WORLD ECONOMICS ASSOCIATION

WORLD ECONOMIC REVIEW

An open access, open peer review journal

Issue 4, February 2015

World Economic Review ISSN: 2049-3517

Published by the World Economics Association

Sister WEA open-access journals: *Economic Thought* and *Real-World Economics Review*

Does Saving Increase the Supply of Credit? A Critique of Loanable Funds Theory

Fabian Lindner[1]
Macroeconomic Policy Institute (IMK), Dusseldorf, Germany

Abstract

The paper presents a critique of loanable funds theory by using simple accounting relationships and standard excess demand analysis. It is shown that many economists identify saving and the credit supply by interpreting the macroeconomic saving-investment identity as a budget constraint. According to that interpretation, more saving through lower consumption (and government spending) leads to a higher supply of credit, lower interest rates and thus more funds to be used by firms for investment. The paper shows that proponents of this theory confuse quite different economic phenomena and commit serious fallacies of composition. In the first step, the concepts of "saving" and "credit" will be clearly distinguished using simple accounting. It will be shown that credit is not limited by anybody's saving and that no one has to abstain from consumption in order for a credit to be provided. Also, it will be shown that financial saving (an increase in net financial assets) through a reduction in expenditures reduces other economic units' revenues and thus their ability to spend and save. Using the concept of excess demand and supply, it will be shown that excess saving does not lead to an excess supply of credit − which would lower interest rates − but to an excess supply of goods, services and/or labor which will lower prices and production. How interest rates change is not determined by excess saving: They could increase, stay the same or decrease. Finally, it will be argued that the identification of saving with the provision of credit is likely to stem from the invalid application of neoclassical growth models to a monetary economy.

Key words: saving, wealth, investment, production, financial markets

JEL-Classification: E210, E220, E230, E440, E500

Introduction

> "According to traditional theory, [...] the amount of credit is strictly limited by the amount of saving. The amount of saving determines the quantity of investment, i.e. the amount of goods production that cannot be directly and immediately consumed. [...] [But] the volume of credit does not at all depend on the quantity of money savings. It depends on banks' ability and willingness to provide credit and on borrowers' willingness to increase their debts" (Gestrich, 1947, pp. 23-24; p.27).[2]

Many economists hold the position that "saving finances investment". They argue that saving − a reduction of consumption relative to income − is necessary for the provision of loans and the financing of investment. This view has been influential both to explain the global economic crisis that began in 2007 and in proposing policies to combat the crisis.

For instance, Bernanke (2005) argues that East Asian and commodity exporting economies exported their scarce savings to the United States, limiting their domestic credit supply and increasing credit to the US. This contributed to reduce US interest rates and led to high housing prices andinvestment. For the Euro Area, Sinn (2010) argues that Germany exported its scarce savings to today's crisis countries, thereby

[1] Contact address: fabian-lindner@boeckler.de. The author thanks Jan Behringer, Dirk Bezemer, Irwin Collier, Sebastian Gechert, Norbert Häring, John Harvey, Erik Klär, Uwe Richter, Alexandros Tokhi, Claudio Zeitz and Rudolf Zwiener for valuable comments. All remaining errors are of course the author's.

[2] The German original reads: "Die Menge des Kredits ist nach der traditionallen Auffassung durch die Menge der ersparten Geldbeträge eindeutig bestimmt. [. . .] Das Kreditvolumen ist demnach durch das Esparnisvolumen starr begrenzt. Die Menge der Ersparnisse bestimmt die Höhe der Investitionen, d.h. den Umfang an Erzeugung von nicht direkt und sofort konsumierbaren Gütern. [. . .] [Aber] das Kreditvolumen ist in keiner Weise abhängig von der Menge der Geldersparnisse. Es hängt von der Fähigkeit und Bereitwilligkeit der Banken ab, Kredit zu gewähren, der andererseits auch eine Bereitwilligkeit zur Kreditverschuldung gegenüberstehen muss."

reducing domestic German credit and investment. Cochrane (2009), Fama (2009) and Ball and Mankiw (1995) argue that higher government spending "absorbs" household savings that would otherwise be available to finance investment.

The present paper shows that the underlying view of the "saving finances investment" doctrine implies that the amount of loans that can be lent in a period is limited by the amount of saving in the same period and that, consequently, an abstention from consumption (and other spending) is necessary for more credit to be available for investors. This view is the basis of loanable funds theory (Robertson, 1936; Ohlin, 1937a, c, b; Tsiang, 1956).

It will be argued that this view is deeply flawed. Using simple accounting rules in the vein of Stützel (1978; 1979) and Lavoie and Godley (2012), the paper will first clarify the difference between saving and the provision of credit: saving changes an economic unit's net worth which the provision of credit does not. The provision of a credit only changes either the composition of a unit's financial assets or the length of its balance sheet, but not its net worth. Second, it will be shown that loanable funds theorists commit a fallacy of composition when they claim that *household financial* saving increases *aggregate* investment in *tangible assets*.

While the debate about loanable funds theory goes back at least to the discussion of Keynes' General Theory (1936/1997), the problem and critique of a limited saving fund has not figured prominently in the Anglo-Saxon debate. The discussion was mainly about whether the supply and demand of the *flow* of credit − identified with loanable funds theory (Patinkin, 1958, 1969; Snippe, 1985; Tsiang, 1956) − or the *stock* of money − identified with Keynesian liquidity preference theory (Keynes, 1937; Lerner, 1944) − determine interest rates.

Hicks (1939) and Patinkin (1958) showed that the implications for interest rate theory are the same whether one uses an analysis of the stocks or flows of credit or money. The loanable funds theorist Tsiang (1956) accepted that view and it is now common to see both theories as being identical (Blanchard, 2000).

But while Patinkin and Hicks are right in their analysis, the question is *not* whether stocks or flows of money or credit determine interest rates, but whether one sees *saving* as the limiting factor for the provision of loans. This issue has hardly been discussed, with a short interchange on the matter between Ohlin (1937a, p. 425) and Keynes (1937, pp. 243-245) being an exception. Recent critiques of loanable funds theory (Bibow, 2001; Hayes, 2010; Keen, 2014) also did not stress this aspect.

However, this fallacy was central to the thinking of the now largely forgotten German economists Wilhelm Lautenbach (1952) and Hans Gestrich (1940; 1947) (see Kindleberger (1999) and Garvey (1975) on some discussion). Both were economists in the 1920s and 1930s in Weimar Germany and closely followed the world economic depression. They were early critics of confusing saving and the credit supply, especially because they were familiar with the workings of financial markets and witnessed the disastrous economic and social consequences of applying the loanable funds doctrine to fiscal policy in Weimar Germany.

The present article will use Lautenbach's and Gestrich's insights and embed them in the analytical framework developed by Lautenbach's pupil Wolfgang Stützel (1978; 1979) (see Schmidt (2009; 2012) for introductions into Stützel's work). In his book "Volkswirtschaftliche Saldenmechanik" ("Financial balance mechanics") (1978) Stützel already in the 1950s developed a stock-flow-consistent (SFC) framework similar to Godley and Lavoies' (2012) later work.

In contrast to Godley and Lavoie, Stützel was not interested in model building, but in using his SFC framework as an analytical tool for the critique of traditional Walrasian theory and the neoclassical synthesis. By clearly defining accounting concepts and their interaction, he often found economists to commit serious "problem entanglements", i.e. confusing quite different economic phenomena and committing fallacies of composition. Most of the analytical framework of this paper builds on Stützel's work. The confusion of the concept of saving which changes units' net worth and the concept of credit provision which only changes the composition or length of units' balance sheets but not their net worth is just such a problem entanglement.

Borio and Disyatat (2011) came to the same conclusions as Lautenbach and Gestrich, using simple accounting as a starting point for their analysis. However, they limit themselves to a critique of Bernanke's "Saving Glut" hypothesis and do not draw the wider implications for the whole of loanable funds theory.

The present paper is close to Post-Keynesian thought, as were Lautenbach, Gestrich and Stützel.

But some of the traditional Post-Keynesian answers to the issues raised by loan able funds theory are not sufficient to counter this theory. It will be shown that neither the endogeneity of money (Moore, 1988; Fontana, 2003; Lavoie, 2013; Keen, 2014) nor the assertion that "investment creates saving" (Palley, 1996; de Carvalho, 2012) are sufficient to counter loanable funds theory.

The article is structured as follows. In the first section, it will be shown that many contemporary economists hold the view that saving is the limiting factor for the provision of credit and that consumption thus has to be reduced to increase credit. In the second section, standard business and national accounting concepts are used to show that saving, finance and investment are different concepts that have to be clearly kept apart. In the third section, the accounting relations between saving, finance and investment in the *aggregate* economy will be clarified. In section four, it will be argued that excess saving is likely to lead to a fall in aggregate income while interest rates might increase, decrease or stay the same. That means that none of the claims of loanable funds theorists is likely to hold in reality. In section five, the implicit assumptions of loanable funds theory will be discussed and how they are related to neo-classical growth models. A final section draws some implications and concludes.

1. The loanable funds theory or: the saving-fund theory of credit

Loanable funds theorists believe that the amount of credit available to finance investment is constrained by saving, and that the identity of saving (income minus consumption) and investment of the national accounts represents a budget constraint. This "saving finances investment" view leads loanable funds theorists to believe that higher saving through lower consumption and lower government deficits (or, ideally, surpluses) leads to a higher credit supply, lower interest rates, more investment and thus a higher capital stock and higher future income.

Unfortunately, it is difficult to find an academic paper that clearly and systematically develops this theory, something Keynes already observed (1936/1997, pp. 175-193). For this reason, I will draw from Gregory Mankiw's intermediate textbook version of the loanable funds theory (1995) (Paul Krugman, too, gives the same account in his introductory economics textbook (2009, ch. 26)). One may object that this could bias the argument since what an author writes for undergraduate audience will likely be oversimplified. But, Mankiw himself says otherwise in an article with Ball (1995):

> "The analysis [consistent with Mankiw's textbook treatment, F.L.] follows the conventional wisdom as captured, for example, in most undergraduate textbooks. In our view, the conventional wisdom in this area is mostly on the right track" (1995, p. 95).

Furthermore, this has the advantage that, for pedagogical reasons, many otherwise implicit assumptions are made explicit (Naples and Aslanbeigui, 1996).

In his textbook, Mankiw starts from the accounting identity of a closed economy with a government:

(1) $$Y = C + I + G$$

Y is income, *C* is consumption, *I* is investment and *G* are government expenditures. Then he introduces taxes, *T*, in order to derive the economy's saving:

(2) $$S = (Y - T - C) + (T - G) = I$$

He goes on to call the first term "private saving" and the second "government saving". According to Mankiw, the sum of private and government savings are the "flows into the financial markets" and investment the flows "out of the financial markets" (Mankiw, 1997, p. 67). That means that he interprets the national accounting identity between saving and investment as a budget constraint: no flows of loanable funds would be available to investors without prior saving.

Mankiw further assumes investment to be a negative function of interest rates, so that both private and government saving determine the interest rate and thus equilibrium investment and saving. Mankiw writes:

> "In fact, saving and investment can be interpreted in terms of supply and demand. In this case, the "good" is loanable funds, and its "price" is the interest rate. *Saving is the supply of loans* [emphasis added, F.L.] − individuals lend their saving to investors, or they deposit their saving in a bank that makes the loan for them. Investment is the demand for loanable funds − investors borrow from the public directly by selling bonds or indirectly by borrowing from banks. [. . .] *At the equilibrium interest rate, saving equals investment, and the supply of loans equals the demand* [author's emphasis]" (1997, p. 63).

Mankiw introduces the loanable funds mechanism in his chapter about "The Economy in the Long Run" but also uses it to derive the short run IS curve of the IS/LM model (1997, p. 260).

From this follows that an increase in saving brought about by a cut in expenditures (less consumption relative to income for households and less government expenditures relative to government revenues) increases the supply of loans. If households and the government (by government surpluses) did not save, not enough finance needed by firms would be forthcoming. Thus, the maximum amount of credit for investment (loanable funds) in a (closed) economy is limited by the fund of current household and government saving.

The whole view is not Mankiw's idiosyncrasy but reflects the opinion of many economists. That credit is constrained by saving in the loanable funds literature is also made clear by D. H. Robertson (1934):

> "And we have a curve SS' representing the rate of new available savings per atom of time − available that is, after deducting new savings absorbed in financing consumption by Governments or individuals" (Robertson, 1934, p. 651).

Woodford (2010) is very explicit in stating that it is a restriction of expenditures that is necessary for credit to be created:

> "The loan supply curve *LS* shows the amount of lending *L* that ultimate savers are willing to finance (*by refraining from expenditure themselves* [emphasis added, F.L.]) for each possible value of the interest rate received by savers [...]" (Woodford, 2010, p. 26).

Bernanke (2005) in his speech on the "global saving glut" also makes the "saving finances investment"-view explicit:

> "All investment in new capital goods must be financed in some manner. In a closed economy without trade or international capital flows, the funding for investment would be provided entirely by the country's national saving. By definition, national saving is the sum of saving done by households [...] and saving done by businesses [...] less any budget deficit run by the government."

One can also open the economy and derive saving by adding net exports, *NX* (exports minus imports):

(3) $$S = (Y - T - C) + (T - G) = I + NX$$

Bernanke (2005) applies the "saving finances investment" view explicitly to the open economy:

> "[...] in fact, virtually all economies today are open economies, and well-developed international capital markets allow savers to lend to those who wish to make capital investments in any country, not just their own. Because saving can cross international

> borders, a country's domestic investment in new capital and its domestic saving need not be equal in each period. If a country's saving exceeds its investment during a particular year, the difference represents *excess saving that can be lent on international capital markets* [emphasis added, F.L.]. By the same token, if a country's saving is less than the amount required to finance domestic investment, the country can close the gap by borrowing from abroad."

In the view of the "saving-fund theory of credit", a trade surplus leads to an export of saving so that loans cannot be used any more in the country realising the surpluses. For instance, this view has been applied to Germany's export surpluses by Hans-Werner Sinn (2010). Sinn argues, consistent with loanable funds theory, that saving which is exported cannot be used to finance domestic investment (see, for an analysis of this point, Horn and Lindner (2011)). Consequently, he claims that there is a direct trade-off between domestic loans and investment and export surpluses:

> "Germany exported its savings instead of using them as loans for investment in the domestic economy. [...] Germany lost a huge amount of capital under the euro even though it urgently needed the capital to rebuild its ex-communist east" (Sinn, 2010, p. 7).

The policy implications of the loanable funds model are straight forward: policies that lead to lower consumption relative to income (higher household saving) and lower government deficits (even surpluses) provide more saving, lower the interest rate and lead to higher investment. Naturally, the reduction of fiscal deficits would be a way to increase national saving:

> "Budget deficits have many effects. But they all follow from a single initial effect: deficits reduce national saving.[...] When budget deficits reduce national saving, they must reduce investment, reduce net exports, or both.[...] A decline in national saving reduces the supply of loans available to private borrowers, which pushes up the interest rate (the price of a loan)" (Ball and Mankiw, 1995, pp. 96-98).

To summarize: According to loanable funds theory, credit is provided by saving. The higher saving is, the more credit will be supplied which ceteris paribus lowers interest rates. Higher saving through lower consumption (relative to income), lower government deficits, and/or lower export surpluses would then induce firms to invest more which would increase long run income through a higher capital stock. But as will be shown in the remainder of this article, this view is untenable.

2. Accounting rules, saving, investment and finance

In the following section, saving, investment and finance will be clearly defined using common business and national accounting concepts and rules (European Communitites, 1996; Lequiller and Blades, 2006; Brümmerhoff, 2007; Möller et al., 2011). A first step in order to clearly distinguish between saving and financing is to understand the concept of income, production, the current account and the financial account. Those concepts are often used for national economies. However, they apply to every economic unit or group of economic units, be it an individual, a household, a firm, the firm sector, the household sector etc. Both Eurostat (1996) and the US Integrated Macroeconomic Accounts (Bond et al., 2007) use the same accounting concepts developed for national economies for every domestic economic sector.

2.1 Balance sheets

First, different balance sheet items will be defined and then the flows that change those items. The balance sheet of an economic unit or group of economic units consists of its assets, its liabilities and its net worth, nw. Its assets are tangible assets, *ta*, like machines, houses, etc. and *gross* financial assets, gfa, like

money, bonds, stocks etc. Its liabilities, *l*, are its debts and equity in the form of stocks held by other economic units:

(4) $$ta + gfa - l = nw$$

Note that we also use stocks as a liability. This is done in accordance with international accounting practice (European Communitites, 1996; Bond et al., 2007) although equity gives rise to a contingent claim (dividends) in contrast to contractual claims (interest) (Dos Santos and Macedo e Silva, 2010).[3]
Net financial assets, nfa, are *gross* financial assets minus liabilities:

(5) $$gfa - l = nfa$$

Gross financial assets can be further split into means of payment − money, m − and all other financial assets, ofa:

(6) $$gfa = m + ofa$$

This distinction is crucial for any financial analysis: all other financial assets – ofa − are promises to receive means of payment. Since some unit's financial asset is another unit's financial liability, financial liabilities are promises to make payments. A payment is defined as the act of servicing a contractual debt (Kaiser, 2008, p. 25). The medium to do that − money − has thus to be clearly distinguished from all assets that only give a right to receive money but are not money as such.

What exactly constitutes money depends on the context (Stützel, 1978, pp. 65-66). For instance, a euro is not accepted to service dollar debts; a deposit at a commercial bank is normally accepted by non-banks as a means of payment but not so among commercial banks themselves.

They only accept central bank money (notes and deposits at the central bank) to service their debts among each other, not their respective liabilities. On the other hand, if somebody accepts a bond or a stock as a means of payment, such financial assets become means of payment etc.

The context-dependence of what constitutes a means of payment makes it so hard to exactly define what kind of financial asset can be called "money". But the distinction between means of payment and other financial assets is at the heart of every financial crisis: in a crisis, debtors have difficulties to make good on their promises to pay money. Even if they held other financial assets but would not be able to convert them into money, they would risk default and creditors had to write down their financial assets.

All balance sheet items are shown in table 1. Depending on the type of economic unit, the balance sheet's composition is different. For instance, non-financial firms normally hold mainly tangible assets and much less gross financial assets. They often have a high net worth and low debts. Private households typically hold both tangible assets (mainly houses) and gross financial assets (deposits, bonds, stocks) and have high net worth. Banks' tangible assets are mostly negligible. They mainly hold loans, bonds, derivatives and other gross financial assets, have very high debts and very low net worth.

[3] If stock options would not be a liability, they would not constitute financial assets for their holders but tangible assets, i.e. assets to which no liability corresponds. Then, the trade of stocks would not be a purely financial transaction but constitute an expenditure or revenue which would change units' net financial asset position.

Table 1: Balance sheet

Assets	Liabilities and Net Worth
Gross financial assets, gfa Money, m Other financial assets, ofa Loans Bonds Stocks, etc	Liabilities, l Debts Stocks
Tangible assets, ta Machines Houses, etc.	Net worth, nw

2.2 Flows

Given those balance sheet items, three flows have to be clearly distinguished: Flows that change a units' overall net worth (income, y, and consumption, c), flows that change a unit's *net* financial assets (revenues, r, and expenditures, e) and flows that change a unit's stock of money (payments and receipts), i.e. its cash flow.

When units consume less than their income, they change their net worth and thus save:

(7) $$y - c = \Delta nw = s$$

Since net worth consists of net financial assets and tangible assets, saving is equal to changes in net financial assets, Δnfa, and changes in tangible assets − i.e. investment ($\Delta ta \equiv i$):

(8) $$s = \Delta nfa + i$$

As this equation clearly shows, investment is just a subcategory of saving. Saving does not "lead" to investment, or investment to saving but investment *is* saving and saving *is* investment. This will be discussed in more detail below.

Units can change their net financial assets (save financially) by spending less than their revenues (and dis-save by spending more than their revenues):

(9) $$r - e = \Delta nfa = \Delta m + \Delta ofa - \Delta l$$

Revenues and expenditures are purchases and sales of goods and services, (exports and imports between countries), wages, capital income (dividends and interest) or taxes and transfers (taxes, subsidies etc.).

Finally, receipts and payments change a unit's stock of money:

(10) $$\textit{receipts} - \textit{payments} = \Delta m$$

Equation (9) is a unit's *balance of payments*. It contains its *current account* which is on the left hand side of equation (9), and its *financial account* which is on the right hand side. On the current account, all revenues and expenditures are booked. On the financial account, all financial transactions are booked, i.e. changes in gross financial assets and gross liabilities (see IMF for details on the balance of payment (2009, p. 9)). Receipts and payments are also recorded in the financial account since they change one category of financial assets, i.e. money. Income/revenues/receipts and consumption/expenditures/payments have to be clearly distinguished:

- **Income but no revenue**: a unit produces a good or service. This constitutes income since a new good can be either consumed or invested but no revenue since the production as such does not change a unit's net financial assets.
- **Consumption but no expenditure**: A self-produced good or service that is consumed does not constitute an expenditure. Also, depreciation decreases a unit's net worth but is no expenditure since it does not reduce a unit's net financial assets. This is why depreciation is also called "consumption of fixed capital".
- **Revenue but no receipt / expenditure but no payment**: If a unit sells goods or services but the purchaser can pay later, the seller increases her other financial assets (the direct credit she provides), but not her money stock. The purchaser decreases her net financial assets by decreasing her liabilities, but not her stock of money.
- **Receipt but no revenue / payment but no expenditure**: If a unit buys a financial asset (for instance, a bond), it reduces its money stock (a payment) but increases its stock of other financial assets. *Net* financial assets do not change so that no expenditure takes place. The same applies to the seller of a financial assets who increase her stock of money (a receipt, but no revenue).

In the Anglo-Saxon literature, the terms *receipts* and *revenues* or *expenditures* and *payments* are often used as synonyms and there is no clear distinction of exactly what economic transaction is meant. For instance, this is the case in Bond et al. (2007) in their presentation of the "U.S. Integrated Macroeconomic Accounts". This makes it sometimes hard to clearly distinguish between different economic phenomena − changes in money holdings and changes in net financial assets − which might also explain some of the confusions of loanable funds theory.

Table 2 summarizes the different possible transactions. It is similar to Godley and Lavoie's transaction flow matrix (2012, p. 37) but differentiates more clearly between production and current account transactions. Lines 1) to 3) capture a unit's production, lines 4) to 7) its current account transactions, lines 8) to 10) its financial account transactions, 11) to 14) captures its income and line 15) its saving.

2.3 Saving

The difference between the types of transactions just enumerated has important implications for the different types of saving as well as for the difference between saving and credit.

Generally, equation (8) shows that there are at least three meanings of saving which have always to be clearly kept apart if one wants to avoid confusion (Stützel and Grass, 1988, p. 365):

> a) Consumption is lower than income: Every economic unit increases its *overall* net worth by consuming less than its income. This definition of saving does not say in what *form* a unit increases its net worth.
> b) An increase in tangible assets: When an economic unit increases its net worth by increasing its tangible assets (invests), it saves. Investment is just one subcategory of the broader concept of saving. A unit can invest by either producing itself a new tangible asset or by purchasing one.
> c) An increase in net financial assets: When a unit spends less than it earns (realizes an expenditure surplus), it increases its gross financial assets relative to its gross liabilities.

To make the differences between those different types of saving clear, three examples are given, illustrated in table 2 (case *a)* to *c)*). First, a restraint in consumption relative to income is not necessarily the same thing as a restraint in consumption *expenditures*: When Robinson, being alone on his island, picks (=produces) coconuts, he can eat (=consume) them or plant (=invest) them. The amount of coconuts he picks is his income (case *a)* in table 2, lines 1) to 3) where the superscript *p* means production).

The more he consumes of his given income (=picked coconuts), the less he can invest and thus save. This is basically the world of neo-classical growth models (discussed in more detail in section 5). No

expenditures, revenues, receipts or payments take place even though there is income, consumption, investment and saving.

Second, if households spend less on their consumption goods than they receive in wage revenues – wages *w* times employment emp –, they increase their net financial assets (case *b)* in table 2). Which *form* the increase in net financial assets takes is however a different matter. Households can either hold their higher net financial assets in the form or higher money holdings, higher holdings of other financial assets or they can decrease their liabilities. They do *not* however provide any credit when they spend less than their wage revenues.

Third, when firms buy machines (investment goods) and realize an expenditure surplus of the same amount, they increase their holdings of tangible assets but reduce their net financial assets by the same amount so that they do not save at all since their overall net worth stays the same (case *c)* in table 2).

Table 2: Production and transactions matrix

		a)	*b)*	*c)*	*d)*
Production	1) Consumption goods 2) Investment goods 3) Sum of production 1)+2)	c^P i^P $c^P + i^P$			
Current account	4) Trade in goods and services 5) Wages 6) Interest and dividends 7) Taxes and transfers		$-c$ $+w \times emp$	$-i$	
Financial account	8) Means of payment 9) Other financial assets 10) Liabilities		$+\Delta m$ $+\Delta ofa$ $+\Delta l$	$-\Delta m$	$+\Delta ofa$ $+\Delta l$
Income	11) Δnfa = 4) + 5) + 6) + 7) = 8) + 9) - 10) 12) Consumption 13) Investment (tangible assets + inventories) 14) Sum of income= 11) + 12) + 13)	 c i^P $c^P + i^P$	$w \times emp - c$ c $w \times emp$	$-i$ $+i$	
Saving	15) = 14) - 12) = 11) + 13)	i^P	$w \times emp - c$	0	0

But the provision of money via a credit does not change a unit's net worth. Credit creation is a pure financial transaction that is only counted in the financial account and does not show up in either the production, the income or the current account.

