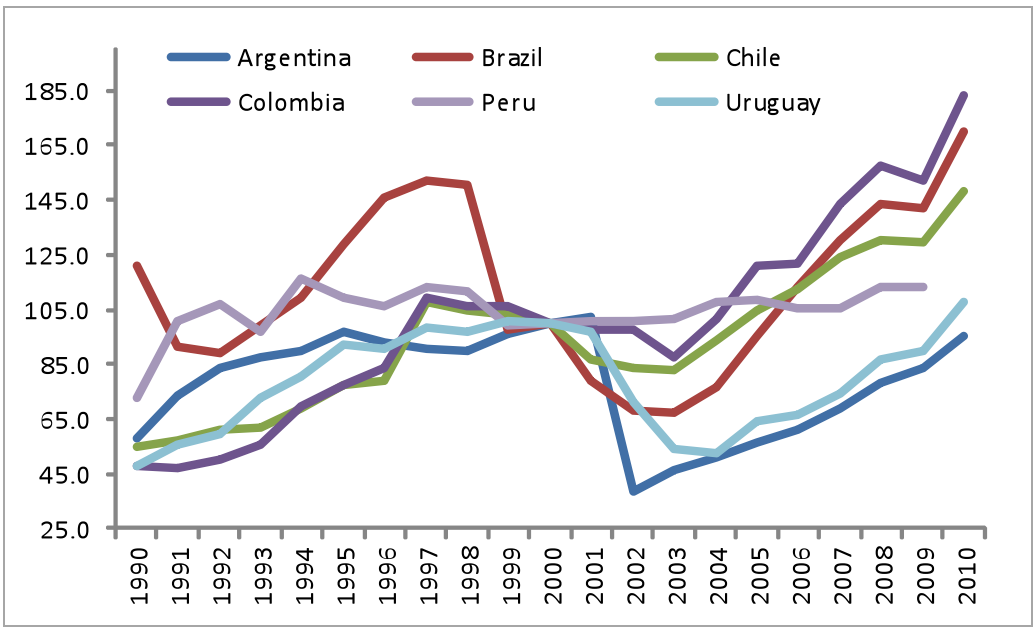
[11] The notion that the RER level that guarantees external sustainability may be more appreciated than the one required to stimulate the development of modern tradable activities was originally formulated by Diamand (1972). A recent elaboration of this argument can be found in Bresser-Pereira (2010).

[12] Formally, ULC$ = $(W / EW^*)(\rho^* / \rho)$, where W represents the nominal wage rate, E the nominal exchange rate, ρ productivity and the asterisk (*) indicates foreign country. Due to the lack of homogeneous data, we calculated ULC$ using the CPI and GDP per capita as proxies of foreign nominal wages and labor productivity, respectively. Consequently, we calculated ULC$ = $(\omega/q)(y^*/y)$, where ω is the domestic real wage rate and y GDP per capita and q, the RER. Constructed this way, the indicator implicitly assumes that foreign real wage remained constant during period of analysis.

euro areas, Asia and Latin America, respectively and thus represent a reasonable approximation of the competitive pressures that the countries in the region face.

Figure 11 shows the evolution of unit labor cost in foreign currency (ULC$) for South American countries between 1990 and 2010. It is apparent that in almost all cases ULC$ have been rising substantially and sustained since 2002-03 when there is a local minimum. In Colombia, Chile and Brazil, the 2010 levels are substantially higher than the maximum levels reached during the 1990s: +68%, +38%, and +12%, respectively. These figures suggest that tradable activities intensive in labor in these countries have been facing increasing challenges in terms of competitiveness and profitability. The trajectories of Argentina, Uruguay and Peru are similar although less accentuated. In the former two, the significant increase in ULC$ since 2002-03 has led to levels similar to the maximum levels reached before the 2001-02 crises. In Argentina, the 2010 level was still 7% lower than the 2001 level and in Uruguay 7% higher than in 1999. In Peru, ULC$ have been rising at a mild pace and its last observation (2009) was still 3% lower than the maximum level of 1994. To facilitate comparisons, Figure 12 shows the 2010 levels, the previous local maximum levels and the 2002-2008 period averages.

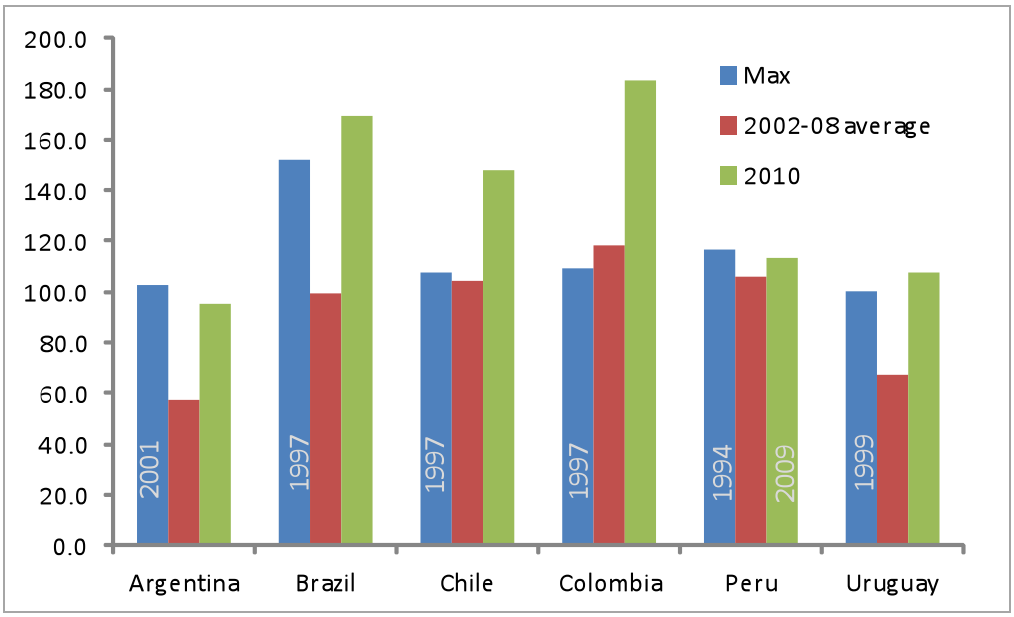
Figure 11: Unit labor costs in US dollars (ULC$), South America (100=2000)



Source: CEPALSTAT, ECLAC[13]

[13] In the case of Argentina, all calculations including the CPI were re-calculated using the IPC-7 series elaborated by CENDA. (See Appendix.)

Figure 12: Unit labor costs in US dollars (ULC$), South America. Maximum levels of the 1990s, 2002-2008 average and 2010 (100=2000)



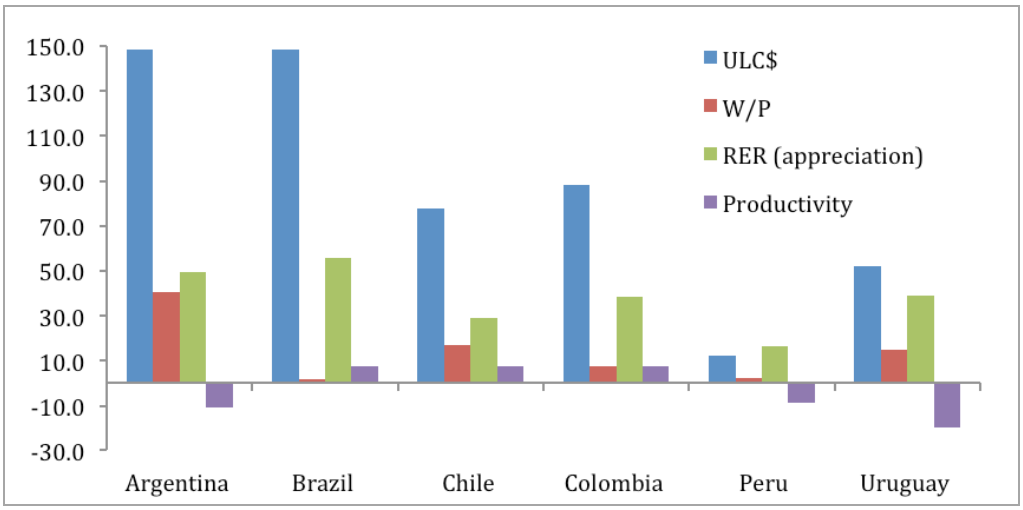
Source: CEPALSTAT, ECLAC[14]

Figure 13 decomposes the rise in ULC$ between 2002 and 2010 in three factors: real wage increases, RER appreciations and variations of relative productivity. The bars indicate the percentage variation of ULC$ and its components[15]. Some aspects are worth noting. First, RER appreciation was a key element explaining the rise in ULC$ in all cases. There is, however, an important difference (not shown in the figure) between Argentina and the rest. In Argentina, RER appreciation resulted from higher domestic inflation relative to foreign inflation whereas in the other South American countries it mainly resulted from nominal exchange rate appreciation, especially in Brazil and Colombia. Second, there are significant differences in terms of productivity growth. Whereas in Brazil, Chile and Colombia productivity grew at a slower pace than in the reference countries (the US, Germany, Brazil and China), in Argentina, Uruguay and Peru, it grew faster. Third, in Argentina real wages increased relatively more than the productivity differential; in Peru and Uruguay, they increased relatively less than the productivity differential. The different behavior of real wages in these countries is a reason why unit labor cost in foreign currency in Argentina rose more than in Peru and Uruguay.

[14] In the case of Argentina, all calculations including the CPI were re-calculated using the IPC-7 series elaborated by CENDA. (See Appendix.)

[15] The sum of all factors adds up to the total variation of ULC$ when the calculation is carried out in continuous time. Because calculations for Figure 13 were done in discrete time, the sum of the parts does not add up to the total. The reported variation of the factors, however, suggests the relative incidence of each of them on the total variation.

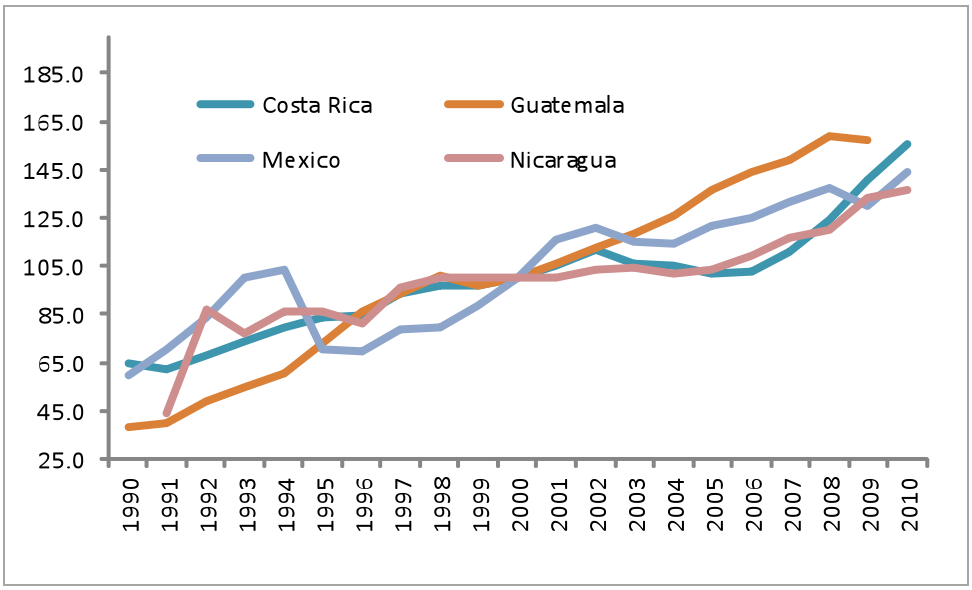
Figure 13: Decomposition of the increase in unit labor cost ULC$ (2002-2010) by explanatory factors, South America (in percentage)



Source: CEPALSTAT, ECLAC[16]

Figure 14 shows the evolution of unit labor cost in foreign currency in Mexico and some Central American countries between 1990 and 2010. ULC$ have been increasing since the 1990s in all these cases. In Mexico, the ULC$ level in 2010 was 35% higher than in 1994, a year before the currency crisis. If that year is taken as a reference of low profitability in Mexican manufacturing, then current levels suggest that the current situation is even worse. The competitive loss in Mexico and the Central American countries is largely a result of the low productivity growth relative to foreign competing countries. This can be observed in Figure 14, which replicates the decomposition of Figure 13 for Mexico and Central American countries.

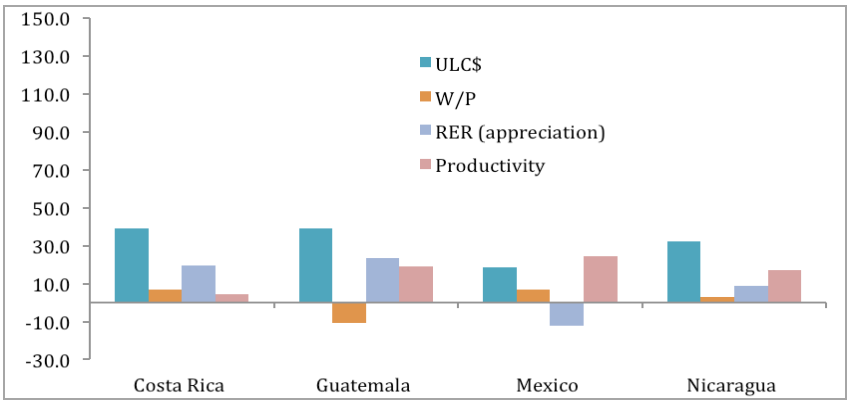
Figure 14: Unit labor cost in US dollars (ULC$), Central America and Mexico (100=2000)



Source: CEPALSTAT, ECLAC

[16] In the case of Argentina, all calculations including the CPI were re-calculated using the IPC-7 series elaborated by CENDA. (See Appendix.)

Figure 15: Decomposition of the increase in ULC$ (2002-2010) by explanatory factors, Central America and Mexico (in percentage)



Source: CEPALSTAT, ECLAC[17]

The evidence gathered so far suggests that – at least in Brazil, Chile, Colombia, Mexico and the Central American countries examined – the profitability in tradable sectors intensive in labor has shrunk substantially in the last few years. One would expect that these developments would affect the performance of these activities and their employment levels. It is, however, difficult to assess the extent to which RER appreciation hurts tradable profitability because their effects are not immediately observed. Facing a profitability squeeze, firms first absorb losses, then reduce the workday or force their workers to anticipate their vacations, then adopt defensive strategies reducing their value added and simplifying the production lines and finally go bankrupt. As a result, the adaptation to RER appreciation takes the form of a gradual contraction of value added and employment that takes time to become apparent. It is documented, for instance, that the effect of RER appreciation on employment in Latin America has operated with a 2-year lag (Frenkel and Ros 2006).

One would then expect that – if the observed RER appreciation and profit squeeze in Latin America continue – modern tradable sectors would gradually decelerate their output and employment growth and that they would eventually start contracting. There are in fact some hints indicating that tradable profit squeeze is negatively affecting the performance of manufacturing activities in Latin America. Table 2 reports the elasticity of industrial value added growth with respect to value added growth in other economic activities, both measures in constant prices for the major South American countries[18]. The analysis is made comparing two periods: 2002-05 and 2005-08. In the former period, countries experienced high GDP growth and had relatively competitive RERs. In the latter period, on the contrary, countries experienced high GDP growth with substantially more appreciated RERs[19]. Although RER levels were even more appreciated in 2010, we did not extend the period until this year because of the contraction in economic activity in 2009 introduces noises in the time series, making the interpretation of results less clear.

If the level of the RER affects positively the development of modern tradable activities (eg. industry) – as argued above – on would expect to observe a worsening in economic performance in these sectors relative to the rest of the economy. In other words, one should observe a reduction in the elasticity of modern tradable sector growth (eg. industry) with respect to other sectors' growth. Evidence in Table 2 is in line with this prediction: all the countries experienced a relative deceleration

[17] GDP per capita data correspond to the *World Development Indicators* from the World Bank. (See Appendix.)

[18] In order to avoid cyclical noises and capture growth trends, we calculated least square growth rates of value added. Specifically, we estimated the growth rate *g* by regressing the following model: $lny_t = a + gt + e_t$, where *y* is value added in constant prices, *a* is a constant, *t* represents quarters and *e* is an error term.

[19] Compared to the 2005-2008 period, in 2002-2005 the bilateral RER with the US in Argentina was 25% higher, in Brazil 48% higher, in Colombia 26% higher, in Chile 21% higher and in Peru 7% higher.

of industrial sector value added growth. In line with this evidence, there is also a positive correlation between the contraction in the elasticities and the degree of RER appreciation. For instance, Peru is the country in this sample with the least degree of RER of appreciation and also the one in which the fall in the elasticity was the smallest.

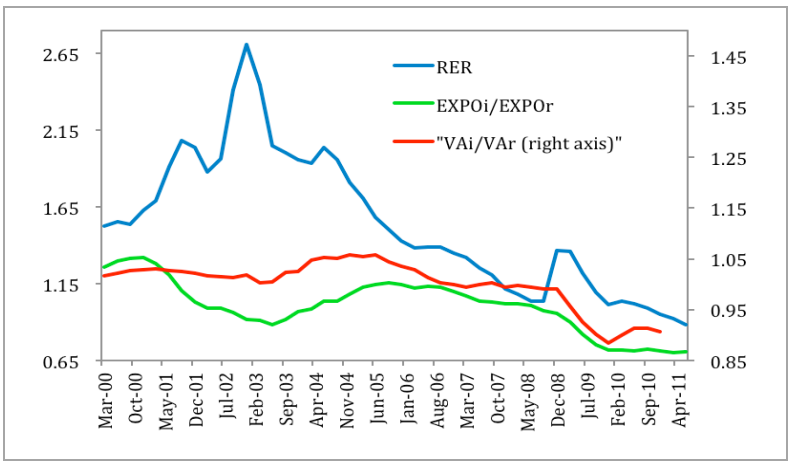
Table 2: Elasticity of industrial value added growth with respect to other sectors' value added growth (in constant prices)

	Argentina	Brazil	Colombia	Chile	Peru
2002-2005	1.48	1.78	1.33	1.14	1.28
2005-2008	0.91	1.23	0.64	0.70	1.07

Source: CEPALSTAT, ECLAC[20]

Figure 16 shows the evolution of the bilateral RER with the US and two indicators that describe the relative performance of the industrial sector in Brazil between early 2000 and early 2011. The green line indicates industrial exports relative to exports of natural resources, both in real terms. The red line does the same for industrial value added (VA_i) relative to the value added of the rest of the sectors (VA_r), also in real terms[21]. Figure 16 suggests that the performance of the Brazilian industrial sector appears to be influenced by the level of RER with a 2-year lag[22]. It is interesting to note how the relative improvement of industrial exports and value added between mid-2003 and mid-2005 was preceded by the significant RER depreciation that started in the first semester of 2001. Similarly, the RER appreciation trend beginning in mid-2004 is followed by a relative worsening in industrial exports and value added starting in late 2005 and early 2006. This relative performance continues until the end of the period of analysis in parallel with the RER appreciation trend.

Figure 16: Brazil: Bilateral real exchange rates with the US, industrial exports/natural-resource exports and industrial value added/value added (rest) (1=average 2000-2010)



(RER – real exchange rate; EXPOi – industrial exports; EXPOr – natural resource exports; VAi – industrial value added; VAr – non- industrial value added)

Source: Banco Central do Brasil

[20] In the case of Argentina, all calculations including the CPI were re-calculated using the IPC-7 series elaborated by CENDA. GDP per capita data correspond to the *World Development Indicators* from the World Bank. The data on Brazilian exports were obtained from IPEADATA. (See Appendix.)

[21] To avoid seasonal fluctuations, both indicators were calculated as annualized values.

[22] The correlation coefficients between the 2-year lagged RER and the export and value added indicators are 0.73 and 0.84, respectively.

The case of the Brazilian manufacturing sector illustrates what in our view is the main threat that Latin American countries are currently facing with the sustained RER appreciation caused by the ongoing wave of massive capital inflows.

5. A development-friendly macroeconomic policy framework

Most Latin American countries have made substantial progress in the conduct of their macroeconomic policies in recent decades. After experiencing high inflation during the 1970s and 1980s, countries in the region managed to stabilize prices during the 1990s. In most cases, price stability was achieved through stabilization programs that used the exchange rate as a nominal anchor. An undesired outcome of these programs was excessive RER overvaluations, which led to external crises[23]. An important lesson from these experiences was that in order to deal with volatile capital movements and to avoid external crises, macroeconomic policy not only needs sound monetary and fiscal management but also an exchange rate policy that avoids RER overvaluation and combines exchange rate flexibility and foreign exchange reserves accumulation.

Current massive capital inflows pose a threat to Latin American countries. In this article, we made the case that these economies do not show signs of external vulnerability and that there is no clear evidence that the external conditions they currently face will dramatically change in the foreseeable future. Hence, we do not see a scenario of crises very likely.

This assessment is based on our conjectures about the future; and we know the future is, by its own nature, uncertain. Conjectures about the future – all of them, including ours – necessarily have to deal with uncertainty. Will the present favorable terms of trade persist? Will current external financial conditions remain in Latin America? About these things, we cannot certainly know. Economic authorities should be especially cautious in the face of uncertainty. In this regard, the design of economic policy should follow two principles. First, it should include all the elements to assure that the proposed goal is achieved in the foreseeable scenarios. The second principle is to minimize the potential damage that an economic policy could provoke if the conjectures in which it is based are finally wrong.

Following these principles, a prudential attitude would suggest implementing measures to offset or mitigate the effects of capital inflows. These measures should be adopted not only to avoid the formation of domestic asset bubbles and control inflation but also to avoid external and financial crises, and consequently a huge damage. Thus, although we have a different perception of the risks involved, we fully agree with the position that the IMF has taken recently about taking a prudential approach about capital inflows. We might be wrong in our assessment and thus countries should take the possibility of crisis seriously. But a prudential economic policy design should broaden the consideration of potential negative effect of capital inflows and include those associated with Dutch Disease. These effects should be taken as seriously as those associated to the risks of external and financial crises because they are largely irreversible. As argued above, it is well documented both theoretically and empirically that a transitory RER appreciation can have long-lasting effects on the manufacturing sector in the form of a permanent destruction of physical, organizational and human capital. Furthermore, a prudent management of the RER is a sound strategy even in the case in which the favorable terms of trade and international financial conditions were perdurable ex-post because conjectures about the effects of the Dutch Disease on long-run growth are also uncertain. For these reasons, we think that macroeconomic policy in Latin American countries should aim to maintain a stable and competitive RER as an intermediate target for economic development. The macroeconomic policy framework we have in mind combines the following features.

[23] One could refer to this period as the 'long' decade of the 1990s, going from the Mexican stabilization program in 1988 to the Argentine and Uruguayan crises of 2001-02.

First, given the multiplicity of policy objectives – inflation, employment and a stable and competitive RER as an intermediate target – the proposed macroeconomic policy framework requires the coordination of monetary, fiscal, exchange rate and wage policies. Exchange rate policy is oriented towards signaling a stable RER trend, which in a managed floating regime is compatible with short-run nominal exchange rate volatility. In context of capital mobility, as in the case of Latin American countries, active exchange rate policy limits the ability of monetary policy to manage the pace of aggregate demand expansion. This does not mean that monetary policy is passive, but that it is not completely independent. Because of this reason, in our proposed framework fiscal policy takes a relevant role in managing aggregate demand and achieving price stability (Frenkel 2008 and Rapetti 2011). This is a key difference with a standard inflation-targeting regime in which monetary policy carries virtually the whole burden of managing the expansion of aggregate demand.

Second, in the proposed regime both monetary and exchange rate policies are active. Their simultaneous conduct requires sterilized interventions and capital controls. Since domestic and foreign assets are imperfect substitutes, sterilized foreign exchange interventions are effective in simultaneously managing the nominal exchange rate and the interest rate in cases of excess supply of foreign currency. These interventions can be thought of as two instruments implemented sequentially. First, intervention in the foreign exchange markets is used to set the exchange rate to the desired level. Then, sterilization is used to absorb the excess of liquidity created in the first step and thus to maintain the interest rate at the desired level. A potential concern is whether these interventions are sustainable over time because they could create explosive quasi-fiscal cost dynamics. In a context in which the domestic interest rate is low enough, sterilized buying interventions are effective and sustainable[24]. The ability to simultaneously conduct monetary and exchange rate policies with sterilized interventions depends on the magnitude of capital inflows. Capital controls can be useful to facilitate the efficacy of these interventions, especially when inflows are large. Additionally, maintaining a fiscal surplus can help absorb the excess supply of foreign exchange and thus operate as a complement to capital controls to moderate the impact on domestic financial markets.

Third, given that each policy by itself is insufficient to neutralize the effects of capital inflows, it seems desirable to implement them jointly and in a coordinated way. In particular, the coordination between central banks and economic authorities – largely absent in Latin American countries – is essential to optimize the efficacy of macroeconomic policy and to neutralize the effects of capital inflows on the RER.

Some observers appear to be skeptical about the efficacy of buying interventions in the foreign exchange market. For instance, recent IMF's documents advice against 'early interventions' and suggest to intervene only when the RER has appreciated substantially to dissipate expectations of further appreciation (IMF 2011a). This view implicitly assumes that agents know the 'equilibrium' level of the RER and that market forces ultimately will take the RER to such a level. Intervention is thus thought as an instrument to avoid excessive but transitory RER misalignments. This is a curious idea. The same recent IMF documents also warn about the possibility of bubbles in domestic assets and the domestic currency is one of them. Why should we neglect the possibility that exchange rate appreciation is the result of a bubble in the foreign exchange market? The observed lack of effectiveness of recent official interventions in foreign exchange markets may be the result of the central banks' inability to change agents' expectation about the future evolution of the exchange rate. Bold interventions carried out by central banks making clear their will to manage the trend of the exchange rate could, on the contrary, influence private sector expectations and thus reduce selling positions and capital inflows. A key goal of central banks' interventions in the foreign exchange market should be to alter market expectations. Interventions should make clear central banks' power

[24] See Frenkel (2007 and 2008) for a formal analysis of the conditions under which sterilized foreign exchange interventions are sustainable.

and their desire to orient the medium-run trend of the exchange rate. An example of this type of intervention is the one carried out by the Swiss central bank that announced in early September 2011 that it would intervene in the foreign exchange market to stop the appreciation of the Swiss franc generated by capitals running away from the euro area and looking for a safer asset to allocate wealth. The Swiss central bank managed to stop the appreciation of the franc because in such a context it has the ability to issue any amount of assets demanded by foreign investors (ie. Swiss franc). If, for instance, capital is flowing into Colombia's and Chile's financial markets looking for peso-denominated assets, what prevents the Colombian and Chilean central banks to issue those assets and stop the appreciation of their currency?

Appendix

Data on balance of payments, external debt, bilateral RER with the US, effective RER, wages and value added in constant prices are all from CEPALSTAT, ECLAC. In the case of Argentina, all calculations including the CPI were re-calculated using the IPC-7 series elaborated by CENDA. GDP per capita data correspond to the *World Development Indicators* from the World Bank. The data on Brazilian exports were obtained from IPEADATA. The quarterly data series of the Brazilian bilateral RER with the US used in Figure 16 is from Banco Central do Brasil. Data on risk premia is from Bloomberg.

References

Aguirre, A and C Calderon (2005), 'Real exchange rate misalignments and economic performance', Central Bank of Chile, Working Paper No. 315.

Bresser Pereira, L C (2010), *Globalization and Competition.* Cambridge: Cambridge University Press.

Diamand, M (1972), 'La estructura productiva desequilibrada argentina y el tipo de cambio', *Desarrollo Económico*, 12 (45), pp. 45-47.

Damill, M, R Frenkel and R Maurizio (2002), *Argentina: Una década de convertibilidad: un análisis del crecimiento, el empleo y la distribución del ingreso*, OIT Press.

Ffrench-Davis, R (2004), *Entre el neoliberalismo y el crecimiento con equidad. Tres décadas de política económica en Chile*. Siglo XXI editores.

Frenkel, R (2007), 'The Sustainability of Monetary Sterilization Policies', *Cepal Review*, 93, pp. 29-36.

Frenkel, R (2008), 'The Competitive Real Exchange-Rate Regime, Inflation and Monetary Policy', *Cepal Review*, 96, pp. 191-201.

Frenkel, R and M Rapetti (2009), 'A Developing Country View of the Current Global Crisis: What Should Not Be Forgotten and What Should Be Done', *Cambridge Journal of Economics*, 33 (4), pp. 685-702.

Frenkel, R and M Rapetti (2010a), 'Economic Development and the International Financial System', in Stephany Griffith-Jones, Jose Antonio Ocampo and Joseph Stiglitz (eds.): *Time for a Visible Hand*, pp 253-268. Oxford: Oxford University Press.

Frenkel, R and M Rapetti (2011), 'Fragilidad externa o desindustrialización: ¿Cuál es la principal amenaza para América Latina en la próxima década?', *Serie Macroeconomía del Desarrollo*, 116, CEPAL.

Frenkel, R and M Rapetti (2012), Exchange Rate Regimes in Latin America: Lessons from History', *Revista de Historia Economica – Journal of Iberian and Latin American Economic History*, 30 (1), pp. 157-188.

Frenkel, R and J Ros (2006), 'Unemployment and the Real Exchange Rate in Latin America', *World Development*, 34 (4), pp. 631-646.

Gala, P (2007), 'Real exchange rate levels and economic development: theoretical analysis and econometric evidence', *Cambridge Journal of Economics*, 32 (2), pp. 273-288.

IMF (2010), *World Economic Outlook*, October 2010, International Monetary Fund.

IMF (2011a), *World Economic Outlook*, April 2011, International Monetary Fund.

IMF (2011b), *Regional Economic Outlook, Western Hemisphere*, April 2011, International Monetary Fund.

Krugman, P (1987), 'The narrow moving band, the Dutch disease, and the competitive consequences of Mrs Thatcher: Notes on trade in the presence of dynamic scale economies', *Journal of Development Economics*, 27, pp. 41-55.

Magud, N and S Sosa (2010), 'When and Why Worry About Real Exchange Rate Appreciation? The Missing Link between Dutch Disease and Growth', IMF Working Papers, 271, International Monetary Fund.

Rapetti, M (2011), 'Macroeconomic policy coordination in a competitive real exchange rate strategy for development', Working Paper 2011-09. University of Massachusetts: Amherst.

Rapetti, M, P Skott and A Razmi (2012), 'The real exchange rate and economic growth: are developing countries different?', *International Review of Applied Economics*, (forthcoming). Working paper version available.

Rodrik, D (2008), 'The real exchange rate and economic growth', *Brookings Papers on Economic Activity*, 2, pp. 365-412.

Ros, J and P Skott (1998), 'Dynamic effects of trade liberalization and currency overvaluation under conditions of increasing returns', *The Manchester School of Economic and Social Studies*, 66 (4), pp. 466-89.

Sachs, J and A Warner (2001), 'The Natural Resource Curse', *European Economic Review*, 45 (4-6), pp. 827-838.

Pension Liabilities: Fear Tactics and Serious Policy

David Rosnick
Economist, Center for Economic and Policy Research, Washington, DC, USA

Dean Baker*
Economist, Center for Economic and Policy Research, Washington, DC, USA

Abstract

There is major national debate over the funding status of state and local pension funds. Many economists have argued that pension funds are being overly optimistic in assuming that pensions will be able to get 8 percent nominal returns on their invested funds. This calculation is based on the historic average of the mix of assets that are held by these funds. They have argued that funds should instead that their assets will get the risk-free rate of return on US Treasury bonds and build up their funds accordingly. This paper applies a funding rule projects returns based on current price to earnings ratios in the stock market. It runs a number of simulations based on the pattern of stock returns since the beginning of the last century. It shows that in all cases this funding rule would imply a more even flow of payments to the pension fund than a funding rule that assume a risk-free rate of return. The implication of this analysis is that if state and local governments want to maintain a relatively even flow to their pensions and not burden taxpayers at certain points in time with excess burdens, pensions should adopt a funding rule like the one described in this paper.

Introduction

The last two decades have generated substantial demand for accounting techniques that paint a dire picture of government finances. In the early 90s we had 'generational accounting' that showed a lifetime tax burden for generations yet to be born of 80 or 90 percent (Auerbach et al 1991; Kotlikoff, 1993, and Auerbach et al 1995). After this became discredited (Baker 1995; Congressional Budget Office 1995) it became popular to express liabilities of programs like Social Security and Medicare over an infinite horizon (eg. Gokhale and Smetters 2003). These projections could show deficits in hundreds of trillions of dollars. A more modest approach is to calculate the gap between the benefits promised current beneficiaries and the current and past taxes paid by these beneficiaries (eg. Government Accountability Office 2008). This produces a deficit in the high tens of trillions of dollars. This methodology conceals both the fact the numbers actually don't seem that large expressed relative to future GDP and that they are driven almost entirely by projections of explosive health care cost growth in the private sector.