2.4 Credit creation

Lenders that cannot create money exchange money for other financial assets (asset exchange). Lenders that are able to create money extend their balance sheets, i.e. they increase both their *gross* financial assets and *gross* liabilities. In neither of those two cases do lenders (and borrowers) change their *net* financial assets, i.e. save.

When a lender cannot create money himself, he decreases his money holdings and increase his holdings of other financial assets (the loan or the bond):

(11) $$0 = +\Delta ofa - \Delta m$$

A lender who is able to create money himself (like a central or commercial bank) increases both its gross financial assets (the new loan that it creates) and its gross financial liabilities (which is the borrower's deposit):

(12) $$0 = +\Delta ofa - (+\Delta l)$$

This is shown in case *d)* in table 2. A borrower always increases his money holdings (be it central bank money or a deposit at the lending bank) but also his liabilities:

(13) $$0 = +\Delta m - (+\Delta l)$$

A lender does not need to abstain from consumption or any expenditures in order to provide a loan. But he needs to abstain from holding his financial assets in the form of money. If he can create money himself, he does not even have to abstain from holding money.

An important caveat is in order here: it is true that commercial banks can create deposits which are normally accepted as means of payment by non-banks. Thus, a bank creates money. However, once the borrower wants to withdraw his money from the bank, the bank has to pay *central bank money* which it can by definition not produce itself – only the central bank can do that. Commercial bank money – deposits – are thus a promise to pay in the form of central bank money. If that promise is no longer believed by depositors, a bank run is likely to follow and banks are likely to go bankrupt if they cannot get enough central bank money to honor their commitments. Or, as Hyman Minsky (2008, p. 255) nicely put it:

> "[. . .] everyone can create money; the problem is to get it accepted."

3. Aggregate income, saving and finance

Until now, only individual economic units and groups of economic units have been analyzed in order to clearly define and distinguish between the economic phenomena "saving" and "credit". However, loanable funds theory is a theory about the interaction of groups within the aggregate economy and aims at deriving statements valid for aggregate variables like aggregate saving, aggregate investment and aggregate credit.

In this section, accounting relations will be presented that are valid for the aggregate economy. Based on this, the fallacy can be tackled that financial saving adds to some fund out of which investment can be financed. This idea is mainly a fallacy of composition.

3.1 Financial saving and investment

Stützel (1978, p. 20-23) developed a formal schema to analyze the relations between single economic units and the aggregate economy. The aggregate economy, which is composed of all economic units – the sum of all firms, households, governments etc. – can be split into a group, *g*, and its complementary group, *cg*. When a group is defined, its complementary group is the rest of the economy. For instance, the complementary group to all households are all non-households, the complementary group to the domestic economy is the rest of the world, the complementary group to lenders are borrowers etc.

Statements valid for the aggregate economy are *global statements*, statements valid for groups are *partial statements*. Very often, the application of a partial statements to the aggregate economy is a fallacy of composition. The link between global and partial statements are *relational statements*. They show how partial statements for a group depend on its complementary group's behavior. This schema can be easily applied to saving and investment for groups and for the aggregate economy.

We have already shown that any individual economic unit or group of economic units can increase (decrease) its net financial assets by realizing a current account surplus (deficit) (equation (9)). This is the partial statement of financial saving. However, since to every expenditure of some unit or group *j* corresponds a revenue of its complementary group (wages are revenues for workers and expenditures for firms, interest are revenues of debtors and expenditures of creditors etc.), the sum of all expenditures of all *N* economic units is necessarily equal to the sum of all revenues (aggregates are written with capital letters):

(14)

$$0 = \sum_{j=1}^{N} \left(r_j - e_j\right) = R - E$$

Note that this is not true for receipts and payments. If a bank creates new money (new deposits), a borrower increases his stock of money and thus realizes a receipt but the bank does not decreases its stock of money and does consequently not realize a payment (until the moment when the bank has to pay central bank money).

Also, every financial asset fa_k is a financial liability l_k so that both the stock of and change in financial assets and financial liabilities is always equal:

(15)
$$0 = \sum_{k=1}^{K}(fa_k - l_k) = \sum_{k=1}^{K}(\Delta fa_k - \Delta l_k) =$$
$$FA - L = \Delta FA - \Delta L$$

Combining (14) and (15) means that the aggregate economy's net *financial* assets are *always* zero. The aggregate economy cannot save or dis-save in the form of net financial assets (global statement):

(16)
$$0 = \sum_{j=1}^{N}(r_j - e_j) = \sum_{j=1}^{N}\Delta nfa_j = \sum_{k=1}^{K}(\Delta fa_k - \Delta l_k)$$

This means that groups (*g*) can only increase their net financial assets to the extent that their complementary group (*cg*) accepts to decrease theirs (relational statement):

(17)
$$r_g - e_g = -(r_{cg} - e_{cg}) =$$
$$\Delta nfa_g = -\Delta nfa_{cg}$$

For instance, when firms reduce their wage expenditures, workers' wage revenues will necessarily fall by the same amount. When workers react by cutting their consumption expenditures, firm revenues will necessarily fall, etc. Financial saving is *always* a zero-sum game. When groups cut their expenditures, they cut their complementary group's revenues. No fund is in any way increased by anyone's act of financial saving.

But if the aggregate economy cannot save financially, how is aggregate saving possible? The only way the aggregate economy can save is in the form of the *production* of new tangible assets, i.e. investment (global statement):

(18)
$$\sum_{j=1}^{N} s_j \equiv \sum_{j=1}^{N}(\Delta nfa_j + i_j) \equiv$$
$$S \equiv \Delta NFA + I \equiv 0 - I \equiv$$
$$S \equiv I$$

The trade of existing tangible assets cancels out in the aggregate since it means the dis-saving of the investment good's seller and the buyer's saving.

This is just a reformulation of equation (14), according to which the sum of all expenditures is equal to the sum of all revenues:

(19)
$$S - I = \Delta NFA = R - E = 0$$
$$\Rightarrow R = E$$

Equations (18) and (14) hold for a closed economy or for the world economy. Open economies are groups which can save both in the form of higher tangible assets and higher net financial assets by realizing current account surpluses.

Furthermore, equation (18) shows that the opposition between the statements "saving creates investment" – attributed to neoclassical economists – and "investment creates saving" – attributed to Keynesian economists (Palley, 1996; de Carvalho, 2012) – is no opposition at all. Both statements express the tautology that *aggregate* saving *is* investment.

However, the accounting identity $S \equiv I$ says *nothing* about a) *how* investment is financed and b) about the size of any group's financial saving. All of investment could be 100 % internally financed by businesses if they had enough liquidity at hand or it could be 100 % externally financed.

Further, there is no necessary relation between the amount of aggregate investment in a period and the amount of financial saving by any group in the economy. Groups could have a current account surplus approaching infinity, provided that their complementary group has the matching deficit, but investment could be zero or even negative (with depreciation being high enough).

On the other hand, investment could approach infinity but financial saving be zero for each group. Thus, there is no necessary relation between financial saving of any one group and aggregate saving. Aggregate saving always is only as high as aggregate investment is, whatever any group's financial saving or dis-saving amounts to.

3.2 Some illustration

A simple example can illustrate the statements made before. Assume an economy consisting only of a firm and a household. The firm wants to build a machine (= invest, I^p) and consumption goods (C^p). In order to produce, it needs workers (it already holds the materials needed itself)[4] but we assume that the firm does not have the money to pay wages. It thus has to borrow money from a bank.

If the bank creates the money anew, the firm increases its money holdings and its liabilities; the bank increases its loan holdings and its liabilities (= the borrower's deposit). This is illustrated in table 3 in which only units' financial accounts are shown since the process is a pure financial transaction. In the aggregate economy (column *d)*), one can see that the amount of money and the amount of loans has increased by the same amount. Since both the bank and the firm have lengthened their balance sheets but not changed their net worth, no saving has taken place, neither for the bank, the firm nor the aggregate economy.

Table 3: Bank credit

		a) household	*b) firm*	*c) bank*	*d) Aggregate = a) + b) + c)*
Financial account	1) Means of payment 2) Other financial assets 3) Liabilities		$+\Delta m$ $+\Delta l$	 $+\Delta ofa$ $+\Delta l$	$+\Delta m$ $+\Delta ofa$ $2\ x\ \Delta l$
Saving	4) = 1) + 2) – 3)		$+\Delta m - (+\Delta l) = 0$	$+\Delta ofa - (+\Delta l) = 0$	0

The newly built machine adds to the aggregate economy's net worth so that the aggregate economy has saved exactly by the amount of the machine. This would of course also hold if the firm had not borrowed but had held enough money itself and used it to hire workers. Produced but unsold consumption goods would also add to the economy's saving since those are inventories ($I^{inv} = C^p - C$ where *C* are sold consumption goods). Table 4 shows all the relations between the different actors.

[4] If we assume a whole business sector with different firms, materials would be traded within the sector so that expenditures and revenues would only take place within the sector and not affect the aggregate sector's or its complementary group's net financial assets.

Table 4: Group saving and aggregate saving

		a) household	*b) firm*	*c) bank*	*d) Aggregate = a) + b) + c)*
Production	1) Consumption goods		C^p		C^p
	2) Investment goods		I^p		I^p
	3) Sum of production 1)+2)		Y^p		Y^p
Current account	4) Trade in goods and services	$-C$	$+C$		0
	5) Wages	$+w\times emp$	$-w\times emp$		0
	6) Interest and dividends		$-int\times l_{t-1}$	$+int\times l_{t-1}$	0
	7) Taxes and transfers				
Financial account	8) Means of payment	$+\Delta m$	$-\Delta m$		0
	9) Other financial assets				
	10) Liabilities				
Income	11) Δnfa = 4) + 5) + 6) + 7) = 8) + 9) - 10)	$w\times emp - C$	$C - w\times emp - int\times l_{t-1}$	$int\times l_{t-1}$	0
	12) Consumption	C	$I^p + C^p - C = I^p - I^{inv}$		C
	13) Investment (tangible assets + inventories)	$w\times emp$	$I^p + C^p - w\times emp - int\times l_{t-1}$	$int\times l_{t-1}$	$I^p - I^{inv}$
	14) Sum of income= 11) + 12) + 13)				$I^p + C^p$
Saving	15) = 14) - 12) = 11) + 13)	$w\times emp - C$	$I^p + C^p - w\times emp - int\times l_{t-1}$		$I^p - I^{inv}$

In order to illustrate the relation between household financial saving and aggregate saving, we will now discuss three cases of the household's saving behavior and the consequences for the firm sector: first, the household saves financially and buys less from the firm than it had received in wage revenues; second, the household spends the same amount on the firm's products than it had received in wages; and third, the household spends more on the firms' products than it had received in wages (For simplicity, we assume that all of the bank's interest income, $int \times l_{t-1}$, is spent on consumption goods by the bankers, where int is the interest rate and l_{t-1}, is the amount of debts the firm owes from the previous period).

First, if households save financially, the firm's revenues in the period will be lower than its expenditures ($C - w \times emp - int \times l_{t-1} < 0$) so that its net financial assets decrease and its complementary group's - the household's - net financial assets increase. It is immediately clear that household's saving has *reduced* the investing firm's financial means. This is quite in contrast to loanable funds theory. The firm would now have to borrow (and accept a higher debt service), *if* it wanted to maintain its initial liquidity position.

Second, if the household did not save financially ($C - w \times emp - int \times l_{t-1} = 0$), exactly the same amount of revenues would flow into the firm that had previously left it in the form of expenditures. Its liquidity position would *not* change (if we assume that revenues and expenditures are equal to receipts and payments) and the firm would not have to borrow if it wanted to maintain its liquidity.

Third, if the household had dis-saved financially, i.e. spent *more* than it received in revenues from the firm ($C - w \times emp - int \times l_{t-1} > 0$), the firm could even repay a part of its initial loan out of its revenues and would have *more* financial means than it started with.

The lesson from this simple example is clear: household's financial saving *extracts* financial means from the firm and does *not* – as loanable funds theorists claim – add to those means. Since net financial assets are zero in the aggregate, financial saving is a zero-sum game: some group's increase in net financial assets always and necessarily is another group's decrease. No fund is filled or emptied by the act of financial saving.

Also, one can clearly see that the credit provision was the *condition* for aggregate and household saving, not vice versa as loanable funds theorists claim. The loan allowed the firm to produce and to invest and thus to save in the aggregate; it also allowed the firm to realize expenditure surpluses which was the necessary condition for the household's revenue surpluses and thus its financial saving (if it realized any).

Now note that in all three cases - i.e. whatever the amount of household saving is – *aggregate* saving is always exactly equal to the newly produced machine, I^p, plus the unsold consumption goods, $C - C^p = I^{inv}$. This illustrates the point that groups' financial saving and aggregate saving are not necessarily correlated. They are rather likely to be negatively related: the more households save, the less financial means the firm has and the less likely is it to invest. Firms are not likely to invest if nobody is willing to buy the products they build with the new machines.

Note that the example also shows that the claim held by some Keynesians that "investment creates *household* financial saving" (Asimakopulos (1983) seems to hold that view) is not valid. In the second case we discussed, household financial saving was zero but investment equal to the newly produced machine.

3.3 Exogenous money and endogenous credit

In the previous example, we had assumed the bank to create money ex nihilo, i.e. that money was endogenous. Keen (2014) argues that endogenous money theory is indeed the opposite to loanable funds theory (Moore, 1988; Fontana, 2003; Lavoie, 2013). According to him, banks' ability to endogenously create new money is an effective counterargument against the exogenous money view of loanable funds theorists. However, that money is endogenously created by banks is neither a necessary nor sufficient condition for countering loanable funds theory.

The inventor of loanable funds theory, Robertson (1934) as well as his follower Tsiang (1956) explicitly acknowledged the endogeneity of money. They argue that interest rates could fall even when no "new saving" would be forthcoming but money would be *newly*, i.e. endogenously, created.

However, analytically, one does not have to resort to endogenous money to show that loanable funds theory is invalid. Even if the stock of money was fixed ex ante and thus exogenous – M –, credit could be infinite since *credit* is always endogenous, even if money is not. This argument was made in 1898 by Knut Wicksell (1936, p. 60-65). This section presents his argument.

With a fixed money stock, a unit's (or group's) payment is necessarily its complementary group's receipt[5]:

(20) $$\Delta m_g = \Delta m_{cg}$$

When lenders, ln, provide a credit in the form of money, they decrease their money holdings and increase their other financial assets. The corresponding borrowers, br, increase their money holdings and their liabilities by the same amount:

(21) $$\Delta m_{ln,br} = \Delta of a_{ln} = \Delta l_{br}$$

On the other hand, if debtors decreased their liabilities by paying off their creditors, creditors' money holdings would increase and debtors' money holdings decrease. With each credit given or debt paid off, money changes hands and other financial assets as well as liabilities change.

If we sum every transaction s over all transactions T in one period, we see that a given stock of money does not limit the number of total transactions. Total transactions are only limited by the frequency in which the stock of money changes hands:

(22)
$$\sum_{s=1}^{T} \Delta \overline{M_s} = \sum_{s=1}^{T} \Delta of a_s = \sum_{s=1}^{T} \Delta l_s$$

Then, the total stock of other financial assets, OFA_t, and liabilities, L_t, at the beginning of the period, t, changes by the frequency by which money has changed hands:

(23)
$$\sum_{s=1}^{T} \Delta \overline{M_s} = OFA_t = L_t + \Delta L_t =$$
$$OFA_{t+1} = L_{t+1}$$

In principle, when the number of transactions T approaches infinity, the stock of other financial assets (and thus necessarily also of liabilities) could also approach infinity (if debt is not reduced). If debts are reduced, money only has to change hands often enough and all liabilities and thus other financial assets could become zero.

The amount of debts and financial assets at any point does *not* depend on the stock of money, $\overline{M}$, but on the frequency of its changing hands, $\sum_{s=1}^{T} \Delta \overline{M_s}$. Thus, in principle, the amount of money in the economy could approach zero while the amount of debts and financial assets could approach infinity.[6]

[5] Receipts and payments are only equal if the money stock would not change, i.e. if no new money would be created. However, when money is endogenous, there are receipts to which no payments correspond. If a bank creates a new deposit by making a loan, the deposit holder has a receipt (his money holdings increase) but the bank does not decrease its money, i.e. it does not have to make payments.

[6] The same reasoning is also valid (and more frequently discussed in the form of the quantity theory of money) for expenditures. What we have derived so far for financial account transactions can also be applied to all current account transactions. The amount of expenditures (and in consequence also revenues) in a period do not depend on the stock of money but on the frequency at which it changes hands. This is the velocity of money. Expenditures and revenues have just to be added to equation (22):

(24)
$$\sum_{s=1}^{T} \Delta \overline{M_s} = \sum_{s=1}^{T} (\Delta l_s - e_s) = \sum_{s=1}^{T} (\Delta of a_s - r_s)$$

The amount of lending (and thus borrowing) does not depend on the stock of money but on the willingness of those holding money to lend it, i.e. their liquidity preference. There is no inherent limit to the expansion of credit even when the stock of money is fixed.

In this sense, Keynes' assumption of a fixed money stock in his General Theory (1936/1997) can be defended (against, for instance, Moore (1988, p. 171-208)) on analytical grounds because it is the best way to show the fallacies of loanable funds theory (Lautenbach, 1937).

The section's finding does not only have theoretical importance, but also practical implications. The liabilities of pure investment banks, insurance companies, investment funds and other nonbank financial institutions (so called "shadow banks") are normally not used and accepted as means of payment. In that sense, non-bank financial institutions are pure intermediaries. But by transferring money between economic units, they create credit. Since those kinds of non-bank financial institutions have become ever more important in recent decades, stressing the ability of commercial banks to create monetary liabilities is not sufficient to explain modern financial markets.

4. Inconsistent plans and changes in interest rates and aggregate income

It is true, as Ohlin (1937a) remarked in his discussion with Keynes, that one needs behavioural assumptions to understand the dynamics of the economy and that accounting – while necessary – will not be sufficient. One has to make explicit the *ex ante* plans that economic units have and how those plans interact. This will be done here by transforming all the variables defined so far into planned variables. Central to the analysis will be excess supply and demand functions for goods, labor and financial assets. The analysis is inspired by Patinkin (1958) with two important exceptions, namely that it is not assumed that full employment holds or that all debt commitments will be honored. That means that default is possible.

With the help of excess supply and demand functions, the loanable funds hypothesis can be summarized as follows: If households decrease their consumption expenditures, there will be an excess supply for consumption goods. Since households supply their savings on financial markets, an excess demand for other financial assets (bonds, loans) will lead to a decrease in interest rates. This will incite firms to increase their borrowing and investment, leading to an excess demand for investment goods which compensates the initial excess supply of consumption goods.

On labor markets, the excess supply of consumption goods will also lead to an excess supply of workers in the consumption goods industry. This will be compensated by an excess demand for workers in the investment industry so that the loanable funds mechanism not only brings goods but also labor markets into equilibrium.

The next analysis will however show that with perfectly flexible prices, an initial excess supply of consumption goods is not likely to lead to an excess demand for investment goods but quite to the contrary to an excess supply of investment goods – even if interest rates decrease. Further, the analysis will show that interest rates are not determined by excess saving.

4.1 Excess supply on the markets for goods, labor and financial assets

Ex ante plans can obviously differ from ex post identities. We can take all the variables defined so far and transform them into planned transactions. Thus, planned expenditures on goods, services and labor can be transformed into demand functions; planned revenues can be transformed into supply functions.

When the demanded and supplied quantity of some good, q^d, q^s, depends on its price, p_q and the demanded and supplied quantity of labor, emp^d, emp^s, depends on wages, w, planned expenditures and revenues on the goods, services and labor market for a given price and wage can be written:

(25)
$$e^{pl} = p_q q^d + w \times emp^d;$$
$$r^{pl} = p_q q^s + w \times emp^s$$

I will use the standard assumption that demand of labor and goods is decreasing in prices/wages, and supply of labor and goods is increasing in prices/wages. Further, all prices are perfectly flexible.

Combining planned expenditures and revenues on the goods, services and labor markets for the entire economy yields excess supply functions, with P_Q being a price index and W average wages:

(26) $$R^{pl} - E^{pl} = P_Q(Q^s - Q^d) + W(EMP^s - EMP^d)$$

One can distinguish between consumption goods, $C \equiv P_C Q_C$, and investment goods, $I \equiv P_I Q_I$, so that:

(27) $$R^{pl} - E^{pl} = (C^s - C^d) + (I^s - I^d) + W(EMP^s - EMP^d)$$

C^s and I^s roughly correspond to the production of consumption and investment goods. Supply and production do not have to be the same because firms may want to decrease their inventories and then supply more goods in a period than they produce. When they want to increase their inventories, they might supply less of their goods than they produce. If firms or households produce goods for themselves, the demanded and supplied quantities are equal.

Further, economic units also plan to pay and receive capital income when they have liabilities and financial assets. This amount is given by the product of the interest rate, int and the face value, fv, of a financial asset/liability so that one can add to equation (27):

(28) $$R^{pl} - E^{pl} = (C^s - C^d) + (I^s - I^d) + W(EMP^s - EMP^d) + INT \times (FV^{pl}_{FA,t-1} - FV^{pl}_{L,t-1})$$

Turning to other financial assets – plans for financial account transactions –, one can formulate supply and demand plans for other financial assets thus, where P_{OFA} is a price index for financial assets[7]:

(31) $$\Delta OFA^{pl} - \Delta L^{pl} = P_{OFA}(\Delta FV^d - \Delta FV^s)$$

Planned changes of liabilities are contained in the supply function of other financial assets: Since a unit's liability is another unit's financial asset ($ofa \equiv l$), the plan to change liabilities is a plan to increase or decrease the supply of other financial assets. If units want to decrease their liabilities, for instance by paying back their debt or buying back equity, their supply of other financial assets (=liabilities) is negative.

[7] Note that planned *changes* of financial assets and liabilities have been used here. As has been discussed in the introduction, much of the debate between Keynesians and loanable funds theorists has circled around the question whether it is *changes* in credit and money or the *stock* of credit and money which matters for interest rates. But as Patinkin (1958) has rightly shown, the implications for interest rate determination are the same whether stocks or flows are used. Differences between loanable funds theory and liquidity preference theory do not stem from this distinction.
This can be easily shown: in terms of *stocks*, an economic unit can make plans about the allocation of its portfolio, i.e. about how much of its financial assets it wants to hold in the form of money or in the form of other financial assets and how much debt it plans to owe:

(29) $$m^{pl} = -ofa^{pl} + l^{pl}$$

When its *actual* holdings of money, other financial assets or liabilities and its *planned* holdings diverge at time *t*, units supply or demand one of those items in order to hold at $t + 1$ the planned amount of assets and liabilities:

(30) $$m^{pl}_t - m_t = m^{pl}_{t+1} - m_t = \Delta m^{pl}$$

$$ofa^{pl}_t - ofa_t = ofa^{pl}_{t+1} - ofa_t = \Delta ofa^{pl} = p_{fa} \times \Delta fv^{pl}$$

$$l^{pl}_t - l_t = l^{pl}_{t+1} - l_t = \Delta l^{pl} = p_{fa} \times \Delta fv^{pl}$$

If the difference between planned and actual holdings is positive, units demand one of the items (Δm^d; $p_{fa} \times \Delta fv^d$); if it is negative, units supply them (Δm^s; $p_{fa} \times \Delta fv^s$). So, expressing plans to hold a certain composition of a portfolio in terms of stocks (equation (29)) or flows (equation (30)) is equivalent. The distinction between stocks and flows should thus not lead to any difference in analysis between loanable funds theorists and liquidity preference theorists.

An excess supply of other financial assets leads to a decrease of financial asset prices and an *increase* in interest rates. And vice versa, an excess demand for other financial assets leads to an increase in financial prices and a *decrease* in interest rates.

Finally, planned receipts and payments can also be expressed as an excess demand for money:

(32) $$\Delta M^{pl} = \Delta M^{d} - \Delta M^{s}$$

An excess supply (demand) for money can be directed both to financial markets and goods/labor markets. All the excess demand functions in the financial and the current account can then be combined thus:

(33) $$\begin{aligned} &\Delta NFA^{pl} = R^{pl} - E^{pl} = \\ &(C^{s} - C^{d}) + (I^{s} - I^{d}) + W(EMP^{s} - EMP^{d}) + INT \times (FV_{FA,t-1}^{pl} - FV_{L,t-1}^{pl}) = \\ &P_{OFA}(\Delta FV^{d} - \Delta FV^{s}) + \Delta M^{d} - \Delta M^{s} \end{aligned}$$

This equation gives the financial saving plans. Expressing aggregate saving (equation (18)) in terms of plans yields:

(34) $$\begin{aligned} &S^{pl} \equiv \Delta NFA^{pl} + I^{pl} = \\ &\Delta NFA^{pl} \equiv S^{pl} - I^{pl} \end{aligned}$$

I^{pl} is the sum of economic units' plans to change their tangible assets. Excess saving takes place when saving plans are higher than investment plans ($S^{pl} - I^{pl} > 0$) which means that the sum of units' plans to change their net financial assets is higher than zero ($\Delta NFA^{pl} > 0$).

From this framework one can already derive a general conclusion: excess saving *necessarily* leads to an excess supply on the goods markets ($(C^{s} - C^{d}) + (I^{s} - I^{d}) > 0$) and/or labor markets ($W(EMP^{s} - EMP^{d}) > 0$) and/or to default and/or lower than expected dividend payments ($INT \times (FV_{FA,t-1}^{pl} - FV_{L,t-1}^{pl}) > 0$). With excess supplies on goods and labor markets, excess saving tends to be deflationary.

Excess saving does however *not* determine how interest rates change. A decrease in interest rates would necessitate an excess demand of other financial assets, i.e. $P_{OFA}(\Delta FV^{d} - \Delta FV^{s}) > 0$. But excess saving only implies that the sum of excess demand for other financial assets *and* money be higher than zero: $P_{OFA}(\Delta FV^{d} - \Delta FV^{s}) + \Delta M^{d} - \Delta M^{s} > 0$. This expression could be higher than zero with an excess supply of other financial assets, with an excess demand of other financial assets or even when there is equilibrium on financial markets.

In order to state that excess saving always leads to an excess demand for other financial assets and thus a decrease in interest rates, one would have to assume that the excess demand for money is always lower than the excess demand for other financial assets.

This is still somewhat abstract. The next section will use a simple example to show that one cannot generally assume that interest rates fall when there is excess saving.

4.2 Consequences of a saving shock

For an illustration of the principles just explained and the fallacies of loanable funds theory, let us play through the case in which households unexpectedly reduce their consumption expenditures in order to increase their net financial assets (this has also been described by Bibow (2001), Schmidt (2012), Lautenbach (1952) and Keynes (1963)). Let us assume that there are only two sectors, a household sector, hs, and a business sector, bs. Let us further assume that there is initially full employment, no government or

central bank and that prices are perfectly flexible. Thus, the following discussion takes place in as perfect a market economy as one can imagine.

In the first instance, households' higher saving leads to an excess supply on consumption goods since the amount of actually purchased consumption goods C is lower than planned sales, C^s.

The excess supply on goods markets has four consequences:

a) businesses' actual revenues fall short of their planned revenues: $r_{bs}^{pl} - r_{bs} > 0$;

b) their actual receipts (cash flow) fall short of their planned receipts: $\Delta m_{bs}^{pl} - \Delta m_{bs} > 0$;

c) their actual net financial saving is lower than their planned net financial saving: $\Delta nfa_{bs}^{pl} - \Delta nfa_{bs} > 0$;

d) and their actual investment is higher than their planned investment due to the accumulation of higher unplanned inventories (unsold consumption goods): $i - i^{pl} > 0$.

At this point, loanable funds theorists assume that households supply their saved money to businesses via financial markets. This would lead to an excess demand for other financial assets and a decrease in interest rates.

By this, however, they assume away risk. By the lower than expected revenues and cash receipts businesses' risk necessarily increases since businesses now have less money to service their debts:

(35) $$\frac{\Delta m}{int \times fv_{bs,t-1}} < 0$$

If businesses would borrow more in order to compensate their lower liquidity, they would further increase their risk since they would increase their future debt service.

How interest rates change in this situation depends on the relative changes in the demand for other financial assets (loans, bonds) by households and the supply of those assets by firms. Both supply and demand are likely to decrease because lenders are less willing to lend and borrowers less likely to borrow when default risk increases – if debtors have not defaulted right away when they were hit by the revenue shock. How interest rates change depends on the relative decline in supply and demand:

a) **Interest rates increase:** Households that increase their net financial assets by reducing their expenditures are more likely to hold their additional net financial assets in the form of money and *not* lend this money to the units whose default risk has increased. This is not an irrational act of "hoarding", but a rational "flight to quality", i.e. the demand for financial assets with lower default risk. Without a government that issues government bonds, money would be the riskless asset. If the units hit by the revenue shock maintained their initial borrowing plans, there would be an excess *supply* of other financial assets and thus an *increase* in interest rates (Bernanke et al., 1996; Kalecki, 1937; Stiglitz and Greenwald, 2003):

(36) $$P_{fa}(FV^d - FV^s) < 0 \text{ and } |\Delta M^d - \Delta M^s| > \left|P_{fa}(FV^d - FV^s)\right|$$

(b) **Interest rates stay the same:** In order to avoid a higher future debt service, potential borrowers could reduce their borrowing plans once their revenues and cash flows fall. If businesses' supply of liabilities fell by the same amount as households' demand, excess demand for other financial assets would be zero and interest rates would not change at all:

(37) $$P_{fa}(FV^d - FV^s) = 0 \text{ and thus } \Delta M^d - \Delta M^s > 0$$

(c) **Interest rates decrease:** Debtors might even be willing to reduce their debts – to deleverage – in order to decrease their debt service. When the supply of liabilities declines more than lenders' demand, interest rates would fall:

(38) $$P_{fa}(FV^d - FV^s) > 0 \text{ and } \Delta M^d - \Delta M^s < P_{fa}(FV^d - FV^s)$$

While this fall in interest rates is in line with the prediction of loanable funds theorists, it stems from an altogether different factor, namely the decline in the willingness to borrow, not an increase in the willingness to lend.

None of the above discussions of interest rates is affected if banks and endogenous money were introduced. While banks can create deposits (normally accepted as money by non-banks) when they make loans, they are still bankrupt when their debtors cannot repay. This is due to the fact that banks can create deposits but no central bank money (or gold under the gold standard). Since debtors and depositors ask for central bank money at some stage (for instance when they withdraw money from their bank), commercial banks have a default risk.

We have framed the discussion in terms of interest rates. However, more generally, one could also frame the discussion in terms of credit conditions. Even if lenders would not want to change interest rates when risk increases, they could tighten non-interest conditions or ration loans (Stiglitz and Weiss, 1981).

In the face of lower revenues, how are firms *likely* to behave? Producers of consumption goods are likely to cut prices, production, employment and/or wages and spending on new investment goods. They will do so for two reasons: first, with lower actual than expected sales, firms have accumulated unplanned inventories. They will thus need less labor, less production and lower production capacities – if they expect household consumption expenditures not to increase again in the next period. Thus, quite in contrast to loanable funds theory, firms are more likely to decrease their investment spending than to increase it. Second, the higher default risk means that they will try to cut their non-interest expenditures – wages – in order to keep honouring their interest payments and avoid default.

The cut in the wage bill (wages and/or employment) will lead to a fall in household income and an excess supply on labor markets. The initial excess supply on consumption goods markets is then likely to spill over to the labor market. Further, households' initial plan to increase their net financial assets has been thwarted due to the reaction of its complementary group. This is the *paradox of thrift* in which higher desired saving does not lead to higher saving but lower income.

If households maintained their initial saving plans, they would again cut their consumption expenditures when their income has fallen. Even if prices for consumption goods have fallen, the further decrease in consumption expenditures will again lead to an excess supply of consumption goods with the same effects as the initial excess supply.

The cut in consumption goods' producers investment spending will lead to an excess supply of investment goods which leads to a decrease in revenues and an increase in default risk of investment goods producers. They are thus also likely to lay off workers and / or reduce wages, their production etc. The excess supply in consumption goods is thus likely to spill over to investment goods.

As long as each group – households, consumption goods producers and investment goods producers – are cutting their expenditures in the face of revenue falls, production and income are likely to fall, in extremis to zero. Only if some group is willing and able to maintain or expand its own expenditures in the face of falling revenues will production and the sum of revenues be stabilized.

But why should groups maintain cutting their expenditures? Some group could just accept to save less than desired and adjust its plans. But remember that financial saving also means that units decrease their debts, i.e. deleverage. The vicious cycle of falling revenues and expenditures is especially likely when households and/or firms have to reduce their debts in order to avoid default. Households might wish to maintain their consumption expenditures but are not able to because they have to repay debt; the same goes for firms which might want to maintain employment but risk default if they do.

One could argue that creditors who receive interest revenues from debtors might not be forced to reduce their expenditures even if debtors are. However, given that the risk of debtors' default increases

when their revenues decrease, creditors' risk of not receiving their planned interest revenues also increases. In case of default, they might also have to reduce their expenditures.

This case is what Koo has termed a "Balance Sheet Recession" (2008) and what Fisher has termed a "debt deflation" (1933). Both approaches – while they differ in detail (see Koo, pages 180-184) – constitute theories of why financial saving plans become inconsistent due to units' need to deleverage and why the paradox of thrift is especially likely to bite in such a situation.

To sum up, none of the predictions of loanable funds theory is likely to hold: lower consumption expenditures are not likely to be compensated by higher investment expenditures when there is excess saving, but both are likely to decline in tandem; interest rates might move in every direction but do not necessarily have to fall; and it is not likely that additional loans are provided when there is excess saving.

5. The likely root of the loanable funds fallacy: the mis-interpretation of neo-classical growth models

The fallacies loanable funds theory commits might be explainable by the mis-application of some ideas and concepts of neoclassical growth models – especially the Ramsey (1928), Solow (1956) and Diamond (1965) models – to the sphere of money and finance. Those models are routinely taught in contemporary graduate economics classes (Blanchard and Fischer, 1989; Romer, 1996).

The Ramsey and Solow models are models of real investment only. Financial markets, financial assets and financial saving do not play any role in those models. There is only one good which, for simplicity, will be called "corn". Corn has three functions: it can be consumed, invested and used as a means of payment since wages and interest payments are made with it. Full employment is assumed.

Without money and other financial assets, the only way units can save is to increase their tangible assets, i.e. to invest. Given that full employment is assumed, corn production is always at its maximum in each period. If the corn is consumed, it cannot be invested; if it is invested, it cannot be consumed. There is thus a real trade-off between consumption and investment. *Only* under this assumption does it indeed make sense to talk about a limited saving fund which is increased when it is not consumed.

However, this trade-off between consumption and investment is not a *finance* constraint, but a *resource* constraint (see Kornai (1979) for this distinction). The fixed fund is no fund of *money* that can be lent and borrowed, but a fund of *newly produced goods* that can be consumed or invested. It is obvious that such a model cannot be used to analyze any aspect of finance, money and other financial assets since none exist in the model.

An indication that the loanable funds fallacy might come from a confusion of the two meanings of saving (financial saving and increases in tangible assets) is Blanchard's (2008) following text in his introductory macroeconomics textbook:

> "As we grow up, we are told about the virtues of thrift. Those who spend all their income are condemned to end up poor. [. . .] The model we have seen in this chapter, however, tells a different and surprising story. [. . .] As people save more at their initial level of income, they decrease their consumption. But this decreased consumption decreases demand, which decreases production. [. . .] This means that as people attempt to save more, the result is both a decline in output and unchanged saving. This surprising result is known as the paradox of saving (or the paradox of thrift). So should you forget the old wisdom? Should the government tell people to be less thrifty? No. the results of this simple model are of much relevance in the short run. The desire of consumers to save more led to the 1990 to 1991 recession. [. . .] But, as we will see later in this book when we look at the medium run and the long run, other mechanisms come into play over time, and an increase in the saving rate is likely to lead over time to higher saving and higher income" (2008, p. 58).

To show that higher saving leads to higher investment in the medium to long run, Blanchard uses Solow's neoclassical growth model (Blanchard, 2008, p. 206-247). But this is not a matter of the short or the long run,

but whether economic units save in the form of higher *net financial assets* or in the form of higher *tangible assets*, i.e. investment.

Since the problems of different *financial* saving plans are not dealt with in Solow's model, the model cannot be used to make any predictions about economic units' financial saving behavior, its inconsistencies and thus about the paradox of thrift – neither in the short, medium or long run. There is no miraculous way of short-run financial saving somehow being transformed into long run investment in tangible assets. The two are simply quite different phenomena that Blanchard confuses with each other.

In contrast to Solow's and Ramsey's model, Diamond's (1965) full employment corn economy allows units to lend and borrow. However, they do not borrow and lend money but again the one good, corn. They face a triple trade-off: they can eat (=consume), plant (=invest) or lend their corn. Here, consumption are not consumption *expenditures* but the actual eating of the corn. When a unit wants to lend its corn, it can of course not eat it so that it has to restrain from consumption to be able to lend it.

How pervasive this approach is, is shown by Eggertsson and Krugman (2012) who add some features of Irving Fisher's, Richard Koo's and Hyman Minsky's theories to a basic neoclassical model. Again, no money but goods are borrowed and lent. Naturally, potential lenders have to save some of their goods before they can lend them to borrowers. But since in the real world money is normally not eaten or planted and keeps circulating in the economy when it is spent or lent, those models cannot be any guide for the analysis of a monetary economy. Specifically, what is true in a one good economy – units have to consume less to lend and invest more – is fundamentally wrong in a monetary economy.

Some implications and concluding remarks

The paper has shown that saving does not finance investment. No saving and abstention of consumption is needed for any lending to take place since lending and borrowing money are pure financial transactions that only affect *gross* financial assets and liabilities.

This result has both important analytical as well as political implications. For instance, Bernanke (2005) argued in his "Global Saving Glut" hypothesis that south-east Asian and oil-exporting current account surplus countries financed US housing investment, i.e. that their increase in *net* financial assets was equal to lending money to the US.

This thesis is again likely to stem from a mis-application of the corn economy to the real world: in a corn economy, China would have to abstain from eating (=from consumption) and planting (=investing) corn in order to lend and export more corn to the US than it borrowed and imported (=net lending = current account surplus) from the US. If it lent corn to the US in the hope of receiving more corn in the future (perhaps due to the US' better technology), but then Americans ate and did not plant the corn, the US would default on its corn payment commitments and the Chinese would suffer.

Reality was different, however. The US did not borrow corn, but money, and more specifically *US dollars* from foreigners. But the US also did not mainly borrow from countries that realized current account surpluses vis-à-vis the US, but from European banks, especially from Germany, France, the UK and Switzerland (Acharya and Schnabl, 2010; Shin, 2012; Borio and Disyatat, 2011). They provided the *gross* means which are not visible in current account *balances*. Acharya and Schnabl (2010) show that there is no systematic relation between a country's current account surplus vis-à-vis the US and its banks' lending to the US.

Since European banks cannot produce dollars, every dollar they lent had to be borrowed beforehand – from Americans. The main lenders to European Banks were US money market mutual funds (Shin, 2012). Since European banks borrowed short and lent long in dollars, they depended on steady refinancing from the US. When the financial crisis hit and mutual funds cut their funding of European banks, the Federal Reserve stepped in and provided the needed short-term dollar funding to European banks (Acharya and Schnabl, 2010; Shin, 2012). Ultimately, the US borrowed from and lent to itself. No "Global/Asian Saving Glut" financed US investment but American and European banks did. It is somewhat strange that Bernanke as the former head of the ultimate dollar producer – the Federal Reserve – believes the US to depend on foreign financing when its liabilities are denominated in US dollars.

Further, China's current account surpluses vis-à-vis the US exist because the US buys more from China than China from the US. The US current account deficits *allow* China to increase its income and saving beyond its own investment. But the US as a whole does not have to borrow money from China or anybody else because China – and many other countries – accept the dollar as a means of payment (McKinnon, 2001). Americans do not need Yuan to build their houses, but dollars.

Because believers in the "Global Saving Glut" variant of loanable funds theory identify lending with current account surpluses (financial saving), they could not anticipate that *European* banks were the first banks to be hit by problematic mortgages, while Chinese banks were hardly hit at all. They could also not anticipate that the Federal Reserve would be the ultimate lender to US households in a financial crisis – not the Chinese government or oil exporters on which loanable funds theorists believe the US to depend on.

Similarly, Sinn's (2010) thesis that Germans exported their savings and could not use them at home stems from the wrong application of the corn economy to the real world. While German banks were indeed among the largest creditors of today's crisis countries, the second biggest creditors were French banks although France had a roughly balanced current account balance since 1999 and then a deficit (Waysand et al., 2010; Lindner, 2012).

The banks in today's crisis countries created deposits anew when they lent to their governments, households and companies. Those in turn used the money to buy German goods which increased German income. This was a much needed stimulus for the otherwise stagnant German economy (Lindner, 2013). Since banks do not accept each others' liabilities as means of payments, the banks in today's crisis countries had to borrow central bank money on interbank markets to transfer its residents' payments to German banks. French and German banks were the biggest interbank lenders (Lindner, 2012).

The money lent by German banks to foreign banks most often came back to Germany in the form of sales revenues for German exporters. This money was again relent on interbank markets and gave foreign banks the means to increase their credit supply so that foreigners could purchase even more German goods etc. Sinn argues that Germany lost liquidity when it increased its current account surpluses. The exact opposite was (and is) however the case (Horn and Lindner, 2011; Bindseil and K¨onig, 2012).

As far as policy is concerned, the doctrine of austerity is related to the doctrine that government saving has to be reduced for private investment to increase. However, this is not likely to be the case. When governments reduce their expenditures, firms' revenues are necessarily reduced and their default risk increases. This will lead to less investment and less borrowing. Loanable funds theorists hold – contrary to all real-world experience – the reverse belief.

Overall, the present paper has not only criticized a core belief held by many economists but also tried to sketch an alternative on the basis of common accounting rules and widely accepted behavioral assumptions. The hope is that such a kind of reasoning will be more widely used in future economic research and inform economic policy.

References

Acharya, Viral V. and Philipp Schnabl, "Do Global Banks Spread Global Imbalances? Asset-Backed Commercial paper during the Financial Crisis of 2007-09," *IMF Economic Review*, 2010, *58* (1), 37–73.

Asimakopulos, Athanasios, "Kalecki and Keynes on Finance, Investment, and Saving," *Cambridge Journal of Economics*, 1983, *7* (2/4), 221–233.

Ball, Laurence and Gregory N. Mankiw, "What Do Budget Deficits Do?," in "Budget Deficits and Debt: Issues and Options," Kansas City: Federal Reserve Bank of Kansas City, 1995, pp. 95–119.

Bernanke, Ben, "The Global Saving Glut and the U.S. Current Account Deficit," in "in" TheFederal Reserve Board 2005.

Bernanke, Ben, Mark Gertler, and Simon Gilchrist, "The Financial Accelerator and the Flight to Quality," *The Review of Economics and Statistics*, 1996, *78* (1), 1–15.

Bibow, Jörg, "The Loanable Funds Fallacy: Exercises in the Analysis of Disequilibrium," *Cambridge Journal of Economics*, 2001, *25* (5), 591–616.

Bindseil, Ulrich and Philipp J. König, "TARGET2 and the European Sovereign Debt Crisis," *Kredit und Kapital*, 2012, *45* (2), 135–174.

Blanchard, Olivier, "What Do We Know about Macroeconomics that Fisher and Wicksell did Not?," *The Quarterly Journal of Economics*, 2000, *4* (115), 1375–1409.,

Blanchard, Olivier, *Macroeconomics (5th edition)*, New Jersey: Prentice Hall, 2008.

Blanchard, Olivier, and Stanley Fischer, *Lectures on Macroeconomics*, Cambdridge, Massachusetts: The MIT Press, 1989.

Bond, Charlotte Anne, Teran Martin, Susan Hume McIntosh, and Charles Ian Mead, "Integrated Macroeconomic Accounts for the United States," *Bureau of Economic Analysis, Survey of Current Business*, 2007, (February), 14–31.

Borio, Claudio and Piti Disyatat, "Global Imbalances and the Financial Crisis: Link or no Link?," *BIS Working Paper*, 2011, (346).

Brümmerhoff, Dieter, *Volkswirtschaftliche Gesamtrechnung, 8. Auflage*, München: Oldenbourg Verlag, 2007.

Cochrane, John H., "Fiscal Stimulus, Fiscal Inflation, or Fiscal Fallacies?," *Working Paper*, 2009.

de Carvalho, Fernando J. Cardim, "Aggregate savings, finance and investment," *Intervention*, 2012, *9* (2), 197–213.

Diamond, Peter, "National Debt in a Neoclassical Growth Model," *The American Economic Review*, 1965, *55* (5), 1126–1150.

Eggertsson, Gauti B. and Paul Krugman, "Debt, Deleveraging, and the Liquidity Trap: A Fisher-Minsky-Koo Approach," *The Quarterly Journal of Economics*, 2012, *127* (3), 1469–1513.

European Communitites, "European System of Accounts. ESA 1995," *Official Journal of the European Communities*, 1996, (30.11.96), 5–469.

Fama, Eugene F., "Bailouts and Stimulus Plans," *Fama/French Forum*, 2009.

Fisher, Irving, "The Debt-Deflation Theory of Great Depressions," *Econometrica*, 1933, *1* (4), 337–357.

Fontana, Giuseppe, "Post Keynesian Approaches to Endogenous Money: A Time Framework Explanation," *Review of Political Economy*, 2003, *15* (3), 291–314.

Garvey, George, "Keynes and the Economic Activists of Pre-Hitler Germany," *Journal of Political Economy*, 1975, *83* (2), 397–405.

Gestrich, Hans, "Kredittheorie und Wirklichkeit," *Weltwirtschaftliches Archiv*, 1940, *52* (2), 277–300.

Gestrich, Hans, *Kredit und Sparen*, Helmut Küpper, 1947.

Godley, Wynne and Marc Lavoie, *Monetary Economics. An Integrated Approach to Credit, Money, Income, Production and Wealth. Second Edition*, Palgrave Macmillan, 2012.

Hayes, M. G., "The Loanable Funds Fallacy: Saving, Finance and Equilibrium," *Cambridge Journal of Economics*, 2010, *34* (4), 807–820.

Hicks, J. R., *Value and Capital*, Oxford: Clarendon Press, 1939.

Horn, Gustav and Fabian Lindner, "Kein Kapitalabfluss aus Deutschland," *IMK Policy Brief*, 2011, (May).

IMF, *Balance of Payments and International Investment Position Manual*, Washington D.C.: IMF, 2009.

Kaiser, Dirk, *Treasury Management. Betriebswirtschaftliche Grundlagen der Finanzierung und Investition*, Gabler, 2008.

Kalecki, Michal, "The Principle of Increasing Risk," *Economica*, 1937, *5* (November), 441–447.

Keen, Steve, "Modeling Financial Instability," *Working Paper*, 2014.

Keynes, John Maynard, *The General Theory of Employment, Interest and Money*, New York: Prometheus Books, 1936/1997.

Keynes, John Maynard, "Alternative Theories of the Rate of Interest," *The Economic Journal*, 1937, *47* (186), 241–252.

Keynes, John Maynard, "The Great Slump of 1930," in "Essays in Persuasion," W.W. Norton, 1963, pp. 135-147.

Kindleberger, Charles, "The Manichaean Character of Economics," *Challenge*, 1999, *42* (5), 21–28.

Koo, Richard C., *The Holy Grail of Macroeconomics. Lessons from Japan's Great Recession*, Hoboken, New Jersey: Wiley & Son, 2008.

Kornai, Janos, "Resource-Constrained versus Demand-Constrained Systems," *Econometrica*, 1979, *47* (4), 801–819.

Krugman, Paul and Robin Wells, *Economics. Second Edition*, Worth Publishers, 2009.

Lautenbach, Wilhelm, "Zur Zinstheorie von John Maynard Keynes," *Weltwirtschaftliches Archiv*, 1937, *45*, 493–525.

Lautenbach, Wilhelm, *Zins, Kredit und Produktion*, T¨ubingen: J.C.B Mohr (Paul Siebeck), 1952.

Lavoie, Marc, *Post-Keynesian Monetary Economics, Godley-Like*, Oxford University Press,

Lequiller, Francois and Derek Blades, *Understanding National Accounts*, Paris: OECD, 2006.

Lerner, Abba P., "Interest Theory - Supply and Demand for Loans or Supply and Demand for Cash," *The Review of Economics and Statistics*, 1944, *26* (2), 88–91.

Lindner, Fabian, "Banken treiben Eurokrise," *IMK Report*, 2012, (82).

Lindner, Fabian,, "Can Germany be an Example for the Crisis Countries?," *Economia & Lavoro*, 2013, *47* (3), 151–162.

Mankiw, Gregory N., *Macroeconomics, Third Edition*, New York: Worth Publishers, 1997.

McKinnon, Ronald I., "The International Dollar Standard and the Sustainability of the U.S. Current Account Deficit," *Brookings Paper on Economic Activities*, 2001, (1), 227–239.

Minsky, Hyman, *Stabilizing an Unstable Economy*, New York: McGraw Hill, 2008.

Möller, Peter, Bernd Hüfner, and Holger Ketteniß, *Internes Rechnungswesen*, Heidelberg: Springer, 2011.

Moore, Basil J., *Horizontalists and Verticalists. The Macroeconomics of Credit Money*, Cambridge: Cambridge University Press, 1988.

Naples, Michele I. and Nahid Aslanbeigui, "What does Determine the Profit Rate? The Neoclassical Theories Presented in Introductory Textbooks," *Cambridge Journal of Economics*, 1996, *20* (1), 53–71.

Ohlin, Bertil, "Alternative Theories of the Rate of Interest: Three Rejoinders," *The Economic Journal*, 1937, *47* (187), 423–427.

Ohlin, Bertil, "Some Notes on the Stockholm Theory of Savings and Investment I," *The Economic Journal*, 1937, *47* (185), 53–69.

Ohlin, Bertil, "Some Notes on the Stockholm Theory of Savings and Investment II," *The Economic Journal*, 1937, *47* (186), 221–240.

Palley, Thomas I., "The Saving-Investment Nexus: Why it Matters and How it Works," *CEPA Working Paper*, 1996, (1).

Patinkin, Don, "Liquidity Preference and Loanable Funds," *Economica*, 1958, *25* (100), 300–318.

Patinkin, Don, *Money, Interest, and Prices*, New York: Harper International, 1969.

Ramsey, F. P., "A Mathematical Theory of Saving," *The Economic Journal*, 1928, *38* (152), 543–559.

Robertson, Dennis H., "Industrial Fluctuations and the Natural Rate of Interest," *The Economic Journal*, 1934, *44* (176), 650–656.

Robertson, Dennis H., "Some Notes on Mr. Keynes' General Theory of Employment," *The Quarterly Journal of Economics*, 1936, *51* (1), 168–191.

Romer, David, *Advanced Macroeconomics*, McGraw Hill, 1996.

Santos, Claudio H. Dos and Antonio C. Macedo e Silva, "Revising "New Cambridge": The Three Financial Balances in a General Stock-Flow Consistent Applied Modelling Strategy," *Levy Institute Working Paper*, 2010, (594).