These scary numbers are useful for those who want to force cuts to Social Security, Medicare and other social welfare programs. By making the government's financial situation appear far more dire than is actually the case, and concealing the extent to which the real problem is the country's health care system, these accounting techniques can make it appear that there is no alternative to substantial cuts in social welfare programs.

The publicity given to the recent spate of papers showing large unfunded liabilities for state and local government pensions must be understood in this context (eg. Novy-Marx and Rauh 2009). (It is important to note that these are state and local government pensions. These governments are

subject to a hard budget constraint. Unlike the federal government, they do not have the opportunity to simply issue money to finance pensions.) These papers purport to show unfunded liabilities for these pension funds in the range of $3-4 trillion as opposed to the roughly $1 trillion in unfunded liability reported using the accounting of the funds themselves. The basis for the difference is that these papers discount pension fund liabilities using either the interest rate on corporate bonds or the risk-free interest rate on Treasury bonds. These interest rates are considerably below the 7.5-8.0 percent return assumed by pension fund managers, which leads to a much higher calculation of future liabilities. (The focus of this discussion is state and local pensions. Under the law, these pensions must be financed by the revenue raised by these government entities.)

Three separate issues have been raised in assessing pension liabilities:

1. The appropriate rate of discount to use in attaching a value to these liabilities
2. The accounting rule that should be used in determining the proper funding level
3. The appropriate mix of assets to be held by public pensions.

This paper does not address the first issue. It is standard to use a risk-free rate of return in calculating future liabilities. Arguably this should be used with pension liabilities as well.

The paper does not directly discuss the appropriate mix of assets for pension funds to hold, although it is worth noting that long-lived entities like state and local governments are far better situated than individuals to absorb the timing risk associated with holding equities. Any argument that state and local pension funds should not hold equities would have to address the fact that individual workers do routinely hold equities in their retirement accounts. It would be difficult to develop a logic whereby individual workers are better situated to bear the timing risk associated with equities than state and local governments.

The focus on this paper is on the second point, the appropriate rule for pension funds to use in assessing their funding situation. This paper argues that pension funds should adopt a funding principle that is consistent with the expected return on their holdings. As will be shown, the expected return used in making this assessment will vary depending on the current ratio of stock prices to trend corporate earnings[1]. This funding rule will lead to a more even flow of contributions into the fund than a rule that is based on a fixed return for assets over time.

It is especially important to make this sort of adjustment to expected returns in periods where price-to-earnings ratios in the stock market get out of line with historic patterns. Such periods virtually guarantee a period of below normal returns. During such periods, a pension fund that does not adjust its expected returns, and therefore its funding levels, will end up substantially underfunded after the bubble bursts.

This rule also avoids the excessive build-up of funds that would result from applying a risk-free discount rate in a context where pension funds actually earned higher rates of return on average. The bad event that is hoped to be avoided from having an underfunded pension is the need to have greater than normal funding – implicitly raising taxes for the governmental unit affected. However, a period of excessive contributions needed to build up reserves to meet a more stringent funding rule also implies higher taxes. It can't make sense to deliberately have a period of higher taxes with certainty in order to avoid the possibility of higher taxes at some point in the future.

An optimal funding rule would maintain a roughly constant ratio of contributions to payouts. If a pension invests in risky assets it will inevitably lead to situations in which the fund has greater or lower than desired levels of reserves depending on actual market returns. An optimal funding rule will

[1] This funding rule was first described in Weller and Baker (2005). At the time, because of the high price to earnings ratios in the stock market, the rule would have implied that pension funds should use a lower rate of discount rate than the 8.0 percent that they were then assuming.

maintain funding in a way that minimizes the frequency and size of the divergences from full funding. Both situations imply greater contributions than necessary: either to build up the surplus or to compensate for the shortfall. If the bad event that we are seeking to avoid in financing pensions is an excessive tax burden (implied by greater than normal funding levels) at a particular point in time, then a surplus or shortfall are both evidence of bad management.

The rest of this paper describes more carefully this funding rule that discounts future obligations based on our calculations for the expected earnings of a standard pension portfolio. It uses data on stock and bond returns over the last 135 years to construct simulations comparing the performance of a pension fund that used this funding rule with the performance of a pension fund that assumed a risk free rate of return on assets and funded accordingly.

Alternative pension funding rules: Risk-free rate of return versus expected rate of return

The simulations in this paper define that a pension as being 'fully funded' if its current assets suffice to cover the next 30 years of projected payouts based on a given discount rate. In keeping with definition we apply a rule that no matter the current funding level, contributions to the pension must be sufficiently large so that in ten years the pension will be fully funded based on the chosen discount rate. (The appendix provides a full description of the basis for the simulations.)

In order to make that work, current assets plus discounted contributions over the next ten years must equal discounted payouts over the next 40 years. For a pension that will pay benefits totaling $1 million this year and increase by 5 percent annually, this calculation is illustrated in Table 1. If we assume a 5 percent discount rate, then the pension must have $30 million on hand to be considered fully funded – more if a lower discount rate is used, and less if a higher discount rate is used. In ten years, however, the pension will need to have $48.9 million in assets. At 5 percent interest, we need $30 million today to reach that goal. Consequently, we need to pay the all benefits over the next ten years without touching the initial $30 million. That is, we need to contribute to the pension at a rate of 100 percent of current payouts.

Table 1: Computing Contributions To A Fully Funded Pension

Discount Rate	Initial Assets	Discounted Payouts		Contributions Years 1-10	Contribution Rate
		Years 1-10	Years 11-40		
(1)	(2)	(3)	(4)	(5) = (3) + (4) - (2)	(6) = (5)/(3)
3%	$40.2	$10.9	$48.7	$19.4	178%
5	30.0	10.0	30.0	10.0	100
8	20.5	8.8	15.5	3.8	43

Note that a contribution rate of 100 percent means that $1 million in payouts must be combined with $1 million in contributions to the pension fund.

Now suppose that we assume an 8 percent discount rate. Because we are imputing a higher rate of interest on assets, then to be fully funded today, we need only $20.5 million. Of this, we need only set aside $15.5 million today to have the required $33.5 million needed ten years from now. This leaves $5 million, which may be used to pay benefits over those ten years. At our 10 percent discount rate, we need $8.8 million today to pay these benefits, leaving us $3.8 million shy. If we make pension contributions equal to 43 percent of each year's payouts, then we will still be fully funded in ten years.

Similarly, a much lower discount rate will require greater initial resources to be fully funded, and also require larger contributions. This would make it appear that a higher discount rate is preferable, but only because we have simply assumed a higher return on assets.

Although the pension manager chooses the discount rate, the actual return on pension assets is determined by the market return on the assets held. In Table 2, we see the results after one year assuming the mix of assets in each case is exactly the same and produces a 5 percent return.

Table 2: First Year Performance of a Fully Funded Pension – 5% Interest

Discount Rate	Initial Assets	Payouts	Contributions	Interest Income	Assets at End of Year	Full Funding in Start of Next Year
3%	$40.2	$1.0	$1.8	$2.0	$43.0	$42.2
5	30.0	1.0	1.0	1.5	31.5	31.5
8	20.5	1.0	0.4	1.0	21.0	21.6

As can be seen in Table 2, a discount rate lower than the actual return on assets results in an overfunding of the pension while a higher discount rate leads to underfunding. These funding issues may be chronic, with contribution rates falling over time. With a lower discount rate the pension may become effectively fully pre-funded, paying for all future benefits out of interest alone. It is hard to imagine it good policy to take money from workers today to fund the retirements of future workers who will enjoy higher real incomes.

While a policy choice that lowers the discount rate may help prevent underfunding, a fully funded pension with a low discount rate will suffer correspondingly larger losses in a bad market and necessitating even greater contributions to rebuild those assets. In Table 3 we see the contribution rates needed in response to an initial underfunding of 20 percent based on various discount rates.

Table 3: Computing Contributions To An Underfunded Pension

Discount Rate	Initial Assets	Discounted Payouts		Contributions Years 1-10	Contribution Rate
		Years 1-10	Years 11-40		
(1)	(2)	(3)	(4)	(5) = (3) + (4) - (2)	(6) = (5)/(3)
3%	$32.2	$10.9	$48.7	$27.5	252%
5	24.0	10.0	30.0	16.0	160
8	16.4	8.8	15.5	7.9	89

At a 5 percent discount rate, the 20 percent shortfall raises the first-year contributions from $1 million to $1.6 million. While this jump in contributions may be undesirable, it is still less than the $1.78 million (Table 1) that would be required for a pension fully funded under a 3 percent discount rate.

Finally, even if a bad market never comes, shifting policy from a higher to lower discount rate requires the immediate building of assets. Suppose we wish to protect ourselves against a situation in which the pension is only 80 percent funded at a 5 percent discount rate – i.e., we have $24 million on hand. In order to avoid a future $6 million shortfall, we lower the discount rate to 3 percent *today*, resulting in an immediate *$10.2* million shortfall. Table 4 shows the required contribution rates that result from raising or lowering the discount rate.

Table 4: Computing Contributions To A Pension With $30 Million In Assets

Discount Rate	Initial Assets	Discounted Payouts		Contributions Years 1-10	Contribution Rate
		Years 1-10	Years 11-40		
(1)	(2)	(3)	(4)	(5) = (3) + (4) - (2)	(6) = (5)/(3)
3%	$30.0	$10.9	$48.7	$29.6	271%
5	30.0	10.0	30.0	10.0	100
8	30.0	8.8	15.5	--*	--*

(* No contributions are required in this scenario)

Addressing the Real Risk of Shortfalls

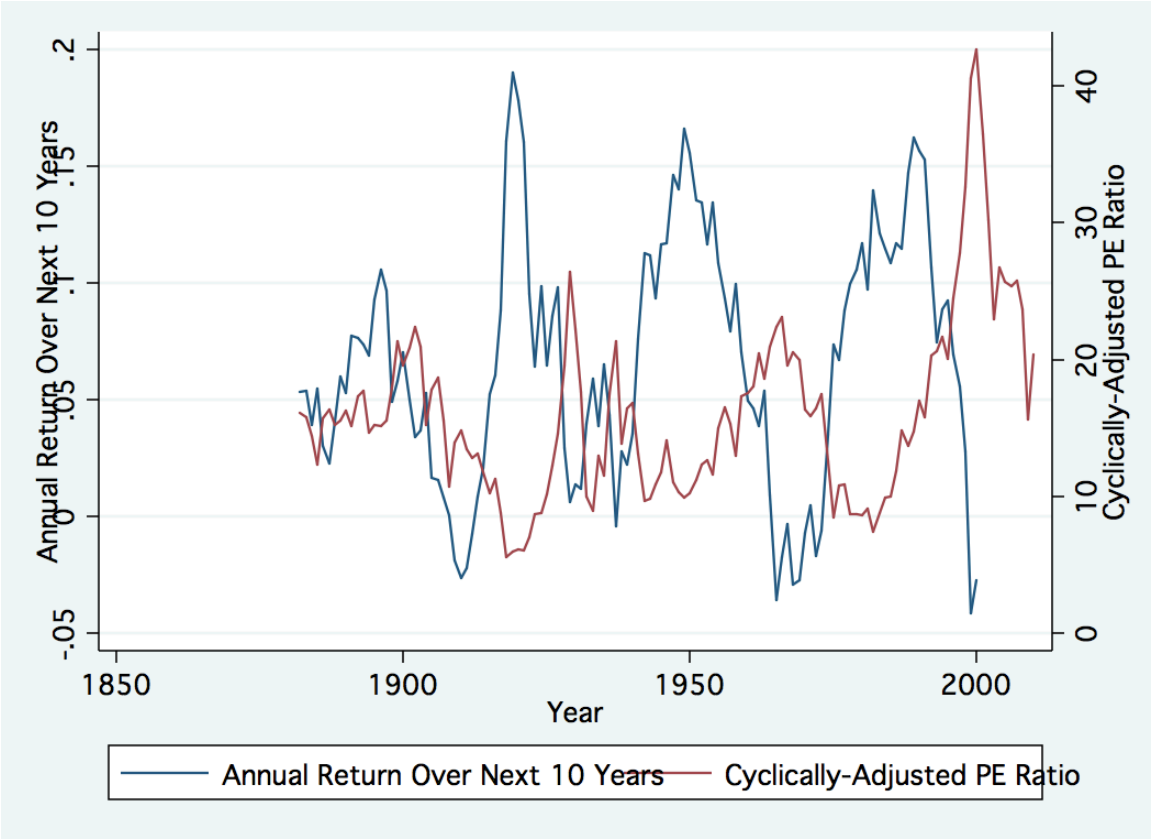
Lowering the discount rate is a strange policy choice that avoids possible large future contributions by insisting on large contributions today. But the greatest risk of adverse market outcomes does not come from a collapse in the price of a fairly valued asset temporarily becoming undervalued. Rather, downside risk comes primarily from the collapse in the price of an overvalued asset suddenly valued according to fundamentals.

In the housing bubble of the 2000s, there was a prevailing insistence that housing prices could not fall. Lenders egged on borrowers to leverage up on the assumption that prices would continue to rise even as rents stayed low, implying an ever higher sales price to rent ratio. Similarly, in the stock bubble of the 1990s, there was remarkable insistence that equities would produce historical returns no matter how expensive stocks became in relation to the earnings the companies had any hope of distributing in dividends.

In short, failure to adjust to the current state of the market caused considerable pain for many people as well as institutional investors – such as pensions. In order to mitigate the effects of a downturn, it is important both to discount at the expected rate of return on assets and to adjust the expected rate of return as market conditions require.

In 1919, the price-to-earnings ratio in the S&P 500 stood at less than 6:1, and over the next ten years produced a real total return of 19 percent per year. By contrast, the PE ratio had skyrocketed to over 40:1 in 1999. From 1999 to 2009 the real total return on the S&P 500 was -4.2 percent per year. In neither 1919 nor 2009 should one have expected a typical return on stocks. Figure 1 below shows the historical relationship between the PE ratio and stock returns.

Figure 1: Ten-Year Forward Real Annual Return to S&P 500 and PE Ratio



Sources: Robert Shiller[2] and authors' calculations

Thus, we propose a rule for adjusting the expected return on stocks so that the PE ratio returns to 15:1 over the course of ten years based on projected earnings.

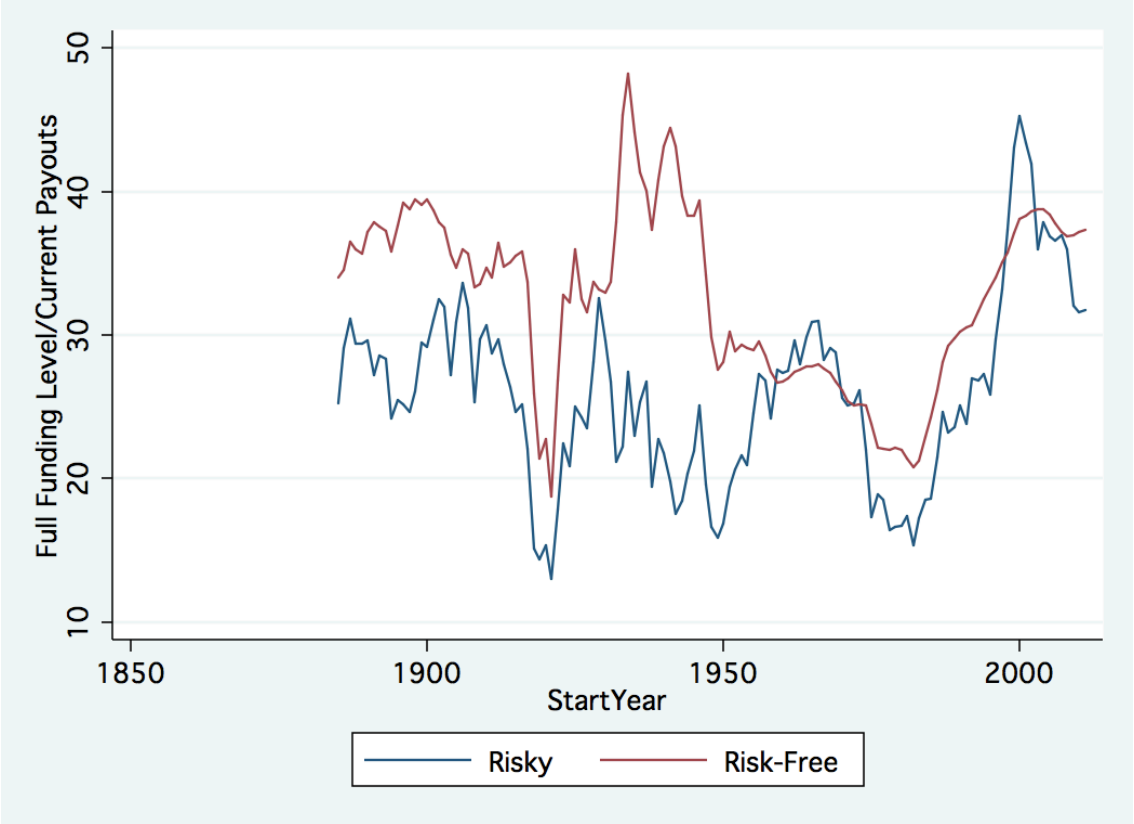
Simulation Under Different Rules

We now compare pension performance under different assumptions for the discount rate: first by assuming a 'risk-free' rate based on ten-year Treasury bonds plus one percentage point, then by assuming a 'risky' rate based on a portfolio of 60 percent S&P 500 and 40 percent 'risk-free'. In each case the actual portfolio is 60 percent stock and 40 percent bonds, but the 'risky' discount rate is adjusted for the current PE ratio as discussed above.

The first difference between the two pensions is the level of assets required in each case to be considered fully funded. As the S&P 500 can be generally expected to outperform bonds, the 'risky' rate discounts more heavily future payouts and therefore carries a lower asset burden than a pension using a 'risk-free' rate. Figure 2 shows that in the typical year some $1.30 in assets are required under the lower rate for every dollar required when setting the discount rate equal to the expected rate of return on the portfolio assets.

[2] [http://www.econ,yale.edu/~shiller/data/ie_data.xls and http://www.econ.yale.edu/~shiller/data/chap26.xls]

Figure 2: Fully Funded Pension Assets



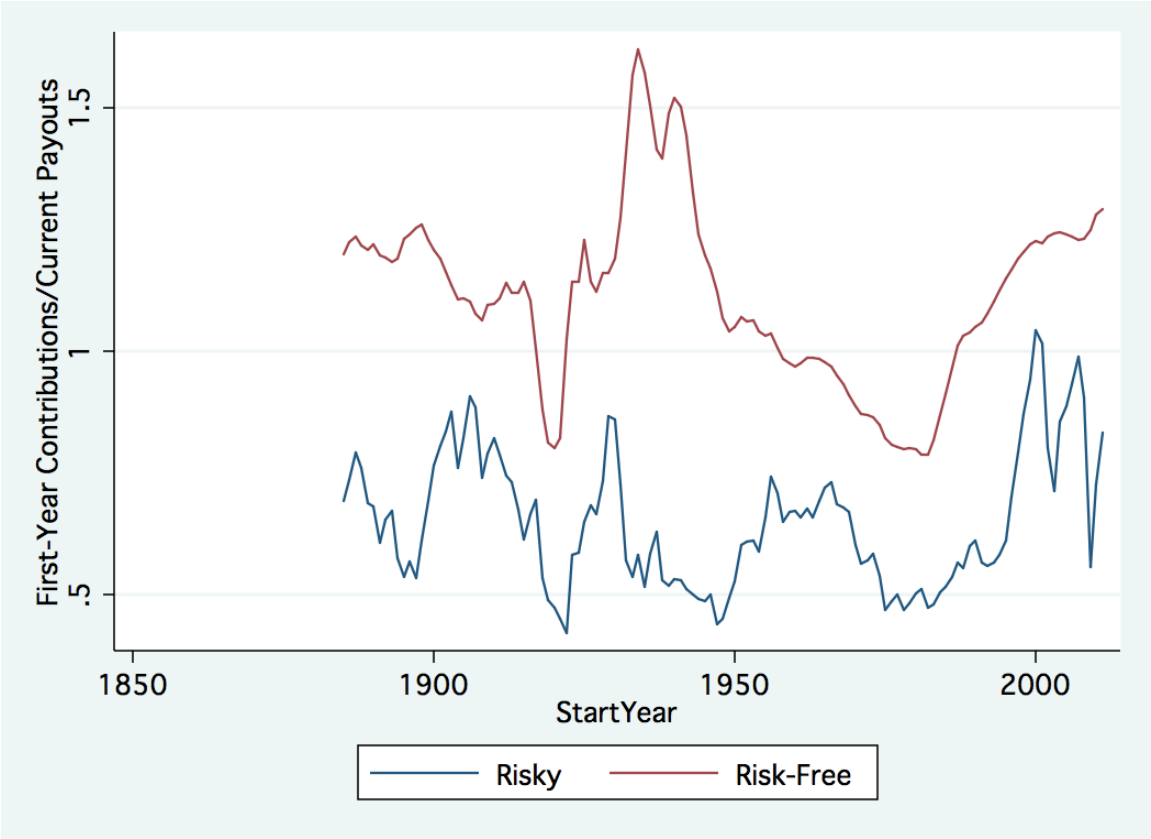
Sources: Louis Johnston and Samuel H Williamson, 'What Was the US GDP Then?' MeasuringWorth, 2011[3], Bureau of Economic Analysis National Income and Product Accounts Tables 1.1.5, 1.1.6, and 1.10, Robert Shiller[4], and authors' calculations

Note in particular the period of the late 1990s when the relationship inverts. As the stock market became overvalued, the expected return on stocks fell well below that of safe bonds. Rather than relying on an already-large stock of assets, the risky funding rule required additional contributions to build up a reserve against the expected fall in stock prices. By construction, this build-up takes place during the run-up of asset prices and so the contribution requirements are not great.

In Figure 3, we see the required contributions under each discount rate, assuming that each pension is fully funded at the start of the year. Typically, the risky discount rate would require contributions to a fully funded pension in the amount of 50 to 100 percent of annual payouts. By contrast, a risk-free discount rate would require much larger fully funded pension contributions.

[3] Available from [http://www.measuringworth.com/usgdp],
[4] Available from http://www.econ.yale.edu/~shiller/data/ie_data.xls and http://www.econ.yale.edu/~shiller/data/chap26.xls

Figure 3: Contribution Rates Toward Fully Funded Pensions



Sources: Louis Johnston and Samuel H Williamson, What Was the US GDP Then?' MeasuringWorth, 2011[5], Bureau of Economic Analysis National Income and Product Accounts Tables 1.1.5, 1.1.6, and 1.10, Robert Shiller[6], and authors' calculations

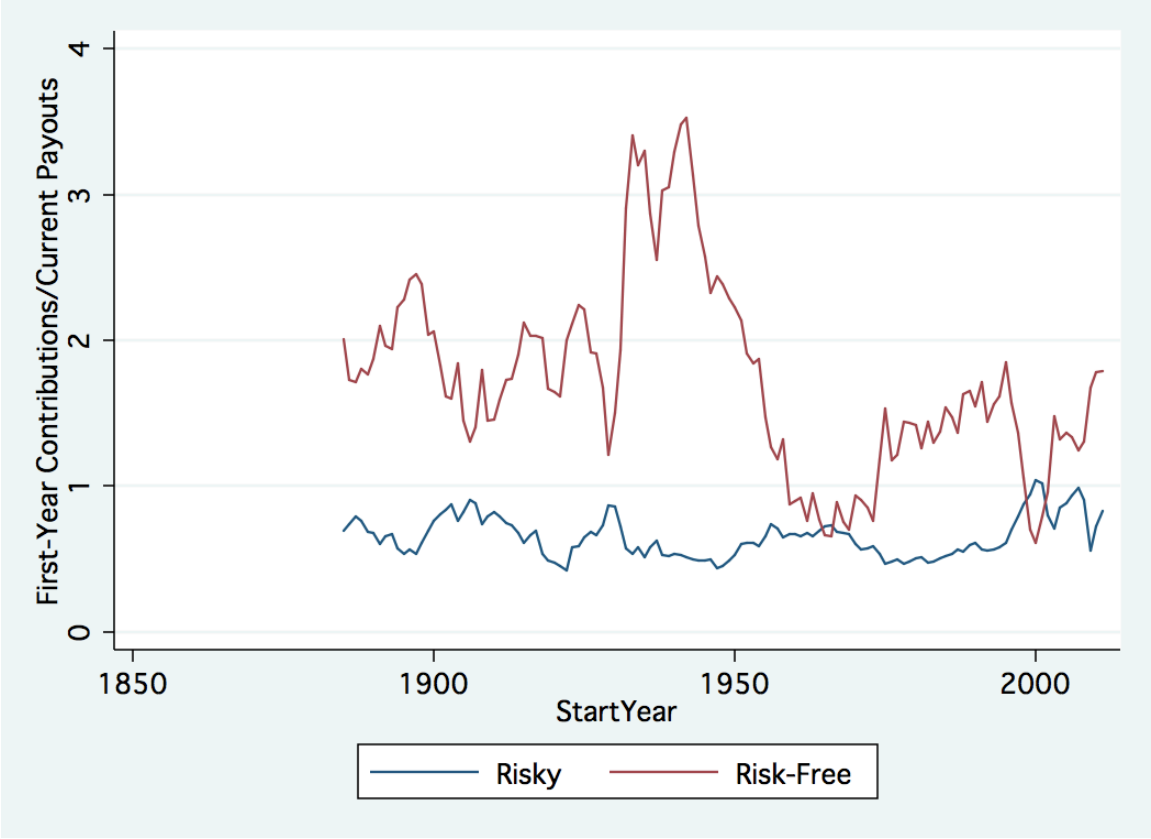
Even in 2000, when a fully funded risky-rate pension would have been at the most risk, the contribution rate needed to sustain full funding was less than was typical for a pension working from a risk-free rate.

Of course, pensions are not always fully funded. During the run-up of an asset bubble, the pension may become overfunded leading to a fall in contributions. Rather than examining fully funded pensions in various years, it is important to observe the performance of pension rules over time.

The first thing to note is that a proposed switch in the discount rate affects the current funding status of the pension. Absent a very large contribution, a cut in the discount rate pushes a pension farther below full funding, as in the examples of Table 4. Figure 4 shows the first-year contribution rates under each discount rate, assuming that the risky pension is fully funded and the risk-free pension must adjust over ten years to full funding.

[5] Available from http://www.measuringworth.com/usgdp
[6] Available from http://www.econ.yale.edu/~shiller/data/ie_data.xls and http://www.econ.yale.edu/~shiller/data/chap26.xls

Figure 4: First Year Contribution Assuming Equal Starting Assets

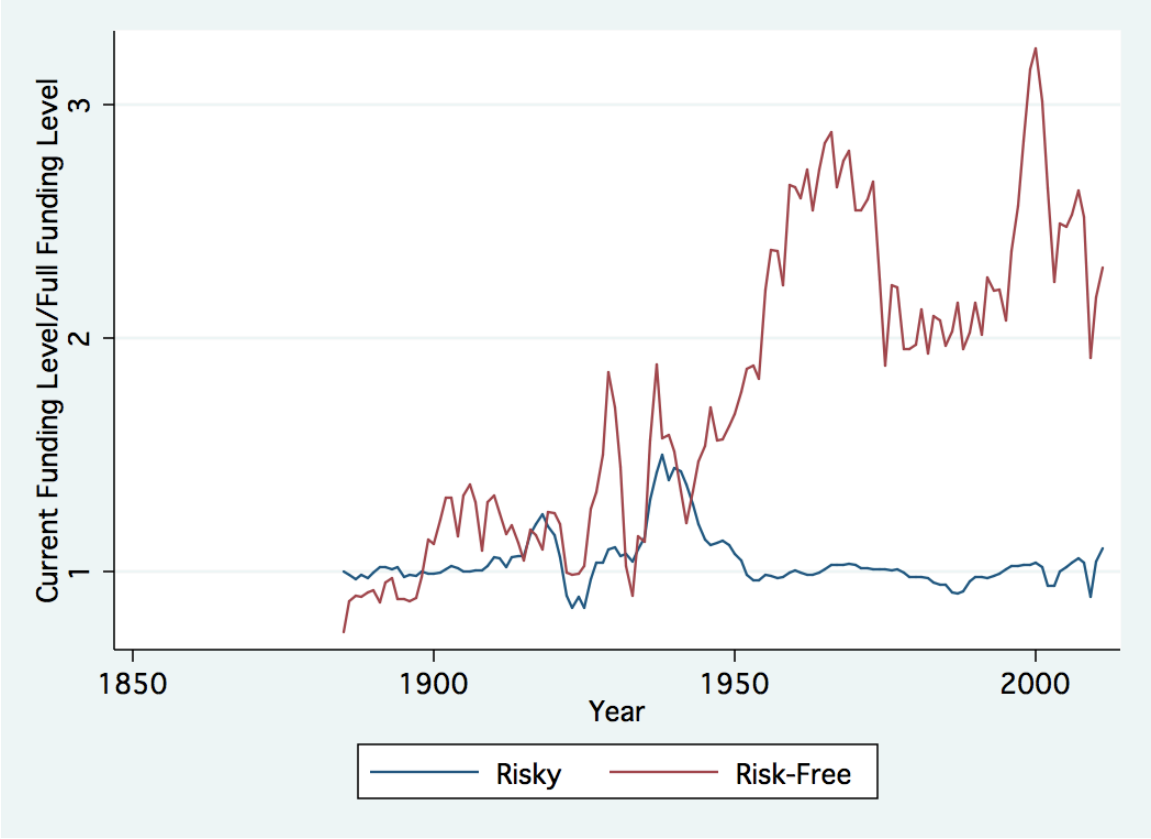


Sources: Louis Johnston and Samuel H Williamson, 'What Was the US GDP Then?' MeasuringWorth, 2011[7], Bureau of Economic Analysis National Income and Product Accounts Tables 1.1.5, 1.1.6, and 1.10, Robert Shiller[8], and authors' calculations

As can be seen in Figure 4, the initial contribution rates are much higher under the risk-free discount rate. On the other hand, the actual portfolio risk is assumed to be identical, so this often leads to overfunding. Figure 5 shows simulated pension assets under each discount rate as a share of full funding. In the figure, both are assumed to begin in the year 1885 with identical assets sufficient to fully fund under the risky discount rate. In that year, the 'risk-free' discount rate implies underfunding, but by 1900 the pension is overfunded.

[7] Available from http://www.measuringworth.com/usgdp
[8] Available from http://www.econ.yale.edu/~shiller/data/ie_data.xls and http://www.econ.yale.edu/~shiller/data/chap26.xls

Figure 5: Simulation of Pension Funding – 1885 Start Year



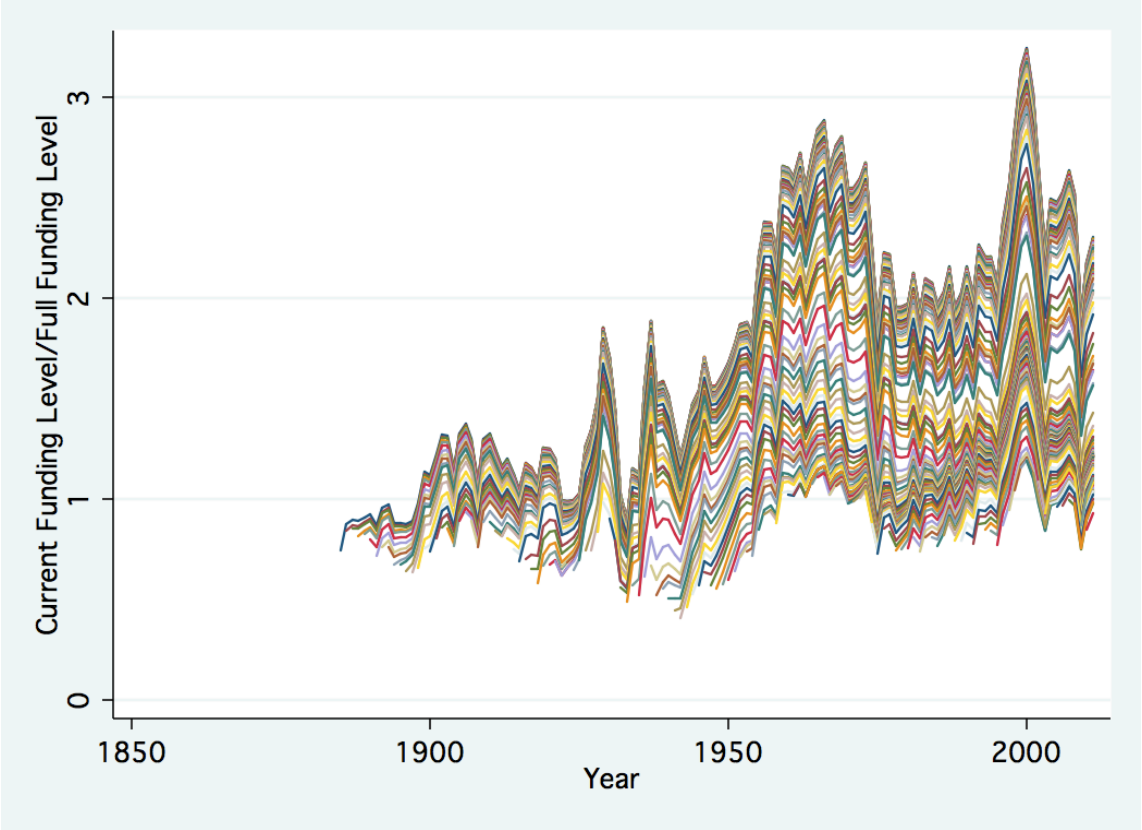
Sources: Louis Johnston and Samuel H Williamson, 'What Was the US GDP Then?' MeasuringWorth, 2011[9], Bureau of Economic Analysis National Income and Product Accounts Tables 1.1.5, 1.1.6, and 1.10, Robert Shiller[10], and authors' calculations

Over time, the risk-free discount rate leads to a tremendously overfunded pension. By the peak of the stock bubble, the pension holds more than three times the required assets. At the peak in 2000, the pension holds assets of more than 120 times the year's payouts and has become entirely pre-funded – having made no contribution since 1943. While it may be comforting to know that future pension obligations may met without making contributions, it is less comforting to know that baby boomer pensions were paid for primarily by workers who were born during the Civil War and then by those who fought in World War II.

The degree of overfunding in and among simulations varies greatly. Figure 6 shows the simulated funding levels under the risk-free discount rate starting with assets sufficient to fully fund under the risky rate.

[9] Available from http://www.measuringworth.com/usgdp
[10] Available from http://www.econ.yale.edu/~shiller/data/ie_data.xls and http://www.econ.yale.edu/~shiller/data/chap26.xls

Figure 6: Simulated Pension Assets – Switch to Risk-Free Discount Rate

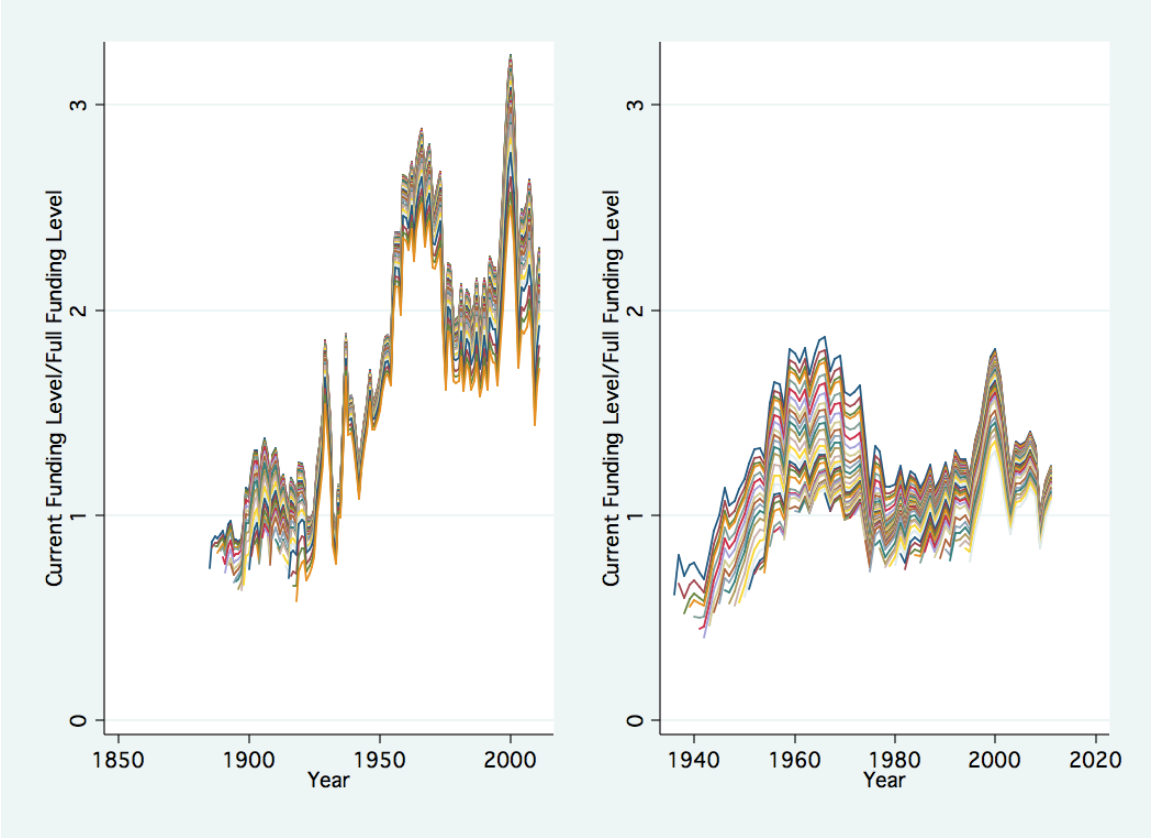


Sources: Louis Johnston and Samuel H Williamson, 'What Was the US GDP Then?' MeasuringWorth, 2011[11], Bureau of Economic Analysis National Income and Product Accounts Tables 1.1.5, 1.1.6, and 1.10, Robert Shiller[12], and authors' calculations

There is less variation among the simulations than Figure 6 seems to indicate. In fact, whether or not a pension becomes overfunded depends greatly on the year in which the simulation starts. Simulations that start in 1885-1918 will greatly overfund, and simulations that start in 1936-1995 less so. This is seen in Figure 7 below.

[11] Available from http://www.measuringworth.com/usgdp
[12] Available from http://www.econ.yale.edu/~shiller/data/ie_data.xls and http://www.econ.yale.edu/~shiller/data/chap26.xls

Figure 7: Simulated Pension Assets – Starts in 1885-1918 and 1936-1995

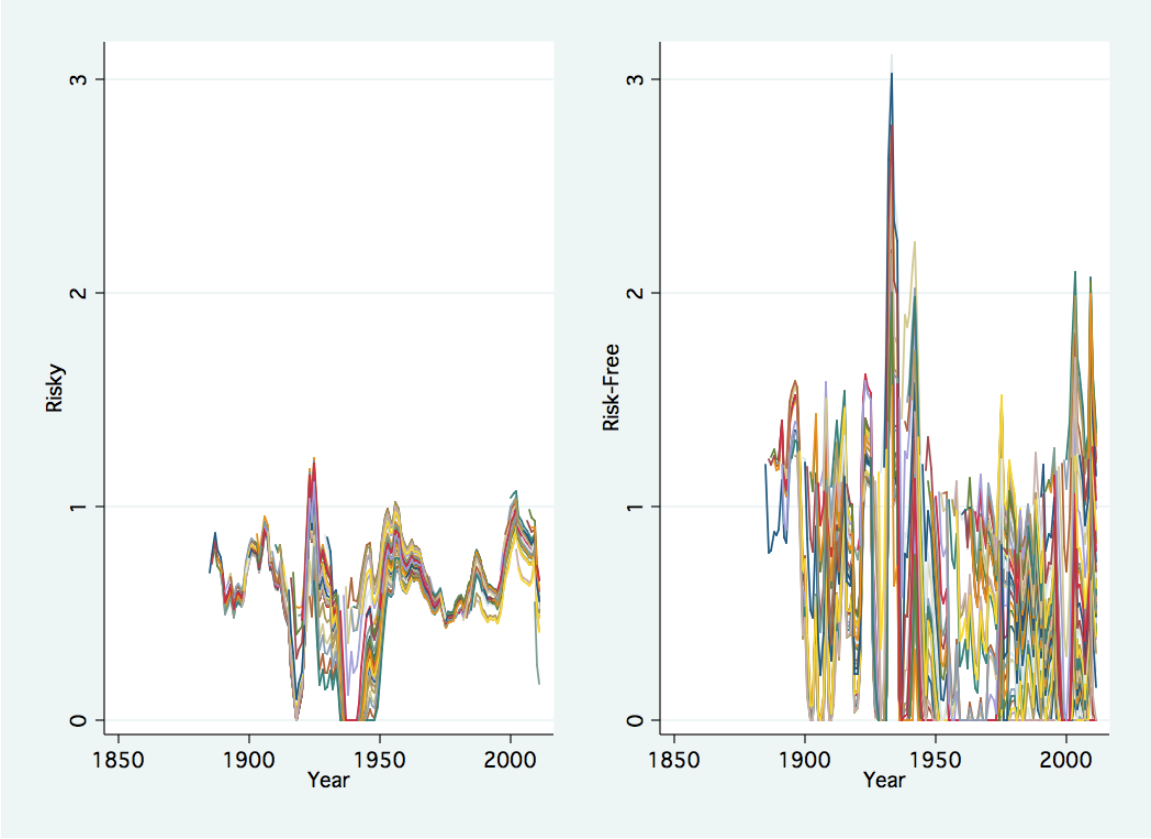


Sources: Louis Johnston and Samuel H Williamson, 'What Was the US GDP Then? MeasuringWorth, 2011[13], Bureau of Economic Analysis National Income and Product Accounts Tables 1.1.5, 1.1.6, and 1.10, Robert Shiller[14], and authors' calculations

Even 'risk-free' pensions that are initially fully funded may be required to make large contributions over time. Figure 8 shows the necessary contributions as a share of payouts in each year for various start years.

[13] Available from http://www.measuringworth.com/usgdp
[14] Available from http://www.econ.yale.edu/~shiller/data/ie_data.xls and http://www.econ.yale.edu/~shiller/data/chap26.xls

Figure 8: Simulated Pension Contribution Rates – Fully Funded At Start of Simulation



Sources: Louis Johnston and Samuel H Williamson, 'What Was the US GDP Then? MeasuringWorth, 2011[15], Bureau of Economic Analysis National Income and Product Accounts Tables 1.1.5, 1.1.6, and 1.10, Robert Shiller[16], and authors' calculations

Under the risky discount rate, pensions rarely must make contributions in excess of payouts. That is, in most years benefits are partially funded out of assets and interest. On the other hand, it is far more common for a pension funded under rule requiring a risk-free discount rate to pay for current benefits exclusively out of contributions. Figure 9 shows the percentage of simulations in which required contributions are zero (green), otherwise less than or equal to payouts (red), and greater than payouts (blue).

[15] Available from http://www.measuringworth.com/usgdp
[16] Available from http://www.econ.yale.edu/~shiller/data/ie_data.xls and http://www.econ.yale.edu/~shiller/data/chap26.xls

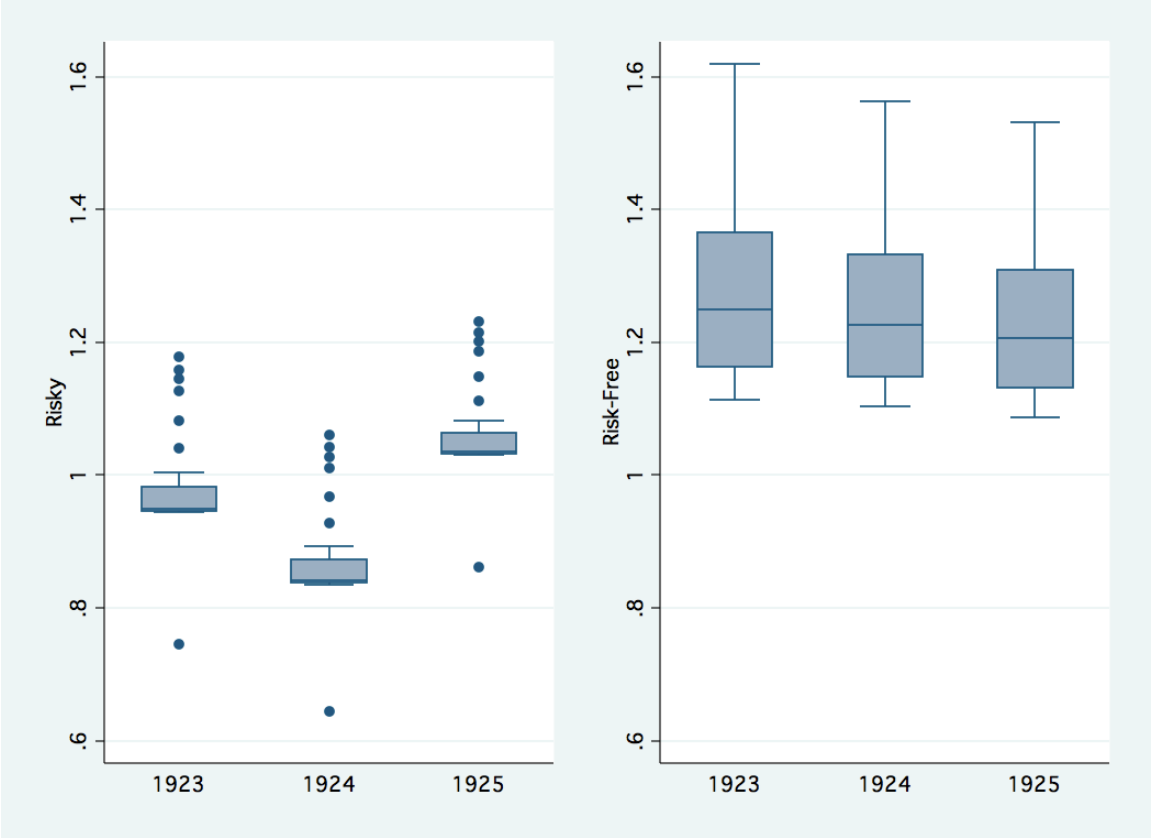
Figure 9: Percentage of Simulations with Zero, Moderate and Large Contributions

Sources: Louis Johnston and Samuel H Williamson, 'What Was the US GDP Then?' MeasuringWorth, 2011[17], Bureau of Economic Analysis National Income and Product Accounts Tables 1.1.5, 1.1.6, and 1.10, Robert Shiller[18], and authors' calculations

We can see in Figures 8 and 9 that the worst outcomes under the risky discount rate (that is, with the greatest need for contributions) come in the early 1920s. Even so, the required contributions under the risk-free discount rate are even larger in these years. Figure 10 below shows the range of simulated contribution rates in the years 1923-25. As above, these simulations all start with fully funded pensions, and therefore do not reflect the need to build up assets in moving to a risky-free discount rate.

[17] Available from http://www.measuringworth.com/usgdp
[18] Available from http://www.econ.yale.edu/~shiller/data/ie_data.xls and http://www.econ.yale.edu/~shiller/data/chap26.xls

Figure 10: Contribution Rates – 1923-25 (Simulations Starting Fully Funded in 1885-1922)

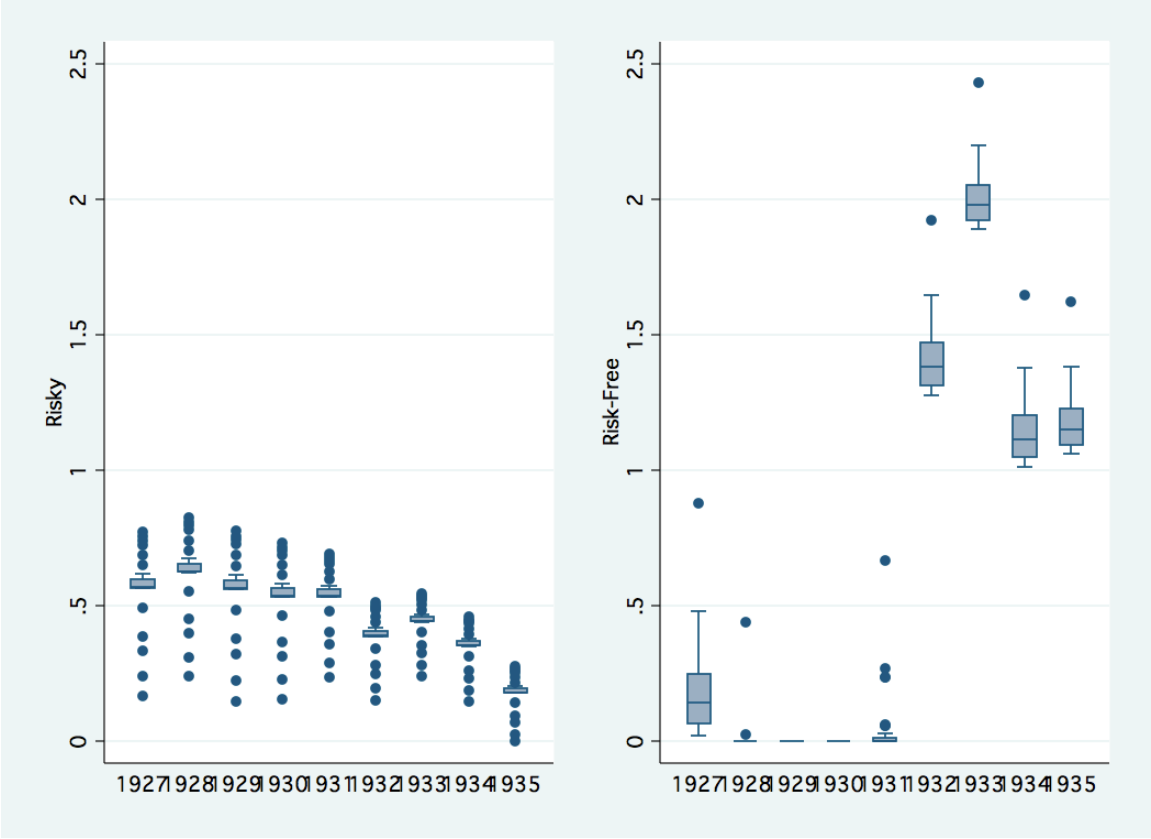


Sources: Louis Johnston and Samuel H Williamson, 'What Was the US GDP Then? MeasuringWorth, 2011[19], Bureau of Economic Analysis National Income and Product Accounts Tables 1.1.5, 1.1.6, and 1.10, Robert Shiller[20], and authors' calculations

Similarly, the contribution rates in the lead-up to and through the early part of the Great Depression are striking.

[19] Available from http://www.measuringworth.com/usgdp
[20] Available from http://www.econ.yale.edu/~shiller/data/ie_data.xls and http://www.econ.yale.edu/~shiller/data/chap26.xls

Figure 11: Contribution Rates – 1927-35 (Simulations Starting Fully Funded in 1885-1926)



Sources: Louis Johnston and Samuel H Williamson, 'What Was the US GDP Then?' MeasuringWorth, 2011[21], Bureau of Economic Analysis National Income and Product Accounts Tables 1.1.5, 1.1.6, and 1.10, Robert Shiller[22], and authors' calculations

Although 'risky' pension contributions were modest prior to the Great Depression, the PE ratio for the S&P 500 fell below 10:1 in 1932-33 – indicating an unusually large expected return. Thus, contribution rates actually fall after the crash. By contrast, the use of a 'risk-free' discount rate meant very low contribution rates as the stock market inflated – recognizing the existence of a very large asset cushion. However, the implicit failure to recognize that the market had subsequently overcorrected and would yield larger returns exaggerated a need for very large contributions at the worst possible time.

In effect, using the risk-free rate of return assumption would have required governments to impose large tax increases to fully fund their pensions in the middle of the Great Depression – even though such contributions were in fact unnecessary to maintain full funding given the actual returns in the stock market.

[21] Available from http://www.measuringworth.com/usgdp
[22] Available from http://www.econ.yale.edu/~shiller/data/ie_data.xls and http://www.econ.yale.edu/~shiller/data/chap26.xls

Conclusion

The participants in the debate over pension accounting have a variety of agendas. If the purpose is to make the situation of these pension funds appear as dire as possible, then using a risk-free rate of return to assess their liabilities can be useful. However if the goal is to actually manage a pension fund holding equities in a way that minimizes the need to increase contributions above the normal level, and therefore implicitly raise taxes, then it is desirable to use a funding rule that is based on the expected rate of return of the assets held by the fund.

This paper uses data on stock returns dating back to 1990 to show that a pension fund that adjusted its expected return assumptions based on the ratio of stock prices to trend earnings would have a much smoother contribution path than a fund that always maintained full funding using the risk-free rate of return as the discount rate. The smoother funding course results not only from avoiding high taxes during the initial build-up period, but also by avoiding the sharp increases in funding that would result from stock market crashes like the ones in 1929-30 and 2000-2002.

If the goal of pension fund managers is to maintain a smooth flow of funding and avoiding temporary tax increases needed to meet funding targets, then the funding rule described in this paper is unambiguously superior to a funding scheme that maintains full funding using the risk-free rate of return as the discount rate.

Appendix: Forecast Methodology

An index historical of risk-free returns is constructed from the Shiller ten-year bond data (Shiller 2011), assuming a one-percentage-point premium reinvested annually.

The GDP deflator is assumed to increase at the rate of the five years prior to and including the base year of the forecast. All base-year prices (including January S&P) are known accurately. All 'real' values are discounted by the GDP deflator. (For 1929-2010, GDP and wage data are taken from the National Income and Product Accounts, Tables 1.1.5, 1.1.6, and 1.10. For years prior to 1929, data was taken from Johnston and Williamson (2011)).

In this paper, several long-term trends are constructed and used for forecasting purposes. In most cases the data is log-transformed, a trend is estimated to be quadratic in time, and over ten years from the most recent available data at the start of forecast. A return to that trend is assumed. In general, forecasts made in January of a base year include historical data through the year prior.

These trends include:

- Population growth starting in 2011. Pension managers are explicitly assumed to otherwise accurately forecast population, rather than project a return to trend.
- Real GDP per-capita.
- Wage share of GDP. The (log) wage share prior to 1929 is assumed equal to the average in the years 1929-71. From 1971 on, the trend is quadratic, but assumed to remain flat over time rather than swing upward.
- A 'risk-free' bond index. The trend in the log index is assumed linear.
- Real earnings. (The nominal earnings and dividends for each year are Shiller's figures for December of the year.)

Pension payouts in each year are assumed to equal 95 percent of the previous year's payouts (adjusted for price level of GDP), plus 0.1 percent of nominal wages (actual or projected) in each of the five prior years.

Price forecasts for the S&P 500 are constructed based on a return to a cyclically adjusted price-to-earnings ratio of 15 over ten years. The PE ratio is constructed identically to Shiller's formulation (except for the use of GDP deflator in lieu of CPI). Specifically, the historical PE ratio is computed as the real S&P 500 index divided by the average real earnings over the prior ten years. Over the next ten years the PE ratio is assumed to move linearly to a value of 15 in the tenth year after the base year of the forecast. The forecast PE ratio is multiplied by the projected GDP deflator and projected ten-year average real earnings in order to arrive at a price projection.

Dividends are forecast in multiple steps. First, the historical dividend-to-earnings ratio is computed by dividing the average real dividends over the prior ten years by the average real earnings over the prior ten years. Initially, the DE ratio is assumed to be unchanged in the forecasts, and therefore the unadjusted real annual dividends are those that maintain the DE ratio given the real earnings. The (log) unadjusted dividends are replaced by their trend over the forecast period.

Pension management

Pensions are assumed to rebalance each January (base year/'year zero') to hold 60 percent S&P 500 stock and 40 percent risk-free bonds. All dividends are reinvested, payouts and (if any) contributions made just prior to rebalancing. The above forecast implies a risky discount rate based on the return on this pension portfolio as well as a 'risk-free' rate based on 100 percent bonds. Regardless of the discount rate applied, the actual pension portfolio is assumed to be risky.

Our rule for a fully funded pension is that it holds assets equal to 100 percent of the present value of payouts over the next 30 years (years 0-29.) Thus, the assets required to meet the funding rule depends on the choice of discount rate. Any pension shortfall must be restored over the next ten years as follows:

Current assets (less) discounted payouts in years 0-9 are subtracted from discounted payouts in years 10-39 to find the present value of contributions which must be made in order to forecast a fully-funded pension in year 9. (Equivalently, the discounted contributions must equal the difference between discounted payouts in years 0-39 minus current assets.) The contribution rate is defined as this present value of required contributions divided by the present value of payouts over years 0-9. If the contribution rate is positive, then current-year contributions are made in the amount of current-year payouts times the contribution rate.

In the first January of a multi-year simulation of pension performance, the 'risky' pension is assumed to hold assets sufficient to make it fully funded with respect to the risky discount rate. Typically, the 'risk-free' pension is assumed to start with the same assets, but in some simulations may instead begin fully funded with respect to the risk-free rate.

References

Auerbach, Alan, Jagdeesh Gokhale, and Lawrence Kotlikoff (1991), 'Generational Accounts: A Meaningful Alternative to the Deficit Accounting'. In David Bradford (ed.) *Tax Policy and the Economy*, Cambridge, Massachusetts: MIT Press.

Auerbach, Alan, Jagdeesh Gokhale, and Lawrence Kotlikoff (1995), 'Restoring Generational Balance in US Fiscal Policy: What Will It Take?', Federal Reserve Bank of Cleveland, First Quarter.

Baker, Dean (1995), *Robbing the Cradle? A Critical Assessment of Generational Accounting*. Washington, DC: Economic Policy Institute.

Congressional Budget Office (1995), 'Who Pays and When? An Assessment of Generational Accounting'. Washington, DC: Congressional Budget Office.

Gokhale, Jagadeesh and Kent Smetters (2003), *Fiscal and Generational Imbalances: New Budget Measures for New Budget Priorities*. Washington DC: AEI Press.

Government Accountability Office (2008), 'US Financial Condition and Fiscal Future Briefing', Presented by the Honorable David M. Walker, Comptroller General of the United States, National Press Foundation, January 17th, 2008. Washington, DC: Government Accountability Office, GAO-08-446-CG.

Johnston, Louis and Samuel H Williamson (2011), 'What Was the US GDP Then?', MeasuringWorth, Available at http://www.measuringworth.com/usgdp

Kotlikoff, Lawrence (1993), *Generational Accounting: Knowing Who Pays, and When, for What We Spend*. New York: Free Press.

Novy-Marx, Robert and Joshua Rauh (2009), 'The Liabilities and Risks of State-Sponsored Pension Plans', *Journal of Economic Perspectives*, 23 (4), Fall 2009, pp. 191-210(20).

Novy-Marx, Robert and Joshua Rauh (2011), 'Public Pension Promises: How Big Are They and What Are They Worth?', *The Journal of Finance*, 66 (4), pp. 1211-1249, August 2011.

Shiller, Robert (2011), 'US Stock Price Data, Annual, with Consumption, Both Short and Long Rates, and Present Value Calculations'. Available from http://www.econ.yale.edu/~shiller/data/chapt26.xls

Weller, Christian and Dean Baker (2005), 'Smoothing the Waves of Pension Funding: Could Changes in Funding Rules Help Avoid Cyclical Underfunding?', *Journal of Policy Reform*, 8 (2), pp. 131-151.

Brain Physiology, Egoistic and Empathic Motivation, and Brain Plasticity: Toward a More Human Economics[1]

John F Tomer
Professor of Economics at Manhattan College, Riverdale, New York, USA

Abstract

The brain physiology research of leading evolutionary neuroscientist, Paul MacLean, has important implications for human economic motivation. Gerald Cory in his research has admirably utilized MacLean's findings and has persuasively explained that humans have two dominant motivations: 1) ego or self-interest and 2) empathy or other-interest, which our brains attempt to balance. This view is clearly important and at odds with mainstream economics in which self-interest is the dominant motivation. The MacLean-Cory view, also known as Dual Motive Theory (DMT), represents a serious challenge to mainstream economics. However, the DMT leaves something to be desired. While understanding the promise of the perspective deriving from brain physiology, some scholars have expressed dissatisfaction with it. Accordingly, the purpose of this paper is to revise DMT utilizing the concept of brain plasticity and argue that the mainstream economic image of the brain is not supported by current knowledge of brain science. Brain plasticity refers to the ability of the brain to change structurally and functionally as a result of input from the environment. Some of this plasticity is no doubt genetically determined but some brain change is a product of individual effort and represents the individual's investment in intangible capital (standard human capital, social capital, personal capital, and so on). In this revised view, the balance that individuals, groups, and societies strike between ego and empathy orientation is to a great extent determined by these intangible investments, not simply by brain physiology.

Key Words: brain plasticity, neuroplasticity, empathy, dual-motive theory, brain physiology, egoistic motivation

1. Introduction

The brain physiology research of leading evolutionary neuroscientist, Paul MacLean, has important implications for human economic motivation. Gerald Cory in his research has admirably utilized MacLean's findings and has persuasively explained that humans have two dominant motivations: 1) ego or self-interest and 2) empathy or other-interest, which our brains attempt to balance. This view is clearly important and at odds with mainstream economics in which self-interest is the dominant motivation. In the view of mainstream economists, all other motivations such as empathy are subordinate to self-interest. The MacLean-Cory view, also known as Dual Motive Theory (DMT), represents a serious challenge to mainstream economics and has begun to attract considerable interest. However, the DMT which traverses the terrain between brain physiology and concrete economic behavior leaves something to be desired. While understanding the promise of the perspective deriving from brain physiology, some students and experienced scholars have expressed dissatisfaction with how the three level modular brain perspective has been used to explain specific, concrete economic behavior. Accordingly, one purpose of this paper is to revise the MacLean-Cory conceptual framework and draw out its implications for economic behavior. The other purpose of the paper is to argue that the vision of the human brain implied by mainstream economics (ME) is not

[1] The author has no personal or professional links that could bias his treatment of his subject.

supported by the current state of knowledge of brain physiology and recent findings from brain science.

An important aspect of the revision of DMT is its focus on a variety of brain capacities. It is useful to think of the three level (triune) brain conception as representing a fundamental human capacity, i.e., the capacity for acting in accord with self-interest, with other interest, or some balance of the two motives. The basic capacity associated with brain physiology is referred to here as the 'hard' brain, which all humans have regardless of age, education, or experience. There is, however, reason to believe that an individual's brain capacity is also determined by one's 'soft' brain, the part of the brain that does not develop the same way in every human. The soft part develops differently depending on the experiences of and influences on the individual. Stating that there is a soft or changeable part of the brain is essentially the same as stating that there is brain plasticity. Some of this changeability or plasticity is no doubt genetically determined but some brain change is a product of individual effort and represents the individual's investment in intangible capital (standard human capital, social capital, personal capital, and so on) (Tomer 2008). Accordingly, the balance that individuals, groups, and societies strike between ego and empathy orientation is to a great extent determined by these intangible investments, not simply by the brain physiology common to all. In other words, it is the soft or plastic aspect of the brain that determines how the capacity of the hard brain gets expressed. Different individuals and societies are very different in this regard.

The broad concept of human development corresponds to a great extent to growth of the soft brain, and this process involves growth in a variety of human mental capacities including empathy. The human development process may also raise the individual's and society's capacity for striking a desirable balance between ego and empathy. Thus, there is reason to believe that intangible investments resulting in soft brain change are a key to how the capacity of the hard brain becomes manifest. An important part of the process of attempting to improve dual motive theory is developing the distinction between the hard and soft parts of the brain and understanding how brain capacities change.

2. The implied model of the brain in mainstream economics

The understanding of brain physiology along with the recent findings of brain science, which will be incorporated in the revised DMT in a later section, constitute an important argument against the implied model of the brain in mainstream economics. The implied model of the human brain can be discerned from the typical assumptions of ME theory. This model features the following:

1. People's one dominant or sole motivation is self-interest
   a. All other emotions or considerations are subordinate to self-interest
   b. People may seem to act in the interests of others but only if it is in their self-interest to do so.

2. People have infinite (or very high) cognitive capacity.

3. People have zero capacity for pure empathetic motivation.

4. Rationality involves applying logic or reason to choose or obtain the best possible outcome(s) (generally satisfaction of wants) for oneself regardless of the nature of the wants.

5. In the pure version of the ME model, no human capacity differences exist. Everyone has an infinite cognitive capacity and zero empathetic capacity. In revised, not so pure,

versions of ME, people's investment in various types of human capital, especially standard human capital, may change some of their human capacities.

3. The importance of brain physiology a la MacLean

Before revising DMT, it is important to review its essential features. The starting point, of course, is the work of Paul MacLean. Paul MacLean's research on brain physiology is extremely important because of the perspective it provides on human motivation. At the heart of MacLean's research is his conception of the human brain as having three interconnected modular levels. It is in a sense three brains in one, a triune brain, which reflects to a considerable degree the evolutionary path from reptiles to early mammals to late mammals (MacLean 1990, pp. 8-9). These three brains are like biological computers operating somewhat independently but of necessity meshing with and functioning with each other (Ploog 2003, p. 489).

The first or earliest of these brains, in evolutionary terms, the innermost core of the brain, is the reptilian complex which 1) governs fundamental physiological operations such as blood circulation, heartbeat, respiration, food-getting, reproduction, and defensive behavior, 2) governs the regulation of an animal's or human's daily master routine and subroutines, and 3) is concerned with self preservation or self-interest (MacLean 1990, pp. 15-16; Cory 2006, p. 25).