Schmidt, Johannes, "Die Bedeutung der Saldenmechanik für die makroökonomische Theoriebildung," in Harald Hagemann and Hagen Krämer, eds., *Keynes 2.0 - Perspektiven einer modernen keynesianischen Wirtschaftstheorie und Wirtschaftspolitik*, Marburg: Metropolis, 2009, pp. 111–148.

Schmidt, Johannes, "Sparen - Fluch oder Segen? Anmerkungen zu einem alten Problem aus Sicht der Saldenmechanik," in Martin Held, Gisela Kubon-Gilke, and Richard Sturn, eds., *Normative und institutionelle Grundfragen der Ökonomik*, Marburg: Metropolis, 2012, pp. 61–85.

Shin, Hyun Song, "Global Banking Glut and Loan Risk Premium," *IMF Economic Review*, 2012, *60* (2), 155–192.

Sinn, Hans-Werner, "Euro-Krise," *Ifo-Schnelldienst*, 2010, *63* (10).

Snippe, Jan, "Loanable Funds Theory Versus Liquidity Preference Theory," *De Economist*, 1985, *133* (2), 129–150.

Solow, Robert, "A Contribution to the Theory of Economic Growth," *The Quarterly Journal of Economics*, 1956, *70* (1), 65–94.

Stiglitz, Joseph E. and Andrew Weiss, "Credit Rationing in Markets with Imperfect Information," *American Economic Review*, 1981, *71* (3), 393–410.

Stiglitz, Joseph E and Bruce C. Greenwald, *Towards a New Paradigm in Monetary Economics* Raffaele Mattioli Lectures, Cambridge: Cambridge University Press, 2003.

Stützel, Wolfgang, *Volkswirtschaftliche Saldenmechanik*, Tübingen: J.C.B. Mohr (Paul Siebeck), 1978.

Stützel, Wolfgang, *Paradoxa der Konkurrenz- und Kreditwirtschaft*, Aalen: Scientia Verlag, 1979.

Stützel, Wolfgang and Wolf-Dieter Grass, *Volkswirtschaftslehre. Eine Einführung für Fachfremde, 2. Auflage*, München: Franz Vahlen, 1988.

Tsiang, S.C., "Liquidity Preference and Loanable Funds Theories, Multiplier and Velocity Analysis," *The American Economic Review*, 1956, *46* (4), 539–564.

Waysand, Claire, Kevin Ross, and John de Guzman, "European Financial Linkages: A New Look at Imbalances," *IMF Working Paper*, 2010, (295).

Wicksell, Knut, *Interest and Prices*, London: MacMillan, 1936.

Woodford, Michael, "Financial Intermediation and Macroeconomic Analysis," *Journal of Economic Perspectives*, 2010, *24* (4), 21–44.

Roots and Fruits of Democracy: Natural Resources, Income Distribution and Social Violence

Matteo Cervellati, Piergiuseppe Fortunato and Uwe Sunde[1]
University of Bologna and IZA Bonn / United Nations Conference on Trade and Development / and University of Munich, IZA Bonn, CEPR London

Abstract

This paper proposes the argument that natural resource abundance and large economic inequality, by shaping the interests of different social groups, are key factors for the determination of the transition scenario from authoritarianism to democracy. In turn, the transition scenario, and in particular the level of violence during democratization, determines the success or failure of a democratic reform. We analyze the historical experience of countries that democratized during the "third wave" of democratization in order to shed some light on the determinants and consequences of current and future democratic transitions.

1. Introduction

In the last decades democracy extended beyond its traditional boundaries in the Western hemisphere and spread around the globe. Between 1973 and 2003 during the so-called “third wave” of democratization more than sixty countries all over the world adopted democratic institutions either after the collapse of larger states (e.g., former USSR), or after the end of dictatorships and authoritarian dominant-party systems (e.g., South Korea). By the beginning of the 21st century, democracies existed in every major world region except for one, North Africa and the Middle East, where not even a single democratic regime was in place. The recent events in the region, however, might put an end to this “Arab anomaly”, paving the way towards a full globalization of democracy.

The mass movement claiming political enfranchisement that ousted President Zine al-Abidine Ben Ali in Tunisia triggered popular pressures in neighboring countries, led to the overthrow of President Hosni Mubarak in Egypt and of Colonel Muammar Gaddafi in Libya, a violent civil conflict Syria, and protests in other Arab countries including Algeria, Bahrain, and Yemen. This new Arab awakening in reference to the spread of democratic ideals in these countries has raised many hopes of a fast political and economic development in a region that lags behind Western economies despite its remarkable abundance of natural resources. It is already apparent, however, that the (attempted) regime transitions are taking very different paths in the different countries. While in some cases the mass movements have succeeded in opening up the possibility of rapid and relatively peaceful regime changes, in others the movement towards democracy is faced with stronger resistance by (parts of) the ruling elites and appears longer, more uncertain, and stained with blood. Will the emergence of democracy be (equally) beneficial to all countries? Will the different transition modes make a difference for the features of the emerging democracies?

The experiences of countries that democratized in the last decades may be interesting not only in an historical perspective but could potentially be insightful for the prospects of the current democratic movements in the Arab world and elsewhere. The data suggest that not all democracies have succeeded in granting the protection of basic individual rights and in fostering the implementation of well-crafted economic policies. According to the Freedom House, out of the sixty-seven countries which experienced a democratic transition between 1973 and 2003, only about one half are classified nowadays as “Free” (countries where there is a compliance with a wide array of political rights and civil liberties). All the others are classified as

[1] Matteo Cervellati, University of Bologna and IZA Bonn (m.cervellati@unibo.it); Piergiuseppe Fortunato (corresponding author), United Nations Conference on Trade and Development – UNCTAD (Piergiuseppe.Fortunato@unctad.org) ; Uwe Sunde, University of Munich, IZA Bonn, CEPR London (uwe.sunde@lmu.de).

either "Partly Free" (countries with some significant limitations on these rights and liberties) or "Not Free" (countries where basic political rights and civil liberties are widely and systematically denied) as summarized in Table 1.[2]

Table 1. Civil Liberties after Democratic Transition

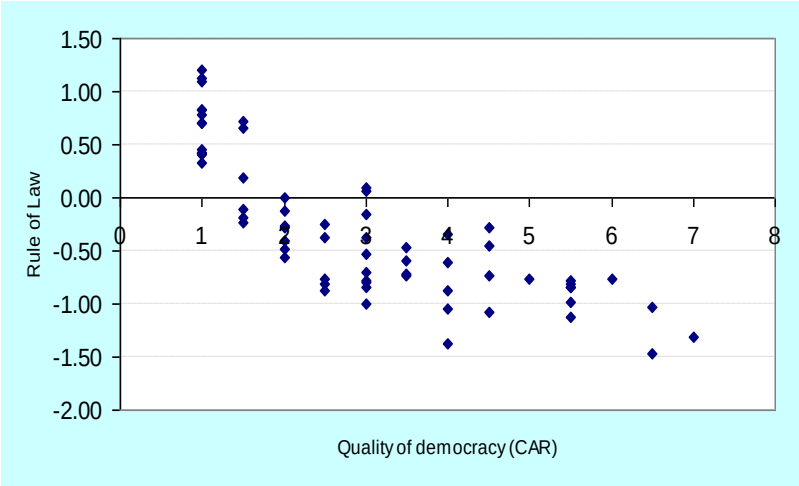
	Free	Partly Free	Not Free
Pre Transition	0	31	36
2005	35	23	9

Source: Freedom House (2005).

Furthermore, while some countries were able to implement stable democratic institutions, democracy has been unstable in others, and some form of autocratic rule has been restored. The data further suggests the existence of a relatively weak correlation between the prevalence and quality of democratic liberties and the implementation of growth-enhancing economic policies and institutions. The quality of rule of law is, on average, higher in democracies. Nonetheless, countries with very similar democratic liberties exhibit large differences in the quality of rule of law. Figure 1 plots, for instance, the relationship between quality of the rule of law and the combined average rating of political rights and civil liberties indices of the Freedom House (so-called CAR index) for the sixty seven countries that democratized during the "third wave". The CAR index takes values on a scale from 1 to 7, with 1 representing the best practice and 7 the worst. Furthermore, whether democratization or an improvement in democratic rights is indeed causal for improving economic policies and institutions in different countries is still debated in the literature. The empirical evidence, surveyed in more detail below, seems to suggest that the interaction between rule of law and political regime runs both ways, implying non-trivial bi-directional feedback mechanisms that, to a large extent, still remain to be explored.

The democracies that emerged in the last years do not only differ in terms of outcomes (the *fruits* of the change of political regime) but have also very different origins (the *roots* of democratization). Historically, authoritarian regimes have collapsed and new democracies emerged under very different scenarios. In some cases, the transition is initiated from the bottom up: the disempowered part of the population forces a change of regime opposed by the authoritarian power-holders threatening revolution and social unrest. In fact, the transition to democracy has occurred in very different ways, involving different levels of conflicts across countries, being essentially peaceful in some case and very violent in others. In other cases, the process is top-down or, even if it initiated from below, essentially consensual and accompanied by the Elite.

Figure 1. Quality of Democracy and Rule of Law. Source: Freedom House (2005) and World Bank (2005).



[2] Table A1 in the Appendix contains the corresponding raw data on transition years and civil liberties.

This paper proposes the view that abundance of natural resources and inequality in the distribution of the generated rents shape the interests and relative fighting power of the different groups and, in turn, determine the type of transition from authoritarianism to democracy. At the same time, the transition scenario may have a significant impact on the success or failure of a democratic reform. The goal of the analysis is therefore to study the link between the contingencies of democratization and the features of the emerging democracies or, in other words, to investigate whether different paths (to democracy) are relevant for the economic and political outcomes or whether the mere emergence of democracy is all that matters. The idea is to look into the historical experience of the countries that democratized in the last decades in order to shed some light on the determinants and consequences of current and future democratic transitions.

The paper is structured as follows. Section 2 provides a discussion of the related literature, focusing on democratization processes and on the political economy of the resource course. Section 3 discusses the transition from authoritarian regimes to democracies in the light of a simple "political game" between different social groups and summarizes the testable hypothesis on determinants and consequences of the different scenarios of democratic transition. Section 4 investigates the empirical validity of these hypotheses in light of the available data from the third wave of democratization. Finally, Section 5 offers some concluding remarks.

2. The literature

This paper contributes to two important strands of the literature. The vast literature on democratization and democratic consolidation, which has been developed since the 1950s by political scientists, sociologists and (more recently) economists, and the literature on the natural resource curse, which attempts to account for the observed negative correlation between natural resource abundance and growth performance. In particular, this article relates to the more recent branch of the literature focusing on the political-economy characteristics of resource-rich countries as a source of the "paradox of the plenty" (Karl, 1997).

Democratization. The idea that growth in income and human capital causes institutional improvement, is the cornerstone of the influential *modernization hypothesis* most closely associated with the work of Seymour Martin Lipset (1959), and also reproduced in many central works on democracy (e.g., Dahl, 1971). Lipset believed that educated people living in economically developed countries with a rather equal distribution of wealth are more likely to resolve their conflicts of interest through negotiations and voting than thorough violent disputes. Education is needed for courts to operate and to empower citizens to engage with government institutions. According to this view, countries differ in their stock of human and social capital, and institutional outcomes depend on a great extent on these endowments. Huntington (1991) argued that the third wave of democratization was indeed facilitated by the "high levels of economic well-being which led to more widespread literacy, education and urbanization, a larger middle class, and the development of values and attitudes supportive of democracy".

The focus on the consequences of modernization as a main determinant of democratization, was challenged by the work of Moore (1966), and more recently by Luebbert (1991), who emphasized the existence of different "paths to the modern of world". According to this view, the class structure and the relative strength of the bourgeoisie ultimately determine the political consequences of a modernization process. In other words, only in those countries characterized by a relatively strong middle class economic modernization paves the way to democratic institutions. Not only structural economic change can lead to political transformations: democratization may also arise from shocks that hit the society. For instance Haggard and Kaufman (1995) suggest that economic crisis may accelerate the democratization process by extending the discontent against authoritarian regimes.

An alternative approach studying the democratic transition sees democratization as a concession by part of authoritarian rulers in order to raise taxation. The more elastic is the tax base, the more difficult it is to raise taxes without agreement, and therefore the higher will be the likelihood of democratic concessions. As a consequence, democratization is less likely to occur in agrarian and natural resource based economies where production factors are easily taxed than in economies where the production structure is centered on (more mobile) human and physical capital. Arguments along this line are proposed in contributions by Bates

(1991), Rogowski (1998) and Tilly (2004), among others.

Apart from the transition to democracy, scholars have also extensively studied the problem of *democratic consolidation* focusing on the challenge of making new democracies secure, of extending their survival beyond the short term, and of making them immune against the threat of authoritarian regression. Linz and Stephan (1996) stress the importance of the regime in place before the transition, distinguishing, in particular, among four types of nondemocratic regimes: authoritarian, totalitarian, post-totalitarian and sultanistic. Differences in the basic characteristics of the regime in place (ideology, degree of pluralism, forms of mobilization and type of leadership) crucially affect the range of paths of transition and in turn the consolidation of the new institutions. For example *pacted* transitions, which take place through extensive negotiations about the character and sequence of political change, constitute one of the paths from authoritarianism, but appear unlikely to emerge from totalitarianism.

Putnam (2002) looks at the level of social capital (mainly trust and cooperation) whose primary engine is the civil society. In his view, only those countries that are characterized by an active and organized civil society are able to consolidate the democratic institutions after the transition. Departing from this analysis, Hibnik (2005) builds a challenging case in favor of an alternative view: social capital is not generated from below, in grassroots organizations and voluntary associations, but is instead the product of leadership and institution building from above, by elites in the state and political parties.

Finally, as mentioned above, the last decade has witnessed a reviving interest in the issues of democratization and democratic consolidation also within the field of economics. Acemoglu and Robinson (2006a) argue that the revolutionary threat faced by the elite and its inability to commit to policy reform are the principal reasons leading to transitions to democracy. In particular, in the absence of the possibility of credible commitments to future fiscal redistribution, political elites are forced to release power once the opposition is able to organize and mount a revolution. Other authors suggest that under certain circumstances the elite may find it profitable to change the regime, indicating the potential for top-down democratization. Bourguignon and Verdier (2000), for example, argue that democracies may represent a better environment for the provision of public education, which is needed to sustain the process of economic development. According to Lizzeri and Persico (2004), the top-down democratic extension in Britain's age of reform facilitated the implementation of policies that benefited the community at large.

The (political economy of) natural resource curse. The idea that political incentives are key to understanding the resource curse has been explored by political scientists and economists alike. Within political science, many authors focused on specific case studies; Karl (1997) and Ross (2001) analyzed, respectively, the dismal effects of oil in Venezuela and timber in several South-East Asian countries. The economic literature, instead, is more theoretically oriented and proposes various mechanisms consistent with the prediction of a lower income associated with abundance of natural resources.

A first wave of economic models studying how the abundance of resources affects policy making by shaping the incentive of politicians consisted of rent-seeking models. These models show how natural resources may foster rent-seeking behavior which has a destructive effect on normal productive investment and hence growth. In general, as long as rent-seeking represents a dead weight loss, anything that encourages rent-seeking will lower steady state income and therefore growth along the transition to steady state. The case studies in Gelb (1988) and Auty (1990) lend support to this hypothesis.

More recently the literature has focused attention on incumbency distortions. Acemoglu and Robinson (2006b) propose a model where underdevelopment results from political elites blocking productive innovations because of a “political replacement effect”. In their model, innovations erode elites’ incumbency advantage, increasing the likelihood that they will be replaced. Fearing replacement, political elites are unwilling to initiate change, and may even deliberately slow down economic development. Such behavior is more likely to arise when the rents from maintaining power are high, such as where public income is derived from natural resources. In a related contribution, Cervellati and Fortunato (2007) study an economy where reforms are non-neutral and inequality in the distribution of resources generates vested interest and slows down economic development.

Other authors suggest that natural resource abundance favor an over-expansion of the public

sector. Robinson et al. (2006) show that with more (or more valuable) resources, the future utility of having political power will increase. An incumbent will therefore tend to employ people in the public sector in order to raise the probability of remaining in power. Similarly, Robinson and Torvik (2009) develop a political economy model of soft budget constraints, where the political desirability of soft budgets increases in resource wealth. Surveying the literature, van der Ploeg (2011) concludes that in non-democratic regimes resource abundance creates corruption and leads to persistence of bad institutions and lack of rule of law.

This paper complements the two broad strands of literature reviewed in this section by suggesting that an unequal distribution of (rents from) natural resources may lead to institutional development traps and slow development in association with the failure of democracy. The paper, therefore, (i) offers a contribution to the literature on natural resource course by proposing a political economy explanation of the course based on the origins of the state rather than policy formation, and (ii) add to the democratization literature a detailed investigation of the effects of different (structurally determined) types of regime transitions.

3. Determinants and consequences of the modes of democratization

3.1 Democratization as a political game

To make sense of diverging democratization experiences and of the existence of many democracies that differ widely in their structure and performance, it may be useful to think of the emergence of democracy as resulting from a *political game* played by different social groups receiving different benefits under democratic and non-democratic institutions.

To lay down the trade-off faced by the different groups let us think of the government as a body with two crucial responsibilities: the distribution of the benefits of the production of income (say through a fiscal redistribution scheme) and the provision of productive public goods (say property rights protection or, more in general, a well-functioning rule of law). Think of these two functions as being linked by a trade-off: the larger is the re-distribution of income among different groups the lower is the efficiency in the protection of property rights and the rule of law. One can think at many reasons behind the existence of such a trade-off. Firstly, both redistribution and provision of public goods are costly and require the capacity to collect taxes from the population. The larger is the level of taxation the larger are the distortions on economic activity. Therefore larger redistribution may come at the cost of fewer resources available for providing public goods. Next, in line with an argument discussed above, extracting resources from the population by means of taxation and implementing a good rule of law requires, or at least is facilitated, by the active cooperation of the citizens.[3] If the population (or a part of it) does not recognize the authority of the government, it may react by, e.g., hiding taxable income and trying to circumvent the legal system. An implication of this view, which is very relevant for the purposes of this study, is that the effectiveness of the action of the government is likely to shrink if the population (or part of it) finds the action of the government in place to be not legitimate.

Let us further assume that the main difference between a democratic and a non-democratic regime lies in whether the state apparatus (the government) is under the control of a minority of the population (e.g., an oligarchy or an autocracy) or a majority of the population (e.g., a democratic regime holding free and contested elections). An authoritarian regime is characterized by the existence of substantial constraints on the participation in the political decision process. Constraints are present in a number of different institutional arrangements. In some cases, as it was the case in the oligarchies of the 18th and 19th century, the franchise is officially restricted; alternatively a dictator or a military *junta* govern the country without the authorization of an electoral process; finally, even when elections are officially held, in many countries the political power remains *de facto* in the hands of a restricted number of people (as in the case of one-party systems). As opposed to an oligarchic regime, in a democracy the franchise extends to the great majority or the entire the population and (at least some) political *voice* is granted to all social groups.

[3] The comparison between the efficiency of democracies and non-democracies in equilibrium and their ability to implement a rule of law is studied in the theory by Cervellati, Fortunato and Sunde (2008). The work by Cervellati, Fortunato and Sunde (2012) further study the role of expectations in sustaining multiple equilibria with different rule of law and propose the view that different transition to democracy may work as a coordination device for individual expectations.

Given these assumptions, the trade-off faced by the group controlling the state apparatus is between attempting to implement a more concentrated distribution of income (in favor of the group in power) at the cost of larger economic distortions, or limit the activities of rent extraction in favor of a higher provision of public goods, and ultimately, higher efficiency. Notice that, in principle, this trade-off might be in place irrespective of whether the state is controlled by a small (and rich) elite or a (poor) majority of the population, that is, irrespective of whether the political franchise is limited or extended. Either the government is able to implement policies that create widespread benefits and, in turn, confer legitimacy to its actions, or the state has little legitimacy. In the last case the quality of the emerging institutions, including the rule of law, will be poor. The point is that the bidirectional feedback between the choices of a central government and the actions of citizens may give rise to failed states in both democracies and non-democracies. Nonetheless, the incentives (and the trade-off) faced by an oligarchic elite may be related to the ones faced by the masses in the emerging democracy.

A natural implication of this political economy representation is, as discussed next, that one should expect that the structural features in place in one economy before the democratic transition should matter for both the modes of the transition and the features of the emerging democracy.

3.2 Natural resources and the democratization scenario: deriving testable hypotheses

Samuel Huntington (1993) provides an analysis of regime transitions during (what he calls) the “third wave” of democratization. He concludes that the implications of a violent transition for the quality of the emerging democracy are not entirely clear, but conjectures that violent uprisings should be expected to lead to worse democracies. This prediction would emerge also from a theoretical framework, like the one described above, interpreting violent conflicts as a result of a game between different social groups that aim at maximizing the utility of their members. Given the assumed features of the prototype democratic and authoritarian regimes described above, one can analyze the preferences of the different social groups with respect to the trade-off they face and derive some hypotheses about the determinants and consequences of the different modes of democratic transitions.

By its very definition, in an oligarchic (non-democratic) regime only the preferences of a minority are taken into account for the selection of public policies. Being unconstrained and free to pursue its objectives, the empowered elite will try to appropriate economic rents, and expropriate the powerless majority of the population. Faced with the trade-off between rent extraction and imposing distortions discussed above, the elite has a larger incentive to expropriate the easier it is to redistribute disposable income to themselves. We should therefore expect that the ruling elites will exploit their power to forcefully extract income at the cost of imposing large distortions in countries where natural resources are abundant and, importantly, easy to loot. In the presence of natural resources like, e.g., oil and minerals, it is easier to exclude the disempowered majority from the benefits of these resources. Under these conditions the disenfranchised masses of people should be less likely to confer legitimacy to the state. As consequence, these authoritarian regimes should be associated with inefficient institutions and large inequality.

In democracies where the majority of poor have a larger political voice, the redistribution schemes tend to be more progressive. However, also the wealthy (former) elites take part in the decision process and universal franchise is usually coupled with institutional mechanisms that protect the minorities. While redistribution is generally limited in democracies, certain forms of expropriation of private property, or nationalization of firms and industries, by part of the newly enfranchised population can be expected to take place during the process of democratization, however. Again, and crucially, we should expect this to be more likely in countries in countries richly endowed with natural resources that are easy to grasp like the ones that do not require investment in expensive equipment, and can easily be sold in the market legally or illegally.[4] Also natural resources, that have higher operational costs associated with them, such as oil, bauxite, or mineral gas, are subject to public expropriation since they are much less mobile across borders than physical or human capital, which makes it more difficult for their owners to escape seizure. Furthermore

[4] Timber, alluvial diamonds, and some drugs like coca or opium poppy are examples of such easily lootable natural resources.

these resources can be made subject to state monopolies which allows the group in control of the government to direct rent extraction and redistribute the revenues.

Violent regime transitions. When natural resources are abundant and their distribution is highly concentrated, it is relatively unlikely for small elites to be willing to give up power. In this case, elites rather prefer an authoritarian regime, which allows them to keep full control over the policy space. This is the scenario characterizing many oil and mineral resource-rich economies where restricted elites control power and resources, and are firmly opposed to any kind of democratic reform. Democratization can be enforced, however, by the disenfranchised population under (the shadow of) conflict and against the will of the ruling elite. If the powerless masses face a window of opportunity, e.g., as they become sufficiently strong, determined and organized, they may revolt and overthrow the existing regime. Historically, many regime changes resulted from the uprising of a politically (and economically) deprived class. The threat of revolution and social unrest, for example, played a decisive role in the establishment of voting rights in many Latin American countries, like in Nicaragua (1979) and Bolivia (1982).

When the transition is forced by the masses, while the elites are still entrenched and unwilling to accept a reduction of their political power, however, it is unlikely that the overthrown elite will accept the new rules of the game and confer legitimacy to the government. Furthermore, the masses themselves should be less likely to favor efficiency and the emergence of a good rule of law since this would reduce their ability to extract rents once in power. New democracies that emerge under such a scenario are therefore not very likely to generate efficient economic institutions and inclusive societies.

In sum, when controlling public rents from natural resources is easier, then we should expect that the trade-off between redistribution and efficiency is more likely resolved in favor of the former even when the (previously disenfranchised) masses manage to get to power. Consequently, if the rents controlled by the incumbent authoritarian rulers are very high, it is more likely that the new rulers will attempt to extract resources as well. Under these conditions the costs of giving up political power and succumbing to a regime change are sizable for the oligarchic elite which might attempt violent repression of upraising and the incentives to get in control of the state apparatus even with violent means is higher for those currently excluded from the rents.

Peaceful regime transition. In an alternative scenario, the transition to democracy can also be accepted, or even in some cases actively promoted, by (a part of) the formerly authoritarian power-holders. The implementation of democratic institutions can play an instrumental role to enable the emergence of effective property rights protection and rule of law. Efficient economic institutions are particularly important in an industrial economy where their implementation contributes to align the economic incentives, favors the adoption of new technologies and spurs productive investments. Oligarchic elites may lack the commitment power that is needed for a government to be legitimate and firmly convince the population that the government action will be aimed at providing public goods by limiting rent extraction. If the level of inequality is low and/or natural resources are not abundant (i.e. rent extraction is not particularly rewarding) the elites may find it profitable to trade-off a certain degree of progressive redistribution in a democracy against the possibility of having an environment more favorable for economic activity.[5] The formerly disenfranchised masses in turn are more likely to confer legitimacy to the new political system since this allows them to have a say in the political arena and to influence policies in their interest, e.g., by implementing a progressive system of redistribution of resources and incomes.[6] Peaceful transitions to liberal democracies followed the industrial take-off in the 19th century in many European countries and Western offshoots. During the "third wave" of democratization, consensual transitions characterized some eastern European countries after the break-down of the former USSR.

[5] The works by Lizzeri and Persico (2004) and Llavodor and Oxoby (2005) discuss the role of conflict within the elite. The parts of the elite that derive a large part of their economic returns from entrepreneurship and human capital rather than natural resources may favor or push a process of democratization to reap the benefits of higher returns to their activities.

[6] The prediction that peaceful transitions are more likely to occur when inequality is low and tend to lead to an environment more favorable for economic activity is consistent with the recent empirical findings of Chong and Gradstein (2007) and Sunde et al. (2008). More specifically, evidence on the differential role of the transition scenario for the quality of democratic institutions and growth, respectively, is provided by Cervellati, Fortunato, and Sunde (2011) and Cervellati and Sunde (2013).

This discussion can be summarized as saying that in economies where natural resources are important, easy to loot and materialize in large inequalities democratization is less likely to take place peacefully, and less likely to lead to high quality democracies. The discussion so far can be therefore summarized in two testable hypotheses:

> **Hypothesis 1: The roots of democratization.** Transitions to democracy can emerge under different scenarios: peacefully and under a broad consensus in the population, or under a violent conflict. When natural resources are abundant and inequality is high a democratic transition is more likely to emerge against the will of the ruling elite and under a scenario of conflict triggered by the disenfranchised masses and fuelled by attempts of repression by part of the elite.
>
> **Hypothesis 2: The fruits of democratization.** The democratization scenario has important implications for the quality and growth prospects of the emerging democracy. A new democracy is more likely to be characterized by a better rule of law and effective protection of property rights if it emerges out of a larger consensus and a peaceful transition. Democratization under conflict is less likely to bring effective protection of property rights and rule of law.

Summing up, these hypotheses imply that democratic institutions may play an instrumental role for the implementation of growth-enhancing policies and institutions, such as a rule of law. Democratization may not be sufficient to implement good institutions, however. The democratization scenario depends on whether the political and economic interests are broadly aligned in the population. Non-democracies characterized by a large inequality and/or abundance of natural resources are more likely to experience violent transitions and lead to low quality democracies.

4. Some evidence from the "Third Wave" of democratization

We now turn to investigate the consistency of the testable hypothesis with empirical evidence. We proceed in two steps. First, we investigate the relationship between natural resource abundance, inequality and the type of democratic transition, and then we study the relationship between the type of transition and the features of the emerging democracy.

Resource abundance, inequality and democratic transitions. It is not straightforward to classify countries according to whether they experienced *consensual* or *conflictual* transitions without necessarily being somewhat arbitrary. In a consensual scenario, when all the social groups agree on the necessity of an institutional change, the transition should occur in relatively smooth and non-violent fashion. In the alternative scenario, the conflict of interest between different social groups and the opposition of the elite to the change of regime may materialize in mass movements and pressure by part of the disenfranchised people. Democratic transitions taking place after social unrest, political pressure, and violent struggles may represent only a subset of the "conflictual" democratization as described above. In fact, democratic transitions could be conflictual even if they take place in the absence of open violence but in the shadow of an imminent, credible conflict. Measuring the intensity of violence and social unrest that precedes a political change, therefore, provides a pragmatic (but potentially conservative) strategy to distinguish consensual from episodes of democratization with open conflict.

The Freedom House (2005) classified the democratic transitions that occurred in the period 1973 to 2003 according to the level of violence which characterized the transition scenario. In particular, the study considers four categories: "High Violence", "Significant Violence", "Mostly Nonviolent" and "Nonviolent" transitions.

As a first investigation of the hypothesis, Figure 2 divides the democracies of the third wave in two groups according to the inequality in the distribution of income measured in the year in which the transition took place. Those countries with a Gini index below the average of the sample are labeled as (relatively) "equal" while the others are defined "unequal". The figure shows that, in line with the arguments presented above, the great majority of equal countries experienced regime changes which are classified by the Freedom House as either "nonviolent" or "mostly nonviolent" (i.e., consensual transitions, which occur without violent social conflict). On the contrary, more unequal countries displayed much higher levels of violence during the transition towards a democratic political system. In particular, seventy percent (25 out of the 37) of the countries that had a Gini index lower than the average in the year before the transition experienced a peaceful democratization. Conversely more than two thirds (20 out of 28) of the countries that had a Gini index above the average did experience a violent transition to democracy.

Figure 2. Inequality and the Mode of Transition to Democracy.

Source: Freedom House (2005) and WIID2 (2005).

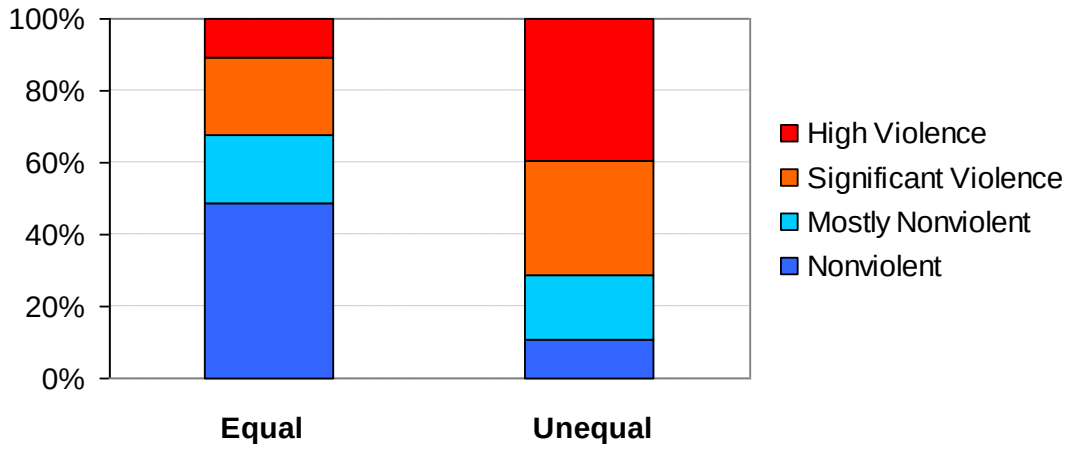


Table 2 investigates the hypothesis more structurally by ways of a multivariate regression analysis. The dependent variable is a binary indicator of whether the transition to democracy was associated with violence. The key explanatory variables are inequality (measured by the share of landless or by the Gini-index) and, alternatively, either a binary indicator of whether a country exports more than 33% of GDP in oil (column 1, 2 and 3) or the interpolated share of GDP made up by fuel (i-fuel, column 4, 5 and 6) to proxy for natural resource abundance. All the explanatory variables are measured in the year previous to the democratic transition. In order to account for potential confounds, we also present specifications that control for institutional quality (in terms of law and order), ethnic tensions, and log GDP per capita, all measured before the democratic transition.[7] Regardless of the specification, the results suggest that the propensity of observing violence during democratization was significantly higher in countries with abundant natural resources and high levels of inequality, consistent with the argument before. Although the inclusion of some controls (particularly land inequality) reduces the sample size considerably, the results suggest a significant and sizable effect of oil (and generally mineral) resources on the likelihood of observing a violent regime transition.

[7] Data for law and order and ethnic tensions are taken from the International Country Risk Guide, which are available from 1984, which implies a substantial loss of observations since democratization events before 1984 have to be dropped from the respective specifications.

Table 2: Inequality, Natural Resources and Violence During the Democratic Transition

Dependent Variable	Violent Transition (0/1)					
	(1)	(2)	(3)	(4)	(5)	(6)
Inequality pre-transition (Gini)	0.016***	0.014**	0.012*		0.016**	
	[0.004]	[0.006]	[0.006]		[0.006]	
Oil dummy (pre-transition)	0.402***	0.389***	0.376***	0.769***		
	[0.075]	[0.095]	[0.096]	[0.114]		
Civil liberties		0.059	0.026		-0.012	
		[0.074]	[0.074]		[0.072]	
log GDP per capita		-0.01	0.012		0.011	
		[0.044]	[0.034]		[0.034]	
Ethnic Tensions			-0.08		-0.079	
			[0.076]		[0.078]	
Share of Landless				0.026***		0.026***
				[0.008]		[0.008]
I-fuel (pre-transition)					0.004**	0.009***
					[0.002]	[0.002]
Constant	-0.153	-0.259	0.081	0.071	-0.007	0.02
	[0.172]	[0.421]	[0.550]	[0.144]	[0.515]	[0.166]
Observations	61	32	32	23	32	22
adjusted R-squared	0.19	0.146	0.181	0.301	0.216	0.322

Notes: Dependent variable is binary, coefficients are OLS estimates, standard errors in brackets.
*, **, *** denotes significance on the 10%, 5%, and 1% level. See the text for details of data description and sources.

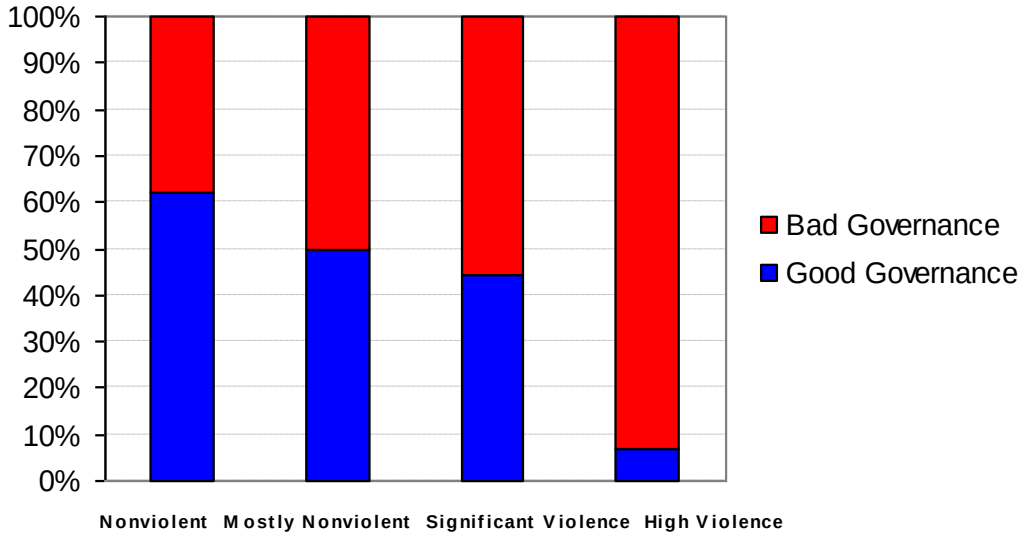
Democratic transitions, rule of law and civil rights. We next turn to investigate the empirical relevance of the second hypothesis, namely whether the level of violence during the transition scenario (i.e., the type of democratic transition) has any bearing on actual outcomes in terms of property rights protection and the quality of democratic institutions.

To measure the actual implementation of property rights, we use an indicator of rule of law proposed by the World Bank. This indicator reflects the quality of institutions and governance and it "measures the success of a society in developing an environment in which fair and predictable rules form the basis for economic and social interactions and the extent to which property rights are protected" (Kaufman *et al.*, 2004). In particular, we classify the countries of our sample in two groups, the "Good Governance" and the "Bad Governance" group, depending on whether they perform better or worse than the average.

A first look at the data suggests indeed that the level of governance in the emerging democracies might indeed be related to the level of violence during the regime change. Figure 3 shows that the majority of countries that implemented democratic institutions peacefully perform above average in terms of rule of law and property rights protection. Conversely, those countries characterized by a violent transition scenario perform, in the great majority, very poorly in terms of the quality of governance. This evidence is in line with the predictions of the hypothesis derived in the previous discussion.

Figure 3. Transition to Democracy and Quality of Governance.

Source: Freedom House (2005) and World Bank (2005).

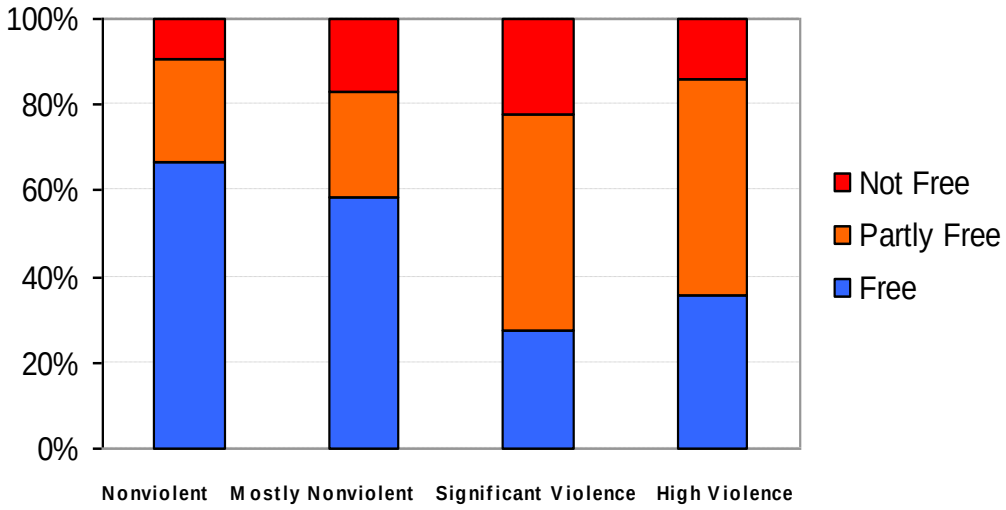


Almost seventy percent (23 out of 32) of the countries that changed regime under a *conflictual* scenario (i.e. significant or high violence) display a rule of law below the average quality within the group of new democracies as of 2005. On the contrary, 19 of the 33 countries that experienced a peaceful political (i.e. nonviolent or mostly nonviolent) change have been able to implement economic institutions above the average quality.

The previous discussion also suggests that consensual and non-violent transitions to democracy should lead to stable democracies that are accepted by the great majority of the population. The degree of civil liberties and the political rights within a country can be informative on the overall quality of democratic institutions. Employing the Freedom House classification into “Free”, “Partly Free” and “Non Free” countries described in the introduction, Figure 4 shows that political and civil rights are granted in the majority of countries that have democratized without social conflict, while the opposite holds true in democracies that originated in a scenario with conflict.

Figure 4. Democratic Transition and Civil Liberties.

Source: Freedom House (2005).



A more structural test of the hypothesis is again performed using multivariate regressions. The dependent variable is institutional quality, measured in terms of law and order, or in terms of civil liberties, as of 2005. The data are again taken from the International Country Risk Guide (law and order) and from Freedom House (civil liberties). The dependent variables are coded such that higher values indicate better outcomes. The explanatory variables are the institutional quality in the year before the transition to democracy, and binary indicators that measure the level of violence during the transition to democracy relative to non-violent transitions. Additional controls are the age of the democracy (in terms of years since democratization as of 2005), natural resource abundance (proxied by the relative importance of oil for the GDP) and inequality before the transition.

The results, presented in Table 3 that violence during the transition reduces institutional quality in the aftermath of the transition. Relative to non-violent transitions, significant or high levels of violence during the democratic transition reduce institutional quality in the long run. The results also provide evidence that institutional quality before the democratic transition, as well as inequality before the transition, have an effect on institutional quality after the transition.[8] The main result of the detrimental effect of violence holds up, however, even when controlling for institutional quality before the transition. The somewhat weaker effects of violence during the transition when controlling for pre-transitional institutional quality and inequality also indicates that these variables affect the likelihood of the occurrence of violence during the transition to democracy and account for some of the effect of violence. Taken together, however, the results are consistent with the hypothesis that violence during the transition represents a key indicator of whether democratization indeed leads to an improvement in institutional quality. These results are not confined to the Freedom House classification of violence during the transition. Using data on the occurrence of civil conflict in the year of, or the year before, the transition to democracy using an alternative coding of democratization than Freedom House, Cervellati, Fortunato and Sunde (2014) also find a significant negative effect of violence during the transition on the institutional quality of the emerging democracies.

The results in Table 3 also show that inequality and natural resource abundance affect institutional quality partly through the type of transition. These findings are fully consistent with the recent literature on natural resource course that has extensively documented the existence of direct linkages between resource abundance and institutional quality (Moehne et al., 2006, and Torvik, 2009). The findings complement these results and show that resource abundance matters for institutions also indirectly by contributing to the determination of the transition scenario which in turn influences institutional quality.

Unreported results finally suggest that an interaction term between law and order and a binary indicator for violence accounts for the potentially differential change in institutional quality in countries that democratized peacefully as opposed to violent transitions.

5. Concluding remarks

The evidence from the successful democratic transitions of the “third wave” suggests that democracy is more likely to flourish when economic and political interests are broadly aligned in the population, thereby creating fertile soil for more inclusive and more efficient democratic institutions. Countries heavily relying on natural resources whose rents are unevenly distributed in the population were more likely to experience violent conflicts during democratization, and in turn display worse institutions after the transition to democracy. The experience of countries like Taiwan is paradigmatic; they grew under *de facto* one-party authoritarian regimes eventually turning peacefully into legitimate and stable democracies after the economic take-off. On the contrary, in unequal societies highly dependent on natural resources, democratic institutions have often been introduced by force and against the will of a substantial part of the society. The democracies that emerged under such a scenario have typically not been able to provide effective protection for political and economic (property) rights, as in the case of many natural resource abundant countries as diverse as Bolivia, Uzbekistan, or Zimbabwe.

[8] The low number of 35 observations of democratization is due to data limitations for the ICRG index of law and order, which is only available since 1984.

Table 3: "Long-Run" Consequences of Violence During the Democratic Transition

Dependent Variable	Law and Order (in 2005)		Civil Liberties (in 2005)		Political Rights (in 2005)		Combined Avg. Index	
Mostly Non-Violent Transition	-0.354	0.059	0.8	0.888	0.436	0.484	0.618	0.703
	[0.373]	[0.471]	[0.500]	[0.568]	[0.615]	[0.714]	[0.548]	[0.631]
Significant Violence	-0.910***	-0.395	1.039**	0.8	1.047*	0.815	1.043**	0.806
	[0.329]	[0.457]	[0.442]	[0.518]	[0.544]	[0.653]	[0.485]	[0.578]
High Violence	-0.729*	0.507	1.386***	1.485**	1.411**	1.392*	1.398**	1.430*
	[0.373]	[0.516]	[0.496]	[0.643]	[0.611]	[0.807]	[0.544]	[0.717]
Age of Democracy (Years)	0.036	0.039	-0.03	-0.048	-0.025	-0.031	-0.028	-0.04
	[0.022]	[0.041]	[0.031]	[0.033]	[0.038]	[0.042]	[0.034]	[0.037]
Oil (pre-transition)		-0.67		1.701		1.644		1.632
		[1.008]		[1.086]		[1.341]		[1.196]
Law and Order (pre-transition)		0.272						
		[0.164]						
Inequality pre-transition (Gini)		-0.026*		-0.001		-0.007		-0.004
		[0.015]		[0.018]		[0.022]		[0.020]
Civil Liberties				-0.039				
(pre-transition)				[0.221]				
Political Rights						0.135		
(pre-transition)						[0.215]		
Combined Avg. Ranking								0.07
(pre-transition)							[0.233]	
Constant	3.511***	3.283***	2.902***	3.196**	2.658***	2.282	2.666***	2.647
	[0.366]	[1.068]	[0.547]	[1.447]	[0.626]	[1.703]	[0.558]	[1.645]
Observations	57	35	61	59	67	59	67	59
adjusted R-squared	0.116	0.243	0.134	0.089	0.045	0.01	0.065	0.039

Notes: Law and order is measured on a scale 0-6, with higher values indicating better outcomes. Civil Liberties and Political Rights are measured on a scale 1-7 with higher values indicating worse outcomes. Reference category for violence is `no violence during the transition'. Coefficients are OLS estimates, standard errors in brackets.

*, **, *** denotes significance on the 10%, 5%, and 1% level, respectively. See the text and Table A1 for details of data description and sources.

Democratic transitions, however, are not always exclusively driven by internal forces. The international community may in fact inflict economic and political sanctions (like embargos or diplomatic isolation) on autocratic states in order to foster democratic transition. The results suggest that these measures might be more likely to be fruitfully applied on those economies with a relatively even distribution of income and wealth and that rely relatively little on the rents from lootable natural resources. These economies might be able to sustain democratic institutions emerging out of external pressures since a regime change pushed only internally might be made difficult and hindered by the generally difficult coordination of the internal civic opposition. On the contrary, external pressures on resource-based economies with rampant inequality and easy rent seeking are likely to pave the way for democratic institutions emerging out of violence and materializing in the persistence of government characterized by a serious deficit of popular legitimacy. The results suggest that the level of violence during the transition might have long term (persistent) effects on the ability to improve rule of law and civil liberties. Under these conditions, pushing a country towards political change by any mean may be a particularly risky strategy and policies aiming at changing the internal socio-economic environment by implementing economic reforms favoring a more widespread distribution of income and equalized economic opportunities may represent a useful preliminary step to increase the pressure towards a peaceful change of regime and to make democracy work.

Finally, the evidence presented here poses more than a shadow on the prospects for a full democratic blossoming in the Arab countries. The unequal distribution of riches and the violent confrontation that characterizes the movement to democracy in some of these countries, following decades of heavy-handed governance, suggest that it could be difficult to reach a democratic consensus in the population at large. In these countries, there is a substantial risk of emergence of democracies characterized by poor protection of civil liberties and poor rule of law.

References

Auty, R. M. 1990, *Resource Based Industrialization: Sowing the Oil in Eight Developing Countries*, Oxford University Press, New York.

Acemoglu, D. & Robinson, J. 2000, *'Why Did the West Extend the Franchise? Democracy, Inequality, and Growth in Historical Perspective'*, *Quarterly Journal of Economics*, 115(4), pp. 1167-1199.

Acemoglu, D. & Robinson, J. 2001, 'A Theory of Political Transitions', *American Economic Review*, 91(4), pp. 938-963.

Acemoglu, D. & Robinson, J. 2006a, *Economic Origins of Dictatorship and Democracy*, Cambridge University Press, Cambridge.

Acemoglu, D. & Robinson, J. 2006b, 'Economic Backwardness in Political Perspective', *American Political Science Review*, 100(01), pp. 115–31.

Alesina, A., Devleeschauwer, A., Easterly. W., Kurlat, S. & Wacziarg, R. 2003, 'Fractionalization', *Journal of Economic Growth*, 8(2), pp. 155-194.

Bates, R. H. 1991, 'The Economics of Transition to Democracy', *PS: Political Science and Politics*, 24: 24-7.

Bourguignon, F., & Verdier, T. 2000, 'Oligarchy, Democracy, Inequality and Growth', *Journal of Development Economics*, 62(2), pp. 285-313.

Cervellati, M., & Sunde, U. 2013, 'Civil Conflict, Democratization, and Growth: Violent Democratization as Critical Juncture', *Scandinavian Journal of Economics*, forthcoming.

Cervellati M. & Fortunato, P. 2007, *Resistance to Reforms, Inequality and Development*, unpublished manuscript.

Cervellati M., Fortunato, P. & Sunde, U. 2008, 'Hobbes to Rousseau: Inequality, Institutions and Development', *Economic Journal*, 118(531), pp. 1354-1384

Cervellati M., Fortunato, P. & Sunde, U. 2012, 'Consensual and Conflictual Democratisation', *B.E. Journal of Theoretical Economics*,12(1) (Contributions), Article 33.

Cervellati M., Fortunato, P. & Sunde, U. 2014, 'Democratization and Civil Liberties: The Role of Violence During the Transition', *European Economic Review*, 66, pp. 226-247.

Chong, A., & Gradstein, M. 2007, 'Inequality and Institutions', *Review of Economics and Statistics*, 89(3), pp. 454-465.

Yun-han, C. 2006, 'Third-Wave Democratization in East Asia: Challenges and Prospect', *ASIEN 100*, pp. 11-17.

Dahl, R. A. 1971, *Polyarchy: Participation and Opposition*, Yale University Press, New Haven.

Fearon, J. D. and Laitin, D.D. 2003, 'Ethnicity, Insurgency, and Civil War', *American Political Science Review*, 97(1), pp. 75-90.

Fearon, J. D. 2005, 'Primary Commodity Exports and Civil War', *Journal of Conflict Resolution*, 49(4), pp. 483-507.

How Freedom is Won: from Civic Resistance to Durable Democracy 2005, Freedom House, New York.

Freedom in the World Survey 2006, Freedom House, New York.

Gelb, A. 1988, *Oil Windfalls: Blessing or Curse?*, Oxford University Press, New York.

Haggard, S. & Kaufman, R.R. 1995, *The Political Economy of Democratic Transitions,* Princeton University Press, Princeton.

Hagopian, F., & Mainwaring, S.P. 2005, *The Third Wave of Democratization in Latin America Advances and Setbacks*, Cambridge University Press, Cambridge.

Huntington, S.P. 1991, 'How countries democratize', *Political Science Quarterly*, 106 (4), pp. 569-616.

Huntington, S.P. 1993, *The Third Wave: Democratization in the Late Twentieth Century,* Univ. of Oklahoma Press, Oklahoma.

Karl, T.L. 1997, *The Paradox of Plenty: Oil Booms and Petro States*, Berkeley University Press, Berkeley.

Kaufmann, D., Kraay, A. & Mastruzzi, M. 2004, 'Governance Matters III: Governance Indicators for 1996-2002',*World Bank Research Paper*.

Inglehart, R. 2004, 'How Solid is Mass Support for Democracy—And How Can We Measure It?', *Political*

Science and Politics, 36(01), pp51-57.

Linz, J.J., and Stepan, A. 1996, *Problems of Democratic Transition and Consolidation*, Johns Hopkins University Press, Baltimore.