The second brain, the paleomammalian complex, sometimes identified as the limbic system, is located largely on top of the reptilian brain. This brain provides for distinctively mammalian features (not found in reptiles) such as warmbloodedness, maternal care, parental responsibility, nursing, infant attachment, audiovocal communication between mother and child, play, and more generally family life and social bonding. It is also associated with emotional behavior and procreation (MacLean 1990, p. 17; Cory 2006, p. 26). In contrast to the self-interest orientation of the reptilian brain, it is oriented to caring, other interest, and empathy.

The third brain, the neomammalian complex or neocortex, is the large mass of hemispherical brain matter which envelopes the other two brains. This brain provides the capacity for problem solving, learning, memory of details, language, communication of thoughts and feelings, and the creation and preservation of ideas (MacLean 1990, p. 17; Cory 2006, p. 26). This brain allows humans to make behavioral adaptations to complex circumstances.

This human brain structure and circuitry evolved over millions of years. The three evolved brain modules or neural assemblies constitute a hierarchy in which the whole is greater than the sum of its parts (MacLean 1990, pp. 8-9; see also Wilson 2006, pp. 628-629). It is important to note that MacLean's contribution is integrative in nature, not reductive. As such he provides an overall evolutionary phylogenetic perspective on brain functioning that may not coincide precisely with findings of scientists operating in reductive and comparative structural functional modes. This is to be expected and does not invalidate his invaluable contribution. Even if it does not link exactly with evolutionary history, MacLean's conception of the triune brain can be seen as a powerful way to organize the functioning of the brain and relate it to the evolution of brain physiology.

MacLean's findings on brain physiology strongly indicate that empathy or other interest deriving from our paleomammalian brain is a core human motivation along with self-interest deriving from our reptilian brain. This is clearly at odds with mainstream economics.

4. The conflict systems neurobehavioral model a la Cory

The conflict systems neurobehavioral (CSN) model developed by Gerald Cory (1999) is based on MacLean's conception of brain physiology, and it elucidates the dynamic balancing behavior of the neomammalian complex. According to Cory, the two core motivations, 1) ego or self-interest deriving from the reptilian brain and 2) empathy or other-interest deriving from the paleomammalian brain, may

at times be out of balance. Due to this, it is up to the executive programming provided by the neocortex or neomammalian brain to attempt to bring about a balance between the two motivations (Cory 2006, pp. 26-27). When one or both of the two fundamental motivations cannot find satisfactory expression due to some unresolved conflict or blockage, behavioral tension and stress are produced in the individual. In response to this tension and stress, the executive program of the individual's neocortex will be enlisted to deal with the difficulty, thereby helping the person make a moral and rational choice, which should restore a sense of balance and reduce the tension (Cory 2006, pp. 27-28).

If the balancing is relatively successful, the outcome will be in the dynamic balance range which is associated with compromise, fairness, justice, and respect for self and others, a relative equality of ego and empathy (Cory 2006, pp. 28-29). Alternatively, the outcome might wind up in the egoistic range or the empathetic range, relatively unbalanced outcomes. The egoistic range is associated with power seeking, domination, assertiveness, competitiveness, and self over others. The empathetic range is associated with self-sacrifice, submission, responsiveness, supportiveness, and others over self.

Highly egoistic behavior involves neglected empathy and may be experienced as a sense of obligation to others. Whereas others to whom the egoistic behavior is directed may experience a sense of imposition, hurt, and desire to 'even the score'. Conversely, highly empathetic behavior may be experienced as neglected self-interest, a need to 'collect our due', resentment, and victimization. Both the above involve substantial tension. On the other hand, dynamic balanced behavior creates feelings of mutuality, shared respect, and a relative absence of tension. This dynamic balance is the central tendency and is associated with reciprocity, a universal norm of behavior. The CSN model tends to produce this balance, and thus cooperation, fairness, and morality, all outcomes facilitating desirable market and other economic behavior.

It should be emphasized that the CSN model is a descriptive model, not a normative model. On the one hand, the dynamic balance outcome tends to be produced by the homeostatic process involved, and it is a better outcome than those in the two unbalanced ranges. On the other hand, the balanced outcome should not be idealized. While it may be a good enough outcome to facilitate trade and other economic activity, it certainly does not imply that the behavior of the people involved is highly moral, saintly or 'enlightened'[2]. Nevertheless, the tendency to achieve dynamic balance does indicate that humans are strongly motivated by both ego and empathy and that they have a strong tendency to achieve a balance between ego and empathy that is favorable for economic activity[3]. This conception of the tendency off our brains to balance the two core motivations is further at odds with the mainstream economic view.

5. The model of the human brain in Dual Motive Theory

The above summary of Dual Motive Theory (DMT), which is MacLean's brain physiology combined with Cory's CSN model, implies a fundamentally different model of the human brain than the one implied by mainstream economic (ME) theory. Below the essential feature of these two models are

[2] In some of the writings of Gerald Cory (eg. 2006, p. 38) and Gary Lynne (eg. 2002; 2006, p. 649), there is a tendency to idealize human behavior that is motivated by a favorable balance of ego and empathy.

[3] In a comment on this paper, Patrick Spread informed me that "support-bargaining" offers an alternative interpretation to Cory's conflict systems neurobehavioral model. According to Spread, "Support-bargaining rests on the psychological premise that human individuals all need the support of their associates for a sense of personal security. But at the same time, they want to pursue their own individual interests. Hence a process of bargaining, in which individuals [act empathetically and] make concessions to the group, but only to the extent necessary to gain the support they need" so that they can get their way. This process results in a "balance" of egoistic and empathetic behavior.

compared. After considering the DMT model of the human brain below, this model is compared to the ME model.

The DMT model of the human brain includes features which contrast sharply from the ME model. These features include:

1. People have two dominant motivations.
    a. Ego or self-interest
    b. Empathy or other-interest.

2. The model seems to imply that people have limited, but adequate, cognitive capacity for most purposes (similar to but not the same as bounded rationality).

3. People have a strong capacity to achieve a balance between the two dominant motivations.
    a. Imbalances tend to be corrected with the rise of tension or stress which directs people's attention to the imbalance.
    b. Some imbalances are persisting and not easily corrected (they involve pathologies, either individual or societal).

4. Rationality has a different meaning than it does in the ME model. Rationality in the DMT model involves a) striving for a balance between self- and other- interest as well as b) application of logic and reason to attain the person's goals or desires. Rationality here is not a matter of simply maximizing the self's utility or satisfaction. It involves doing well for self and doing well by others, ie. living a successful well-balanced life. This rationality is necessarily a matter of both ends and means. Behavior is not rational unless the person is attempting to help others (ends) as well as making choices (means) conducive to one's own satisfaction.

5. Strictly speaking, the DMT model is silent on human mental capacities. The model, however, does imply that people have mental capacities such as the following:
    a. Empathetic capacity
    b. Cognitive capacity
    c. Integrative capacity, capacity for achieving a desirable balance between ego and empathy
    d. Achievement capacity, capacity for pursuing and attaining goals.

6. The DMT model emphasizes the particular human capacities that arise from the structure of the brain or brain physiology. The essentials of brain physiology are the same for all humans, ie. they are determined by human genetic makeup. In this model, differences among individuals in brain capacity would be due to individual genetic differences and differences in how these capacities change over different life stages (also determined genetically). The DMT model is silent on difference in capacity among individuals on account of differences in investment in human capital. This aspect, however, could be added to the model[4].

[4] Although the ME model of the brain is quite different from the DMT model, there is a sense in which the ME model can be considered a special case of the DMT model. This is the case where 1) empathetic motivation is zero, 2) cognitive capacity is very high, 3) rational behavior only involves striving for self, and 4) other brain capacities are low or nonexistent.

Implications for the mainstream economic brain model

Quite clearly, due to its fundamental assumptions, the ME model depicts 1) humans' cognitive capacities as unrealistically high and 2) a number of other human intangible capacities (empathetic, integrative, achievement, and perhaps others) as unrealistically low or nonexistent. With the perspective afforded by the knowledge of brain physiology from MacLean, Cory, and others, the ME model's depiction of human behavior appears to be very biased and unbalanced (see Rosenau 2006). For certain types of economic analysis, the ME model's bias might not matter much. But for other types of analysis where human motivation and intangible capacities are critical, the ME model's biases are likely to prevent us from achieving a satisfactory level of understanding. For the latter types of analysis, the DMT model is clearly advantageous.

It should be noted that the critique of the ME model deriving from DMT is similar to the criticisms of ME that have been made by a number of feminist economists. For example, England (2003, p. 34) explains that the individual in ME has a "separative self", a self which is autonomous, impermeable to social influences, and lacks the kind of emotional connection necessary to feel empathy. This self has been characterized as masculine because this is not a self capable of being caring, nurturing, altruistic, or attentive to family relationships and obligations (all associated traditionally with a feminine conception of self) (pp. 35-36). In a similar vein, Folbre (2008, p. 2) refers to love as being "as powerful an economic motive as self-interest" especially for work that takes place outside the marketplace.

6. Going beyond the basic dual motive theory

Paul MacLean and Gerald Cory have paved the way, but if DMT and knowledge of brain science are to be incorporated into economics, a great deal of work remains to be done. The rest of this paper is intended to make a start on this. Part of the task involves marshalling relevant knowledge about brain science and empathy. The other part of the task involves sketching a framework useful for explaining how this knowledge can be useful, ie. useful for providing a model which is based on the core features of DMT but which adds new features to make the model more human and more helpful in gaining a better understanding of important economic endeavors.

Findings of Brain Science: Brain Plasticity

The first important insight from brain science goes under the heading of brain plasticity (or neuroplasticity). Brain plasticity refers to the ability of the brain to change structurally and functionally as a result of input from the environment. This input could take many forms including external stimuli which lead to learning or events causing brain damage. It is widely recognized that there is a critical period in early childhood when the brain is extremely plastic. However, after the critical period, the consensus of neuroscientists has been that brain structure is relatively immutable. Recent findings challenge this conventional wisdom and suggest that many aspects of the brain continue to be plastic throughout one's life. So even though every human has basically the same brain physiology, and every human has particular genetically based brain capacities, a person's brain functioning is shaped as well by the individual's unique path through life.

The conventional wisdom of an unchanging brain anatomy after childhood is associated with the notion of localization. According to localization, the brain is like a machine made of many parts, "each performs a single function, so that if one of those parts was damaged, nothing could be done to replace it" (Doidge 2007, p. 13; see also Begley 2007, pp. 31-32). Localization implies, for example, that each of our six senses are processed in a different part of the brain and can only be processed in that part of the brain. Similarly, localization implies fixed specialization of the brain hemispheres

(Doidge 2007, p. 278). "One function, one location" expresses this hardwiring of the brain (p. 17). It is important to note that the view of an unchanging brain anatomy implies that human nature is relatively fixed (see xviii).

Much of the findings from recent research supports brain plasticity and rejects localization. There are three main types of brain plasticity. First is the brain change that occurs as the physical and chemical environment of individual neurons changes in response to external and/or internal events. These changes result in changes in neuronal connections or wiring, changing their strength and their current circuit patterns. Second is the anatomical brain change that occurs allowing a particular function to be performed by a different structure or set of brain modules (Doidge 2007, p. 276, p. 295). Third is neurogenesis, the birth of new neurons (Begley 2007, pp. 52-72).

The following insights are related to brain plasticity. First, "culture is not just produced by the brain; it is also by definition a series of activities that shape the mind" (Doidge 2007, p. 287). Second, "every sustained activity… including physical activities, sensory activities, learning, thinking, and imagining – changes the brain" (p. 288). Third, the "signature activities of a culture… require training and cultural experience and lead to the development of a new, specially wired brain" (pp. 290-291). Fourth, "'perceptual learning'… occurs whenever the brain learns how to perceive with more acuteness… or in a new way and in the process develops new brain maps and structure" (pp. 299-300). In short, our brains change to adapt in a variety of ways to the lives we lead (Begley 2007, pp. 8-9).

Brain scientists also find that being the recipient of prosocial behavior fosters positive brain change and that being the recipient of antisocial behavior contributes to harmful brain change (Eisler and Levine 2002, p. 18). This is particularly true for children whose early experiences determine to a great degree their later patterns of behavior (pp. 18-19). Among these determinants of later behavior are a child's family, cultural, and societal experience. One interesting finding is that damage due to the experience of "uncaring behavior is reversible when there is sufficient social support" (p. 40). Relatedly, psychotherapy has been shown to produce detectable changes in the brain. "Recent brain scans done before and after psychotherapy show both that the brain plastically reorganizes itself in treatment and that the more successful the treatment the greater the change" (Doidge 2007, p. 233). This all indicates that "we are not prisoners of our genes" (Eisler and Levine 2002, p. 39) and that humans can overcome poor experience and learn better behavior, changing their brains in the process[5] [6].

There are some other interesting findings that point in the direction of brain plasticity. For example, the cortical areas of meditators' brains have been found to be thicker than nonmeditators and are less subject to decline in thickness with age (Baime 2011, p. 47; see also Begley 2007, p. 8, pp. 212-242). There is also much evidence that the "human brain adjusts to environmental stress via gain or loss of cells and synapses" (Wilson 2006, p. 627). Another very significant consideration involves the distinction between personal temperament (largely inherited) and character (largely formed through experience) (Levine 2006, p. 622). Thus, character development is one aspect of brain plasticity, and it has three components: self-directedness, cooperativeness, and self-transcendence (p. 623). When all three character components are simultaneously at high levels, that has been found to be related to the healthiest mental functioning, and when all are at low levels, that is associated with the lowest levels of mental health. Two of the character components (self-directedness and cooperativeness) are strongly and directly associated with successful economic behaviors (p. 623).

[5] Cory (2006, p. 27) himself has acknowledged the role of life experience.

[6] According to Eisler and Levine (2002, p. 26), while genes no doubt affect human behavior, the inhibition or expression of those genes may be a product of conditions that people experience.

Findings of brain science: empathy

As explained earlier, we humans have an empathetic nature by virtue of the old mammalian part of our brains. Thus, a number of scientists have remarked that the capacity for empathy is a fundamental part of being human (see, for example, Eisler and Levine 2002, p. 25 and Singer 2009, p. 254). As the Swedish psychologist Ulf "Dimberg demonstrated… we don't decide to be empathetic – we simply are" by virtue of our humanity (de Waal 2009, p. 66). The German psychologist Theodor Lipps called empathy an 'instinct' in the sense that we have it from birth (p. 67). This is in contrast to the discredited view that empathy is a voluntary process requiring role-taking and higher cognition (p. 78). This latter view implies that "individuals faced with others in need decide whether to help, or not, by mentally tallying up costs and benefits" (p. 115). Even though empathy is apparently a natural, robust human trait that can be counted on, many societies make substantial efforts to foster and grow it. These efforts make sense in the light of the growing evidence of brain plasticity. It is also important to note that how an individual expresses empathy is determined by how one's empathetic quality has been developed and the degree to which it has been developed, both of which depend on the existence of brain plasticity (Begley 2007, p. 9)[7]. And as Singer (2009, p. 253, p. 265) points out, there is evidence that individual differences in empathy are a good predictor of whether or not a person is high or low in selfishness and whether or not he/she is engaged in prosocial behavior.

Eisler and Levine (2002, p. 12) emphasize that although empathetic behavior has a genetic basis, there is much evidence that experience influences development of the brain circuits involved in empathy (and many other behaviors). Moreover, they believe that experience interacts with inherited temperament and personality to determine behavior by altering brain chemistry and structure. For example, positive and caring experiences strengthen neural circuits associated with positive emotions and bonding, and this happens in the presence of dopamine and oxytocin (p. 13). Dopamine inputs are stimulated by generosity, creativity, and good moods; whereas oxytocin inputs are stimulated by positive emotions relating to social and family connections. The upshot is that our experience along with our long evolutionary inheritance jointly determine our very important capacity for empathy.

Historical perspective on empathy

Although empathy is part of the very nature of humans, it has manifested differently through the "great stages of human history – forager/hunter, hydraulic agriculture, and the First, Second, and emerging Third Industrial Revolutions" (Rifkin 2009, p. 612). Through these stages:

> "human consciousness expanded to encompass the complex energy/communications structures we created. Mythological consciousness, theological consciousness, ideological consciousness, psychological consciousness, and now dramaturgical consciousness mark the evolutionary passages of the human psyche. And with each successive reorientation of consciousness, empathic sensibility reached new heights" (pp. 612-613).

Whereas in forager-hunter societies, empathetic feelings rarely extended beyond those in one's tribe, "today empathy is beginning to stretch beyond national boundaries to biosphere boundaries" (Rifkin 2010, p. 2). According to Rifkin (2009, p. 452), this marks the "greatest surge in empathetic extension in all human history".

[7] According to Rifkin (2009, p. 177), "While primitive empathic potential is wired into the brain chemistry of some mammals, and especially the primates, its mature expression in humans requires learning and practice and a conducive environment."

The hard brain and the soft brain

To progress beyond basic DMT and gain greater understanding of economic behavior, what is needed is clearer conceptual links among 1) brain science, 2) the DMT model, and 3) economic concepts related to intangible capital formation. To start, we need a clear distinction between the part of the brain that is our genetic inheritance and the other part that develops in response to experience or as a consequence of our intentional efforts. Recall that the former part of the brain, which is associated with brain physiology, has been labeled the hard brain. It is the part all humans have regardless of age, education, or experience. The other part of the brain is the soft or plastic brain. Depending on people's activity, the soft brain becomes larger or smaller and reorganizes or restructures itself; it cannot be expected to develop or function in the same way in every person. The soft part also develops differently depending on one's family, culture, socio-economy, experience, and so on. As MacLean has explained, the hard brain of humans has evolved over many millions of years and has a large role to play in governing our relatively automatic or instinctive behavior. The soft brain, on the other hand, is to a large degree associated with the gray matter of the neocortex and is to a much greater degree concerned with intentional decision-making, planning, and problem solving in response to external situations. Some of the soft brain's development is a product of efforts by people to improve themselves, perhaps economically, and represents these persons' investment in intangible capital (standard human capital, social capital, personal capital, and so on). As indicated earlier, it is the soft brain that determines how the capacity of the hard brain gets expressed. Different individuals and societies are very different in this regard.

7. The revised DMT model of the human brain

Based on the scientific findings regarding brain plasticity and the brain's capacity for empathy, it makes sense to revise the DMT brain model. The following revised model adds two features (5 and 6) not present in the original model:

1. People have two dominant motivations (ego and empathy)

2. People have limited cognitive capacity

3. People have a strong capacity to achieve a balance between the two dominant motivations

4. People tend to be rational in two senses: applying reason to attain goals and striking a balance between self- and other-interest.

5. People have many different soft brain mental capacities, and people are often highly motivated to raise some of these capacities (especially empathetic capacity) as part of their human development efforts

6. People and societies differ in their mental capacities based on the soft brain (or intangible capital) investments that they have made. Their mental capacities are typically below their potential level.

Figure 1 below shows a side view representation of MacLean's triune brain with the three familiar modules. What is added is that the individual is making change efforts in response to the encountered external situation. As a result, the neocortex region of the brain experiences growth and restructuring

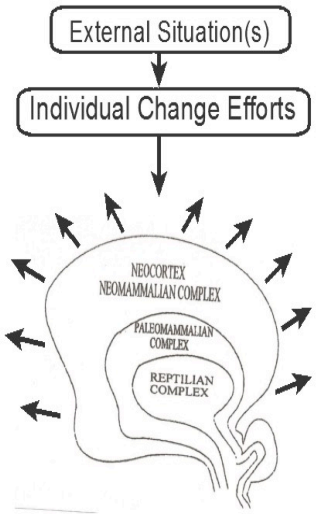
as shown by the outward arrows emanating from the neocortex. That growth reflects investments in intangible capital that produce change in the soft brain, more specifically growth, largely in the neocortex.

8. Implications of the revised model

Does the revised model which incorporates the possibility that people may grow in their soft brain mental capacities make a difference? Yes! When people, intentionally or not, grow mentally, becoming more competent to handle life situations, their brain's neural pathways develop. Following Goleman (1995, 1998), they develop emotional intelligence involving a variety of emotional competences. People's efforts in this regard can be interpreted as an investment in personal capital (eg. Tomer 2003) or as investment in noncognitive human capital (eg. Heckman and Rubinstein 2001). A person's investments in personal capital may enable them to manage their emotions better, raise their empathy levels, and attain a more balanced lifestyle. The particular personal capital investments that an individual is likely to make will presumably reflect their family, cultural, and societal background. In many cases, people will be motivated to make personal capital investments in areas where they are relatively deficient. Based on societal recognition of common mental capacity deficits, societal leadership may try to influence the overall direction of people's investments in personal capital, thereby attempting to overcome the perceived societal weakness. In general, it seems fair to say that people's overall development of their mental capacities typically falls far short of their potential. Thus, although this intangible investment activity may enable people to achieve a higher level of empathy and a better balance between empathy and ego, for most people it is unlikely to achieve anything close to an ideal. This is so despite the tendency toward balance, which is a prominent feature of the basic DMT model.

Figure 1

Growth of Neocortex in Response to Individual's Efforts to Deal with External Situation(s)



Source: Cory (2006)

Arguably, when an individual invests in intangible capital such as personal capital, this might contribute to the quality of the individual's empathy, improve his ego/empathy balance, and thereby improve the socio-economy's economic performance. When many people have greater empathy and a better ego/empathy balance, there will be more reciprocal behavior (following Cory 2006), enabling increased market economic activity. This is likely to lead not only to increased economic growth and development in poor, backward economies, but greater realization of the economic potential of advanced socio-economies. As Adam Smith seemed to recognize, "we achieve the true wealth of a nation only with… [a] symbiotic and integrated balance" of ego and empathy (Lynne 2002, p. 585; see also Sheeder and Lynne 2011, p. 443). Moreover, as Rifkin (2009) has suggested, the coming of the Third Industrial Revolution, which is associated with advanced, distributed information and communications technology and a distributed, renewable energy regime, will make possible further growth in empathy. To take full advantage of this possible empathy growth will require further investment in personal capital[8].

Sheeder and Lynne's (2011) research related to the conservation decision making of farmers in Nebraska and Kansas indicates that farmers' decisions are far from being purely self-interested. The evidence indicates that farmers by and large temper their pursuit of self-interest with consideration for the interests of others. Statistical tests of the decision making model are consistent with the hypothesis that farmers are motivated by both financial self-interest and empathetic concern for downstream water users with respect to their choices regarding farm processes and conservation. Further, the data indicate that the farmers' other-interest motivation is actually greater than their self-interest motivation. According to Sheeder and Lynne, "farmers seek to balance and integrate, seeking pragmatic "peace-of-mind" in contrast to strictly utilitarian solutions, finding what works best at the time to resolve and satisfy their dual interests" (p. 34). In other words, it is not only that in their decision-making efforts farmers are attempting to balance ego and empathy but that there is much learning effort (or investment) going into their empathizing and creating the right balance of ego and empathy.

It is interesting to speculate on the role of and importance of the different kinds of investment in intangible capital with respect to economic performance. Surely, some socio-economies can benefit more from intangible investment than others. Some socio-economies no doubt manifest typical pathologies that are associated with a lack of ego/empathy balance and too little investment in empathy augmenting intangible capital. Presumably, there are characteristic tension/stress patterns in these socio-economies that are associated with either an overly egoistic or overly empathetic pattern. Probably these unbalanced, and thus, poor performing socio-economies suffer from different types of market or economic failure. It should be noted that although investment in appropriate types of intangible capital might be the best prescription to deal with a situation involving a deficient ego/empathy balance, there are no doubt other ways to handle the problem. A socio-economy might for instance resort to developing norms, values, institutions, and government regulation to lessen their economic pathology.

9. An implication for government policy

One obvious general implication of the DMT model is that behavior reflecting a desirable balance of ego and empathy is good for both the economy and society. The implication for government policy is that government leaders who recognize a problematic ego/empathy imbalance occurring might want to try to exert a positive influence on the situation. There is a significant recent example that illustrates this. This is the situation that occurred in the aftermath of the shooting in Tucson, Arizona on January 8th 2011 that left six people dead and Representative Gabrielle Giffords severely injured from a

[8] Rifkin (2009, pp. 613-614) seems to recognize this. He states "The empathic predisposition that is built into our biology is not a fail-safe mechanism that allows us to perfect our humanity. Rather, it is an opportunity to increasingly bond the human race into a single extended family, but it needs to be continually exercised."

gunshot wound. President Obama discerned that in the US in the period prior to the shooting there had been an increasingly vitriolic public atmosphere. Accordingly, in a public talk, he asked, "What has gone so terribly wrong with America? Why are we becoming more aggressive, violent, self-interested and intolerant as a society?" (Rifkin 2011). He spoke of the rising plague of intolerance that is spreading across the land. Thus, President Obama called on Americans to "sharpen our instincts for empathy". This was a clear plea for a better national balance between ego and empathy.

While President Obama's response to the Tucson tragedy seemed appropriate, this action by itself is unlikely to be more than a short-run corrective. There is clear need for a sustained, long-run governmental response to the apparent lack of empathy. Thus, it is noteworthy that Obama's Tucson action is apparently part of a larger policy initiative. According to Rifkin (2009, p. 177), President Obama "has made empathy the core of his political philosophy and the centerpiece in his political decisions, from the conduct of foreign policy to the selection of Supreme Court Justices". Hopefully, this Obama administration policy initiative will lead to growth of the nation's empathic capacity.

10. The failure of mainstream economics – toward a more human economics

The development of a revised DMT model that utilizes not only brain physiology a la MacLean but important recent findings on brain plasticity and empathy makes the deficiencies of mainstream economics very clear. The failure of mainstream economics stems first from not recognizing the fundamental importance of empathy and second from overemphasizing ego, ie. it stems from self-interest motivation unbalanced by concern for others. This is not just an academic or theoretical problem. Mainstream economic conceptions affect society overall because these conceptions influence the design of institutions, the expression of values, and the development of norms, all of which, if relatively unbalanced, lead to excessive tensions and stress in the socio-economy and corresponding harmful consequences for the citizenry. If economics were to incorporate empathy as a fundamental motivation, it would mean putting less emphasis on incentives appealing to self-interest and putting more emphasis on developing institutions, norms, values, and ethics consistent with an appropriate balance of self- and other-interest. Further, it would mean that the man in economics is a full-brained man. And that would make it possible to conceive of creating not just a more efficient economy but a better society in which good human relations, win-win relations, predominate, ie. a more humane society.

With the rise in importance of behavioral economics and neuroeconomics, it is now incumbent on economists to be aware of what knowledge of brain functioning implies for economics. Economists need to recognize that human brains change, grow, and develop and that this change might be intentional and involve intangible investment. Economists also need to be aware of human differences in brain capacities and potentials as well as the limitations of our brains. In other words, the very basic assumptions about human behavior in mainstream economics need to change.

8. The Promise of the Revised DMT Model

In light of the revised DMT model, it is clear that simply relying on the innate human tendency for ego/empathy balance is unlikely to produce anywhere near optimal economic performance. To realize its economic potential, a socio-economy must make substantial investments in the types of intangible capital necessary to raise its empathetic capacity and its capacity for achieving ego/empathy balance. Economics (or socio-economics) might be quite different if it were based on a more realistic conception of human brain capacities. Here's a list of questions that a more human economics might be able to find better answers to:

1. In what way is human development involving growth in brain capacity an essential part of economic growth and development?

2. To what extent does growth of well-being require not just growth in cognitive capacity and tangible capacity but growth in a variety of intangible human capacities?

3. In what ways have unbalanced human development associated with societal pathology retarded economic growth?

4. Is it possible to anticipate, and thereby encourage, the growth in brain capacity that will be necessary in the future to take advantage of new economic growth possibilities?

There is reason to believe that an economics that is based on an accurate knowledge of the human brain, and thus human development potentials, will be able to speak more authoritatively regarding issues of economic development and be better prepared to answer the kind of questions listed above.

5. Conclusions

Gerald Cory has done a great job in bringing economists' attention to MacLean's perspective on brain physiology and in developing its implications for economic motivation. The resulting DMT has made a very important contribution insofar as it explains why humans are not motivated solely by self-interest but by a shifting balance of ego and empathy. While recognizing this contribution, this paper has sought to go beyond the basic version of DMT and to recognize that human behavior is not simply determined by the inherited triune brain structure. Further, this paper has sought to revise DMT by introducing brain plasticity, especially with regard to empathy. The resulting brain model is different in some important respects from the basic DMT brain model, dramatically different from the ME model, and has important implications for economics and policy[9]. The revised model is important because it highlights the failure of the mainstream model to incorporate empathy as a core human motivation and its failure to incorporate important new insights concerning brain functioning.

References

Baime, M (2011), This Is Your Brain on Mindfulness, *Shambhala Sun*, July, pp. 44-48.

Begley, S (2007), *Train Your Mind, Change Your Brain: How a New Science Reveals Our Extraordinary Potential to Transform Ourselves*. New York: Ballantine Books.

Cory Jr, G A (1999), *The Reciprocal Modular Brain in Economics and Politics: Shaping the Rational and Moral Basis of Organization, Exchange and Choice*. New York: Kluwer Academic.

Cory Jr, G A (2006), Physiology and Behavioral Economic. In: Altman, M (ed.), *Handbook of Contemporary Behavioral Economics: Foundations and Development.* Armonk, New York: M E Sharpe.

[9] There is reason to believe that the Adam Smith who wrote:

> "How selfish soever man may be supposed, there are evidently some principles in his nature, which interest him in the fortune of others, and render their happiness necessary to him, though he derives nothing from it, except the pleasure of seeing it." (Smith 1966, p. 3)

in *The Theory of Moral Sentiments* would approve of the revised DMT model.

de Waal, F (2009), *The Age of Empathy: Nature's Lessons for a Kinder Society.* New York: Three Rivers Press.

Doidge, N (2007), *The Brain That Changes Itself: Stories of Personal Triumph from the Frontiers of Brain Science.* London: Penguin Group.

Eisler, R and D S Levine (2002), 'Nurture, Nature, and Caring: We Are Not Prisoners of Our Genes' in *Brain and Mind*, 3, pp. 9-52.

England, P (2003), 'Separative and Soluble Selves: Dichotomous Thinking in Economics'. In M A Ferber and J A Nelson (eds.), *Feminist Economics Today: Beyond Economic Man.* Chicago: University of Chicago Press, pp. 33-59.

Folbre, N (2008), *Valuing Children: Rethinking the Economics of the Family*. Cambridge: Harvard University Press.

Goleman, D (1995), *Emotional Intelligence.* New York: Bantam Books.

Goleman, D (1998), *Working with Emotional Intelligence.* New York: Bantam Books.

Heckman, J and Y Rubinstein (2001), The Importance of Noncognitive Skills: Lessons from the GED Testing Program', *American Economic Review*, 91 (2), pp. 145-149.

Levine, D S (2006), 'Neural Modeling of the Dual Motive Theory of Economics', *Journal of Socio-Economics,* 35, pp. 613-625.

Lynne, G D (2002), 'Book Review: The Reciprocal Modular Brain in Economics and Politics', *Journal of Socio-Economics,* 31, pp. 581-586.

Lynne, G D (2006), 'Toward a Dual Motive Metaeconomic Theory', *Journal of Socio-Economics*. 35, pp. 634-651.

MacLean, P D (1990), *The Triune Brain in Evolution: Role in Paleocerebral Functions.* New York: Plenum Press.

Ploog, D W (2003), 'The Place of the Triune Brain in Psychiatry', *Physiology and Behavior*, 79, pp. 487-493.

Rifkin, J (2009), *The Empathic Civilization: The Race to Global Consciousness in a World in Crisis.* New York: Jeremy P Tarcher.

Rifkin, J (2010), 'Empathic Education: The Transformation of Learning in an Interconnected World', *The Chronicle of Higher Education*, May 30th, pp. 1-5.

Rifkin, J (2011), 'Will We Heed President Obama's Call for a More Empathic Society?', Huffington Post internet article, January 13th.

Rosenau, P V (2006), 'Is Economic Theory Wrong about Human Nature?', *Journal of Economic and Social Policy,* 10, pp. 1-17.

Sheeder, R J and G D Lynne (2011), 'Empathy Conditioned Conservation."Walking-in-the-Shoes-of-Others" as a Conservation Farmer', *Land Economics*, 87 (3), pp. 433-452.

Singer, T (2009), Understanding Others: Brain Mechanisms of Theory of Mind and Empathy. In: Glimcher, P W et al (eds), *Neuroeconomics: Decision Making and the Brain.* Amsterdam: Elsevier, pp. 251- 268

Smith, A (1966/1759), *The Theory of Moral Sentiments.* New York: Augustus M Kelley.

Tomer, J F (2003), 'Personal Capital and Emotional Intelligence: An Increasingly Important Intangible Source of Economic Growth', *Eastern Economic Journal*, 29 (3), pp. 453-470.

Tomer, J F (2008), *Intangible Capital: Its Contribution to Economic Growth, Well-Being and Rationality*. Northampton, Massachusetts: Edward Elgar.

Wilson, D R (2006), 'The Evolutionary Neuroscience of Human Reciprocal Sociality: A Basic Outline for Economists', *Journal of Socio-Economics*, 35, pp. 626-633.

Revisiting Arab Socialism[1]

Ali Kadri
Middle East Institute/National University of Singapore

Abstract:

The class in charge of development under Arab socialism was an alliance of the military with intermediate classes. Intermediate classes are sections of the working class differentiated by skill, higher income or occupation. In this relationship, the army held sway, and insofar as it placated populist aspirations for more egalitarian distribution, its role was progressive. The army and the intermediate class form the state bourgeois class, and exercise a collective ownership of the state. Hussein and Abdel Malek characterised the Arab socialist state as a surrogate bourgeoisie. Under Arab socialism, the predominant relationships remained capitalist and the repression over the labour process necessary for value creation persisted. However, by creating the financial space for the expansion of state-led industrial investment, investing in social infrastructure and undertaking vast land reform and redistribution measures, the Arab socialist model had outperformed the ongoing neoliberal model. Under the neoliberal model, the military re-allied itself with merchant/rentier capital and global financial capital. Generals became merchants, weakened the national front considerably, and the old state bourgeois class transmuted into a fully fledged comprador class. In this new ruling class alliance, the army no longer held sway. Global financial capital became the dominant player to which generals prostrated. Prior to the ongoing revolutionary phase, the Arab neoliberal state compressed wages, lifted its protection of national industry, set single exchange and interest rates and opened up trade and capital accounts with the intention of readying cheapened national resources for transfer abroad. This essay traces the metamorphosis of the state bourgeois class in Arab socialist countries, namely Egypt, Iraq and Syria, into the neoliberal/comprador class. With the exception of Iraq, which collapsed from without by the force of outright occupation, the other states experienced gradual social erosion leading to massive uprisings from within. Under the weight of successive Arab defeats, the state bourgeois class structurally fulfilled the terms of surrender and underwent a transformation from surrogate national bourgeoisie to surrogate international financial bourgeoisie.

1. Introduction

Early on in the post-independence years, 'Socialist Arab States', principally Iraq, Syria and Egypt, undertook massive measures of land reform, nationalisation of principal industry and financial institutions, provided universal health and education and clamped the circuit of resource usurpation. This class of Arab regimes sought self-sufficiency in production, endorsed import substituting industrialisation, and effected public investment in heavy industry when raising capacity. Many consider that the Arab socialist model has outperformed the ongoing neoliberal model, which began in the early eighties. Although standards of living are historically determined, the period 1960-1980 represented an epoch in which Arab countries exhibited a dynamic performance in terms of income growth, more equal income distribution, and improvements in infant mortality, life expectancy and other social indicators (World Bank 1995). Among Arab countries, Arab socialist regimes, in particular, have outshone the rest as a result of land reform, socialised health and education and, on average, higher per capita real income growth (Bush 2004). In comparison, the implementation of

[1] I am grateful to an anonymous referee for comments. This essay benefited also from significant contributions from an anonymous contributor. All shortcomings, however, are mine. The author declares no conflict of interest.

neoliberalism in later years, however, was to strip the working population of its previous social gains, all the while, raising the dose of repression.

Development outcomes depend on the commitment of a social class in power to real-resource mobilisation. This essay revisits the issue of Arab socialism and investigates the structure of the class formation that shaped development in the post-independence years. The class of military officers and its adjunct ally, the intermediate strata, including sections of the small land holding peasantry, hereinafter the state bourgeois class, by virtue of their control over state resources, assumed the role of development agent during the Arab socialist phase. The intermediate stratum is a differentiated and class conscious section of the Woking class. It is distinguished from the less privileged working strata the basis of salary, education, skill and, more decisively, by the degree of control over the means of production delegated to it by the military. The state bourgeois class is a surrogate national bourgeoisie (Hussein 1971 and Abdel Malek 1971). Account taken of the detailed aspects of the interface between national policies and their developmental outcomes, experience shows that the broad determinants of the relative success of the Arab socialist model hinge on the fact that they joined together security and developmental concerns. Security is a totality that encompasses, communal and national securities, with working class security being essential to national sovereignty. In the Arab region, imperial control of oil is pivotal to maintaining the global accumulation process. In this context, Arab autonomy over territory, resources and policy become *sine qua non* for development. Subsequently, reverse development or the de-development incurred by the Arab world under the ongoing neoliberal phase can be mainly attributed to the structural terms of surrender to imperialism. The neoliberal phase of development has not ended with the beginning of the Arab Spring. If anything, failure to redistribute assets and adherence to free market policies indicate that many of the dispossessing policies of the past will gain momentum under newly elected Islamic governments.

The turnaround from Arab socialism to neoliberalism occurred as a result of successive Arab military defeats and the rise of a neoliberal ideology bolstered, soon after, by the defeat of social ideology attendant upon the demise of the Soviet Union. These relevant historical events imparted the ideological foundations for the state bourgeois class in power to undergo an apostasy and lead the assault on working-class security, which would later compromise national sovereignty. The defeat of the Arab socialist project had more to do with a shifting class alliance that structurally absorbed the terms of surrender and the global ideological move away from socialism than with mass dissatisfaction with socialist measures. Egypt succumbed first as it signed the Camp David Accords, Syria followed with its second phase of neoliberal reform introduced in 2007 and, in the case of Iraq, despite many concessions offered by the regime, the precipitous destruction of the social formation was more crucial to capital than its gradual descent into the neoliberal quagmire. The metamorphosed social class that presided over the current neoliberal social disaster was no longer a surrogate national bourgeoisie; it became compradorial or an alliance with international financial elites and a proxy thereof.

To be sure, heavy state intervention and state-led investment during the sixties characterised the path of Arab socialist development. The state bourgeois class supplanted the national entrepreneurial class, promoted investment in heavy industry and, more generally, built productive capacity. The rise of this class is premised on the absence of a bourgeois class *proper,* which could have acted as an agent of development (Amin 1976 and Petras 1976). Not that the case of the Arab world is peculiar; in most of Africa and the Middle East, the post-colonial capacity debacle and the absence of an effective ‘entrepreneurial business’ paved the way for the rise of a ruling class made up of small land-holding peasantry and middle classes (Kalecki 1972, p. 164). This was a instance of an in incapacitated bourgeoisie in a capacity wanting context. The state bourgeois class became the subject of the post-colonial development project in many of the newly independent Arab states.

However, in much of the immediate post-independence Arab world, there was no total absence of an entrepreneurial class. In Iraq, Syria and Egypt, prior to independence, sections of the national bourgeoisie partook in virulent anti-colonial struggle. Yet, as soon independence was won, the social frailty of the national bourgeoisie came into evidence, especially as the umbilical cords with the colonists were severed. Upon the ashes of the old classes, there arose a set of military dictatorships competing in populist and genuine anti-imperialist positions. The old, especially the colonially reared sections of the bourgeois classes, were politically subjugated in the process of strengthening national independence. With the old bourgeois class weakened, the state in this group of Arab socialist countries became a substitute bourgeois class with a more expanded welfare function. In a tidal ideological atmosphere of equity and national liberation, the inequity of colonialism was scorned upon and egalitarian measures were implemented to redress odious development under colonialism.

Whether it is a bourgeois class per se or a state bourgeois class, the common denominator 'bourgeois' implies a similar substance to both of these classes. In an oil or geopolitical rent driven accumulation context and, given the pull of a cross-border class alliance, the state bourgeois class is innately predisposed to surrender under imperialist pressure. As a bourgeois class, it can furnish concessions to the working classes, but it cannot provide it with the civil liberties that would compromise capital as a social relationship. As early as 1959, Braverman foresaw the bifurcation in development lying ahead and warned that unless the working population politically participates in the making of socialism, the whole process of egalitarian distribution, undertaken by Arab socialism, could be easily reversed (Braverman 1959). In a paraphrased version of his expression, these Arab socialist regimes were dictatorships masquerading as revolutions. The authoritarianism and repression attendant upon the labour process implied socialisation and egalitarian distribution could be easily overturned from above without much resistance from below once historical conditions ripened. When frequent military routs instilled a state of defeatism in the upper layers of society and when working class and civil society organisations remained appendages to one-party regimes, the class in power easily overturned the gains acquired by working people under Arab socialism. That it is impossible to militarily defy nuclear powered imperialism was the thought that encapsulated the *état d'esprit* motivating the process of de-socialization. In preponderance of oil, imperialism is not drawn by trade and doux-commerce with the Arab world; it seeks destruction, resource grab and, ultimately control.

By the time the uprising occurred in 2011, the state of social devastation in the Arab world was so severe such that the masses poured into the streets without any long term revolutionary plan. Their principal aim was changing the status quo. In the state of retreat in social ideology, these were revolutions without a party vanguard, strategic social programs and where the only organised groups were free-marketeer islamists already operating under the auspices of imperialism. This was the outcome of a capital accumulation process commandeered by international financial capital.

The political relationships of differing social classes to imperialism, their relationship to the means of production and informal forms of property control via the medium of the state underscore the transition from socialisation to privatisation. The social mutation of the ruling elites from national bourgeoisie to agents of global financial capital occurred as a result of successive imperialist military victories over the Arab world and the receding zeal of socialist ideology. Under the clout of international financial capital, it is arduous for an alliance of national classes in the Arab world to maintain a nationalist predisposition in accumulation. Gradually, the Arab ruling elites internalised the terms of surrender dictated by the Bretton-Woods institutions and became modern suzerains for empire.

In exploring the subject of this essay, the argument will pursue the patterns of shifting class structures and the changing mode of appropriation in a class determined capital accumulation process. On its own, the topic is broad and a single essay will not do it justice. But the purpose of this contribution is to supplement the debate by bringing into consideration the class determinants of development. Capital accumulation in a

developing context cannot be read using a neoclassical framework of full employment. A full employment hypothesis does not apply in a developed context, leave alone a developing context. Although necessary, demand-side approaches are also inadequate on account of frail capacity in the developing world. What is there to demand when there is little capacity to put to work. Capital accumulation is not solely about the incremental rise in the capital stock over time, it is primarily about the social relationship that underpins the decision to build and re-circulate the social product nationally. Hence, the point is to explore the primacy of the social class in charge of development. When the state bourgeois class represented an alliance of the working class with the military, one witnessed a marked developmental impact. However when the class alliance became that of the military with comprador capital, one notices developmental retrogression. The essay contends that changes to the class structure in the subject group occurred under the weight of consecutive Arab defeats and a concomitant rise of neoliberal ideology globally. The state bourgeois class structurally fulfilled the conditions of surrender and underwent a transformation from national bourgeoisie to surrogate international financial bourgeoisie.

The line of thought tracing the changing nature of the state bourgeois class from nationalist to comprador will be structured as follows. In section one, I outline the relevance of the concept of state capitalism and state-led development in a developing Arab context. In sections two, I investigate the state bourgeoisie, its frailty and predisposition towards capitulation to imperialism. The allegedly weak entrepreneurial class in the Arab World is dealt with in section three. No discussion of this specific Arab mode of development can proceed without exploring the dominant role of the military in power. In section four, the rise of the military to power, its transformation into a comprador segment of the ruling class and the versatility it exhibited when faced with successive defeats are exposed. Section five deals with the historical achievements of the Arab socialist experience and the reasons behind it collapse.

2. State-led development

In characterising socio-political entities resembling Arab socialist states, Kalecki employed the concept of 'intermediate regimes'. He describes those regimes as being neither strictly capitalists, as foreign influence on them was limited – nor socialists. On an international level, intermediate regimes obtained credits from both the socialist and imperialist countries. In the case of the Nasser regime for instance, the 1958-1961 archives of the state department illustrate how Nasser acquired Hawk missiles in addition to being taunted with credit from the US in an effort to dissuade him from alliance with the Soviet Union[2]. Gramsci (1978, p. 409) also tackled similar class formations in peripheral Eastern and Southern European societies and noted the conspicuous rise of an intermediate stratum. His comment relayed how in peripheral countries… a broad spectrum of intermediate classes stretches between the proletariat and capitalist-classes which seek to carry on… policies of their own, with ideologies which influence broad strata of the proletariat, but which particularly affect the peasants and rural communities[3]. But the intermediate strata in the Arab world cohabitated with the military in a context of war and did not own significant property – apart from that for personal use. The power afforded to the intermediate strata during the Arab socialist age was delegated to them by the military. Theirs was a subordinate position to the military, but they effectively became part of the ruling nomenclature by their indirect control of the economy via the state. When, at the beginning of the neoliberal age, the lingering socialist stance of certain sections of these intermediate strata contradicted the terms of structural submission to imperialism,

[2] http://www.state.gov/r/pa/ho/frus/kennedyjf/xvii/17717.htm

[3] The state bourgeois class, reasoning that food security is part of national security, glorified farming and invested substantially in agriculture. When the neoliberal age dawned, investment in agriculture fell and so did agricultural output. Nearly 100 million people left the countryside to urban areas between 1980 and 2010. The uprising cohort is principally composed of pauperised rural migrants. http://www.globalresearch.ca/index.php?context=va&aid=30519

the military retracted its delegated authority and allied itself with the local merchant class and, more decisively, it became a subordinate partner of global financial capital.

Within the class of intermediate regimes, a narrower band of regimes closely resembling the Soviet model were dubbed state-capitalist regimes, (Petras 1976; Binns 1986; Binns and Hallas 1976; Cliff 1974/1955; and Burnham 1945[4]). Cliff developed the concept of state capitalism in relation to his characterisation of the Soviet state. Extensive socialisations in land and social resources, which were carried out under Arab socialist regimes, would indeed bear resemblance to the Soviet model. However, Arab socialist regimes did not uproot the old bourgeois class. They never went as far as the Soviet Union in extirpating the private owning national bourgeois class. In their five-year plans, their inter-industry prices were not fully set at cost (shadow price), and where the private sector was involved in the production chain, a price mark-up generating profits was permitted up to a certain level. A cap on the prices of goods provided by the private sector mirrored the depth and the power of this sector's involvement in the broadly socialised economy. Until 1980, these price caps were only slightly above cost and restricted private sector expansion. In the neoliberal age, price-capping was removed except in Iraq, which was at war and rationing food. When, for instance, in 2007, Assad liberalised prices and removed price caps, the private sector, which is mainly the president's kin, assumed full control of the Syrian economy. In Egypt, Sadat began the process of liberalisation as early as 1980. In the age of socialisation between 1960 and 1980, however, the private sector in these states survived as a repressed component of the Arab socialist economies.

The very notion of state-capitalism, however, can be vacuous or a sort of one size fit-all concept. It is unlikely that any institution under capitalism could escape the totality of capital as a social relationship. Repression/regimentation of wage labour and resource allocation on the basis of profit are ubiquitous. The abolition of capital as a social relationship under state ownership of national resources, although possible, if only on account of structural shifts in history, is unlikely to occur precipitously and in a single developing country. This is so even if civil liberties and working class participation in the political process were effected. Imperialist assault aimed at snuffing successful socialist models and the unevenness inherent in capitalism would necessitate redistribution under which surplus value would be generated and unequally allocated in the early phases of socialisation. Either equalisation payments between sectors or industrialisation would imply a distorted surplus value distribution. This would also entail exploitation, regimentation and protracted aspects of repression pertaining to a capitalist labour process – these salient characteristics will emerge all the more so in developing formations whose aim is to industrialise. A working class consciousness where public property is fully public ownership and state owned, is a historically contingent process that is dependent upon several parameters including the level of advancement of productive forces and the successes of socialising appropriation. Thus, for capital as a social relationship to promptly self-disband after state ownership of the means of production is not at all a plausible assumption even under a working class participatory democracy. Not that it is impossible to make historical transformations, for the very idea of impossibility is a form of thought that is irrelevant to social agency within a historical process where change is a compromise between the real and ideal.

To boot, the traditions and symbols of oppression are not immutable, but they are historically cast and would, *pari passu*, thwart the progress of socialisation. The incalcitrant force of habit alone can prove intractable. For this reason, the use of the term state capitalism tallies with any state under capitalism or those, which are in transition to socialism and, hence, its meaning is only discernible within a specific structure of a

[4] The concept of state capitalism arose in relation to the Soviet Union, which was isolated in the midst of a hostile capitalist world, and did not sui generis possess the means and capabilities for the construction of a socialist economy. Completing socialist construction contingently relied upon the unfolding proletarian revolution in the more advanced capitalist economies of Western Europe, which would later supply Russia with the prerequisites needed to develop a socialist workers' state (Cliff 1974/1955 and Mandel 1951). For several years after the October revolution, the official language of the Comintern remained German in the hope that Germany would become the next socialist revolution.

given discourse. For the purpose of this essay, the class that owns the means of production through the state and is in charge of development in the subject group of countries will be referred to as the state bourgeoisie.

The very act of socialising represents a self-negating process of capital. The social relationships holding capital together as a state of exploitative existence tear at each other. In the intermediate term, socialisation disciplines but does not abolish capital. Nationalising assets, combined with a democracy for the working class, is a requisite step for a socialism growing in the midst of hostile milieu. However, within the class of socialist Arab states under question, the socialisation of the means of production was incomplete and, working class repression abounded. 'Arab socialism' was intrinsically capitalistic and it could not have bonded the national front in anti-imperialist struggle, which is the quintessential condition for a raw material-third world developmental model. Save the interface between policy and outcome, an anti-imperialist position, stemming from a sovereignty substantiated with working class security, provides the autonomy over policy that would mobilise resources in line with social requirements. Arab socialism, in this regard, was moribund but, a model nonetheless which had put into place the necessary measures that were to redress the catastrophe of mal-distribution and underdevelopment inherited under colonialist rule.

Arab socialist regimes undertook three vital steps in consolidating their post independence position. The first step was the confinement of policy to the remit of the state and its control over natural resources, which would harness the surplus to be redeployed in national development projects (Petras 1976 and Amin 1978). The second was agrarian reform, which concurrently limited the political power of the traditional landlord class. The third was economic nationalization of large-scale financial and industrial institutions. Typically, these regimes supported import-substitution strategies, controlled the capital account and invested in heavy industry and infrastructure. The state arose as the chief owner of the means of production and appropriator/allocator of the social product. In all of this, the private sector was not wholly superseded. It absorbed a minor proportion of the labour force in services and traditional activities. State ownership existed side by side with a constrained private sector, where with the certitude of hindsight, the private sector's scope for expansion would be reignited when the state bourgeois class in power underwent a metamorphosis into a fully fledged comprador bourgeoisie.

Economic planning and government intervention in relative autarky represented the means by which the foremost binding constraint of underdevelopment, which is the financing needed to galvanise national resources, was to be overcome. The national currency circulated without the trepidation of the international financial market. The current account real and capital balances were devised to suit industrial policy. Several interest and exchange rates were at play to attenuate the impact of foreign exchange shortages on the national issuance of currency. Situated historically, this was a time when the idea of free markets was ebbing relative to the advance of socialism and national liberation movements around the globe. Government ownership and massive intervention were based on the inadequacy and inefficiency of market mechanisms, the need for social control in investment strategy, and a more egalitarian redistribution of income (Kalecki 1976). This was all to change under the weight of several Arab defeats that ended with the occupation of Iraq, the rise of monetarism and the deepening of the crisis of socialist ideology assuming catastrophic depth upon the collapse of the Soviet Union. The very state bourgeois class in power shifted its allegiance from the intermediate strata and peasantry to merchant-rentier class and, more decisively, the global financial elite. From around 1980 onward, the age of US/Saudi neoliberalism eclipsed the Nasserite period of Arab history.

3. The state bourgeois class

The state bourgeois class, composed of the military and intermediate strata, spun bureaucratic control over state and economy via a one party system controlled at the helm by the military. The state bourgeois class in symptomatic terms consists of the party bosses who are at once the top ranks of the military forces, the upper

level of the state bureaucracy and the various regime-allied civil society organisations led by professionals from the intermediate classes. The litany of scholarship on stratification in the Arab world, however, pursues a more disciplined approach in explaining the autocratic nature of Arab regimes relating it to the 'cultural mentality' of Arabs. The cultural paradigm, which is a disguised racial differentiation collated to the practice of imperialism, presupposes inherent traits that are inbuilt into Arab culture. At a recent workshop at Harvard University, two researchers were running separately two fifteen hundred years regressions to show that either by the length of time a ruler stays in power or by the number of elected councils, the Western world appears more stable and more democratic across fifteen centuries[5]. The book entitled 'the Arab mind by Raphael Patai', the manuscript of choice for informing State Department personnel about the Arab world, is of course the epitome of the cultural stance.

But in more historically concrete definitions of social classes, however, there is to begin with a prevalence of class relationships over a social class. For Marx, in particular, social classes cannot exist outside class relationships that tie them together and analysis of class in diverse formations must begin at this point. It is within these relationships that class structures, including their history and evolution, are constituted. What a class is and does depends on where it is situated in relation to other classes, not only in terms of relationship to property, but also in terms of its forms of organisation and ideological leanings in relation to imperialism. In other words, classes are a macro-sociological personification of relations between classes including their subjective, cultural and symbolic dimensions. These relationships between classes are founded upon social relations of production, which are capitalist and will have to be defined in their specificities and development. Here, the specificity in the Arab world is the newly independent formation which came under imperialist assault as soon as it proclaimed its independence. To the extent that these relationships between classes rest upon relationships of production, they are essentially, but not exclusively, relationships of domination and exploitation, not only within the national borders, but also in relation to the global accumulation process. This latter point encapsulates the key relationship of national classes to imperialism. Eurocentric thought regularly fails to observe the intersection at which national peripheral classes meet with imperialism and are accordingly defined by this relationship. When gauged in relation to imperialist control of oil, the lines of demarcation separating differing national classes in a post-colonial Arab context would be best determined in relation to the formation of an anti-imperialist national front. A notional ruling class is not ruling in relation to the working classes in the Arab world alone. It is also ruling in proportion to the strength of the bonds tying it to imperialism. Seen from this perspective, for instance, the House of Saud as a constituent of a social class and its Wahabism are modern constructs of British/US hegemony and, pointedly, have little or no referent in Arabic-Islamic history prior to the modern age of imperialism.

The somewhat elusive intermediate class namely comprises skilled professional, eg. schoolteachers, university professor, civil servants, accountants, military officers, medical doctors, engineers, and lawyers, whose status is not dependent on the ownership of property and wealth, but based on training and performance (Petras 1976). Its elusiveness stems from its wavering position vis-a-vis imperialism. This stratum is a differentiated section of the working class that supplies recruits to the elites. It is considered a subsidiary force in the government of society. The stability of a post-independence Arab political entity partly depended on the power delegated to this strata by the military. The military personified the dominant faction within the ruling class alliance. Petras (1976, p. 439) defined the 'intermediate strata' as a class conscious and independent social stratum – apart from workers and from the traditional landowners – that is horizontally and vertically linked to the salaried middle income strata and, which has its own political and economic agenda and, whose promotion for market relations and capitalist development is extended under the expansion and

[5] 'The Role of Economic Institutions in the Organization of Middle Eastern Economic Life in the Modern and Pre-Modern Periods', Centre for Middle Eastern Studies, Harvard University, March 2011.

shadow of state enterprise. The new Arab ruling post-independence class reinforced its grip on power by populist appeal to pan-Arabism and socialism.

The alliance of the intermediate strata and the military contributed to the partial erosion of some traditional classes (landowners and colonialist-bred bourgeoisie) and the promotion of others, such as the new and more expanded intermediate strata as a result of free universal education and health. The expanded intermediate strata were the professionals who sprung around the vast social programs of the state. While an industrial bourgeoisie was crippled politically either because its ties to the previous colonist were severed or, because the post-independence crisis was too steep, it became conceivable for the state bourgeois to fill the class vacuum and expand a bureaucracy around the state (Abdel Malek 1967). In these new social relationships of production, the state-bourgeoisie organised around the state, maintained not only their relations to the means of production through the sate itself, but also distributional arrangements measured against the pace of capital accumulation and regime stability. This is the organised dimension of capital or a political process mediating the inevitable capitalistic growth of the state bourgeois class. Solidifying control over state power becomes a goal meant to entrench the newly fattened class. The state, in return, would promulgate the law in tandem with the new-class' own expansionary pace. Under more stable conditions, working classes in relation to the dominant military class would have maintain their organisation, struggles and experience, which would influence the cohesion of the working class a whole, including the peasantry. In conditions of war, these workers rights were subjugated by the alibi of national defence. The degree of egalitarian redistribution of the social product was proportionate to internal and external power considerations meant to maintain regime stability. In the Arab socialist age, national security and developmental objectives were relatively intertwined and ran parallel to each other. The first major departure of development from security was to occur when Egypt signed the Camp David Accords in 1978 and flew in the US orbit. This step weakened an already frail Arab security arrangement, uncovered working class security and opened a window for imperialist intervention in much of Africa as well as the Arab world.

While under privately ordered capitalism, the process of capital accumulation is mediated through various legal and political institutions organised by individual capitalists owning shares or cadastral property titles; the process under Arab socialism is organised by officials of the state or their hired managers (Abdel-Malek 1971). As the build-up of contradictions inherent in class society emerged as a result of regressive mal-distribution and successive defeats to imperialism, an internal clash unfolded between the state capitalist class, which owned state resources, and an increasingly socialised labouring class. This phenomenon has also been witnessed elsewhere. Kalecki observed in respect to the general case of intermediate regimes that the demands of a growing urban labour force has brought pressure to bear on the state bourgeoisie and repression against groups representing rural and urban paupers rose (Kalecki 1976, p. 35).

The class fault lines under Arab socialism progressively became pronounced. In the latter years of the Arab socialist project, access to foreign exchange in the form of petro-rents permitted sections of society to visibly display signs of wealth – first in Egypt and then in Syria. Iraq is a particular case, which was too engrossed in conflict for foreign earnings to weigh in on class divisions. Under Saddam's regime, the force of the law dealt quickly with black-market currency traders. The initial populist euphoria, which accompanied state-bourgeois led development, was followed by more repressive measures against the working class, especially as the walls protecting the national industry came tumbling down under stealthily infiltrating petro-dollar from the oil states. Dollars and petro-Dinars, traded in the black markets, allowed similar labour which is denominated in petro-currency abroad to acquire much more wealth at home as a result of repatriation and black market dealings. The lax attitude of the regime to black market operations indicated that certain segments of the military and state bourgeoisie were beneficiaries of the illicit activity. Gradually, the multiple exchange rate regime protecting national assets from being bought at prices set by the imperialists came tumbling down.

These multiple exchange rate structures came into effect to redress the prices consigned to differing quality of products or national versus imported goods. Apart from regulating imports, the multiple exchange rates disproportionally raised the price of imported commodities in order to protect national products. Multiple exchange rates addressed a lopsided power structure that trailed from the days of colonialism in which the discrepancy between the real valuation of currency and its nominal one were set according to the desires of colonialism. Trade unions and less favoured sections of the working class bore the worse of the assault on the currency when the cohort of the professional strata veered to the right in pursuit of petro-dollars and class alliances began to shift (Hussein 1971, pp. 188 and 281). But a caveat is in order here, the intermediate stratum and the military's gradual shift to the right was preceded by several military defeats and a Camp David Accord.

At a general level, the case may be that the emergence of this class of Arab socialist structures is a movement from one type of exploitation to another without a radical shift in the value-extraction process attendant on accumulation (Petras 1976, p. 442). The increasing mechanisation of production meant that there was more of a relative surplus being produced vis-a-vis the previous era, while the additional product was more equally distributed. More importantly, there was resource retention within the national economy. The act of re-circulating national resources arrests the monetary and non-monetary elements that form value from being transmitted to the ex-colonists. Prior to independence and the rise of etatism, the colonial circuit of value transfer, by which national resources were grabbed at minimal prices, prevailed. The transition from the post-colonialist bourgeois formation to state-ownership led formation occurred with only formal changes to the fundamental construct of capitalist property relations. The appropriation of value proceeded via class control through the state as distinguished from control through private/titled ownership. Industrialisation drove up the degree of relative exploitation; however, the additional national wealth was distributed in a more egalitarian fashion. Reinvestment in the social infrastructure, particularly, in health and education, was significant and recycled value back to the working classes. Arab socialism was an era in which a country, such as Iraq, had gone from an eighty percent illiteracy rate and one of the highest income inequality profiles in the world to a state of near complete literacy and relatively fair wealth distribution (Todaro 1979).

Social change, in particular, the leap from colonisation to Arab socialism, is to be gauged as a matter of degree and, specifically, in the overall betterment of living condition. Although a fundamental breach with the past in terms of welfare occurred, capital as an exploitative relationship stood its grounds. Despite a reduction in the degree of income inequality under Arab socialism, workers and peasant remained passive and disengaged participants in the social and political process. One party rule and corporatism disengaged the impetus that internalises the culture of resistance and social change. It created, by the degree of repression, workers prohibited from organising or erecting the defence mechanism against the erosion of the achievements acquired under Arab socialisation. Insofar as fundamentally revamping exploitative social relationships, the post independence experience of this class of Arab countries is not a structural transformation *per se* but rather a transition, in the sense that no radical changes to the totality of capital as a social relationship were observed (Abdel-Malek 1971; Pfeifer 1979; Petras 1976). Most significant in this process is the practical banishment of autonomous working class organisations by state suppression partly on account of conflict and under the pretext of defending the homeland. Notwithstanding the fact that the state-led development experience of these developing formations lifted their economies out of their decrepit postcolonial status, the construct of Arab socialist structure rested principally on a single determining moment: an autocratic state bourgeois class rose to become the collective owner of the means of production. However, the rising dosage of autocracy represented a readymade conveyor belt that had cut the costs of the transition to neoliberalism in the early stages of social transition. As welfare dwindled, repression rose.

In the colonial age, as elsewhere, Arab industrialisation was curtailed. A classic example of that is Alduri's illustration of the British suffocation of Egyptian textile in the mid nineteenth century (Al-Duri 1969). Not

that colonial plunder requires proof, the splitting up of Tripoli, Aleppo and Mosul (1917), which had formed an embryonic industrial hub in the early twentieth century, further corroborates colonial anti-developmental bias. Merchant capital activity thereafter represented the pervasive activity for the capitalist class. Ever since independence, conditions of uncertainty associated with a combination of imperialist assault and institutions devoid of working class representation rendered short-term profiteering the principal undertaking of private investors. Private investment until today is centred on short-term gestating capital and commercial undertakings. Colonialists did not breed an autonomous bourgeoisie. The bourgeois class that they weaned in the immediate post-colonial years was financially and, more pertinently, in terms of real resources, incapable of promoting developmental investment. When development is considered akin to industrialisation, it then follows that a national class incarnating the capabilities of the state, other than the national bourgeoisie, should have assumed the responsibility for development. Only state resources can challenge post colonial dilapidation.

There is quite a difference between colonial plunder and the late developer syndrome. Late developer syndrome imposes enormous burdens on newly independent countries such that the possibility of solving problems by means of individual entrepreneurial activity all but disappears (Buick and Crump 1986, p. 46). But late developer sounds innocuous and bereft of politics. Colonial plunder, however, never receded in the Arab world and the imperialist assault regained momentum after the fall of the Soviet Union. A constant state of war or the serious threat thereof were endured by Arab late 'developers', which not only forbade the emergence of a national bourgeois industrial class, but also increased the risk to private investment to the point where it became almost futile. The role played by the bourgeois class had to be passed on to other social classes, which were no less committed to the accumulation of capital than a typical industrial bourgeoisie. The new class alliance of military and intermediate strata sought to achieve development by social means as distinct from individual entrepreneurship (Abdel-Malek 1971). Such a process paved the way for major economic, social and political changes to be carried out from above, either through the state or through an alliance with the soviet bloc (Turner 1984, pp. 61-62 and Petras 1976, p. 440). But it is relevant to recall the order of priority. It is not only the vacuum of an apt national bourgeoisie that accounts for the agent of development to be transposed in and around the state; conditions of imperialistically infused debilitation wrought upon the Arab world have all contributed to the development of this environment. Key among these factors, were successive Arab defeats to Israel. A spectre of war had since haunted the region. The dangers to national security provided an alibi for the state bourgeois class to repress and exploit at will, and to later position itself to ascend as an imperialist satellite of the world capitalist order.