Lipset, S.M. 1959, 'Some Social Requisites of Democracy: Economic Development and Political Legitimacy', *American Political Science Review* 53(1), pp. 69-105.

Lizzeri, A. & Persico, N. 2004, 'Why did the Elites Extend the Suffrage? Democracy and the Scope of Government, with an Application to Britain's `Age of Reform'*, Quarterly Journal of Economics*, 119(2), pp. 707-765.

Luebbert, G.M. 1991, *Liberalism, fascism, or social democracy: Social classes and the political origins of regimes in interwar Europe*, Oxford University Press, New York.

Moene, O., Mehlum, H. & Torvik, R. 2006, 'Institutions and the Resource Curse', *The Economic Journal*, 116(508), pp. 1-20.

Moore, B. 1966, *The Social Origins of Dictatorship and Democracy*, Beacon Press, Boston.

Putnam, D. 2002, *Democracies in Flux: The Evolution of Social Capital in Contemporary Society*, Oxford University Press, New York.

Robinson, J.A. & Torvik, R. 2005, 'White Elephants', *Journal of Public Economics*, 89(2), pp. 197–210.

Robinson, J.A., Verdier, T., & Torvik, R. 2006, 'Political Foundations of the Resource Curse', *Journal of Development Economics*, 79(2), pp. 447–68.

Rogowski, R. 1998, 'Democracy, Capital, Skill, and Country Size: Effects of Asset Mobility and Regime Monopoly on the Odds of Democratic Rule', *The Origins of Liberty: Political and economic liberalization in the modern world*, Princeton University Press, Princeton.

Ross, M.L. 1999, 'The Political Economy of the Resource Curse', *World Politics*, 51(02), pp. 297–322.

Sunde, U., Cervellati, M. & Fortunato, P. 2011, 'Are all democracies equally good? The role of interactions between political environment and inequality for rule of law', *Economics Letters*, 99(3), 552-556.

Tilly, C. 2004, *Contention and Democracy in Europe, 1650-2000*, Cambridge University Press, New York.

Torvik, R. 2009, 'Why do some resource-abundant countries succeed while others do not?', *Oxford Review of Economic Policy*, 25(2), pp. 241-256.

Van der Ploeg, F. 2011, 'Natural Resources: Curse or Blessing?', *Journal of Economic Literature*, 49(2), pp. 366-420.

Table A1: Raw Data: Democratic Transitions, Violence, and Civil Liberties

Country	Transition Year	Pre-Transition: Liberties* and Inequality						Post-Transition (2005): Liberties*				Transition Scenario
		PR	CL	CAR	Status	Gini	Year	PR	CL	CAR	Status	
Albania	1992	7	6	6.5	NF	29.3	1989	3	3	3	PF	Mostly Nonviolent
Argentina	1983	6	5	5.5	NF	43.8	1981	2	2	2	F	High Violence
Armenia	1998	6	5	5.5	NF	28	1988	4	4	4	PF	Significant Violence
Azerbaijan	1991	6	5	5.5	NF	31.7	1988	6	5	5.5	PF	Significant Violence
Bangladesh	1991	4	4	4	PF	33.6	1989	4	4	4	PF	Significant Violence
Belarus	1991	6	5	5.5	NF	24.2	1988	6	6	6	NF	Nonviolent
Benin	1991	7	7	7	NF		1989	2	2	2	F	Mostly Nonviolent
Bolivia	1982	7	5	6	NF	51.5	1982	3	3	3	PF	High Violence
Bosnia-Herz.	1995	6	6	6	NF	32.88	1994	4	3	3.5	PF	High Violence
Brazil	1985	3	3	3	PF	57.7	1984	2	3	2.5	F	Mostly Nonviolent
Bulgaria	1991	7	7	7	NF	20.9	1988	1	2	1.5	F	Nonviolent
Cambodia	1993	7	7	7	NF	46	1990	6	5	5.5	NF	Significant Violence
Cape Verde	1991	5	5	5	PF		1990	1	1	1	F	Nonviolent
Chile	1990	6	5	5.5	PF	47	1987	1	1	1	F	Mostly Nonviolent
Croatia	2000	4	4	4	PF	30	1998	2	2	2	F	Nonviolent
Czech Rep.	1993	7	6	6.5	NF	19.8	1987	1	1	1	F	Nonviolent
El Salvador	1994	3	4	3.5	PF	50.6	1991	2	3	2.5	F	High Violence
Estonia	1992	6	5	5.5	NF	27.8	1988	1	1	1	F	Nonviolent
Ethiopia	1995	7	7	7	NF	52.7	1990	5	5	5	PF	High Violence
Gambia	2001	7	5	6	NF	50.2	2000	4	4	4	PF	Mostly Nonviolent
Ghana	2000	3	3	3	PF	40.7	1999	2	3	2.5	F	Nonviolent
Greece	1975	7	5	6	NF	41.3	1973	1	2	1.5	F	Mostly Nonviolent
Guatemala	1996	4	5	4.5	PF	55.3	1995	4	4	4	PF	High Violence
Guyana	1992	5	4	4.5	PF	51.5	1989	2	2	2	F	Nonviolent
Hungary	1990	5	4	4.5	PF	26.8	1988	1	1	1	F	Nonviolent
Indonesia	1999	7	5	6	NF	36	1997	3	4	3.5	PF	High Violence
Iran	1979	5	6	5.5	PF	46	1978	6	6	6	NF	High Violence
Kazakhstan	1991	6	5	5.5	NF	29.1	1988	6	5	5.5	NF	Nonviolent
Kyrgyzstan	1991	6	5	5.5	NF	31.2	1988	6	5	5.5	NF	Nonviolent
Latvia	1991	6	5	5.5	NF	25	1988	1	2	1.5	F	Mostly Nonviolent
Lithuania	1991	6	5	5.5	NF	24.4	1988	2	1	1.5	F	Significant Violence
Macedonia	1991	5	4	4.5	PF	32.22	1989	3	3	3	PF	Nonviolent
Madagascar	1993	5	4	4.5	PF	62.5	1989	3	3	3	PF	Significant Violence
Malawi	1994	7	6	6.5	NF	62	1991	4	4	4	PF	Significant Violence
Mali	1992	6	5	5.5	NF	54	1990	2	2	2	F	Significant Violence
Mexico	1997	3	4	3.5	PF	54.5	1999	2	2	2	F	Nonviolent
Moldova	1994	6	5	5.5	NF	26.4	1988	3	4	3.5	PF	Significant Violence
Mongolia	1993	7	7	7	NF	33.2	1989	2	2	2	F	Nonviolent
Mozambique	1994	6	4	5	NF	39.4	1991	3	4	3.5	PF	Mostly Nonviolent
Nepal	1990	4	5	4.5	PF	54.6	1989	5	5	5	PF	Significant Violence
Nicaragua	1990	5	5	5	PF	55.7	1989	3	3	3	PF	High Violence
Nigeria	1999	7	6	6.5	NF	50.2	1997	4	4	4	PF	Significant Violence
Panama	1994	7	6	6.5	NF	56.5	1988	1	2	1.5	F	High Violence
Paraguay	1993	6	6	6	NF	45.1	1988	3	3	3	PF	Significant Violence
Peru	1980	5	4	4.5	PF	50.5	1999	2	3	2.5	F	Mostly Nonviolent
Philippines	1987	4	3	3.5	PF	45.5	1985	2	2	2	F	Significant Violence
Poland	1990	5	5	5	PF	24.5	1988	1	1	1	F	Nonviolent
Portugal	1976	5	6	5.5	NF	40.1	1973	1	1	1	F	Mostly Nonviolent
Romania	1990	7	7	7	NF	31.2	1988	3	2	2.5	F	Significant Violence
Russia	1993	6	5	5.5	NF	23.9	1988	6	5	5.5	NF	Mostly Nonviolent
Senegal	2000	4	4	4	PF	29.3	1999	2	3	2.5	F	Mostly Nonviolent
Serbia-Mont.	2000	5	5	5	PF		1999	3	2	2.5	F	High Violence
Slovakia	1989	7	6	6.5	NF	18.6	1988	1	1	1	F	Nonviolent
Slovenia	1992	5	4	4.5	PF	21.9	1989	1	1	1	F	Nonviolent
South Africa	1994	6	5	5.5	PF	59.5	1989	1	2	1.5	F	Significant Violence
South Korea	1987	4	5	4.5	PF	34.54	1986	1	2	1.5	F	Significant Violence
Spain	1978	5	5	5	PF	34.1	1974	1	1	1	F	Nonviolent
Taiwan	1992	5	5	5	PF	29.7	1991	2	2	2	F	Nonviolent
Tajikistan	1991	6	5	5.5	NF	31.8	1988	6	5	5.5	NF	Significant Violence
Tanzania	1995	6	5	5.5	NF	38	1993	4	3	3.5	PF	Nonviolent
Thailand	1992	6	4	5	PF	49.8	1991	2	3	2.5	F	Significant Violence
Turkey	1983	5	5	5	PF	57	1980	3	3	3	PF	Significant Violence
Uganda	1986	5	4	4.5	PF	37.7	1984	5	4	4.5	PF	High Violence
Uruguay	1985	5	4	4.5	PF	40.4	1983	1	1	1	F	Nonviolent
Uzbekistan	1991	6	5	5.5	NF	30.6	1988	7	6	6.5	NF	Mostly Nonviolent
Zambia	1991	6	5	5.5	PF	48.4	1989	4	4	4	PF	Nonviolent
Zimbabwe	1979	6	5	5.5	NF	62.9	1975	7	6	6.5	NF	High Violence

PR: Political Rights, CL: Civil Liberties, CAR: "Combined Average Rating" (average of FIW Political Rights and Civil Liberties scores. The scores are based on a 1-7 scale: 1 represents the highest level of freedom and 7 the lowest.) Status: F Free, PF Partly Free, NF Not Free.
Source: Freedom House (2005) and WIID2 (2005).

Incrementum ad Absurdum: Global Growth, Inequality and Poverty Eradication in a Carbon-Constrained World

David Woodward[1]
United Nations Conference on Trade and Development (UNCTAD)

Abstract

The paper seeks to assess the timeframe for eradication of poverty, defined by poverty lines of $1.25 and $5 per person per day at 2005 purchasing power parity, if pre-crisis (1993-2008) patterns of income growth were maintained indefinitely, taking account of the differential performance of China. On the basis of optimistic assumptions, and implicitly assuming an indefinite continuation of potentially important pro-poor shifts in development policies during the baseline period, it finds that eradication will take at least 100 years at $1.25-a-day, and 200 years at $5-a-day. While this could in principle be brought forward by accelerating global growth, global carbon constraints raise serious doubts about the viability of this course, particularly as global GDP would need to exceed $100,000 per capita at $1.25-a-day, and $1m per capita at $5-a-day. The clear implication is that poverty eradication, even at $1.25-a-day, and especially at a poverty line which better reflects the satisfaction of basic needs, can be reconciled with global carbon constraints only by a major increase in the share of the poorest in global economic growth, far beyond what can realistically be achieved by existing instruments of development policy – that is, by effective measures to reduce global inequality.

Keywords: economic growth, inequality, carbon constraints, poverty, income distribution, poverty eradication

The United Nations High-Level Panel on the Post-2015 Development Agenda (2013) has proposed as a central development goal for the post-2015 period the eradication by 2030 of extreme poverty, as defined by the poverty line of $1.25 per person per day at 2005 purchasing power parity. This follows a number of studies seeking to estimate the extent of poverty reduction over this period as a basis for discussions of post-2015 development goals (eg Karver *et al.*, 2012; Chandy *et al.*, 2013; Ravallion, 2013).

The author has argued elsewhere (Woodward, 2013) that the shift from poverty reduction to poverty eradication is very welcome and entirely appropriate, but that the latter is qualitatively different from the former, and has major implications for our approach to development, particularly in relation to the relative importance of global economic growth and the global distribution of income. This article develops this theme by considering the likely timeframe for the eradication of poverty, based both on the $1.25-a-day line and a poverty line of $5-a-day, which is considered as a more realistic reflection of requirements to fulfil basic needs, on the basis of pre-crisis (pre-2008) trends. Adapting and improving the methodology developed in an earlier paper (Woodward and Simms, 2006, whose results are also updated), the article extrapolates pre-crisis (1993-2008) trends in global GDP growth and in income growth in the lower deciles of the world population. Since poverty reduction has hitherto been largely driven by China, whose greatly reduced incidence of poverty will considerably diminish its effect on poverty trends in the future, it bases its findings primarily on analysis of the world excluding China.

Adopting this as a baseline scenario, the article considers the implications of variations in the two basic variables, the rate of global economic growth and the growth rate of per capita income of the poorest decile relative to global GDP per capita in the context of environmental (and particularly carbon) constraints to long-term growth of the global economy.

[1] The author is Senior Adviser in the Division for Africa, Least Developed Countries and Special Programmes in the United Nations Conference on Trade and Development (UNCTAD). However, this article was written in its entirety, and submitted for publication in an individual capacity, prior to his beginning work with UNCTAD. The author is grateful for comments on previous drafts by Roberto de Vogli and Nazrul Islam. Any views expressed in the paper are entirely the author's own and do not reflect the position or policy of UNCTAD; and any shortcomings remain solely the author's responsibility.

All poverty data used in this article are taken from PovcalNet, the on-line tool for poverty measurement developed by the Development Research Group of the World Bank.[2] The methodology for estimating decile average incomes is described in the Annex to this paper.

1. "Growth Isn't Working" revisited

Woodward and Simms (2006) presented estimates of the share in global economic growth accruing to those below the (former) "$1-a-day" and "$2-a-day" poverty lines, at 1985 purchasing power parity, between 1981 and 2001. Since then, not only have later data become available (up to 2008), but the entire PovcalNet database has been updated to 2005 purchasing power parity.

Table 1 and Figures 1 and 2 present an updated version of the estimates presented in Table 4 and Figures 7-8 of Woodward and Simms (2006), following the World Bank in changing the lower poverty line to $1.25 per day, while leaving the upper poverty line unchanged at $2 per day at 2005 purchasing power parity.

Table 1: Changes in Incomes of Poor Households

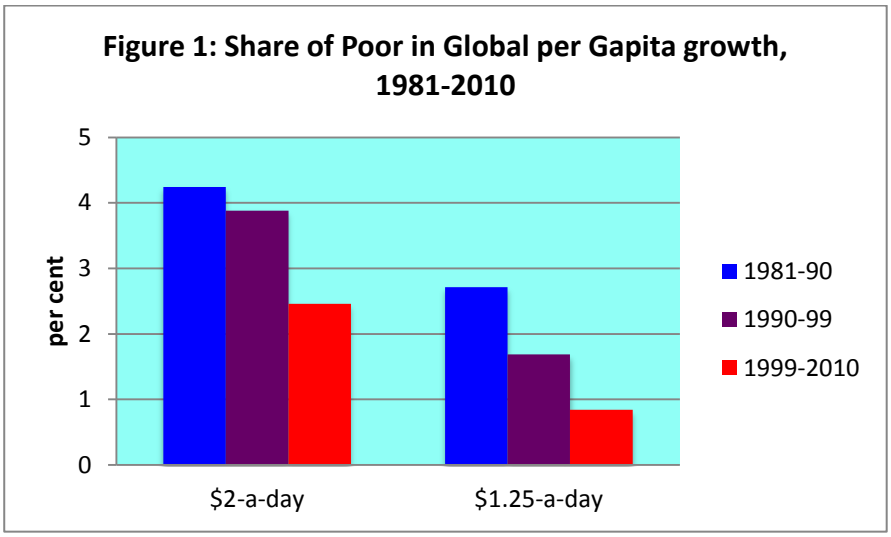
		$1.25-a-day			$2-a-day		
		1981-90	**1990-99**	**1999-10**	**1981-90**	**1990-99**	**1999-10**
without population adjustment	poverty reduction ($bn)	216	164	193	377	364	532
	change in world GDP($bn)	9,132	9,963	21,717	9,132	9,963	21,717
	poverty reduction as % of change in GDP	2.37	1.65	0.89	4.13	3.65	2.45
	income of poor as % of GDP	1.95	1.59	1.17	3.32	3.11	2.67
	ratio of share in growth to share in GDP	1.22	1.04	0.76	1.24	1.17	0.92
with population adjustment	poverty reduction ($bn)	106	62	84	165	143	246
	change in world GDP ($bn)	3,898	3,689	9,983	3,898	3,689	9,893
	poverty reduction as % of change in GDP	2.71	1.69	0.84	4.24	3.88	2.46
	income of poor as % of GDP	1.95	1.59	1.17	3.32	3.11	2.67
	ratio of share in growth to share in GDP	1.39	1.06	0.72	1.28	1.25	0.86

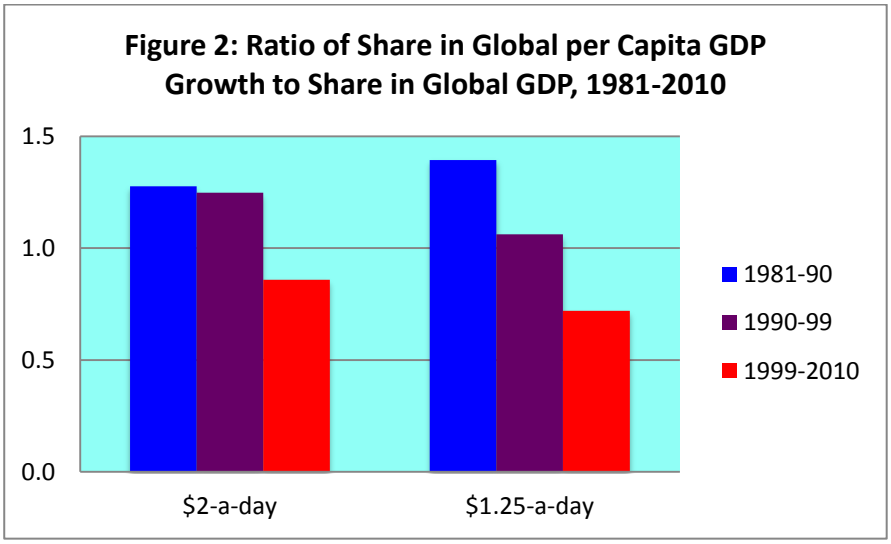
[2] http://iresearch.worldbank.org/PovcalNet/index.htm?0,2

The historical results are significantly different from those presented in Woodward and Simms (2006), and appear to present a significantly less pessimistic picture of poverty in the 1980s and 1990s: both the share of poor households in global GDP growth and the ratio of this share to their initial share in GDP are higher in both the 1980s and the 1990s (except at the $2-a-day level in the 1980s, where the two figures are approximately equal), and the reduction between the 1980s and 1990s is less marked; and the ratio remains greater than unity at both the $1.25 and the $2 level.

However, there is again a marked reduction in both the share of global growth and the ratio to share in GDP at both poverty levels from 1999 to 2010. It will be noted, however, that the ratio at the $1.25 level in 1999-2008 remains substantially higher than in the earlier analysis, at 0.76 (as compared with 0.51).

To update one of the widely cited findings in the earlier analysis, the greater share of the $1.25-a-day poor in global per capita growth in 1999-2010 than of the $1-a-day poor in 1990-2001 (0.89% cf 0.60%) suggests that $111 of additional growth in global GDP is required to achieve each $1 reduction in poverty, based on the $1.25-a-day poverty line, as compared with $166 (based on the then $1.08-a-day line) in the earlier analysis. It should be noted, however, that this change essentially reflects the much greater proportion of the population categorised as "poor" by the $1.25 poverty line in 1999 than by the $1.08-a-day line in 1990 (the first year of the final sub-period in each case) – 34.1%, as compared with 23.2%. Adopting a poverty line equivalent to the $1-a-day line (in terms of the corresponding poverty headcount) would clearly give rise to a figure well above $111.





There are, however, a number of reasons for caution in comparing the two sets of historical results. First, while the poverty lines defined at 1985 and 2005 PPP are not directly comparable, a comparison of the poverty headcount ratios to which they give rise indicates that the $1.25 line at 2005 PPP effectively represents a higher line than the $1-a-day line at 1993 PPP. In fact, by this measure, the $1.25-a-day line at 2005 purchasing power parity might be considered broadly equivalent to the $2-a-day line at 1985 PPP at the beginning of the period: the headcount ratio given by the $1.25-a-day line in 1981 gives a headcount ratio of 52.2% in 1981, only marginally below the 54.3% indicated by the earlier $2-a-day line.

If, accordingly, we compare the current $1.25-a-day results with the earlier $2-a-day results, we find that the share in growth is now little more than half the earlier figure, and exhibits a similar proportional reduction between the 1980s and 1990s (from 2.7% to 1.7%, as against a reduction from 5.3% to 3.1%). However, the ratio to GDP share is slightly higher, and falls somewhat less (from 1.39 to 1.06, as against a reduction from 1.25 to 0.8), consequently remaining greater than unity until 1999.

A second consideration is the change in periods covered: while the second period in the earlier analysis was from 1990 to 2001, that in the current analysis ends in 1999. While poverty reduction between 1999 and 2002 does not appear to have been substantially out of line with the overall pattern for the period as a whole, a minor effect on the results cannot be ruled out.

Third, at any point in time, the most recent poverty estimates (as those for 2001 in the earlier analysis) are partially based on extrapolation from past trends in countries for which no recent survey data are available. As subsequent data become available, these extrapolated estimates are replaced with actual survey data or (more reliable) interpolated estimates. Again, this may affect the earlier estimates for the 1990s to some extent. However, to the extent that extrapolated estimates exhibit an optimistic bias (as argued in Woodward and Abdallah, 2006, pp 16-18), one would expect this effect to operate in the opposite direction to that observed, rendering the earlier results artificially optimistic.

Finally, the change from 1993 to 2005 purchasing power is itself likely to have had a substantial, though indeterminate, effect on the historical results, as such revisions give rise to serious issues of consistency (Woodward and Abdallah, 2010, pp. 10-11).

At first sight, this might appear to be a case for preferring the current results to the earlier ones, as they are based on a more recent PPP. However, it is far from clear that this is justified. In assessing historical trends, it would seem appropriate to adopt a base year for purchasing power parity as close as possible to the mid-point of the period under consideration. For the historical (1981-99, mid-point 1990) estimates, this rule of thumb would suggest a preference for 1985 rather than 2005 (although the converse would clearly apply to the 1999-2008 results).

In this particular context, however, there is arguably a further reason to prefer the earlier base-year. Since PPP estimates are based on GDP weights and not consumption patterns of poor households, the question is whether the composition of poor households' consumption between 1981 and 2008 is closer to 1985 GDP weights or to 2005 GDP weights. While the difference in both cases is likely to be considerable, there would seem to be a strong intuitive case for preferring weights from 1985, when GDP per capita was substantially lower, and thus less far removed from the $1/1.25- and $2-a-day poverty lines.

In summary, while the revised estimates presented here may appear indicative of a significantly less pessimistic picture of poverty reduction in the 1990s than that suggested by the results in Woodward and Simms (2006), this appears to arise in large measure from the change in the definition of poverty associated with the updating to 2005 purchasing parity. The change in PPP base year itself may also have some effect; but in the particular context of poverty analysis, there is no clear reason to prefer a later to an earlier base-year; and, for historical analysis, there would seem a strong *prima facie* case for preferring a base year during the period analysed to a subsequent base year. Thus the historical results presented here should not necessarily be viewed as superseding or invalidating the earlier results.

2. Limitations of the "Growth Isn't Working" methodology

The foregoing discussion has adopted the methodology used in Woodward and Simms (2006) in order to reassess and update its results on the basis of subsequent data. However, while this approach has the merit

of direct relevance to policy discussions, in that it is based on internationally accepted poverty lines, this basis itself gives rise to two significant shortcomings.

First, while poverty is falling, the proportion of households defined as poor falls. Any observed reduction in the proportion of global growth contributing to poverty reduction will at least partly, and may wholly, reflect this change. This is the reason for the use of the ratio between growth shares and GDP shares, rather than shares in growth *per se*, as an indicator of changes in performance over time.

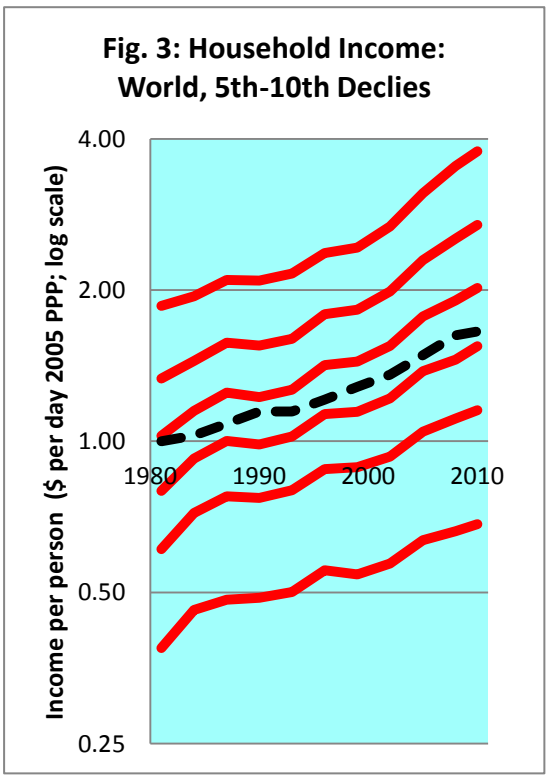
Second, however, the use of fixed poverty lines means that any increase in the income of a households beyond a particular poverty line is excluded from the contribution of growth to poverty reduction with respect to that poverty line. While reasonable in assessing the proportion of growth which reduces poverty defined by a particular poverty line, this is more problematic in assessing the ratio between the share in global growth and the share in global GDP, as the household's entire income is included in the GDP share, but only part of the increase in its income is included in its share in growth. This also makes estimated ratios sensitive to the length of the period considered: the longer the period, the more households will rise above the poverty line, and the greater the proportion of additional income which will be excluded.