In the transition from state-led to privately-led development, there occurred a rising genie coefficient (income inequality gaped away), a trade policy suitable to WTO standards of liberalisation, the annulment of multiple exchange rates, the liberalisation of basic commodity prices (removal of price caps) and the pegging of the exchange rate to the dollar while facilitating capital transfer. Seen as a totality, these measures imply one sordid condition which is to facilitate the transfer of value or non-monetised assets and value forming resources to the ex-colonial powers all anew. Additionally, as a result of the receding power of the Soviet Union and its eventual collapse, the reign of neoliberal ideology became yet more profound. The state bourgeois class, which had guarded the seminal social relationship of capital under the guise of socialism, leapt into the condition of fully fledged capitalism. The military altered its class alliance from the intermediate strata to the merchant class and, more importantly, it became a subsidiary of global financial elites. In the financialisation phase of imperialism, the disarticulation endured by the Arab social formation as a result of a gravely imbalanced power structure and resource usurpation wrought havoc upon class alliances and structures. At the level of working class consciousness, the divide between the state of being and an historical grasp of that state forked away as social ideology took a dive. The Arab state bourgeois class of the neoliberal age assumed a fully compradorial role and merged with global capital. Inherent capitalistic tendencies, pitting the private against the social, in this genre of Arab capital egged on integration with global capital. However,

the outcome was not inevitable, until military and ideological defeats amassed against socialising states in the eighties and were too grave to be withstood.

4. A frail Arab bourgeois class

In the two decades that followed World War Two, more than one-half of the population remained in the countryside of the Arab world and the industrial working class amounted to a relatively insignificant proportion of the working class (Turner 1984, p. 54). The commercially engrossed ancien bourgeois class circulated capital in a mercantilist fashion. It was money, returned to money without significant value added to commodities, and with heavy reliance on imports and few goods produced by local means[6]. The entrepreneurial skill swung towards trade as opposed to industry (Turner 1984, p. 53). Both Berger (1958) and Turner (1984) agree that the weakness of entrepreneurship stemmed from the fact that the merchants and small retailers represented a large proportion of the bourgeois class. Moreover, the small number of manufacturers emerged from the ranks of the merchant class itself. The vast majority of industrial enterprises were founded by traders or financiers, generally merchants, engaged in foreign trade. No class on its own had the capital necessary for post-colonial developmental undertaking (Issawi 1955, p. 131).

In the uncertainty related to weak states and intermittent wars, to shy away from long term industrial activity represented a reasonable individual choice. State collapse in Lebanon, Syria and Iraq are evidence that no animal spirit can endure this structural shift. It is not entrepreneurial psychology per se that inhibits investment in long term industrial developments, it is war and the prospect of state collapse. Bearing in mind that imperialist aggression is premised upon encroachment, in as much as socialist Arab states redressed the balance of forces with the imperialists through redistributing and solidifying the internal class alliance of workers and peasants, they were capable of development. Working class security, which substantiated sovereignty, was a *sine qua non* of development; a lapse in one means a lapse in the other. The dynamic of class polarising capital accumulation under Arab socialism was accentuated by imperialist meddling meant to fragment and weaken resisting Arab formations.

Over the past thirty years, with the commencement of the neoliberal age, the social conditions in the Arab world deteriorated to a point where roughly half the population was spending half of its income on basic food consumption[7]. Investment fell from 30 percent in 1980 to around 18 in 2010 (WDI, various years). The regional rate of unemployment was one of the highest globally and the share of labour in the form of wages fell to around a quarter of national income (KILM-ILO, various years). These were but a few of the results of the structural terms of surrender dictated by imperialism via the interlocution of the World Bank and IMF policy. At the same time oppression became more expansive. Apart from the gruelling human rights record of Arab regimes, a 2007 report by the International Trade Union Confederation, for instance, indicates that 'workers in the Arab region still have fewer trade union rights than anywhere else in the world'[8]. Imperialism, eying oil control, drives an agenda of anti-development in the Arab world. Thus, one is well advised to recall the definitive historical context before judging on the basis of psycho-behavioural assumptions whether a risk taking entrepreneur exists in an Arab context. As Keynes was apt to differentiate, risk is calculable but

[6] The role of a capitalist and a merchant class are not to be confused, as the merchant class acts as a circulator of money rather than convertor of money into physical elements, used for producing wealth (Fine and Saad-Filho 2004). That is to say, the process is that of distribution of goods rather than production of goods and creation of new value. The profits attained from the process are effectively a transfer of value from members of society to the merchants.

[7] http://www-wds.worldbank.org/external/default/WDSContentServer/WDSP/IB/2011/05/27/000001843_20110601143246/Rendered/PDF/P126506000AWIFS000PID000Concept0Stage.pdf

[8] Annual Survey of Violations of Trade Union Rights (2007).

uncertainty related to war is wholly unpredictable and could imply structural realignment[9]. The Arab region is as highly uncertain now as it was throughout the second half of the twentieth century. The imminent prospects of war have essentially thwarted development.

5. The military in power

The post-independence capture of state power by socialist Arab regimes occurred mainly by coup d'état. Not that it is unusual, in most modern Middle Eastern countries the colonially trained military emerged as a social force pursuant to the disbanding of previous colonial military formations (Owen 2004, p. 180). They expanded after independence and most officers were recruited from intermediate strata or small-holding peasantry. Given the state of tension that reigned in relation to imperialist assault and the formation of the state of Israel, armed forces played a central role in the politics of Arab states. Coups and the recycling of military regimes were a common feature of postcolonial Syria and Iraq. Although coups were not symptomatic of the Egyptian case given the popularity of Nasser, the May-corrective movement (1971) undertaken by Sadat was in many respects a putsch. The higher frequency of coups in Syria and Iraq may be attributed to fiercer international struggle for both states as well as self-induced factors of instability (Picard 1988). While in some developing regions the army appeared to act outside class relations as a moderator, or in a way so as it intervene in civilian affairs only to put 'things in order' then 'return to the barracks' (Khuri 1982), in the Arab world, the armed forces have resumed authoritarian power. They set up one-party rule, intervened directly in civilian affairs and practically commandeered the economy. In Egypt, Syria and Iraq, the military rose as the principal political actor and instrument of government, but its public visibility varied in relation to regime stability.

Once the disappointing developmental performance of the national bourgeoisie of non-military postcolonial regimes came to the fore, the army took the leading role in 'radical' socioeconomic reform through a revolution from above. Vatikiotis (1972, pp. 12-13) argues that political insurrections, uprisings, and rebellions witnessed in the Middle East region, or what is known as the 'revolution from above' cannot be considered as revolutions and only represented 'middle class dissent' with Western intervention. However, if revolutions were to imply social and political change, then these coups were indeed revolutions. Army officers broadened and cemented their hegemony over the social base by espousing the aspirations of the broad section of the masses. Nasserism in Egypt and Ba'thism in Syria and Iraq undertook extensive land reforms, infrastructural projects and heavy industry development favouring the betterment of conditions for lower strata (Khuri 1982, pp. 17-21). The pedantic tweaking with revolutions as some sort of platonic ideal to which real world processes have to measure, is more theology than social science.

The wide consensus in the literature presents the military as a progressive social force and an instrument of social change and modernity (Vatikiotis 1972; Abdel-Malek 1971). It is also acknowledged for partially meeting the hopes and the aspirations of the middle class (Petras 1976, p. 440) and to have shrunk the influence of the traditional classes of the *ancien* regime, landlord aristocracy and Colonialist bred bourgeoisie. Unlike the old elite whose origins remained in part strongly feudal, the army acted as a social 'revolutionary' force in the Arab World enacting laws that revamped the foundation of civil rights, including, the rights of women. The army rank and file were composed, by and large, from well-to-do sections of the working class and small land owning peasants. The case may be that the vestige of petty property ownership desists from social progress, however, the élan of modernity and Soviet influenced social programs subjugated certain socially regressive traditions. At a later stage when the army capitulated to imperial diktat, it was not the resurrection of some 'inert' cultural value that swayed its position on the formation of class alliance in compradorial ways; it was pliancy in relation to a state of military defeat. Although the alleged inertness of

[9] http://membres.multimania.fr/yannickperez/site/Keynes%201937.PDF.

reactionary traits pertinent to Arabs occupies a huge space in the Western imaginative, this approach cannot conceptually be admitted into social science because the abstraction of the Arab as a socially regressive type is ahistorical.

To assess the shifting class allegiance of the military, one needs to examine the key alliances that the armed forces undertook at different historical periods. The firm articulation of the military and the newly expanding intermediate strata in the socialist radicalisation phase of the sixties strengthened the hegemony of the armed forces over the social base (Leca 1988). As of the early eighties, weakened national security implied weaknesses at all security levels including working class security or security from want. The military internalised the conditions of defeat, and managed the transition to neoliberalism by aligning itself with the newly revitalised business community. By the early eighties, the rent petro-dollars began to seep into the semi-autarkic economies of Arab socialism and signs of conspicuous consumption emerged. The state bourgeois class became noticeably capitalistic and parted with its populist working class rhetoric. It started pursuing economic liberalisation so as to reverse the process of socialisation. At first sight, it appears as if this new predisposition is driven by the dynamic of inequity under any capitalistic mode of development. However, upon a closer look, these changes were decisively determined by the onslaught of global militarised capital onto this crucial oil region. Egypt and Syria were left with debilitating costs of wars and Iraq had just re-entered another war. Evidently, these routs tallied with the potential for expanding accumulation based on dislocation for certain sections of the state bourgeois class. However, the determining moment for the transition had little to do with the tendency of national capital to grow under more unequal conditions and, a lot to do with the implicitly enshrined conditions of surrender meant to pauperise and dis-empower working people. The mode of integration of the Arab world through the channel of oil requires the articulation of the Arab world with the global economy via a mode of destructiveness and disengagement of national assets that would not afford the population with sovereignty over national resources[10].

Although to a lesser degree, in Iraq the military absorbed the conditions for surrender pursuant to successive military defeats and projected them in piecemeal neoliberal policies that structurally swung the economy towards a privatised mode of accumulation. Apart from being in a permanent state of war, Iraq's form of nationalistic capitalism, the intransigence of its leadership and its vast reserves of oil, implied that no matter how malleable its leadership became, its terms of surrender could only come about by direct re-colonisation. At one point prior to the second Gulf war in 2003, the Iraqi leadership practically passed over its oil field ownership to French and Russian oil companies; however, to no avail. The subjugation of Iraq, a country too strategic for imperialism to be left standing, its dissolution into several antagonistic social entities and continuous internal strife, represented a model that stifles adequate popular sovereignty. In retrospect, the collapse of Iraq as a state appears to have been necessary to tilt the balance of power securely in favour of US imperialism. It is fallacious to comprehend the US's goal in terms of a desire to set up a stable social formation in Iraq. The continued state of violent flux militarises the Gulf and hollows out the social formation so that it is no longer in a position to reproduce itself on better terms. That latter objective is development as a process of bettering living standards and the rights of people to fulfill their potential by the exercise of politics through the state. Ultimately, the destruction of the Iraqi social formation as a viable entity and the rise of Israel as a regional superpower intrinsically weaken the capacity of working people in the Arab world in confronting American imperialism.

At different stages in their development, the Arab military re-allied itself with global financial capital and the merchant class, and slowly de-industrialised and de-socialised their respective social formations. By the time of the uprising, one out of three children in Egypt was malnourished and nearly thirty percent of Syrian

[10] Articulation by the barrel of the gun, lecture by Soula Avramidis, Historical Materialism Conference 2005, http://mercury.soas.ac.uk/hm/pdf/2006confpapers/papers/Avramidis.pdf

children suffered from stunting[11]. In Iraq, however, the pace of pulling down the social structure was determined by military bombardment. To date, more than one million orphaned children are roaming the streets of Baghdad[12]. It is not bombastic to state that conditions in the Arab world are appalling; it is an understatement of the impact of bombardments. It was during this neoliberal period, however, that the military's intelligence agency, the *mukhabarat* – became the most powerful institution and main instrument of control. The suppression of organized labour and other civil society organisations assumed new heights (Pfeifer 1979). Praetorian regimes that before had capped affluent consumption and funnelled resources into public investment, came to mimic western consumption patterns and, ultimately, to rely more heavily on coercion. The post-colonial crisis of underdevelopment, which was bolstered by a crisis of inequitable distribution, and which was partly redressed by Arab socialism, re-emerged under the rule of the newly established military-global financial elite-merchant class alliance. In Egypt, one year prior to the uprising, a major Arab newsaper[13] headed with an article entitled: Egypt is back to King Farooq's days in terms of inequality when two percent of the population owned 98 percent of the wealth. Income inequality rose at an increasing rate. In this new order, the army concentrated more power its hands and became strictly the subject of development. The army's rule in Egypt in the post uprising days remains in evidence. The army's top ranks have intertwined themselves with vested interests in the commerce of their respective economies. They did so in relation to an import-led mode of development. Upon retirement, the brigadier was destined to enjoy a state patent over a certain imported product. The public sector was fated to become privately owned.

6. Arab socialism

During the socialisation phase, the ruling class in Arab socialist states was specified as an alliance between a dominant military and an intermediate stratum. However, a formalised characteristic of a 'middle class', borrowed from a Western context (in how it is used as a separate category of middle class from the working class in the US) and implied upon Middle Eastern formations was such that:

> "the propertied middle class could not compare in terms of capital, skill, and organisation with the resources and power which governments muster for rapidly overcoming economic backwardness and keeping pace with social change in the Middle East... only the new, salaried middle class, clustered around a core of civilian and military politicians and administrators, seem capable of leading the quest for status, power, and prosperity by taking control of the state apparatus" Halpern (1962, p. 279).

Although well parameterised, this distinction between strata within the working class does not read into a social class in relation to other classes and, more specifically, in relation to imperialism. The degree of cohesion of a working class depends on its position vis-a-vis the ideological power of capital. It is not the reasoned ideological discourse of capital that conquered the Arab world; winning hearts and minds was never the purpose of US aggression anywhere in the region. It is militaristic imperialism in an oil/Arab context that exemplifies a formidable force. The military bases strewn around strategic resources are the embodiment of international financial elites in actual physical form. None of the Arab class structures can be said to exist outside this militaristic relationship to imperialism. In the sixties, with colonialism retreating, it may not have been possible to assess imperialist clout in respect of national social classes. However, in as much as these

[11] http://mrzine.monthlyreview.org/2009/irin081109.html

[12] http://www.aljazeera.com/video/middleeast/2011/05/201151041017174884.html

[13] http://www.alquds.co.uk/ (in Arabic).

modern states were created by imperialism, so were their social classes. As soon as the Arab socialist project succumbed under a combination of imperialist assault and internal fracturing resulting from capitalist accumulation, the ties of national social classes with imperialism were re-established.

It may as well have been the frustration of the new educated 'middle class' that emerged after the colonial period that escalated the sentiments of anti-Westernism, but the real reasons for anti-western stance are far complex and multi-layered. They cannot be reduced to the undulating feelings of a sub-stratum. Anti-imperialist positions gained momentum in response to the practice of imperialism and the abject conditions that were created by the colonialist forces. The rising tide socialist ideology and national liberation movements in this epoch earmarked anti-imperialist struggle. The Arab bourgeois class – composed of merchants and large retailers – was deeply integrated into the capitalist system during pre and post-colonial period. It acted as an appendage to dominant imperialism (Amin 1978). It served foreign interests by colluding with the West and stifling resources (al-Hamsh 2004, p. 40; Springborg 1993, p. 3). In the more general case, Petras (1976) argues that the global expansion of imperial capital into the developing societies has incorporated the national bourgeoisie into its international network through trade, joint ventures, patents, loans, and credit. These linkages strengthened the dependency of host economies on foreign capital in financial and technical support.

In the immediate postcolonial phase, the Arab comprador class failed to garner the legitimacy and sufficient financial and real resources to build productive capacity. The inherited inequality laid the groundwork for discharging the old land-owning and comprador classes once the tide of socialist ideology took root in the Arab world. With the Soviet Union exemplifying the social model of state-led development *par excellence*, the alliance of military and intermediate strata emulated soviet style socialism. While the expanse and form of state intervention in political, economic, and social activities varied from one country to another, the trend of the post independence years was for growing state ownership of the economy. The Arab Socialist states acted as engines of growth (Anderson 1987, p. 11). The state bourgeois class led the process by focusing on heavy industry and infrastructure. In the mid-sixties, The Syrian Ba'ath carried out massive land reforms. In Iraq, the Ba'ath redistributed land, oil revenues in more equitable ways. In Egypt, state-owned enterprises accounted for about 60 per cent of value-added in manufacturing and Syria's accounted for about 55 per cent. The output of these newly established public enterprises recorded 13 per cent of GDP in Egypt and 11 per cent of GDP in Syria (Richards and Waterbury 1990, p. 192). Land reform tagged along and evidence of increasing agricultural productivity assisted in partly stemming urban migration. Until 2006, Syria was exporting cereals. With extensive social investment, standards of living rose significantly (Ayubi 1995; Anderson 1987; Richards and Waterbury 1990, pp. 187, 255 and 416).

The growth rates of the Arab socialist period as compared to the more recent neoliberal phase were higher. GDP and GDP per capita growth rates of the Middle East region registered 7.5 per cent and 4.8 per cent, respectively for (1964-1974) period. Notice that the rate of population growth is nearly steady. These rates then dropped to 2.9 per cent and 0.3 per cent for (1985-1995) period (World Bank 2009). In the uncertain post-independence environment, the state acted as a guarantor of long term investment in plant and equipment. Industrial and agricultural state owned banks lent to national projects at concessional long term rates. A so-called black list protected the national industry from foreign competition. A tightening of the capital account and a multiple interest and exchange rate policy galvanised national resources and provided exchange rate stability. Subsidies and price controls in essentials raised the standard of living for the rural population. Land reforms, which have since been rolled back, raised food production considerably.

Throughout this process of socialisation, the private sector remained active. Cooper (1983) for instance argues that there is a tendency among scholars to dismiss public-private oscillations that characterised 'state capitalists' regimes of the Arab region. He concludes that 'state capitalist' regimes failed in transforming the fundamental structure of their societies, either into a dynamic capitalist form or into a non-capitalist one, thereby they oscillated between various mixes of public and private sectors, ie. mixed or joint

sector structure in which the public sector was inefficient. Moreover, Cooper (1983, p. 458) presents Egypt as a case model, whereby 'state control did not obviate the role of the private ownership'. Cooper notes that the incentive to accumulate was always present and strong, both in the agrarian – where the incentive was to escape from state control – and the non-agrarian sectors of the economy – where efforts were taken to maximise one's ability to gain from state's economic development. The issue with this doctrinal line that dubs inefficient the public sector in a developing context is the looseness associated with branding anything state capitalist and then assuming that privately motivated accumulation is incongruous with the status quo. The pursuit of private interest pitted against public welfare is conditioned by the form of social organization which is capital and, therefore, it is common place for rent seeking around the state to counter public interest. When the so-called inefficient public sector was curbed under neoliberalism, the whole of the social structure went into a tailspin. The public sector served as a welfare cushion as job creation declined and in the absence of unemployment insurance. In more specific terms, the Arab world is an imperialistically aggressed region and the criterion for efficiency is class inspired. Under capitalism, in whichever shape or form it appears, private incentives are not going to disappear. The indelible fact remains that no individual incentive framework of the post-colonial private sector could have carried out the task of development given its structural incapacitation in terms of resources and finance. Moreover, the comprador-class' tolerance of imperial intervention in national affairs ideologically discredited it as a subject worthy of carrying out the task of development.

When these socialised countries implemented the neoliberal mantra of free markets and private-led investment, the investment rate in general fell from a high of over thirty percent in 1980, to a low of around 18 percent in 2010 (WDI, various years). That private investors cannot lead in capacity building is not only related to penury of finance or to geopolitical uncertainty but also to the broader imperialist objective of controlling oil by disempowering and pauperising the population in abundant-oil region. In the socialist age, the private sector was free riding upon the success of the state interventionist model – public investment crowded in private investment. In Syria, what remained of the private sector after massive socialisation in the mid-sixties grew in tandem with growth in public investment. Downstream involvement in commerce and industry in the socialist age was later to engulf the whole economy under the neoliberal-openness age. The private sector piggybacked the developmental process at all stages of the socialist project. This very private sector would later backstop a new capitalist class that would metamorphose into the newly emergent owning class of the neoliberal age. The private sector typified the formal model that would constitute a stepping stone for the new capitalist class, which arose out of the state bourgeoisie to join the circles of global financial capital. That a more radical socialist transformation to the social structure was not in evidence, especially when the working folk were disengaged from political life, does not mean that the socialist state-led developmental experience did not forebear structurally and historically relevant social results. The welfare outcome of socialisation is not wholly irreversible, but to date it has proven difficult to reverse.

The case of Iraq is, however, different because higher oil revenues provided the socialising state with leverage to finance welfare with foreign exchange. Iraq is also different in the way its terms of surrender were imposed by military devastation. Whatever criticism of the regime that arose from a left perspective appears retrospectively to have been a gross mis-assessment of the situation. Al-Khafaji (1986), for instance, argues that the presence of the 'socialist' state in Iraq strengthened the private sector and the economy's transition towards a capitalist state. He describes the Iraqi genre of Arab socialism throughout the seventies and eighties as the state's incubation of Iraqi capitalism by which the state supported and monitored the development of Iraqi capitalism in a controlled manner. That may as well have been a case of Syria or Egypt, but the Iraqi state in which the old state bourgeois class presided is not the Iraqi state of today. If that had been incubation of the private sector in Iraq, then it hatched re-colonisation. Arab socialist regimes, which are in essence capitalistic, cannot be expected to make transitions to socialism without democracy for the working classes. Their social reforms were carried out by the international élan of post-independence restructuring. In Iraq, the conditions of

surrender were not laid down by the Washington consensus, they were imposed by real devastation and by Paul Bremer[14]. The latter 'freed' the economy and tore down national industry and agriculture. The majority of the left in Iraq purposefully underestimated imperialist intent and tempered their critique of imperialism in order to demonise Saddam, which also justified imperial expansion[15]. A nationalist social class is nationalistic by the working class alliances it keeps and the distance it holds from the imperialist centre. Undererating the significance of the destruction of Iraq as a state to imperialism or, worse yet, relegating the whole US assault on Iraq to a mistake based on misinformation, channels the scope of research into what is historically counterfactual. The historically relevant condition rests in the deepening crisis of capital that rides on sturdy trends which call for further imperialist expansion, oil control and resource grab. To condemn the preservation of the private sector in Iraq, is to assume that with more radical forms of socialisation, Iraq could have escaped the oil control agenda of imperialism. Deconstructing capital is not an analytical process but a real historical process in which class struggles, especially those stemming from the struggle against imperialism, politically pressure capital into making further reforms and concessions in favor of working people.

In all Arab socialist countries, the state-centred 'socialist' economy protected the niches occupied by the private sector. Private sector activities were basically concentrated on retail and construction rather than large scale manufacturing. Their inputs were state subsidised and the prices they passed on to consumers were capped to moderate profits. The tariff walls that were designed to protect public enterprises also protected private enterprises from foreign competition. The state extended supply and construction contracts to the private sector and private sector profits were amassed through subcontracting. The latter gave rise to patron-client activities between state officials and businessmen. However, Batatu (1986) notes that the leveraging power that the private sector enjoyed under the regime of Saddam Hussein was minimal. This particular point could be drawn for all the Arab socialist regimes by the degree to which they upheld an anti-imperialist position that required a fortification of social cohesion through redistribution. The post-independence imperialist assault on these formations also conditioned domestic economic and social policies. Until their capitulation, these socializing states were at risk of war or in a state of war and national defence required the engagement of all social forces.

Notwithstanding Iraq, which was totally devastated by imperialist aggression, the remnants of the private sector in Arab socialist states did indeed serve as the embryo for broader private sector-led development at a later stage. The very class in power under socialism moved into the private sector and formalised its control of state property in cadastral form during the neoliberal stage. The wealth of Mubarak and Assad families are irrefutable testimonies to this public to private wealth transfer. President Assad's cousin is said to own sixty per cent of the country's economy through a complex web of holding companies[16]. The same cannot be said of Saddam's regime, however, probably since it was not allowed to last long enough to experience neoliberal transformation. In the case of Egypt and later Syria, adhering to the diktat of imperialism proceeded in a way that may have pre-empted imperial aggression Iraqi-style upon their territory. In the case of Iraq, its demise as a state can be attributed to its ideological stance or refusal to surrender early on to the condition of capital; it was rather necessary for imperialism to structurally turn the balance of power in its favour by demolishing the whole of the social formation. Iraq as a weakened social form is crucial to the standing of US Empire[17]. Along the same logic, the current protracted process of destruction in Syria follows a similar rationale from the point of view of empire.

---

[14] http://www.alternet.org/story/19293/

[15] 'I did not want to be a collaborator' Isam al-Khafaji, a former member of the Iraqi reconstruction council, explains his decision to resign: http://www.guardian.co.uk/world/2003/jul/28/iraq.comment

[16] From Financial Times article, April 21st 2012, titled 'Assad cousin accused of favouring family'.

[17] Downloadable PDF available from: http://www.networkideas.org/featart/aug2011/Ali_Kadri.pdf

By the late nineties, in both Egypt and Syria, the process of class restructuring was completed. The state bourgeois class became a typical compradorial bourgeoisie. The alternating debate concerning equitable distribution, which existed between the state and the national bourgeoisie in the sixties, mirrored the crisis of inequity inherited from colonialism. In the highly unequal conditions of post colonialism, the state had to force through equalisation measures such as land reform. However, the very idea that was flaunted by the state bourgeoisie that the private sector constitutes an indispensible constituent of the national front in the struggle against imperialism may have concealed an ulterior motive. The state bourgeoisie later used this private sector to absorb and effect the conditions of capitulation to imperialism and launched itself as fully fledged capitalism. The defeat of Arab socialist states was manifold, but the further retreat in socialist ideology pursuant to Soviet collapse provided the elites with the ideological pretext for the *volte face*. It denudated elements in these regimes that had acquired wealth via state control and that subsequently needed to actualise this wealth in titled and private form. Without the condition of Soviet collapse, which resembles a force majeure, the ease with which these social transformations and class restructurings were carried out by the autocratic regimes may not have been unproblematic, as they evidently appear so in hindsight.

For structural reasons, the private sector, in the initial stages of post-colonial reconstruction, could not have carried out egalitarian redistribution required to align the national forces into a defensive anti-imperialist position. The private sector in the neoliberal age also failed miserably in the task of development as witnessed by the appalling social conditions prior to the Arab uprisings. Pursuant to independence, the private sector was submerged in crisis, its reputation tarnished by the stigma of a shifty national allegiance, and its structural shortcoming, which is its lack of capacity to tap into the substantial resources needed for development, was in evidence. During the Arab socialist phase, the determining moment in the course of development lied not only in superficial squabbles between party bosses and merchants, but more so in outside imperial pressures, which were kept up by military force to drain the capacities of newly liberated and noncompliant states – especially so, concerning those around huge oil reserves. A state of constant war or the threat of war, whether real or perceived existed, which disallows significant privately led investment in plant and equipment to gestate over the long term. Unless the state intervened, it was doubtful that any relevant net additions to high capital output ratio capital could materialise. The lower ratio of overall investment to output and the ephemeral nature of investment during the neoliberal age provide ample proof of this point (as mentioned earlier investment rates declined from 30 percent in 1980 to around 18 percent in 2010).

It makes for an intractable task to dissociate the interface between policies and development outcomes by attributing them either to internal national class formations or the external meddling of the imperialist camp in national affairs. The external involvement in national affairs forms part of the class structure that constitutes the state. To question whether the Arab developmental disaster is the fault of national elements or that of imperialism sinks social science to the whim of empire. This issue is not an analytical condition per se; it is a process of history. Whatever class alliance existed in the socialist Arab state, it was developmental or anti-developmental by the space it kept from the imperialist centre. Depending on the centripetal strength of the working class, national class formations shape the condition that combines the struggle against imperialism and development outcomes in real and ideological forms.

In one glaring example of faulting the national forces alone for their mal-development, Jomo (2005) analytically proposes in no uncertain terms that it is the national bourgeoisie that is responsible for poor developmental outcomes. In a work titled the origins of development, structuralism is faulted for stating that subordination to the world market seals the fate of nations, and that it is squarely the doing of the national class in charge of development (Jomo, 2005). The fate of nations in an increasingly interconnected world is contingent upon an anti-imperialist international class alliances supported by socialist ideological zeal. The national bourgeoisie is determined in relation to its mode of integration with global capital and in relation to appropriation from the national economy. Unfettered expansion denominated in dollars will inevitably convert

the national class into a compradorial one. The upshot of this becomes: is a comprador class in any way national? The very being of the nation state mediates the category of capital and not vice versa. A social class cannot be defined in terms of the passport it holds. Historically, a social class supersedes the modern form of the nation state. A bourgeois organises the pursuit of profit through the state and the state is only means to an end. Development in relation its agent, which is a social class that is not defined in terms of ethnicity or national colour but in terms of its relation to the appropriation of surplus from self-expanding value, is the subject of social science. Accordingly, to expect that a national bourgeois class will oppose imperialism on national grounds or because it is committed to its own national working class is absurd and runs counter to its very raison d'être.

From the point of view of a third-worldist under American bombardment, structure appears immutable. The kernel of the issue, however, is neither structuralist nor atomized-individualistic. Change is class determined and the social class with hundreds of military bases strewn across the Middle East forms the determining constituent of Arab states. This structure has so far escaped deconstruction and, historically, Western anti-war protest movements have been unsuccessful. They nationalised, as opposed to internationalised, the struggle and sought an end to war as if it was a form of charity by the rich North to the poor South. The creation of social wealth by more productive means in the West became possible in part by destroying the productive of means of the Arab world. Wealth creation was given a Western nationality proper. This Eurocentric approach represents a civilisational distortion of value, one wherein Western lives are dearer than those of the rest by the criteria of a nationalism that is invariably laced with racism as is the case in any form of nationalism. But one presumes that the impact of Hollywood culture on social science reduces the social agency of class to colour, sect, and ethnicity or, worse yet, good and bad guys in order to exonerate imperialism.

The state bourgeois class, which paraded itself as an ardent foe of imperialism under Arab socialism, was later to become more alienated from its own social base under Infitah (openness or the Arabic equivalent of neoliberalism) and a close ally of imperialism. Infitah was a model of value usurpation and an opening up of national markets to geopolitical rents, which destroyed the protective dual-exchange rate arrangement and enacted the Saudi model of earning without effort as the leading social model. US/Saudi administered geopolitical rents did not come unaccompanied with ideology, they brought along the retrogression of Wahabism. Weakened by lack of endurance to stave off imperialist pressure and drawn by a cross border class alliance, these transformed state bourgeois elites were socially predisposed to implement the terms of surrender. The blame game, in the sense that it is the fault of the third world or the fault of the centre that caused developmental disasters, is a surreptitious argument that misses the point of global accumulation is an organically interconnected social process, which is monetised subject to the rapport de force in international relations.

7. Closing comment

From a culturalist perspective limited industrialisation and development in the Third World has been ascribed to the lack of the spirit of entrepreneurship or the 'problematic of the missing middle class' (Turner 1984, p. 44). However, structuralists such as Baran (1973) and Poulantzas (1973), think little of the problematic of the missing middle class. They argue that although the capitalist class plays a dominant role in capital accumulation, it nonetheless is not to be found in a vacuum. To them, the totality of the mode of production, in which the state is central, constitutes the determining moment of social and economic transformation. It is difficult to identify social classes without prior identification of the mode of production (Poulantzas 1973). The case for the Arab world is especially acute since the capitalist system incorporated the Arab region in terms of production, exchange and cultural relations, via colonisation (Amin 1978). For Baran, the entrepreneurial spirit

is said to be bred within the capitalist structure, where in the 'absence of industrial capitalism there are no industrial capitalists' (Baran 1973, p. 385). 'The existence of "entrepreneurial role" is sociologically determined, not generated by the creative activities of individuals. Entrepreneurship is an outcome of capitalist structures. The contemporary class formation in the Arab World is the product of an oil-determined articulation, which is carried out by military force (Avramidis 2005). It is this articulation that forbade not only the evolution of a vibrant entrepreneurial class, but also precipitated the planned abortion of development in the Arab world. This is not a condition which is specific to the Arab world, it occurs at the point where the metabolic rate of the reproduction of capital imposes social dislocation to a degree which is satisfactory to value grab.

Real developmental achievement leaves its imprint in real time, which is one continuous whole and not the formalised short-term/long-term travesty it is often made up to be. The positive developmental impact of the socialist Arab state was real and lasting. Lumping Arab socialist and their later mutations, the neoliberal regimes, together on the grounds that they were both repressive dictatorships is somewhat academically complacent. These Arab socialists may have been capitalists parading under the logo of socialism, but the egalitarian redistribution undertaken in the early stages of independence left an everlasting impact on welfare. To have drastically reduced the illiteracy rate in Egypt at first, is an accomplishment that neoliberalism found difficult to roll back. Different starting points of development lead to different paths of development – China's lead over India is a case in point. Thus, a quasi-socialism emerging in a less developed Arab formation, bombarded by belligerent imperialism, is in some respects a victory of labour over capital and a collusion of an ideologically inclined social theory and practice. Social theory and practice in the Arab world, to be sure, meet and depart from each other more so by the outcome of international class struggle than nationally confined class struggle.

The reading of class transformation in this essay is based on the premise that it is submission to imperialism that revamped class structure in order to absorb the terms of surrender. After several routs, Arab socialist regimes went from being led by an alliance of the military and the intermediate strata to an alliance of the military with the more domineering international financial capital. There is another underlying hypothesis that underwrote the Arab path of development and that needed to be brought forth at the very start, but it is equally relevant to resurrect it now. Arab development is determined in relation to contradictions in global capital accumulation mediated by international relations manifest in oil-control wars upon the Arab region. Arab development, if it were to happen, imparts security to working people and, by implication, popular sovereignty and a sturdier Arab national sovereignty. Therefore, Arab development is unlikely to be wrought by peaceful means, since it will inevitably contribute to weakening imperialist hold over a geostrategic region whose control is central to empire. Imperialist wars aimed at the selective deconstruction of Arab social formations exemplify the mode of praxis of the foremost constituent of the subject of history, which is US Empire, in respect to the Arab world.

A telling drawback of Arab socialist regimes, however, is that the working population was not afforded with the requisite civil liberties to participate in the political process. This repression came handy when the turn to neoliberalism took place at a later stage. Critiques of this experience from the far left have a moral overtone that indicts the Arab socialist model on the grounds that it maintained the relationship of capital under a more collective ownership of the means of production. Ethical evaluation begins at the intersection of theory with practice and not a theory with theory[18]. Capitalist social relations are entrenched, and radical change would have required more than simple commitment to socialism in some corner of the third world, but a social ideological avalanche. On its own, the level of development in Arab states with traces of semi-feudal despotism would not, sui generis, attenuate the repression attendant upon the labour process. Nevertheless, anti-imperialist positioning represented a necessary step in the formation of global anti-imperialist fronts and, the

[18] From Barrows Dunham, Thinkers and Treasurers, *Monthly Review* 7 (8), December 1955.

consequent, retention and redeployment of real resources in the national economies that resulted in comparatively positive developmental outcomes. This post colonial stage in Third World development with its political landmark achievements in Bandung and the rise of the non-aligned movement was possible because of concerted international working class alliances. Dislodging the Arab development debacle would require more than just an Arab spring, it requires a World spring, else one will continue to bear witness to an elected Islamic brotherhood submitting World Bank wishes and a Syria engulfed in baleful fratricide.

References

Abdel-Malek, Anouar (1967), The Crisis in Nasser's Egypt, *New Left Review*, 45, (Sep-Oct), pp. 67-82.

Abdel-Malek, Anouar (1968), *Egypt: Military Society; the Army Regime, the Left, and Social Change under Nasser.* New York: Vintage Books.

Abdel-Malek, Anouar (1971), *Al-jaysh wa-al-harakah al-wataniyah* [*The Army and National Movement]*, translated from Arabic. Beirut: Dar-Ibn Khaldoun.

Amin, Samir (1976), *Unequal Development: An Essay on the Social Formations of Peripheral Capitalism.* Hassocks: The Harvester Press.

Amin, Samir (1978), The Arab Nation: Some Conclusions and Problems, *MERIP Reports*, 68 (June) pp. 3-14.

Amin, Samir (1992), Contribution to a Debate: The World Capitalist System and Previous Systems. In: Fawzy, Mansour (ed), *The Arab World: Nation State and Democracy.* London: Zed Books, pp. 1-29.

Anderson, Lisa (1987), The State in the Middle East and North Africa, *Comparative Politics*, 20 (1), pp. 1-18.

Avramidis, Soula (2006), *Articulation by the 'Barrel of a Gun': Development Under the Threat of War in the Near East.* London: Historical Materialism Conference.

Ayubi, Nazih N (1995), *Over-Stating the Arab State: Politics and Society in the Middle East.* London: Tauris.

Baran, Paul A (1973), *The Political Economy of Growth*. Harmondsworth, Middlesex: Penguin Books.

Batatu, Hanna (1986), State and Capitalism in Iraq: A Comment, *Middle East Report*, 142 (Sep-Oct), pp. 10-12.

Berger, Morroe (1958), The Middle Class in the Arab World. In Walter Z Laqueur (ed.), *The Middle East in Transition: Studies in Contemporary History.* London: Routledge & Kegan Paul, pp. 61-71.

Binns, Peter (1986), State Capitalism, Marxism and the Modern World, *Education for Socialists,* 1 (March)

Binns, Peter and Duncan Hallas (1976), The Soviet Union: State Capitalist or Socialist? *Marxist Internet Archive,* (Part 2). Available from http://www.marxists.org/archive/hallas/works/1976/09/su2.htm#top

Braverman, Harry (1959), The Nasser Revolution. *American Socialist Internet Archive,* (January). Available from http://www.marx.org/history/etol/newspape/amersocialist/amersoc_5901.htm

Buick, Adam and John Crump (1986), *State Capitalism: The Wages Systems under New Management.* London: Macmillan.

Burnham, James (1945), *The Managerial Revolution.* London: Penguin.

Bush, Ray (2004), *Civil Society and the Uncivil State: Land Tenure Reform in Egypt and the Crisis of Rural Livelihoods.* Geneva: United Nations Research Institute for Social Development.

Cliff, Tony (1974/1955), S*tate Capitalism in Russia*. London: Pluto Press.

Cooper, Marc N (1983), 'State Capitalism, Class Structure, and Social Transformation in the Third World: The Case of Egypt', *International Journal of Middle East Studies*, 15 (4), pp. 451-469.

Al-Duri, Abd al-Aziz (1969), *Muqaddimah fi al-tarikh al-iqtisadi al-Arabi*. Beirut: Dar al-Ṭaliiah.

Gramsci, Antonio (1978), *Selections from Political Writings (1921-1926; with additional texts by other Italian Communist leaders, (ed),* Quintin Hoare. London: Lawrence and Wishart.

Halpern, Manfred (1963), *The Politics of Social Change in the Middle East and North Africa*. Princeton: Princeton University Press.

al-Hamsh, Mounir (2004), *al-Fiqr al-Iqtissadi fi al-Khitab al-Siyassi al-Souri fil al-Karn al-Ishreen* [*Economic Thinking of the Syrian Political Speech in the 20th Century*]. Beirut: Bisan.

Hussein, Mahmoud (1971), *La Lutte de Classes en Egypt*, *1945-1970*. Paris: Maspero.

Issawi, Charles (1955), The Entrepreneur Class. In: Nettleton Fisher, Sydney (ed.), *Social Forces in the Middle East.* New York: Cornell University Press, pp. 116-136.

Jomo K S and Erik S Reinert (2005), *The Origin of Development Economics: How Schools of Economic Thought have Addressed Development.* New Delhi: Tulika Books.

Kalecki, Michal (1972), The Structure of Investment. In: Kalecki Michal (ed.), *Selected Essays on the Economic Growth of the Socialist and the Mixed Economy*. Cambridge: University Press, pp. 102-112.

Kalecki, Michal (1976), *Essays on Developing Economies*. Hassocks: Harvester Press.

al-Khafaji, Isam (1986), State Incubation of Iraqi Capitalism, *Middle East Report*, 142 (Sep-Oct 1986), pp. 4-9, 12.

Khuri, Fuad (1982), The Study of Civil-Military Relations in Modernising Societies in the Middle East: A Critical Assessment. In: Kolkowicz, R and A Korbonski (eds), *Soldiers, Peasants and Bureaucrats: Civil-Military Relations in Communist and Modernising Societies*. London: Allen and Unwin, pp. 9-28.

Leca, Jean (1988), Social Structure and Political Stabilty: Comparative Evidence from the Algerian, Syrian, and Iraqi Cases. In: Dawishi, A and I W Zartman (eds), *Beyond Coercion: The Durability of the Arab State*. New York: Croom Helm, pp. 164-196.

Owen, Roger (2004), *State, Power and Politics in the Making of the Modern Middle East*, 3rd Edition. Oxford and York: Routledge.

Petras, James F (1976), State Capitalism and the Third World, *Journal of Contemporary Asia*, 6 (4), pp. 432-443.

Pfeifer, Karen (1979), 'Three Worlds or Three World Worldviews?: State Capitalism and Development', *MERIP Reports*, 78, pp. 3-11, 26.

Picard, Elizabeth (1988), Arab Military in Politics: From Revolutionary Plot to Authoritarian State. In Dawishi, A and I W Zartman (eds), *Beyond Coercion: The Durability of the Arab State*. New York: Croom Helm, pp. 116-147.

Poulantzas, Nicos Ar (1973), *Political Power and Social Classes* (translated from French). London: Sheed and Ward.

Richards, Alan and John Waterbury (1990), *A Political Economy of the Middle East: State, Class, and Economic Development*, 1st edition. Boulder, Colorado: Westview Press.

Springborg, Robert (1993), The Arab Bourgeoisie: A Revisionist Interpretation, *Arab Studies Quarterly,* 15 (January), pp. 13-40.

Todaro, Michael P (1979), *Economic Development*. Englewood Cliffs, New Jersey: Prentice-Hall.

Turner, Bryan (1984), *Capitalism and Class in the Middle East: Theories of Social Change and Economic Development*. London: Heinemann Educational Books.

Vatikiotis, P J (ed) (1972), *Revolutions in the Middle East and Other Case Studies*. London: Allen and Unwin.

World Bank (1995), *Claiming the Future: Choosing Prosperity in the Middle East and North Africa*. Washington, DC: The World Bank.

World Bank (2009), *World Development Indicators*. Washington: The World Bank.

Grass Roots War on Poverty

Alice H Amsden
Barton L Weller Professor of Political Economics at Massachusetts Institute of Technology, Cambridge, Massachusetts[1]

Abstract

Sub-Saharan Africa's failure to slay the dragon of poverty is due to a logical flaw in its policies: the remedies to reduce poverty don't address the causes. Poverty is caused by unemployment, owing to a scarcity of jobs that pay above bare subsistence, but grass-roots poverty alleviation measures are exclusively designed to make job-seekers more capable although no jobs are available. The 'appropriate' technologies of the grass roots movement that dominates anti-poverty policies are oriented towards consumption, ignoring production jobs. Poverty persists from low productivity in agriculture or outright landlessness. Irrigation and rural electrification are required to facilitate economic diversification into non-agricultural work. Yet irrigation and electrification require central political coordination and application of modern science and technology. Centralized decision-making is low on the agenda of the anti-poverty movement, with deep roots at the local level. To create employment requires capital investments to expand entrepreneurial opportunities and increase productive jobs. The most successful countries to grapple with poverty have 'scaled up,' not down; Big, not Small, is Beautiful. The statistical evidence for a large number of developing countries strongly supports the hypothesis of a trickle down effect, not a bottom up effect as the best way to beat poverty.

Key words: poverty, grass roots, Africa

1. Introduction

Sub-Saharan Africa's failure to slay the dragon of poverty is due to a logical flaw in its policies: the remedies to reduce poverty don't address the causes. Poverty is caused by unemployment, owing to a scarcity of jobs that pay above bare subsistence, but grass-roots poverty alleviation measures are exclusively designed to make job-seekers more capable (healthy, educated, mobile), although no jobs are available. Employment might rise if wages fell, but wages are already at the subsistence level and cannot fall much further, if at all. The 'appropriate' technologies of the grass roots movement that dominates anti-poverty policies are oriented towards consumption, ignoring production jobs altogether if only out of ignorance of what industries could be created in the global South. Poverty persists from low productivity in agriculture or outright landlessness, requiring at minimum irrigation and rural electrification to facilitate economic diversification into non-agricultural work, and to lower production costs. Yet both projects, irrigation and electrification, require central political coordination and application of modern science and technology. Centralized decision-making is low on the agenda of the anti-poverty movement, with deep roots at the local level.

To create employment requires action on the demand side, in the form of capital investments to expand entrepreneurial opportunities and increase productive jobs. Yet the anti-poverty programs that have swept through Africa take a 'Capabilities Approach'. They focus exclusively on the individual's supply side, as though a greater supply of better-qualified job-seekers (in terms of education, consumption, mobility housing and human rights) will automatically stimulate the demand to employ it, as in the Eighteenth-century 'law' of the conservative French economist Jean Baptiste Say. The expectation that more job seekers who are qualified will

[1]Alice Amsden was the Barton L Weller Professor of Political Economics at MIT, Cambridge, Massachusetts. She passed away on March 15th 2012, while this paper was under review. No conflicts of interest are known to the editors.

generate the demand to employ them, based on Say's Law, is irrational in the presence of soaring unemployment and underemployment rates among people that already have schooling and access to modern infrastructure but cannot find remunerative work. Volunteers from the global North – people of faith, students, scientists and engineers – strive to invent appropriate technologies for Africans to consume, ie. use, like the corn shucker, not to produce in large volumes for export or for a mass domestic market; this is in spite of the fact that Africa's gravest long-term problem is unemployment, not costly consumption. Compared to investment rates, consumption rates are already high, so why the irrational focus on consumption?

In fact, the most successful countries to grapple with poverty – China is key given its huge population – have 'scaled up,' not down; Big, not Small, is Beautiful[2]. The statistical evidence for a large number of developing countries strongly supports the hypothesis of a trickle down effect, not a bottom up effect as the best way to beat poverty, although the same GNP growth rate may serve the poor more in some economies than in others. Far Eastern manufacturers and Middle Eastern energy suppliers drastically decreased poverty with top-down approaches that created new industries and professionally managed firms, which then moved beyond the micro-enterprise and cultivated small- and medium-size firms as subcontractors. In the Twentieth Anniversary issue of the *Human Development Report*, for 2010, the greatest progress over a 40-year period in health and education was found in Asia. Algeria, Morocco, Oman, Saudi Arabia and Tunisia – blessed either with oil or with remittances from workers linked to the oil business – ranked in the top ten in terms of progress. "Life expectancy in the Arab countries generally increased from 51 years in 1970 to almost 70 today, the greatest gain of any region in the world, while infant mortality rates plummeted from 98 deaths per 1,000 live births in 1970 to 38 in 2008, below the current world average of 44 per 1,000. School enrolment in the Arab states nearly doubled over the past four decades, rising from 34 percent in 1970 to 64 percent today. The average years of education for the current adult population of the Arab countries is now estimated at 5.7 years; less than the world average of 7.4 years, but significantly above the levels of sub-Saharan Africa and South Asia, with 4.5 and 4.6 years, respectively" (UNDP 2010, p. 3362).

2. Internal inconsistencies

Africa's grass roots poverty movement operates with two internal inconsistencies that undermine its effectiveness in peasant export economies. In such an economy, there is a delicate balance between land and labor, and units of production may be clustered in villages, but these are spread out geographically. This formation has suffered from rapidly rising populations and a weak capacity to absorb foreign knowledge in the form of technological change. The way private and public incentives for innovation are structured, these incentives, in the form of prizes and awards, are irrationally biased in favor of inventors who focus on small-scale 'appropriate technologies' for consumption using cheap local labor and materials. If incentives were neutral, and did not serve to keep such inventors afloat, more resources would probably be invested in technologies for mass-produced public goods. Rural electrification is emblematic of this set of goods, and is also democratic; it tends to be high on the list of local consumer preferences. Graffiti in a poor Pakistani village read: "Give us electricity and we'll give you a vote" (Tavernise 2010). The World Bank's Independent Evaluation Group observed "The people who live in rural areas greatly appreciate the benefits [of electrification] and are

[2] Scaling-up has become a preoccupation of grass roots movements. An assessment on the scaling-up of community efforts notes: "Limited research has been conducted into the nature of scaling-up efforts undertaken by community groups and initiatives and there is no single definition that clarifies what scaling-up means to, or for, communities. At a broad level, scaling-up refers to efforts 'to bring more quality benefits to more people over a wider geographical area more quickly, more equitably, and more lastingly' (there is no mention of costs). At a finer level of resolution, however, four distinct types of scaling-up have been identified – quantitative, functional, organizational and political." Quantitative refers to the expansion of an organization or program by enlarging itself or increasing its membership base (economies of scale). Functional refers to expansion in the types of an organization's activities (economies of scope). Organizational refers to the expansion of an organization's efficiency and effectiveness (economies of restructuring?). Political refers to a grass roots movement forging relations with a national state (Hooper and Jafry 2009).

willing to pay for them at levels more than sufficient to cover the costs"[3]. But the evaluation of these and other benefits is "sparse" (World Bank 2008)[4]. Big projects, with broad environmental impact and little participatory democracy tend to be denigrated by the grass roots as dinosaurs.

Incentives are also irrational in economies with high unemployment because they do not reward inventions 'to be made' but rather 'to be bought'. Sub-Saharan Africa is regarded by inventors as a market for a new product to raise welfare by improving consumption, not as a manufacturing base for a new product that increases jobs. This is because prizes reward inventors' knowledge. Since most inventors operating in Sub-Saharan Africa are foreign-born, their knowledge of local consumption problems is greater than their knowledge of local production possibilities, satisfied through import-substitution or export promotion in small- or medium-size firms. As a general rule, most innovations fail, but most appropriate technologies that fail carry no warranty or after-sales service. If bought with microcredit, they create debt[5]. Instead of doing good, or causing no harm, they "ambush the community's spare change" (the term is Thorstein Veblen's), especially if equally well-designed and cheaper products are available from foreign markets, usually emerging markets, as in the case of wheelchairs from war-torn Vietnam.

In sum, inside the market economy incentives are irrational because they encourage innovation but not necessarily 'economic development' – defined as a trend towards full employment, rising productivity and stability in the balance of payments. Like the Enlightenment movement of the eighteenth century, the grass roots anti-poverty movement idealizes the entrepreneur, but he or she is fettered by exposure to only 'appropriate' technology and consumer wants (or supposed wants, since some technologies – the corn shucker, for example – save time rather than money, although time is not in scarce supply at the grass roots).

The grass roots movement has also failed to reduce poverty, let alone to trigger a growth momentum, because it embraces the defunct eighteenth century market law (of Jean Baptiste Say, discredited by Keynes), to the effect that supply creates its own demand (Amsden 2010). Here the Enlightenment movement and the anti-poverty bottom-up movement overlap. Grassroots activists fall into the lap of Say's Law in believing that if the poor can secure their individual human rights broadly construed – to include better healthcare, housing, education and training – then they can attract the demand necessary to earn a living wage, either by working for others in paid employment or employing themselves as small entrepreneurs. Human development supposedly emerges from the gentle market laws marveled at by Enlightenment thinkers like Adam Smith. Yet in the presence of historically unprecedented rates of educated unemployment and underemployment; characteristic of the de-colonized generation's poverty, it is a leap of faith to move from better-qualified job-seekers on the supply side to above-subsistence job-holders on the demand side. If costs could be reduced and profits increased by more qualified job seekers, there would be no unemployment. Investment in new industries has to be undertaken on the demand side to create more jobs and entrepreneurial opportunities, but the grass roots is generally hostile to government interference. Investments in education create more educated unemployment and underemployment, or brain drain from rural to urban areas or abroad.

Estimated rates of youth unemployment for Egypt (see Table 1) tell a story that appears to be applicable to other developing countries: urban unemployment is higher than rural unemployment, total unemployment is higher for the better-off than the poor (who are too indigent to sustain open unemployment),

[3] For centrally organized electricity projects that join the grassroots, see ASADI 2010.

[4] According to a World Bank study, "lighting alone brings benefits such as increased study time and improved study environment for school children, extended hours for small businesses, and greater security. But electrification brings more than light. Its second most common use is for television, which brings both entertainment and information" (World Bank 2008) Available from: lnweb90.worldbank.org/oed/oeddoclib.nsf/DocUNIDViewForJavaSearch/EDCCC33082FF8BEE852574EF006E5539/$file/rural_elec_full_eval.pdf

[5] A study of 1,800 households in Bangladesh found only 'very marginal improvements' for borrowers. For a review of microcredit, see Chowdhury (2009). Given a typical rate of return, a $250 one-year loan would raise a borrower's income by $2.50/year, or about $0.03/day. For someone living on $2/day, this is a 1.5 percent increase (Roodman and Queshi 2006).

unemployment rises as education rises, and on average it is as high as nearly 20 percent for all youths actively seeking work[6].

Table 1. Egypt, youth unemployment rate (15-24), by education and poverty, 2005 (%)

	Urban		Rural		All Egypt	
Education	**Better-off**	**Poor**	**Better-off**	**Poor**	**Better-off**	**Poor**
Illiterate	1.8	4.9	0.5	1.1	0.8	1.7
Read & write	1.9	10.1	0.7	1.6	1.2	3.6
Basic education	7.1	9.3	2.1	4.0	4.3	5.5
Secondary*	31.9	37.2	21.1	25.7	25.4	28.6
> Secondary**	36.8	48.8	32.2	34.2	35.1	39.3
University+***	45.3	53.0	42.1	37.9	44.2	43.4
All	**26.0**	**24.9**	**13.7**	**13.5**	**18.7**	**16.1**

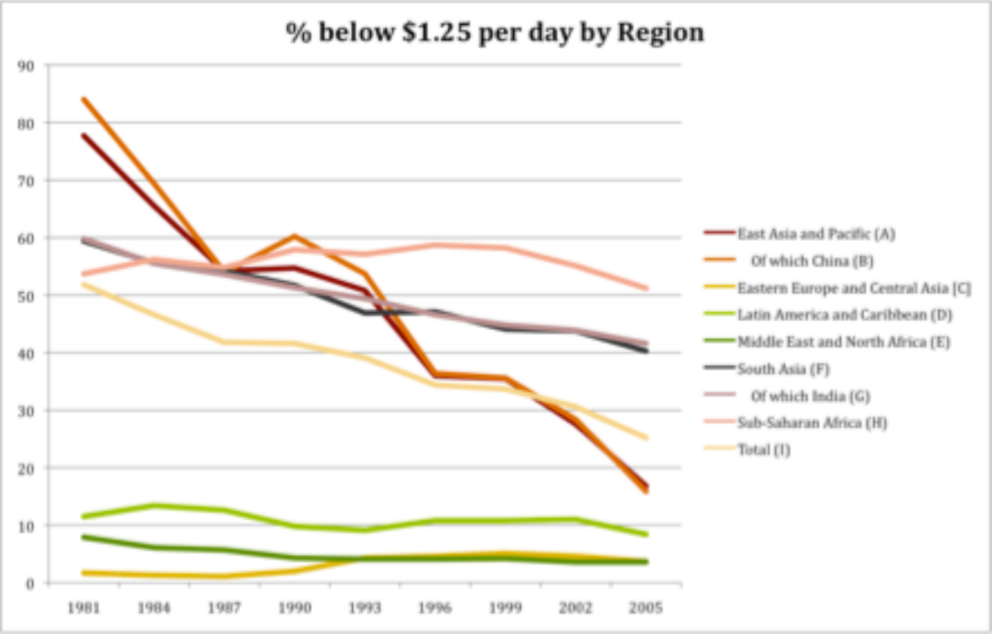
(*Secondary degree or equivalent, **Higher than secondary degree but below university degree ***University degree and higher)

Source: Adapted from World Bank and Ministry of Economic Development of Egypt 2007.

3. The persistence of poverty

The ineffectiveness of the grass roots anti-poverty movement, centered on technology and entrepreneurship, is suggested by country data on abject poverty, defined as the equivalent of an individual's consumption of $1.25 or less a day, or the size of its shortfall with that of an average income earner (see Figures 1 and 2).

Figure 1. Percentage below US$1.25 per day by region

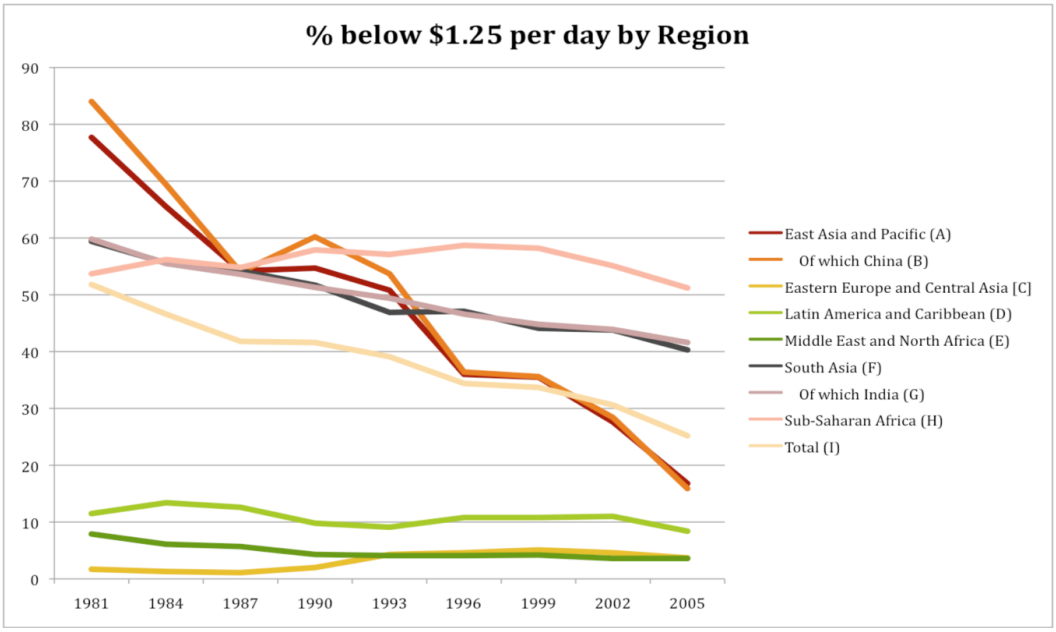


Source: All data are adapted and taken from Chen and Ravallion 2010

[6] Duflo (2001) has primary school data for Indonesia that indicate that primary school is positively correlated with (small) wage increases; no information is presented on secondary or tertiary education.

Despite bottom-up initiatives, despite social enterprises and heroic entrepreneurs, despite movie stars who publicize the human plight, the percentage of Africa's population consuming the equivalent of $1.25 a day or less remained roughly the same in 1981 and 2005, at around 55 percent. If the share of the population in abject poverty fell thereafter, this was largely due to large export commodity price increases (Beny and Cook 2009)[7]. There was no poverty reduction at the bottom-most end of the income ladder for twenty-five years, and no probing into the mere possibility that the grass roots approach was partly to blame.

Figure 2. Poverty gap index (x100) by region



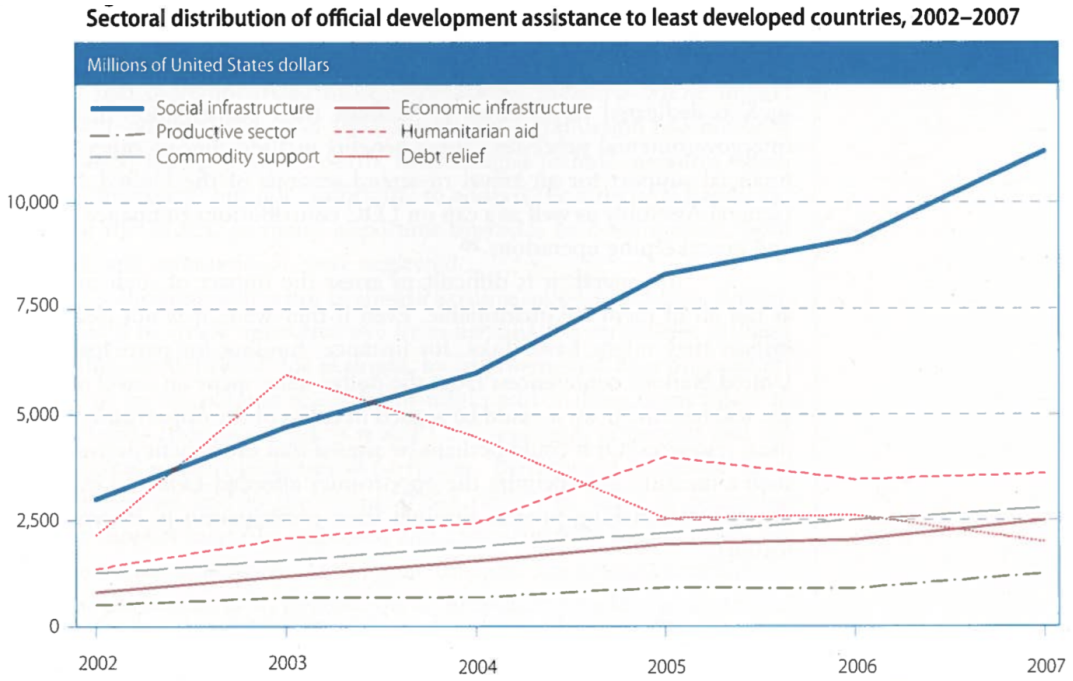
Source: All data are taken from Chen and Ravallion 2010

Note: Poverty gap index = mean distance below the poverty line as a proportion of the line where the mean is taken over the whole population, counting the non-poor as having zero poverty gaps.

Expanding international trade could, in theory, extricate small producers from poverty. Yet the situation in Africa in recent decades is particularly worrisome, where farmers have lost 25 per cent of their purchasing power due to relative declines in their export prices; average yearly farm incomes have fallen to below $200 per head (UNCTAD 2010). The percentage in world trade of the 49 'least developed countries,' a UN category established in 1971, has since then shown a downward drift from 1.6% to 0.47% in 1995. The percentage hovered around only 0.33% for the entire time period if oil exports are excluded (Cornia 2010). This is not surprising given that only the smallest proportion of aid to the least developed countries (2002 to 2007) was allocated to the 'productive sector', where new export products might have been created (see Figure 3).

[7] During the period 2002-2007, according to an UNCTAD report, the 'least developed countries' as a group experienced high gross domestic product (GDP) growth rates, surpassing 7%. However, about a quarter of the LDCs continued to experience very sluggish growth or economic regression. Moreover, even in the more successful countries, growth was associated with a pattern of insertion into the global economy based on high commodity export prices, low-skill manufactures and tourism, which meant that they were highly vulnerable to external shocks. Omitting oil-exporting countries, there was little improvement in domestic investment and savings, and very slow technological progress in the LDCs. Agricultural productivity growth lagged and there was widespread de-industrialization rather than a progressive structural transformation. Most significantly, the form of economic growth was not associated with broad-based improvements in human well-being, but rather very slow poverty reduction. In 2008 and 2009, there was a sharp though very heterogeneous slowdown in growth in the LDCs. The LDCs did not fare as badly as other developing countries, partly because commodity prices recovered in 2009 and partly because multilateral institutions provided increasing official flows (TWN 2010).

Figure 3: Distribution of official development assistance



Source: Based on the Organization for Economic Cooperation and Development *Development Cooperation Report 2009* (Paris, OECD, 2009).

In poor countries where the average farm size is less than two hectares (1 hectare = 2.5 acres), "boosting production in these units is hampered by imperfect factor markets, high input prices, poor infrastructure, restricted access to credit and inadequate research and development (R&D). These farms, moreover, are among those most vulnerable to climate change" (UNCTAD 2010). Africa's newest exports, such as fresh-cut flowers for European markets, are grown in Kenya and Ethiopia, on large-scale commercial farms[8]. In the absence of scale, which would require a land reform to consolidate rather than sub-divide holdings and possibly worsening rather than improving income distribution, risk of experimentation is high, technological change is slow, and expanding exports is difficult. When Sub-Saharan Africa was opened to trade with Europe in the late nineteenth century, Africans were quick to specialize in cash crops for export but not to introduce new production technologies (Myint 1964/1995, p. 312). New technologies were slow to diffuse partly because of minimal amounts of European immigration. Later, of all Third World regions, Africa participated the least in the Green Revolution (Myint 1964/1995). In its heyday in the mid-1970s, the Green Revolution's investments in high-yielding varieties catered to the big staple crops, the wheat fields of the Near East and Latin America and the rice paddies of Asia. Africa, with fragmented markets, small-holdings and multiple types of staple foods other than wheat and rice, was least transformed (see Table 2). Because appropriate technologies cater almost exclusively to local consumers, small-holder exports are unlikely to expand.

[8] The importance of scale for agricultural exports was recognized as early as the nineteenth century, in the Middle East, for example. "The great increase in sea-borne trade with Europe involved a fundamental restructuring of many parts of the Middle East economy. In some cases such a restructuring was necessary if the area developed to the production of cash crops was to expand in the first place. Lines of credit had to be established to provide producers with working capital, methods of transport improved and new systems of irrigation introduced where necessary. Later, the growth in trade itself began to produce important effects in many areas, leading to an intensification of monetary relations, an increase in agricultural specialization, and important changes in the relations between producers and those who controlled the cultivated land" (Owen, 1981).