(This is the reason for ensuring equal periods in the current analysis. It also suggests that the difference in sub-periods in the earlier study, at 9 years and 11 years, though inevitable given the format of the data, may have made a modest contribution to the estimated reduction in ratios.)

In the remainder of this paper, we therefore adopt an alternative approach which does not rely on the use of fixed poverty lines, but rather estimates the average per capita income of income deciles. This has the added advantage of side-stepping the issue of whether, and if so how, the results should be adjusted to for global population growth.

3. Evolution of incomes by decile of world population

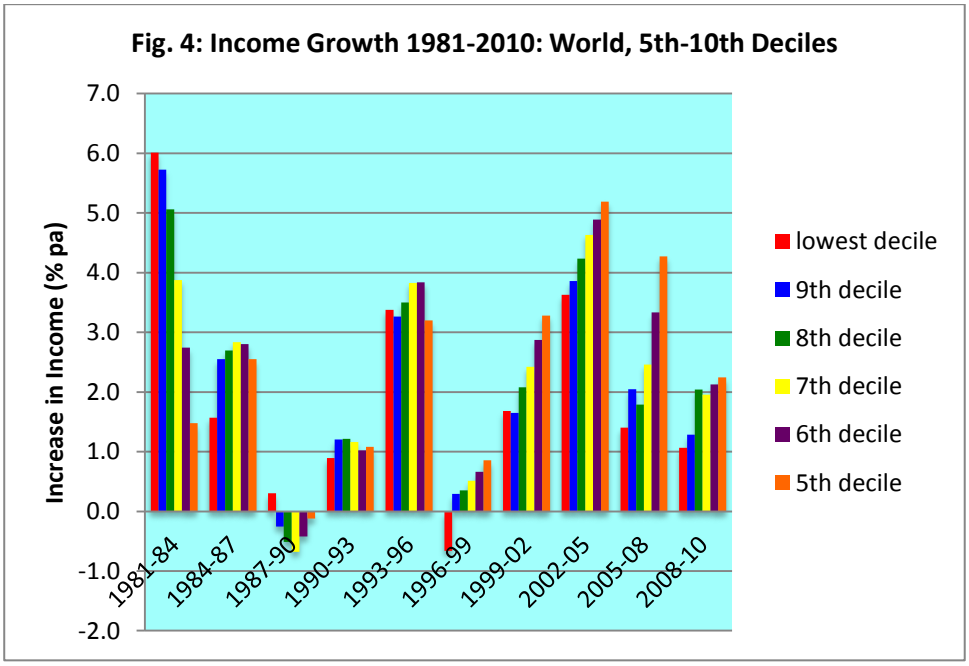
At first sight, progress in reducing poverty appears to have been substantial over the last 30 years. Despite the 1980s debt crisis, the Asian financial crisis of the late 1990s and concerns about the social impact of globalisation, incomes rose consistently in all income bands throughout the period, except in 1987-90 (when incomes fell in all except the lowest band) and 1993-6 (when they fell in the lowest band, with limited increases in all other bands). (See Figure 3.)



On closer inspection, however, there are some issues for concern. First, the rate of increase of incomes has been very erratic, with growth greater than 1.6% pa in only three three-year periods (1981-4, 1993-6 and 2002-5).

Second, given their very low initial levels, the average rate of growth suggests a very long timeframe for poverty eradication. The average per capita income for the bottom decile was $0.68 per day in 2010, and the average growth rate 2.0% pa between 1981 and 2010 (1.4% pa from 1984), indicating a doubling every 35 years (51 years based on the post-1984 growth rate). It would thus take nearly this long for the average income of the poorest 10% to reach the $1.25 a day poverty line, and considerably longer – well beyond the life expectancy of a child born into poverty today – for those at the bottom of this band. This issue is discussed in greater detail in the final section of this paper.

Third, the higher income bands performed significantly better than the lower bands. Over the period as a whole, the growth rate of incomes rises consistently across income bands, from across bands, from 2.0% pa in the lowest to 2.4% pa in the highest. More worryingly, this trend strengthened considerably through the period. While incomes in the lower bands grew much faster in the lower than in the lower bands in 1981-4, growth rates were broadly similar between 1984 and 1996, after which the gap between the two widened at an accelerating rate until 2008, moderating only partially in 2008-10. (See Figure 4.) From 1996, the average income growth again increased consistently across deciles, from 1.5% pa in the lowest decile to 3.7% pa in the fifth decile.

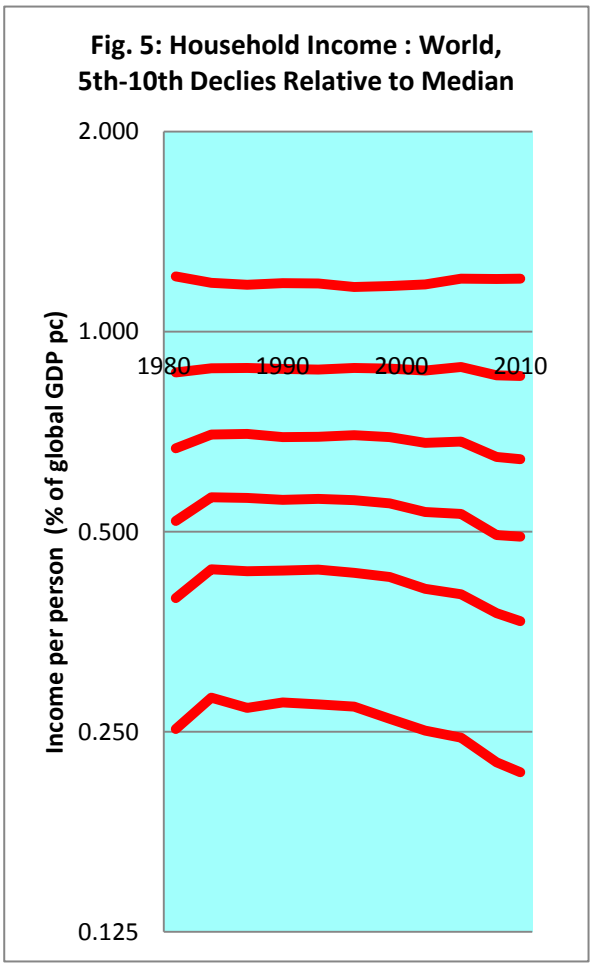


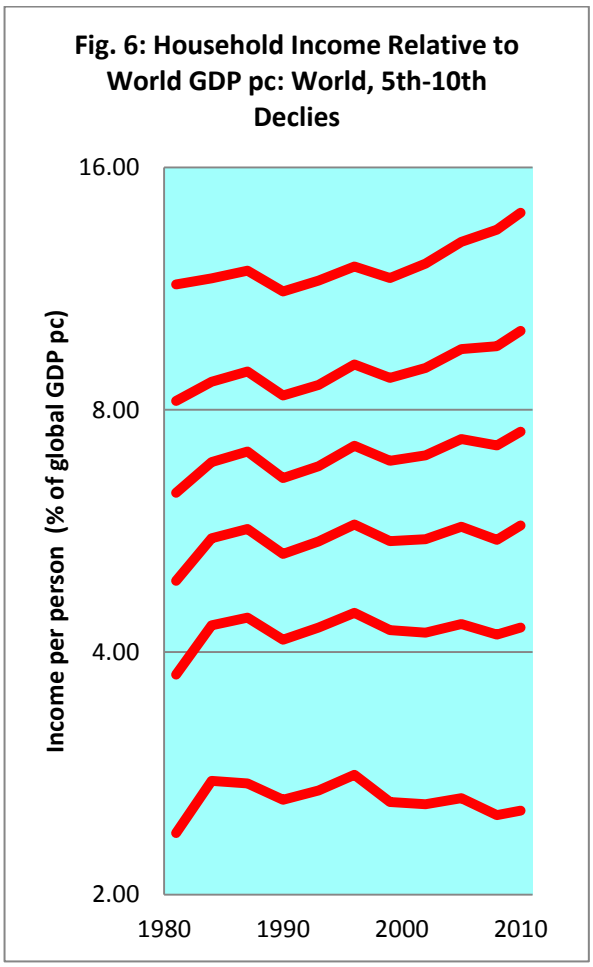
Moreover, this pattern of divergence is extremely consistent: the average income in every income band fell relative to every higher band in virtually every three-year period since 1996 (except for increases of 0.09% in the tenth decile relative to the ninth between 1999 and 2002, and of 0.04% in the seventh decile relative to the sixth between 2002 and 2005).

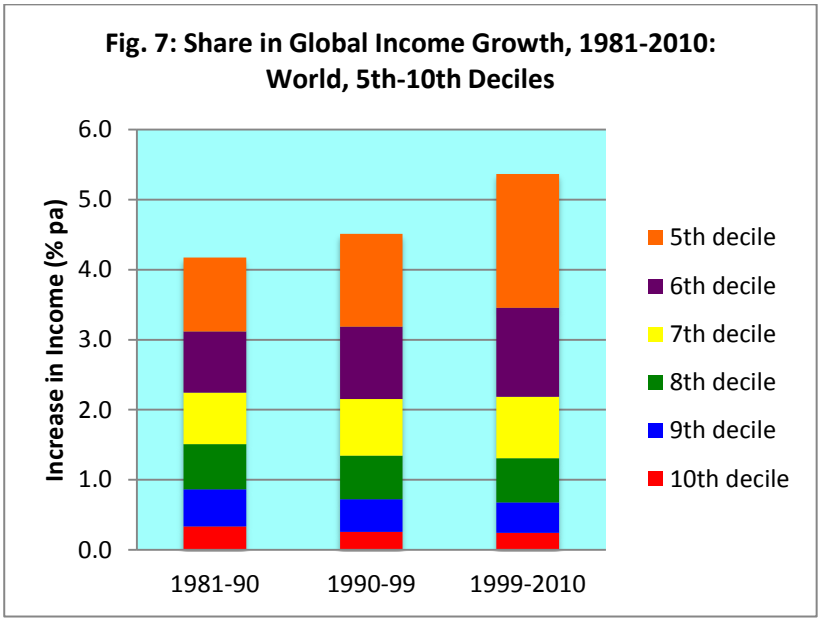
This gives rise to a marked divergence in the lower 60% of the world income distribution from 1996 onwards. The ratio of the average income in the lowest decile to the median rose from 0.252 in 1981 to a consistent 0.281 in 1984, declining slowly to 0.273 in 1996 (at 0.0007 pa), then much more quickly to 0.217 in 2010 (at 0.004 pa) in 2010. (See Figure 5.)

A third issue for concern is the relationship between global economic growth and the incomes of the poorer 60% of the world population. As discussed in *Growth Isn't Working* (Woodward and Simms, 2006), this is a critical issue in the context of potential carbon constraints to future global growth. While the picture presented by the estimates here is less pessimistic than in those presented in *Growth Isn't Working*, it is far

from positive, showing a significant reduction in per capita incomes in the bottom two income deciles relative to global GDP per capita since 1996. (See Figure 6.) The picture for the lowest decile is particularly bleak, with a cumulative reduction of 9.8%.

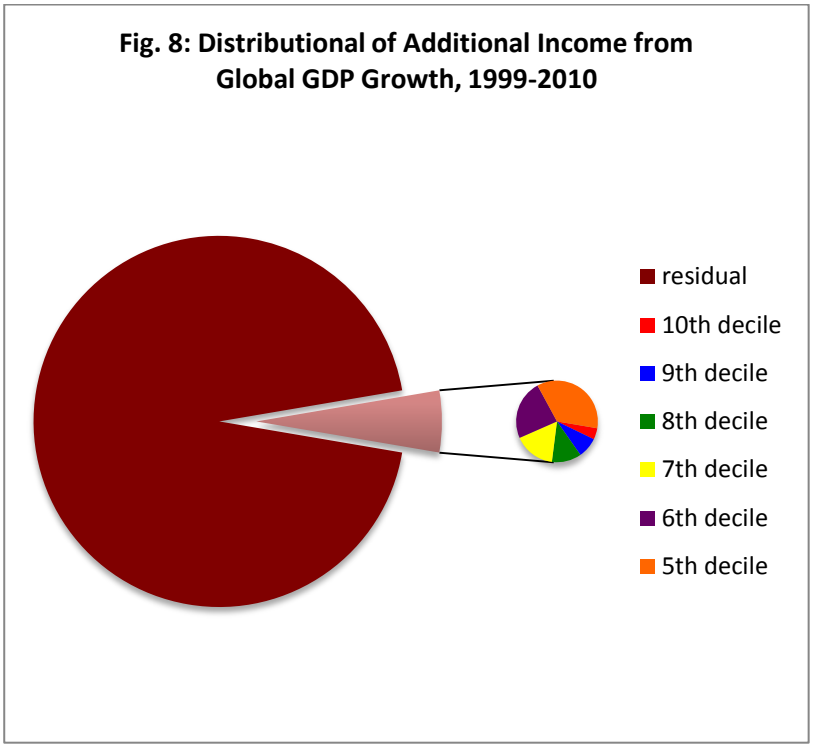






Unsurprisingly, the share of the proceeds of global growth attributable to the lower income deciles is very small. It also shows a significant regressive shift since 1981: the share of the bottom decile in the additional income generated by global growth fell by nearly a third, from 0.33% in 1981-90 to 0.24% in 1999-2010, with somewhat smaller falls (from 0.53% to 0.44% and from 0.65% to 0.63%) in the eighth and ninth deciles, while the shares of the fifth, and particularly the sixth and seventh deciles increased. (See Figure 7.)

To set this in context, the poorest 30% of the world population received just 1.2% of the additional income generated by global GDP growth between 1999 and 2008, the poorest 60% received 5.0%, while 95% accrued to the richest 40%. (See Figure 8.)



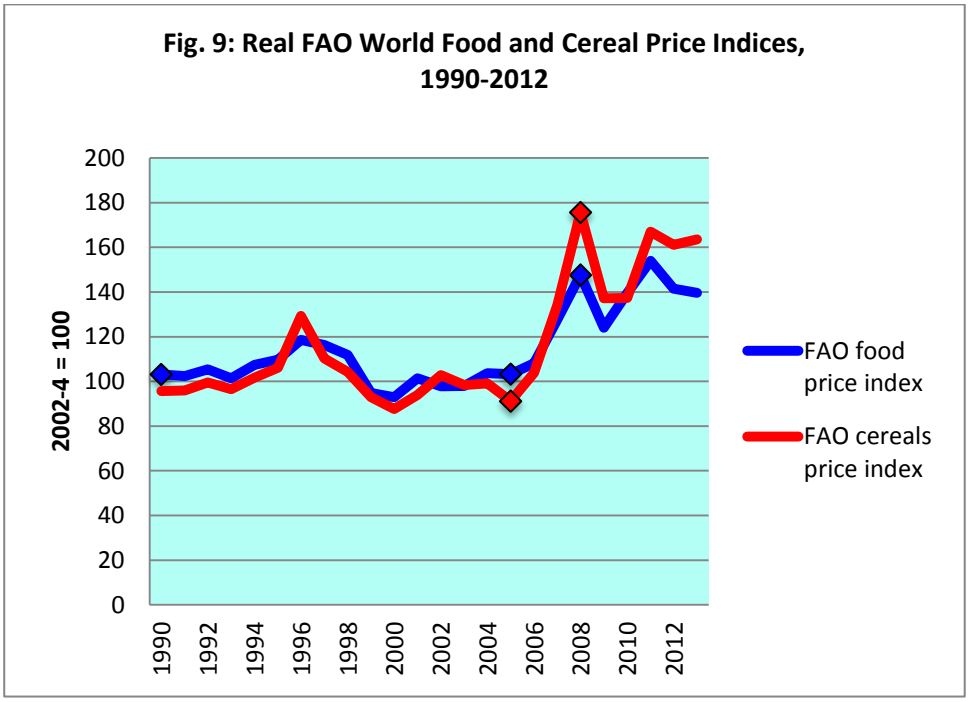
The fourth issue for concern is the effect of the major increases in food prices since 2005. Global cereal prices rose very considerably between 2005 and 2008 (see Figure 9); and this will inevitably have been reflected in the prices paid for basic foods by poor households. This renders the interpretation of poverty estimates based on PPP exchange rates extremely problematic, as these are based on GDP weightings (ie overall consumption) rather than on the consumption patterns of poor households themselves (Pogge and Reddy, 2006; Woodward and Abdallah, 2010).

PovcalNet income estimates attempt to adjust for this effect.

> The steep rise in food prices in 2008 has been taken account by re-weighting the CPI whenever possible to accord with the food share in a neighborhood of the poverty line. We have done this for as many countries as these data are available, and where food price increases much faster than the national inflation rate.[3]

While helpful, however, this would appear inadequate because it does not differentiate between the *composition* of food expenditure at very low incomes and in overall consumption. The caveat "where possible" suggests that this adjustment is only partial; and, while this is not specified, it also seems likely that this adjustment is based on consumption at a single poverty line (presumably $1.25 per day), which would make the adjustment still more inadequate at lower incomes.

[3] http://iresearch.worldbank.org/PovcalNet/index.htm?0,2, accessed March 2013.



Source: FAO (2013).

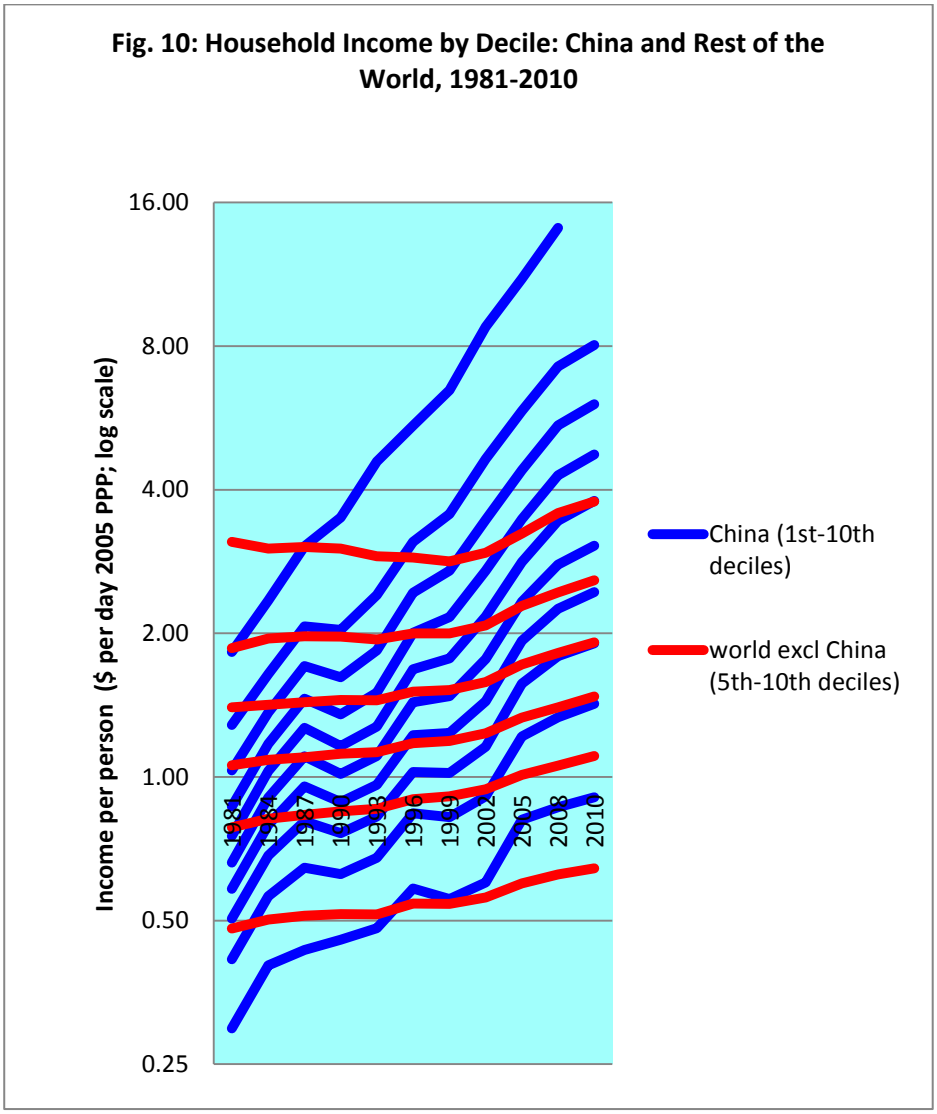
More fundamentally, the raw staple content of food expenditure (eg of maize or rice) at very low incomes is typically much greater than at higher incomes, where diets are more varied and various forms of value-added comprise a much greater proportion of food expenditure. Between 2005 and 2008 (the points highlighted in Figure 9), cereal prices rose by 92.6% in real terms, more than double the increase in overall food prices (42.9%). It also seems inevitable that non-commodity value-added costs (for processing, packaging, transportation, trading, etc), other than fuel costs, will have risen considerably more slowly then food commodity prices, and more particularly cereal prices, during this period.

The clear implication is that changes in overall national food price indices between 2005 and 2008 will considerably under-estimate the increase in food costs to the poorest households, rendering the adjustment of the PovcalNet poverty estimates inadequate.

The scale of cereal price changes, and of the difference between the increases in cereal and overall food prices, in this period suggest that this effect is likely to be very considerable. (See, for example, Ivanic and Marti, 2008; Ivanic *et al,* 2011.) While the scale of the discrepancy not readily quantifiable, the picture in terms of the actual living standards of those at the bottom of the global income distribution is thus substantially more negative in recent years even than the above estimates suggest. This should be borne in mind throughout the discussion which follows.

The final issue for concern is the geographical distribution of changes in income. It is generally recognised that the reduction in poverty in recent decades is disproportionately attributable to increasing incomes in China. Due to the combination of a very large population and very low estimated incomes, developments in China inevitably have a disproportionate impact on any global estimates of changes in poverty over the last three decades; and the markedly different pattern of income growth in China as compared with the rest of the world during this period means that assessing the evolution of decile income averages for the world as a whole may give a misleading (and, overall, artificially optimistic) impression of developments in poverty across the rest of the developing world.

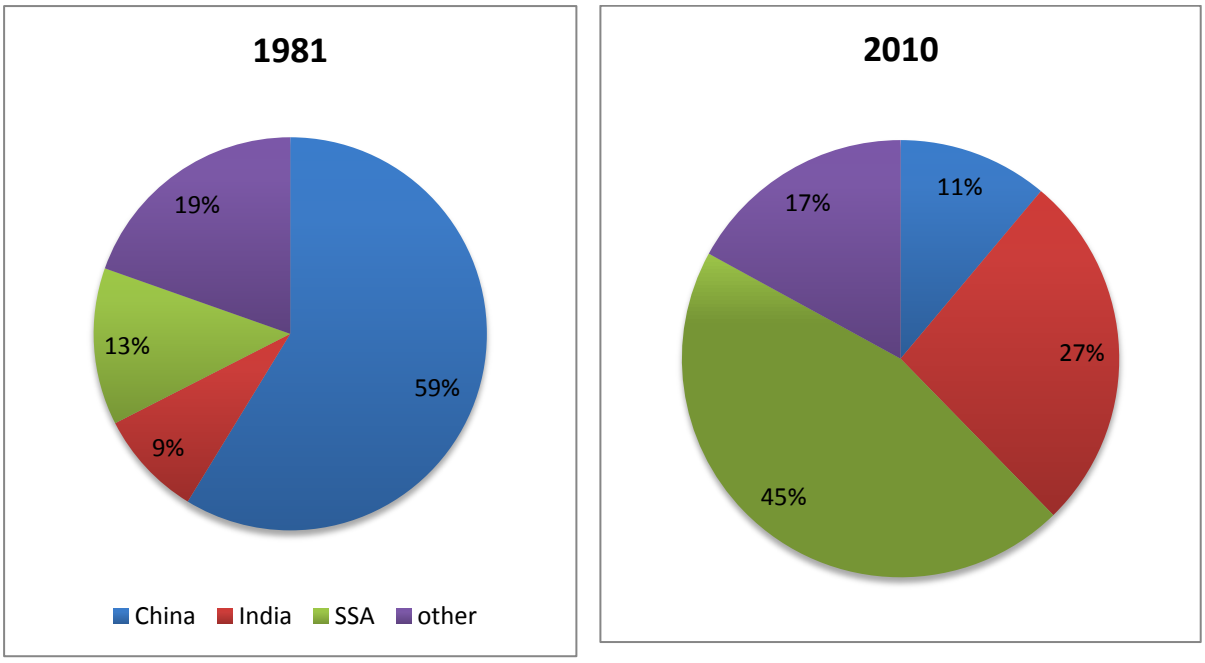
The scale of the difference in trends between China and the rest of the world – and its potential implications for the above analysis – can be gauged from Figure 10. While incomes in the second to tenth deciles of the Chinese population has been much faster than the growth rate of global GDP per capita since 1981, that of the 6th-10th income deciles in the remainder of the world has been much slower.



This has two implications. The first is that the above estimates of incomes for world income deciles are likely to over-state rates of increase for the world excluding China by a substantial margin, and thus to present a much more optimistic (or less pessimistic) picture than is representative of the developing world outside China.

The second, however, with regard to the finding of an increasingly regressive trend in the distribution of income among the bottom six deciles of the world population, is a more nuanced one. As incomes in China have grown rapidly relative to incomes in the rest of the world, so the proportion of the lowest income bands of the world population represented by China has fallen markedly. (See Figure 11.) By 2008, China accounted for only 11.5% of the bottom world income decile, compared with 58.7% in 1981. By contrast, the combined share of Sub-Saharan Africa and India increased from 25.0% to 73.1%.