Table 2. The Green Revolution: High-Yielding Varieties of Wheat and Rice, 1976/77
(Estimated Area and Proportion of Crop Area Planted to High-Yields, (1000s Hectares, %))

Region	Wheat ('000 Hectares)	Rice ('000 Hectares)	Total ('000 Hectares)
Asia (South & East)	19,672.3	24,199.9	43,872.2
Near East*	4,400.0	40.0	4,440.0
Africa**	225.0	115.0	340.0
Latin America	5,100.0	920.0	6,020.0
Total	**29,397.3**	**25,274.9**	**54,672.2**
Region	**Wheat (%)**	**Rice (%)**	**Total (%)**
Asia	72.4	30.4	41.1
Near East	17.0	3.6	16.5
Africa	22.5	2.7	6.5
Latin America	41.0	13.0	30.8
Total	**44.2**	**27.5**	**34.5**

(*West Asia and North Africa, **Sub-Saharan Africa)

Source: Dana G Dalrymple, Development and Spread of High-Yielding Varieties of Wheat and Rice in the Less Developed Nations, as cited in Rockefeller Foundation 1979

Starting from furthest behind, poverty has fallen the fastest in China and the rest of East Asia, and at an impressive rate in India, where bottom-up and top-down approaches co-exist (see Figure 2)[9]. As a share of the world's absolute poor in 1980, one out of every 10 persons lived in Sub-Saharan Africa, compared with one in three in 2000, the lowest incomes in 1980 were found in Asia. In Latin America and the Middle East, abject poverty by 1981 was already low, but in the next 25 years it still more than halved in the Middle East.

For generations, the problem of poverty has been regarded as "primarily an industrial one" (Atkinson 1887 in Tawney 1913). Yet like the benefits of electrification, the relationship between poverty alleviation and employment creation in 'industry' (modern manufacturing, services and agriculture) remains obscure. Despite championing the cause of poor people around the world and dramatizing the human condition, the United Nation's Millennium Development Goals make no mention of employment generation as a means to battle poverty; there is silence on using economic policies — employment, fiscal industrial, monetary and trade — to create more industries that could provide additional jobs above the subsistence level. A World Bank history that starts in 1944 and covers over half a century mentions employment only seven times in a total of 1234 pages (Kapur et al 1997). A popular book on poverty by Paul Collier (2007), *The Bottom Billion*, is highly praised by Larry Summers, George Soros and the Economist magazine but it fails in its nine-page, small-print index to make any reference to employment, unemployment, self-employment, jobs, or work. In India, to be self-employed in informal services like auto-rickshaws, transport portage, and ready-food stalls is to earn less than an agricultural worker, although the self-employed entrepreneur is idealized compared to the wage earner, who may be disparaged as being part of a 'labor aristocracy' that enjoys government support (Banerjee 2010).

[9] China's Minister of Commerce in 2010, Chen Deming, noted China's strategy towards "livelihood-related undertakings." First in order was "a more proactive employment policy.. .to create more jobs and expand employment in every possible way." Second mentioned was a social security system, then investment in social services, and finally an improved social safety net (UNIDO 2010).

Africa's sectoral distribution of employment as late as the 2000s was highly skewed, with roughly 65% in agriculture (mostly small-scale), 25% in services (mostly traditional) and only 10 percent in manufacturing (mostly the processing of raw materials such as sugar refining, beer brewing and flour milling). This minimalist modern-sector employment compares with a 38% average manufacturing share for East Asia (ILO 2008). In what follows, we argue that Sub-Saharan Africa is poor because few paid jobs have been created in modern industry; financial incentives tend to reward appropriate technologies *for use*, to make the lives of the poor easier, rather than *for production and export*, to make the incomes of the poor sustainable.

4. Population

Because rural areas tend to have higher population growth rates than urban areas, a lack of urban jobs perpetuates underdevelopment by indirectly sustaining a high population growth rate (Easterlin 2004). The 30 countries out of 223 with the highest estimated population growth rates in 2005-2010 included 24 small producer-type economies (see Table 3). At the other extreme, the slowest-growing populations tended to be in the former Soviet Union's mass production societies, whose peasantries at one time were large.

Fast population growth rates have destroyed the peasant economy's delicate equilibrium – a fragile balance of land and labor. W. Arthur Lewis noted that in the early stage of this type of economy, de-colonization and incipient growth created conditions for higher living standards, especially in rural areas. True enough, unemployment in Lewis' example, Nigeria, became the "most serious social problem due to the excessive drift away from agricultural employment, mainly by school leavers". But Lewis goes on to recognize an important fact about the countryside's high-wage potential that no longer holds: "the co-existence of *labor shortage* in rural areas (with its promise of higher wage incomes) with unemployment in the towns" (Lewis 1967). With rising population growth rates, unemployment and underemployment spread everywhere, and the peasant-export economy drifted into becoming a labor surplus economy, but one without Asia's competitive labor costs or large domestic markets.

In the 1970s, "the World Bank's population advocacy had done much to legitimize the developmental significance of family planning" (Kapur et al 1997, p. 344-5). But this stopped after negative fall-out from Indira Gandhi's forced sterilization program and China's one-child family, the two countries where poverty has fallen the most. As the importance of participatory democracy in grass roots anti-poverty programs rose, the World Bank took a more "roundabout" approach to population, one in which "family planning was integrated with and partly disguised by health and nutrition projects" (Kapur et al 1997, p. 344-5).

Table 3: Estimated population growth rate
(highest 30 and lowest 30 of 230 countries, 2005-2010)

Highest Rank	Country	Population Growth Rate (%)	Lowest Rank	Country **Former Soviet Union	Population Growth Rate (%)
1	Liberia*	4.50	**201**	Slovakia**	0.03
2	Burundi*	3.90	**202**	Grenada	0.02
3	Afghanistan*	3.85	**203**	Slovenia**	0.01
4	W. Sahara	3.72	**204**	Aruba	0.01
5	East Timor	3.50	**205**	Pitcairn Isl.	0.00
6	Niger*	3.49	**206**	Cuba**	-0.01
7	Eritrea*	3.24	**207**	Japan	-0.02
8	Uganda*	3.24	**208**	Tokelau	-0.03
9	DR Congo*	3.22	**209**	US Virgin Isl.	-0.03
10	Palestine Terr.	3.18	**210**	Czech Rep.**	-0.03
11	Jordan	3.04	**211**	Germany***	-0.07
12	Mali*	3.02	**212**	Croatia**	-0.09
13	Benin*	3.02	**213**	Poland**	-0.15
14	Guin.-Bissau*	2.98	**214**	Armenia**	-0.21
15	Yemen	2.97	**215**	Guyana	-0.22
16	Somalia*	2.92	**216**	Montenegro**	-0.27
17	Burkina Faso*	2.89	**217**	Dominica	-0.29
18	Chad*	2.88	**218**	Hungary**	-0.29
19	UAE	2.85	**219**	Estonia**	-0.35
20	Angola*	2.78	**220**	Romania**	-0.45
21	Rwanda*	2.76	**221**	Russia**	-0.51
22	Madagascar*	2.66	**222**	Latvia**	-0.52
23	Kenya*	2.65	**223**	Lithuania**	-0.53
24	Togo*	2.65	**224**	Belarus**	-0.55
25	Gambia*	2.63	**225**	Bulgaria**	-0.72
26	Malawi*	2.57	**226**	Ukraine**	-0.76
27	Mauritania*	2.53	**227**	Georgia**	-0.79
28	Syria*	2.52	**228**	Moldova**	-0.90
29	Ethiopia*	2.51	**229**	Niue	-1.85
30	Tanzania*	2.47	**230**	Cook Isl.	-2.23

(*Peasant economy with a majority of the working population engaged in small-scale agriculture or herding.
**Former Soviet Union ***East Germany only)

Source: Adapted from United Nations World Population Prospects (rev.), 2006.
Estimates for the period 2005-2010 using medium variant.

With fast increases in population and its redistribution to urban areas, the economic structure of the peasant export economy has collapsed, becoming a hybrid of fragmented markets co-existing alongside mass markets, the latter largely ignored by the grass roots technology movement except to approve widespread use of the 'cell' (according to World Bank estimates for January 2000, the number of internet hosts per 10,000 people was

higher in Sub-Saharan Africa (2.73) than in the Middle East and North Africa (0.55), South Asia (0.22) and even East Asia and the Pacific (2.73), where it was the same (World Bank 2000/2001, cited in Huff 2003).

5. The incentive system for prize-giving

The fight against poverty that idealizes local entrepreneurship and foreign technical assistance typically starts with a 'fact' without a footnote: three million children die everyday from malnutrition, a billion households lack toilets, 70 million people have water-borne diseases, and 90 percent of the handicapped population has no wheelchair and doesn't attend school (which may be true). As awareness of the sheer magnitude of these problems spreads, large financial incentives become available to innovators in the global North and South from donors, increasingly private, who believe in technological solutions for sustainable development.

A strong recommendation for prize-giving came out of a high-level workshop in 2003 on Invention and Innovation for Sustainable Development sponsored by the Lemelson-MIT Program and the National Science Foundation. It recommended:

> "Awards and prizes with large cash sums should be established to motivate inventors and innovators everywhere to focus on sustainable development. Prizes could be sponsored by well-known institutions, and should be given high visibility through media channels. Prizes should focus on serving as incentives for solutions to *large* problems, and the prize money should also be applied to the commercialization and dissemination of the new solution" (Invention and Innovation for Sustainable Development 2003).

By 2007, Congress had passed the COMPETES Act which amended the Sec.7023 of National Science Foundation Act of 1950 to permit the NSF to receive and use funds donated to the NSF for specific prize competitions for 'basic research'.

Below the Nobel Prize level, prizes have proliferated for poverty reduction in an age of 'corporate social responsibility,' convinced that as long as rich and poor 'partner' with each other in 'social enterprises,' advanced technologies can be adapted and applied everywhere on planet earth. Professor Abdul Hussam, of George Washington University, was awarded the Grainger Prize for Sustainability for his work on the Sono water filter to remove arsenic from Bangladesh's well water. The prize included a $1 million award, "much of which was generously donated to the further development and distribution of the filter". Pure Home Water, led by an MIT professor, Susan Murcott, also a prizewinner with a simpler technology to remove arsenic from well water in a region adjacent to Bangladesh in Nepal, was threatened with a lawsuit from SONO for infringing on its patent. Later, the 2010 St Andrews Prize for the Environment was awarded to a team from Queen's University, Belfast, also for an innovative method to remove arsenic from groundwater in West Bengal, next door to Bangladesh.

The Skoll Awards for Social Entrepreneurship, offering an undisclosed sum for three years, support social entrepreneurs whose work has the potential for large-scale influence on critical challenges of our time: climate change, nuclear proliferation, global pandemics, conflict in the Middle East and water scarcity. Goldman Sachs runs a program to help 10,000 women grow their own businesses and reinvest their success back into their communities (for which it won its own prize: a CEPC Excellence in Philanthropy Award). Kopernik awards a $100,000 grant from ExxonMobil to continue efforts to promote economic advancement for Indonesian women through improved energy access. An INDEX Prize of €500'000 in 2010 is regularly awarded by a Copenhagen-based non-profit network for "designs to improve life". The UN Development Program awards $10,000 to winners of the Equator Prize, to communities which approach sustainable development innovatively. Another Equator Prize recognizes local efforts to reduce poverty through the conservation and sustainable use of biodiversity ($30,000 in 2006). The Tyler Prize for Environmental Achievement made $200,000 available in

2009 for eco-friendly technologies to supply water in Southeast Asia and to sustain bio-diversity in Latin America. The Ashden Award for Sustainable Energy carries the award of being presented by Prince Charles. The Schwab brokerage Foundation gives an annual award to a "social entrepreneur", working in a "global social network", whose innovation has a "social impact"; the form of the award is recognition at the World Economic Forum. One of *Time* magazines "Heroes of the Environment", Tulsi Tanti, an Indian wind power mogul, captured the spirit of the age: "Yes, green business is good business, but it's not just about making money. It's about being responsible" (Baker 2007). Three teenage Palestinian girls were selected along with students from 50 other countries to compete in Intel's International Science and Engineering Fair, for a grand prize of $75,000. The girls made a breakthrough by wiring walking canes to detect pavement bumps (much like laser canes do). The first president of Singapore's public-funded University of Technology and Design, Thomas Magnanti (former dean of MIT's Engineering School), said "we will focus on trying to create what we call technically grounded leaders – people who have technology in their hearts and their soul and who will go out to the marketplace and… lead all of us, economically and socially, in Singapore, the region and beyond, to the future" (Joomcool.com 2008).

Prize giving for innovation has a long history, and President Thomas Jefferson calculatedly chose Benjamin Henry LaTrobe, an innovative architect to design two fire-proof buildings which depended on vault designs (Wermiel 2000). In 1714, the British government offered cash prizes to develop a way to determine a ship's longitude (as portrayed in a novel and motion picture). Napoleon offered a prize for new methods to preserve food for his army, and the winner came up with the idea of canning in 1809. The Zecker Prize to Louis Pasteur ultimately led to his discovery in 1865 of pasteurization, which raised exports of French wines. Napoleon III offered a prize for an inexpensive substitute for butter (in scarce supply), and a French chemist won it in 1869 with margarine. London's Royal Agricultural Society's nineteenth-century annual exhibition awarded medals for technological innovation (Bays and Jansen 2009).

Intuitively, medals, awards, appointments, commissions, prizes and other prestigious and lucrative incentives may be expected to enhance competition, creativity and communication among scientists and engineers – however imperfectly (no prize for longitude was ever awarded; "a vaulted loggia that connected the Treasury Building with [President Jefferson's] house collapsed… In 1808, the vaults over the Supreme Court chamber in the Capitol collapsed, killing the clerk of works" (Wermiel 2000)). The $100 laptop designed by the MIT Media Center was upstaged by India, who unveiled a Rs. 1,500 (around $30) laptop designed specifically for students and made in India (India Semiconductor Association 2010). Incentives are especially important for the survival of 'social enterprises' in today's developing world because many technologies intended to reduce morbidity and raise living standards have proven to be technically difficult to design, at a low price, with reliable after-sales service.

The 30-year quest for an ecologically friendly cooking stove is illustrative of the difficulties of devising appropriate technology (by comparison the gas and electric stove seem relatively easy to design and use for multiple cuisines). The average wood-burning cooking stove in a developing country produces about as much carbon dioxide as a car, deforests large tracts of land, and kills more people than almost any other environmental hazard, outdoor or indoor. Therefore, in 2009 a short section in the U.S House of Representatives' Waxman-Markey climate bill called on the Environmental Protection Agency to identify ways to provide safe stoves to twenty million households in five years. The stakes were getting higher in Stove Camp, an annual meeting in Oregon of stove hippie-experts (their research center supports itself with sales of their as one that:

1. reduces fuel use by more than 50 percent.

2. reduces black carbon by more than 60 percent.

3. reduces childhood pneumonia by more than 30 percent (all EPA standards).

4. is affordable ($10 retail or less).

5. is loved by cooks.

Yet "building a good cheap stove (about the size and shape of a stockpot, with a cylindrical combustion chamber and a cooking grate on top) can drive an engineer crazy… Fire is a fickle, non-linear thing, and seems to be affected by every millimeter of a stove's design". Moreover, local people like the taste from their own stove. As Stove Camp ended in 2009, an engineer said, "these designs usually take months and you're still scratching your head". Earlier, he had flown to London to collect the Ashden Award from Prince Charles (Bilger 2009).

Whatever the virtue of prizes for technologies for sustainable economic development, the way incentives are structured encourages *innovation* but not necessarily *economic development*; the two can easily diverge. An acclaimed innovation may not necessarily be the best available; the best may be buried in fragmented markets that contain imperfect information, making discovery costly (Japan's international aid agency had worked for years on a stove for Zambia, and one of Japan's private chemical companies had invented a smokeless briquette). To win a prize, technologists must create a new and better product; no incentive exists for them to search in tens of countries to identify an even better one that already exists, cheaper and more reliable for a poor community. The *appropriate* technologist wins an award for manufacturing an ingenious invention that uses local labor and materials (eschewing foreign trade) – as in the wood burning stove, or the locally made leather sandal, with soles from worn rubber bicycle tires. The aim of economic development, on the other hand, is to create jobs and to make the best products accessible even if they don't use appropriate technology, as in laying on electricity for an electric stove, or manufacturing a synthetic rubber flip-flop in batches of millions using petrochemicals, not local materials: buyer preferences suggest that in hot humid climates the flip-flop is a cheaper, cleaner and cooler footwear, and in hot arid climates the plastic sandal tends to predominate over the leather sandal. Even interpreting the motives of innovators generously, they win awards in the highest circles despite ineffective search routines.

The Rockefeller Foundation gives credit in its 2008 *Annual Report* to the methodology of 'Positive Deviance' in helping to reduce malnutrition in Vietnam. A breakthrough came from a psychologist, Jerry Sternin, who worked in 1970 in Vietnam with the 'Save the Children' Fund, an American NGO. The Rockefeller Foundation's *Annual Report* states:

> "We… came to understand that social sector problem solving could achieve tremendous impact if the individuals and communities with the problems are included in the process of developing the solutions, an idea pioneered by a number of private sector and design companies (and many academics)[10]. To assess this concept, we partnered with a group called Positive Deviance, which identifies behaviors that enable outliers or 'positive deviants' to succeed and then encourages others within the community to adopt these same behaviors. For instance, Positive Deviance initially cast its eyes on malnutrition in Southeast Asia. Researchers visited a low-income Vietnamese village and immediately noticed that children in a scattering of families were in exceptionally good health. Upon closer examination, they discovered that, in these households, providers did not wash away shrimp and crabs found in rice paddies but, instead, cooked them along with their rice – adding protein to a carbohydrate-based diet. This technique

[10] For the role of participatory democracy in a social program's success, see Tendler (1997).

> was embraced in villages across the country. It was a small, user-generated innovation that made a big impact" (Rockefeller Foundation 2008).

In fact, the discovery of the nutritional value of fish in rice paddies spread to Southeast Asia from India 1500 years ago: "Widespread over the world, fish culture in rice fields is mainly concentrated in South East Asia where it has been known for centuries" (Coche 1967). Sternin makes no mention of this body of scientific research in his 2010 book, with a chapter on Vietnam nutrition (Pascale et al 2010).

To reward the positive deviance method rather than the optimum search method probably did Vietnam no harm. But rewarding the second objective, of optimum search, might have helped diffuse the co-farming of rice and fish faster outside South East Asia, through a global network of government officials, agro-engineers and science-trained agronomists. Not to award prizes for 'due diligence' also has escalating costs because, as discussed below, a 'third way' is increasingly available if only it can be recognized and rewarded. The 'third way' involves neither the advanced technology of the developed world, and its high prices, nor the appropriate technology of the destitute world, and its many failures. New 'appropriate' products are in demand by poor countries *and* are produced commercially, in large volumes, at highly attractive prices, in these economies. Examples are the pharmaceutical and wind-power industries of India, the sewing machine and bicycle parts industries of China, the bio-fuel and natural plants industries of Brazil, and the aqua-fish and wheelchair industries of Vietnam.

6. Social enterprise: NGO or EGO?

The incentive structure for innovations to reduce poverty distorts economic development further by being biased towards fresh start-ups; usually an innovation ('new' locally) entails the establishment of yet another social enterprise, without the advantage of accumulated field experience, a professional staff, and stable employment[11]. Typically the new enterprise creates competition for a local business already struggling to survive[12]. If the social enterprise uses prize money or subsidies from its founder to give away its product for free, then this drives the incumbent out of the market, leading to job losses. A comparable experience occurred in the 1950s, when American aid under Public Law 480 took the form of surplus agricultural crops such as rice. This supply tended to drive down world prices, which then hurt commercial rice farmers in Thailand and Burma.

An incumbent's survival may be assured by what has become common practice – a new social enterprise is expected to partner with a local individual, university or NGO. Instead of becoming bankrupt, the incumbent becomes part of the new enterprises' social network. Partnering, however, is threatened in any social enterprise that is struggling. Social enterprises that operate in the red depend for their survival on their founder's philanthropy or fund-raising skills. In times of crisis, an established NGO (Oxfam) may buy water filters at a given price from a start-up (Pure Home Water) and 'sell' them for free to save thousands of lives (and place Pure Home Water's operations in the black). Alternatively, a founder may himself finance a needy person's purchase of his product. Frederick K W Day, one of two owners of Chicago-based Sram, the world's second largest bicycle-components manufacturer, attempted to put "millions of sub-Saharan Africans aboard special heavy-duty bikes designed to withstand the continent's rugged roads while carrying 200 pound cargos". After Asia's tsunami struck in 2004, Day and his wife rushed to Sri Lanka and had 24,450 of his heavy-duty

[11] To my knowledge, there is no precise definition of a 'social enterprise,' which may or may not be 'for profit' versus 'non-profit,' or self-financing versus dependent on external contributions. For the confusion, see Acs and Sany (2009). For a discussion of the difficulties of defining a 'social enterprise' in Puritan America, see Frey (2010) .

[12] For the reasons behind the reluctance of poor consumers to switch products, see Beaudry and Francois (2007)..

prototypes assembled, priced at $100 but given away free (with donations from Sram customers like Trek Bicycle), the object being "to help people move quickly through the countryside" (Fitch 2010).

Such social enterprises have difficulty attracting capable local partners, who may reasonably doubt their sustainability once the White Man leaves his Burden behind[13]. There is little empirical evidence on local partnering, although many home web pages of social enterprises mention their number and praise their loyalty.

In the case of local partners in start-ups with scarce resources and an uncertain future, it is very difficult for them to afford after sales service, and the more fragmented the market, the more costly it is to provide such service and usually the greater is its need. The problem in poor countries of bad maintenance entangled with bad design is hard to quantify, but is suggested by archaeology: rusty relics of corn shuckers are strewn throughout the African countryside. A carpenter at stove camp admitted: "Here's the deal. The world is absolutely littered with failed stoves" (Bilger 2009). Frederick K W Day, the bike-parts mogul, recalls his first long car ride in Zambia in 2006. '"We're looking out the window and we keep seeing bike carcasses piled up alongside the road,' he says. 'It was like something out of The Andromeda Strain'" (Fitch 2010). Pure Home Water, which guarantees the effectiveness of its process, attributes discarded water filters to poor up-keep.

A popular image of the social enterprise is that it does no harm, and in a harsh world miraculously does a little good[14]. But if it pays no taxes and ambushes the community's spare change, then it does a few people harm and a lot of people no good. Little wonder measured poverty rates do not change if their reduction depends on a faulty incentive system and the 'social enterprise', whose definition for prize giving should depend on doing a lot of people a lot of good[15]. USAID changed its requirements for its Development Innovations Venture grants for 2011 where winners would have to demonstrate an impact of their innovation on at least 75,000 people[16].

7. Conclusion: The Flaw of Failure

With no actual success story to follow, the grass roots anti-poverty movement has been stuck; the poverty rate in Sub-Saharan Africa for 25 years, before an explosion in raw material prices, has remained unchanged. It was no different in 2005 from what it was in 1981. The reason no successful example has evolved, I would argue, is because in the context of the poor peasant export economy, the movement's working theories are logically flawed. Africa's entrepreneurial spirit may be alive and well, but with fast population growth and sky-high rates of unemployment and underemployment, the hardiest entrepreneur cannot open a small business because other small entrepreneurs are already exploiting the same idea and making a loss. There is a shortage of opportunities for self-employment and paid employment at above subsistence, but the demand side of the equation is not addressed by the grass roots. The anti-poverty movement suffers from its belief in the fallacious 'Say's Law' that the supply of better educated, healthier workers can create its own demand. Despite a firm faith

[13] It is interesting to speculate on the response of the South's own professional elites to grass roots initiatives; it must be multi-faceted. The sour note is sounded by a Kenyan water engineer about a "turning waste into profit" foreign project: "I have read the story and I am shocked. The group installed only one toilet in Kibera (Nairobi's largest slum) a few months ago and now it is writing about it to the entire World. This toilet is receiving only 10-25 visitors a day. The design and materials used are very wanting. (The entire structure is pre fabricated concrete labs that weigh up to 90 kgs.) The business model is not tested (only dreamt up). I had the opportunity of meeting the team when they were here. My honest assessment is that there is nothing to write home about in this...." (interviewed, November 28th 2010, by S Gulyani, World Bank, Nairobi Office).

[14] Nicholas D. Kristof of the New York Times is responsible for highly favorable editorializing about the social enterprise. See, for example, Kristof and WuDunn (2009) and Kristof (2010).

[15] One of the most successful social enterprises worldwide, Whirlwind Wheelchair International, developed a cheap, comfortable, repairable wheelchair for rugged terrain, crowded streets and tiny dwellings in a business with a reputation for a highly diversified product that caters to a handicapped population with differing individual needs (Krizack and Rogin 2003).

[16] See http://www.grants.gov

in the powers of technology, technology is not raising labor demand by inventing better goods and services since most innovators in Africa are foreign, and do not have a deep knowledge about what Africa could produce in modern manufacturing, agriculture or services for overseas trade, or for the mass markets that domestic population growth is creating. Instead, innovators design 'appropriate technologies' which could be bought mostly for consumption, which creates almost no income stream for the poor consumer. Rural electrification is widely in demand at a price that could almost certainly cover costs, and could broaden the scope for entrepreneurship for millions, but the supply and demand of mass-produced public goods is higher up the supply chain and closer to government intervention than the grass roots can stomach. An illogical incentive system for technology and entrepreneurship creates an inefficient use of time, and which has fostered a 'role model failure.' Prizes for new technology have proved their worth over time in different cultures, but incentives in peasant export economies are biased towards solving the consumption problems of small fragmented communities, when the employment problems of mass consumption communities are not being addressed at all.

These debates, however, are being eclipsed by a possibly transformative change: emerging economies are buying more raw materials and world prices are rising. Such economies are also supplying investment capital to new extractive industries, thereby broadening their choice of investor. New discoveries are being commercialized, and these have the potential to make Africa resource blessed, not cursed. The debates among elites about poverty are being overshadowed by debates about how poor countries should manage their natural resources and avoid ecological disasters. The role of teacher, however, is moving away from the old entrepreneurial elite into contested terrain. Should/could oil-rich West Africa follow the OPEC role model of resource nationalism; considering the fact that Angola is OPEC's newest member as of 2007, or should it follow a new Enlightened order, one of foreign multinationals practicing 'corporate social responsibility'? The answer will reveal the relative strengths of a good adaptive role model (OPEC) and a respected age-old theory of the Enlightenment.

References

Acs, Z and J Sany (2009), Measuring the social value of innovation: The cases of Muhammad Yunus, Grameen Bank and Bill Gates, Microsoft. In: Liebcap G D (ed.), *Measuring the Social Value of Innovation: A Link in the University Technology Transfer and Entrepreneurship Equation (Advances in the Study of Entrepreneurship, Innovation & Economic Growth), Volume 19*, Emerald Group Publishing Limited, pp. 143-170.

Amsden, A H (2010), Say's Law, poverty persistence and employment neglect. *Journal of Human Development and Capabilities,* 11 (1), pp. 57-66.

ASADI (2010), Turning science on: Improving access to energy in sub-Saharan Africa, Sixth Annual Meeting of the African Science Academy Development Initiative, November 7-10.

Atkinson, A B (1987), Poverty. *The New Palgrave Dictionary of Economics,* J Eatwell, M Milgate and P Newman (eds.) Macmillan, p. 3.

Baker, A (2007), 'Heroes of the environment', *Time, October 17th.*

Banerjee, D (2010), *Economic and human development in contemporary India: Cronyism and fragility.* New York and Oxford: Routledge.

Bays, J and P Jansen (2009), *And the winner is... Capturing the promise of philanthropic prizes*, McKinsey: March. Available from: http://mckinseyonsociety.com/capturing-the-promise-of-philanthropic-prizes/

Beaudry, P and P Francois (2007), *The economics of inefficient technology use.* NBER Working Paper 13500.

Beny, L N and L D Cook (2009), 'Metals or management? Explaining Africa's recent economic growth performance', *American Economic Review*, 99 (2), pp. 268-74.

Bilger, B (2009), 'Hearth surgery: The quest for a stove that can save the world', *New Yorker*. December 21st, p. 85.

Chen, S and M Ravallion (2010), 'The Developing World Is Poorer Than We Thought, but No Less Successful in the Fight Against Poverty', *Quarterly Journal of Economics*, 125 (4), pp. 1577-1625.

Chowdhury, A (2009), *Microfinance as a poverty reduction tool – A critical assessment*, United Nations DESA Working Paper.

Coche, A G (1967), 'Fish culture in rice fields A world-wide synthesis' *Hydrobiologia,* 30 (1) pp. 1-44.

Cornia, G A (2010, *Report of the expert group meeting on special international support measures for the Least Developed Countries*, CDP2010/PLEN/5, United Nations.

Duflo, E (2001), 'Schooling and labor market consequences of school construction in Indonesia: Evidence from an unusual policy experiment', *American Economic Review,* 91 (September). Working Paper Version available online.

Easterlin, R A (2004), *The Reluctant Economist: Perspectives on Economics, Economic History and Demography*. Cambridge, UK: University Press.

Fitch, S (2010), 'Can this bicycle save lives in Africa?', *Forbes*.

Frey, D E (2010), Book review: Heavenly merchandise: How religion shaped commerce. In: Puritan America, by Mark Valerie, *Book Reviews in Economic and Business History.*

Harris, J R (1970), Nigerian entrepreneurship in industry. In: C K Eicher and C Liedholm (eds.) *Growth and Development of the Nigerian economy.* Michigan: Michigan State University Press.

Hooper, M, et al (2009), *Scaling-up community efforts to reach the MDGs – An sssessment of experience from the Equator Prize*, United Nations Development Programme.

Huff, T E (2003), *The rise of early modern science*. Cambridge: Cambridge University Press.

ILO (2008), *Global employment trends.* International Labour Organization.

India Semiconductor Association (2010), 'India unveils Rs. 1,500 laptop for students', *ISAWinWire,* July 23-30. http://www.ciol.com/News/Product-Features/News-Reports/India-unveils-Rs1,500-laptop-for-students/139174/0/.

Joomcool.com (2008), 'Singapore 4th university named as University of Technology and Design'.

Kapur, D et al (1997), *The World Bank: Its first half century*. The Brookings Institution.

Kristof, N D (2010), 'DIY Foreign-aid revolution', *New York Times Magazine*. October 20th. http://www.nytimes.com/2010/10/24/magazine/24volunteerism-t.html?pagewanted=all

Kristof, N D and S WuDunn (2009), 'The women's crusade'. *New York Times*. August 17th. http://www.nytimes.com/2009/08/23/magazine/23Women-t.html?pagewanted=all

Krizack, M and J Rogin (2003), 'Whirlwind Wheelchair International launches industrialization project', *Disability World,* 21.

Lewis, W A (1967), *Reflections on Nigeria's economic growth*, Development Centre, OECD.

Myint, H (1964/1995), *The Economics of the Developing Countries*. Santa Barbara, California: Praeger.

Owen, R (1981), *The Middle East in the world economy 1800-1914*. York, UK: Methuen.

Oxford Institute for Energy Studies (2010), 'Suleiman J. Al-Herbish rejects the gloom and doom surrounding energy policy' *Oxford Energy Forum* (May).

Pascale, R, J Sternin and M Sternin (2010), *The power of positive deviance: How unlikely innovatorrs solve the world's toughest problems,* Harvard Business Press.

Rockefeller Foundation (1979), *Annual Report*. Rockefeller Foundation.

Rockefeller Foundation (2008), *Annual Report*. Rockefeller Foundation.

Roodman, D and U Qureshi (2006), *Microfinance as business.* Center for Global Development.

Tavernise, S (2010), 'Upstarts chip away at power of feudal Pakistani landlords', *New York Times,* August 20th.

Tawney, R H (1913), *Poverty as an industrial problem.* London School of Economics.

Tendler, J (1997), *Good government in the tropics.* John Hopkins University Press.

TWN (2010), 'How to develop LDC's productive capacities', *SUNS* No. 7028, October *28th*, Third World Network.

UNCTAD (2010), *Hungry for change: building new paths to food security in LDCs.* UNCTAD Policy Briefs, UN Conference on Trade and Development, June, No 15.

UNDP (2010), Human Development Report 2010 – 20th Anniversary Edition: 'The real wealth of nations: Pathways to human development'. United Nations Development Program.

UNIDO (2010), Going global in great strides. Interview with his Excellency Chen Deming, Minister of Commerce of the People's Republic of China. In: United Nations Industrial Development Organization (ed.) *Making It Industry for Development*.

United Nations (2006), *World Population Prospects, The 2006 Revisions.* United Nations.

Wermiel, S E (2000), *The fireproof building: Technology and public safety in the Nineteenth Century American city*. Johns Hopkins University Press.

World Bank (2000/2001), *World development report*. World Bank.

World Bank (2008), *The welfare impact of rural electrification: A reassessment of the costs and benefits An IEG Impact Evaluation*. Independent Evaluation Group. Washington DC: World Bank.

World Bank and Ministry of Economic Development of Egypt (2007), *Arab Republic of Egypt, A poverty assessment update*: Volume 1 Main Report, June 17th.