Figure 11: Geographical Composition of World Lowest Income Decile, 1981 and 2008



Given the much faster increase in incomes in China than in the remainder of the developing world (taken as a whole) during the period under consideration, it thus seems likely that the increasingly regressive distributional shift observed above may be attributable, partly or wholly, to a shift in the composition of lower income bands from Chinese households with rapidly rising incomes to non-Chinese households with more slowly increasing incomes.

In any event, it seems clear that we cannot adequately understand the evolution of the incomes of poor households without separate consideration of developments in China and in the remainder of the developing world.

4. Evolution of incomes by decile: world excluding China

The evolution of average incomes for the fourth to tenth income deciles of the population of the world excluding China are shown in Figure 12. The most striking feature is the much slower income growth across all income bands when China is removed from the analysis. While this shift appears relatively limited in the lowest band (apart from the much more limited increase in 1981-4), it is nonetheless substantial: between 1984 and 2008, the average income rose by an average of 0.9% pa in the lowest decile for the world excluding China, compared with 1.5% pa for the world as a whole.

The difference is still more marked in the higher income bands, particularly in the first half of the period. For the fifth decile of the world population as a whole (the highest considered here), average income actually fell between 1981 and 1999, although it grew relatively strongly in 2002-10, giving rise to an average rate of increase for the period as a whole of 0.7% pa. This compares with an average increase for the corresponding (ie fifth) decile of the world as a whole of 2.5% pa.

Comparing deciles with similar incomes in 1981 gives rise to the same pattern. The sixth, seventh, eighth and ninth deciles of the world excluding China have approximately the same initial incomes respectively as the fifth, sixth, seventh and eighth deciles of the world as a whole. Comparing income growth in 1981-2010 for the corresponding deciles demonstrates the dramatic effect of removing China from the analysis. The annual income growth rates for the world excluding China bands are respectively, from the highest (with those for the corresponding bands in terms of initial income for the world as a whole in brackets) 1.2% (2.3%), 1.2% (2.4%), 1.1% (2.5%) and 1.1% (2.5%).

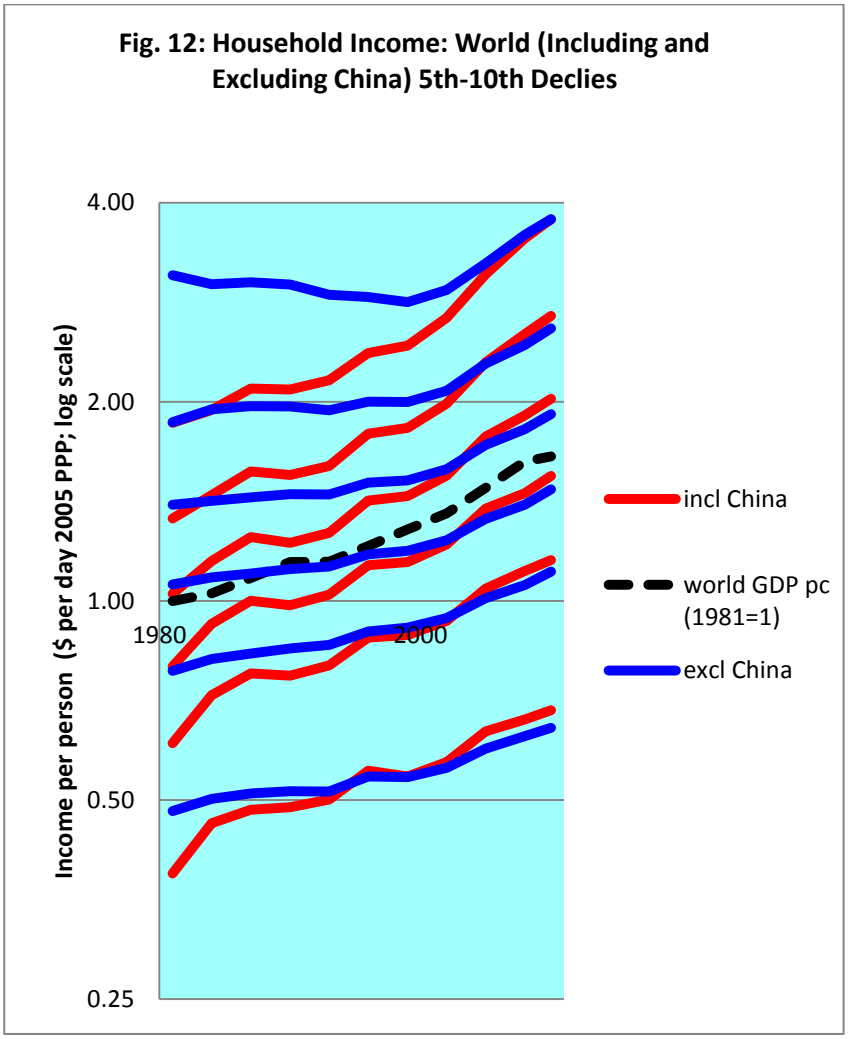


Fig. 12: Household Income: World (Including and Excluding China) 5th-10th Declies
Income per person ($ per day 2005 PPP; log scale)
4.00
2.00
1.00
0.50
0.25
1980
2000
incl China
world GDP pc (1981=1)
excl China

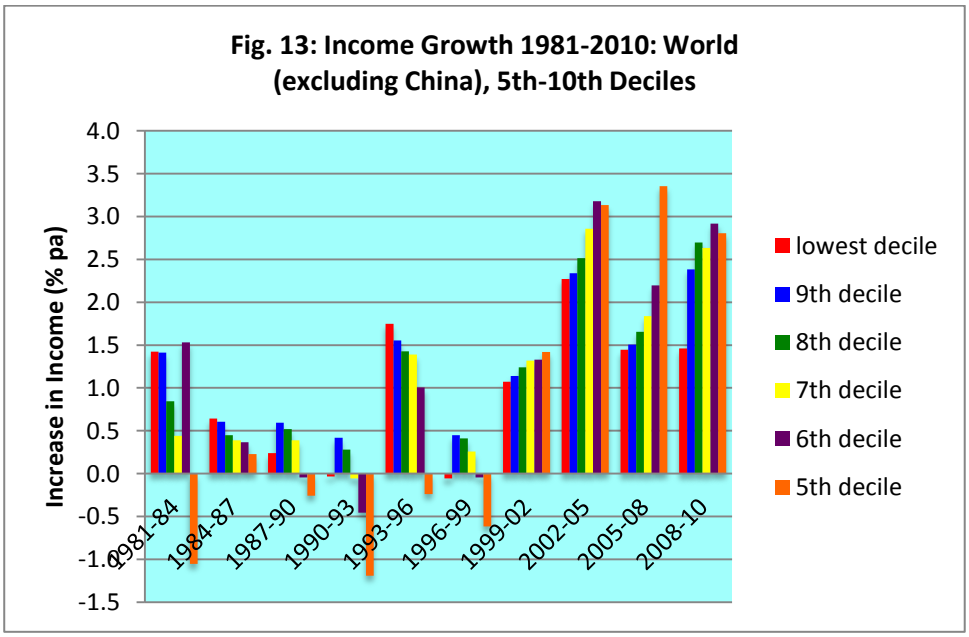
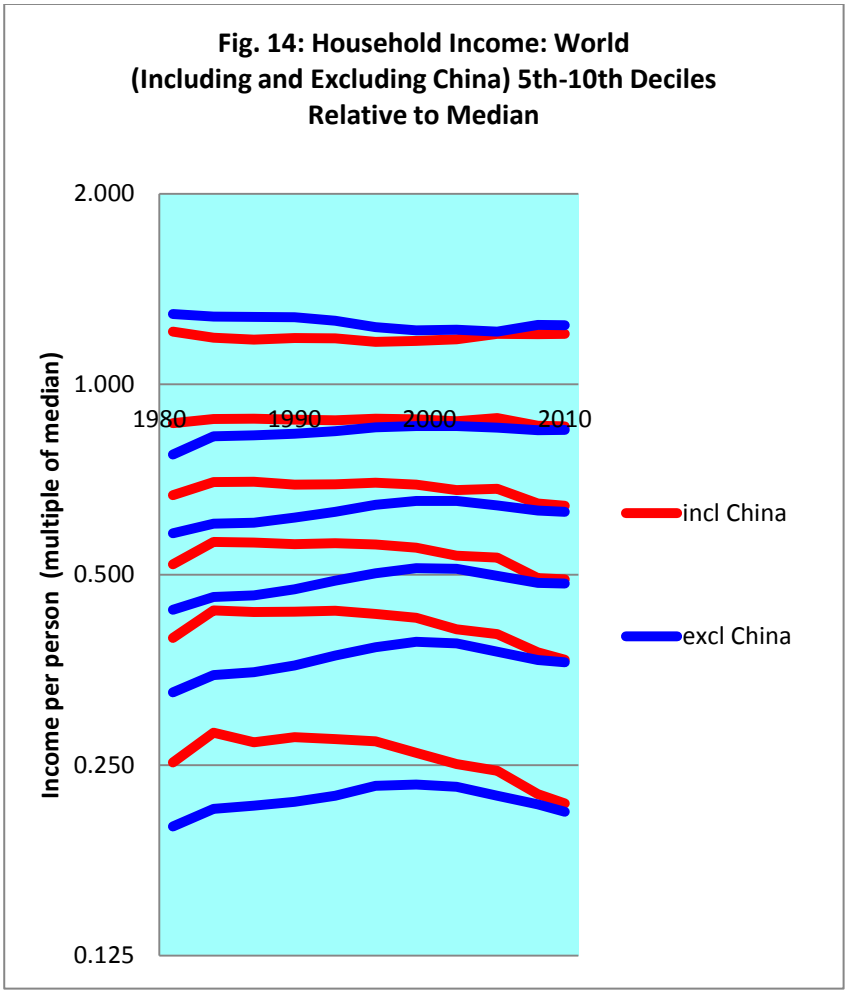


Fig. 13: Income Growth 1981-2010: World (excluding China), 5th-10th Deciles
Increase in Income (% pa)
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
1981-84
1984-87
1987-90
1990-93
1993-96
1996-99
1999-02
2002-05
2005-08
2008-10
lowest decile
9th decile
8th decile
7th decile
6th decile
5th decile

Considering income growth rates by decile over the course of the period to 2010 also suggests two more positive conclusions. First, there has been a significant acceleration in income increases since 1993: between 1981 and 1993, average growth rates ranged from -0.6% to +0.8%; between 1993 and 2008, they were in a much narrower range, from +1.0% to +1.1%. (See Figure 13.) Second, the initial convergence in incomes persists much longer before giving way to divergence, continuing until 1999 rather than 1984 (Figure 14). The rate of divergence is also slower, even after 1999.



However, the much slower income growth rates than for the world as a whole also signify a much more consistently negative pattern relative to global GDP per capita. (See Figure 15.) In contrast with the results for the world as a whole, the average incomes of all deciles decline continuously as a percentage of World GDP per capita, with a marked downward convergence until 1999, after which there is a partial levelling-off in the upper deciles (and a more general upward tick in 2008-10, as the marked slowdown in global GDP growth was not reflected in incomes in the 5th-10th deciles). The average income in the tenth decile declined from 3.0% to 2.4% of world GDP per capita between 1981 and 2008, and that of the fifth decile from 19.1% to 13.6%.

If we compare instead with per capita GDP for the world excluding China, a broadly similar pattern emerges. (See Figure 16.)

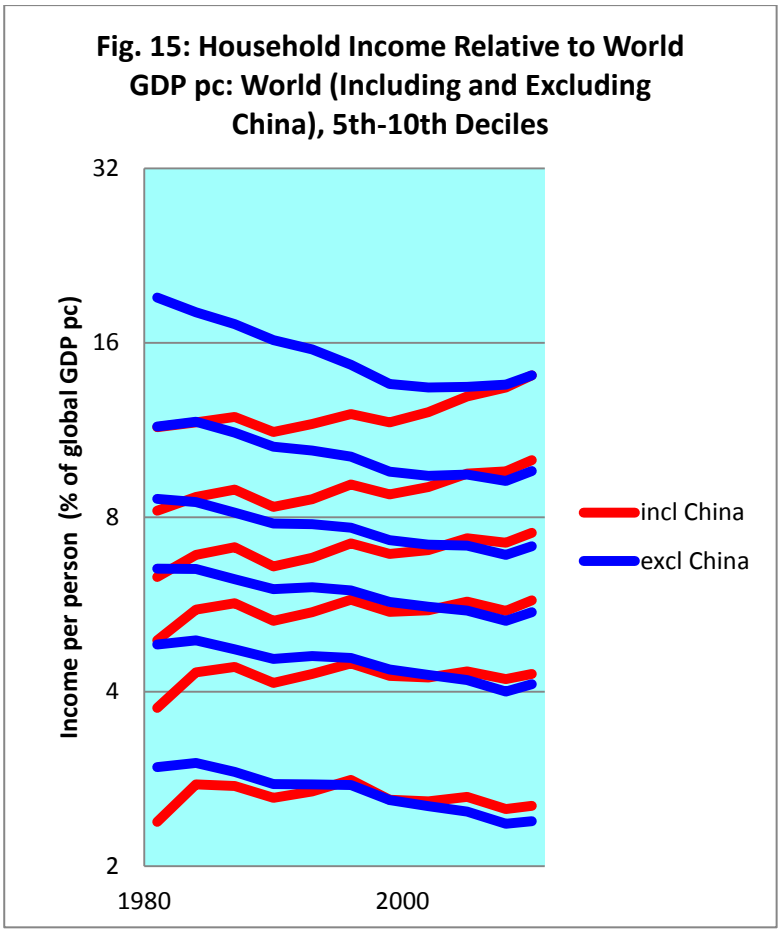


Fig. 15: Household Income Relative to World GDP pc: World (Including and Excluding China), 5th-10th Deciles
Income per person (% of global GDP pc)
32
16
8
4
2
1980
2000
incl China
excl China

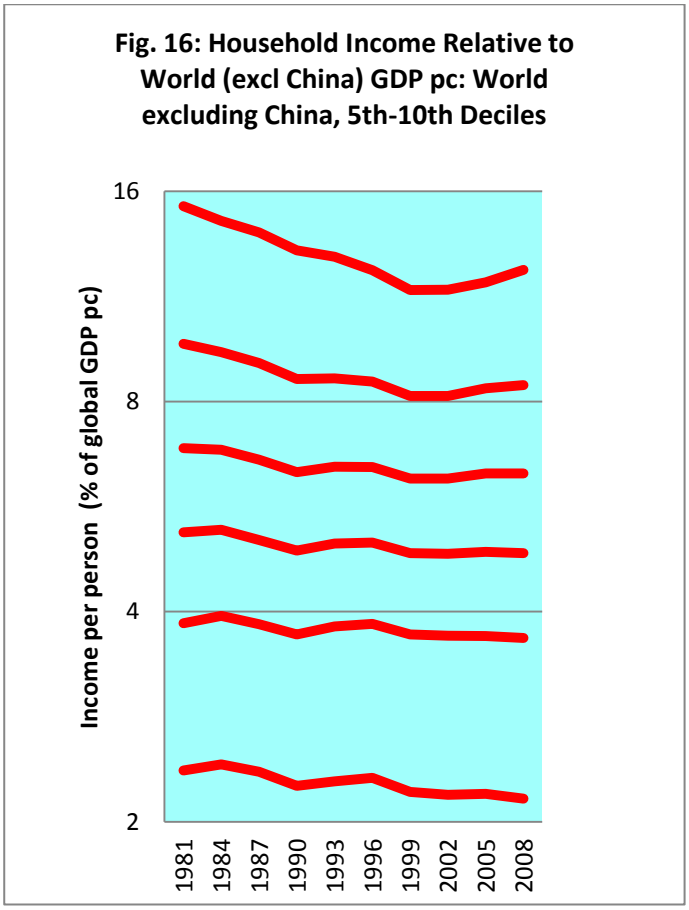
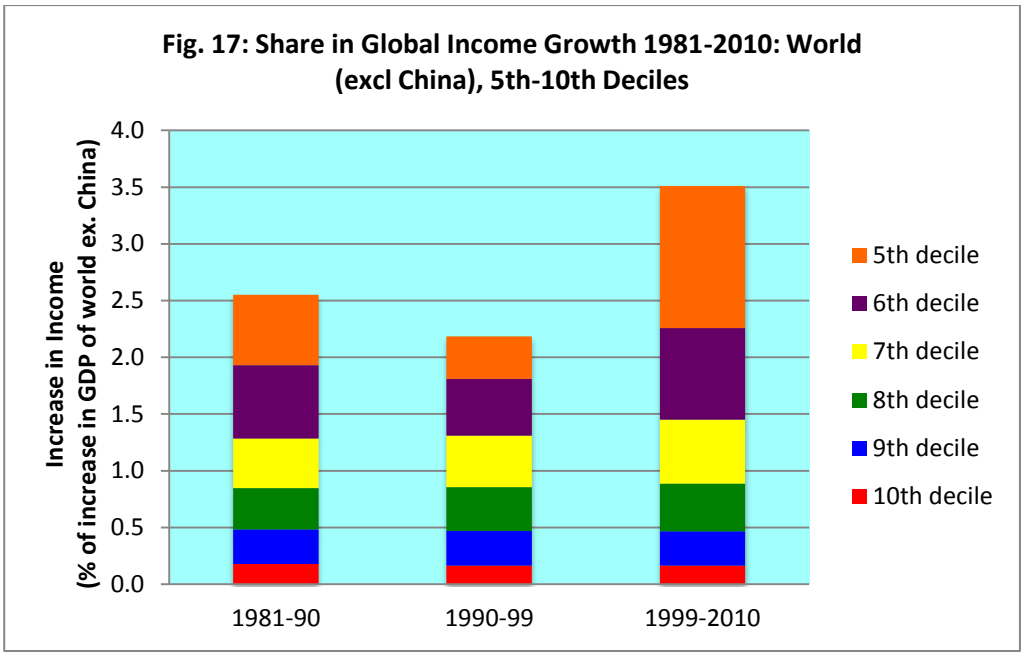


Fig. 16: Household Income Relative to World (excl China) GDP pc: World excluding China, 5th-10th Deciles
Income per person (% of global GDP pc)
16
8
4
2
1981
1984
1987
1990
1993
1996
1999
2002
2005
2008



The decline in incomes relative to global GDP per capita is also reflected in still smaller shares for the lower income deciles in the additional income generated by global growth. (See Figure 17.) Even in the period from 1999, the poorest 60% of the population of the world excluding China (equivalent to 48.0% of the world population at the end of the period) accounted for only 3.5% of the proceeds of global growth. The shares of the fifth and sixth deciles declined sharply in the 1990s, but recovered strongly after 1999, while that of the seventh decile increased slightly throughout the period. However, the shares of the lowest three deciles remained virtually unchanged overall, at just under 0.5%, with a degree of divergence: while the share of the eighth decile increased from 0.36% to 0.42%, that of the poorest decile fell from 0.18% to 0.16%.

5. "Always with us"? Implications for poverty eradication

The following discussion considers the implications of the above analysis, and specifically how long it would take to reduce poverty to zero (or near zero) if recent trends in income growth were sustained indefinitely. Two poverty lines are considered: the $1.25-a-day poverty line, as that currently most widely used; and a $5-a-day line, which might be considered a more realistic reflection of the income at which basic needs may be met, and social and economic rights minimally fulfilled.

At the simplest level, the growth rate of per capita incomes in the lowest world income decile from 1981 to 2010 (2.0% pa), if sustained indefinitely, would raise the average income for the band from its 2008 level ($0.68) to the $1.25 per day poverty line in 2041; but it would reach $5 per day only around 2115. However, this represents only the *average* income for the lowest decile. To raise the *lowest* incomes to this level – that is, to eradicate poverty by this criterion – would take considerably longer.

To assess this, we assume, *ex hypothesi*, (a) that the poorest households have incomes of half the average for the lowest decile, and (b) that their incomes increase at the same rate as the decile average. It should be emphasised that both of these assumptions err on the side of optimism. In 2008, 66m people (1.0% of the world population) had incomes below this level, and 16.7m below half of this level (although it seems likely that the PovCalNet estimates become unreliable at such low incomes). Equally, if the divergence between the income bands noted below were replicated within bands, the incomes of those at the bottom of the tenth decile might be expected to grow more slowly than the average for the band.

Nonetheless, even on the basis of such optimistic assumptions, extrapolating the rate of income growth for the tenth decile would imply as a best-case scenario that poverty would not be eradicated until 2077 – well beyond the life expectancy of those born today in the poorest countries – based on the $1.25 a day line, and 2148 based on the $5-a-day line.

As discussed above, however, the global figures are potentially misleading because of the exceptional performance of China. Clearly, the much slower income growth rates for the world excluding China than for the world as a whole imply a much longer period for the eradication of poverty. To maintain the optimistic bias, we consider the period of most rapid growth in the incomes of the poorest decile, from 1993 to 2008, excluding the slower growth of 1981-93 and the crisis (2008-10) period. It should also be noted that no adjustment is made for the potential under-estimation of food price effects in 2005-8, as discussed above, or for any slowdown in income growth after 2010 as a result of the financial crisis.

Again taking half of the mean income of the lowest decile as the minimum income, extrapolating the 1993-2008 income growth rate from 2010 would result in poverty being eradicated only in 2116 (according to the $1.25-a-day poverty line) or 2224 (using a $5-a-day line). In other words, ***even if we were to resume the pre-crisis pattern of income growth immediately and sustain it indefinitely, extreme poverty would persist for more than a century, and absolute income poverty by a broader definition for more than two centuries***. It should also be noted that the use of half the mean income in the lowest decile as a proxy for minimum incomes means that around 1% of the world population would remain in poverty by the respective definition. This represents some 85-90 million people, more than the current population of Germany or the Eastern Seaboard of the US from Maine to South Carolina.

This impasse arises fundamentally because the incomes of the poorest are growing much more slowly than global GDP: between 1993 and 2008, the average per capita income of the poorest decile of the world population excluding China increased only half as much as world GDP per capita, by 21.3% compared with 41.7%.

As noted above, 1993-2008 is the period of maximum growth in the incomes of the poorest decile. This undoubtedly reflects a number of favourable shifts in development policy during this period, as piecemeal debt cancellation limited to commercial and bilateral debts gave way to comprehensive debt reduction under the Enhanced HIPC Initiative; structural adjustment programmes gave way to poverty reduction strategy papers; the Millennium Development Goals were adopted in 2000; per capita aid receipts more than doubled between 2001 and 2008 (DAC, 2012); and increasing attention was devoted to aid effectiveness following the Paris Declaration after 2005.

By extrapolating income growth rates from the 1993-2008 period, we are thus implicitly assuming a continuation of this pro-poor shift at a similar rate into the indefinite future. This, again, is a very optimistic assumption.

Thus, even if we were to succeed in returning to the pre-crisis trajectory of the global economy, continuing the pro-poor shift in development policies and assuming no adverse effects from climate change, and on the most optimistic assumptions, poverty eradication in any meaningful sense is not a matter of decades or even generations, but of lifetimes or centuries. The prospect of poverty eradication based on a $5-a-day poverty line is nearly as far off in the future as the French Revolution and American Independence are in the past.

6. Alternative scenarios: growth, distribution and global carbon constraints

In principle, there are two ways in which the process could of poverty eradication be accelerated: by increasing the growth rate of the global economy as a whole; or by increasing the share of global growth which goes to the poorest households (that is, by changing the global income distribution to the benefit of the poorest).

However, global carbon constraints raise serious questions about the former option. The baseline scenario outlined above entails an increase in global GDP (in 2005 PPP terms) to nearly 15 times its 2010 level by the time $1.25-a-day poverty is eradicated in 2115, and 173 times its 2010 level by the time of $5-a-day poverty eradication in 2222. The contribution of population growth to this figure (based on current UN projections) would be limited in the former case, and minimal in the latter – a factor of 1.29 and 1.25 respectively. Rather, the increase is almost entirely due to growth in world GDP per capita, which would reach around $112,800 in 2115 and $1,358,000 in 2220, respectively 3.3 and 40.2 times the current figure for high-income OECD countries.

Table 2: Poverty Eradication Scenarios: Summary of Results

		Growth assumptions (% pa)		"poverty eradication"		
		poorest	GDP pc	year	GDP pc ($)	GDP (2010=1)
$1.25-a-day	Baseline	1.29	2.35	2115	112,800	14.9
	High-growth	2.20	4.00	2072	111,600	14.8
	Proportional	2.35	2.35	2068	37,800	5.0
	Pro-poor	4.20	2.35	2042	20,600	2.6
$5-a-day	Baseline	1.29	2.35	2222	1,358,000	173.3
	High-growth	2.20	4.00	2136	1,374,000	176.0
	Proportional	2.35	2.35	2128	152,580	19.8
	Pro-poor	4.20	2.35	2076	45,500	6.1

If global GDP per capita were to increase at 4% pa, keeping other assumptions as in the baseline scenario (and maintaining the ratio between GDP growth and the growth rate of incomes in the poorest decile as in the baseline scenario), the time required to eradicate poverty would be reduced by 43 years at $1.25-a-day and by 64 years at $5-a-day (although eradication would still occur only in 2072 and 2136 respectively).

However, since the ratio between the income growth rate for the poorest decile and global GDP growth would (*ex hypothesi*) remain unchanged, and world population is projected to be relatively constant between 2072 and 2222, the level of both world GDP and GDP per capita at the point of poverty eradication would remain virtually unchanged from the baseline scenario. Thus the time available for technological adaptation to neutralise the environmental impact of such high levels of consumption and production would be greatly reduced, raising serious questions about the viability of this scenario, particularly in a context of increasingly binding global carbon constraints.

In this context, unless we simply dismiss the problem by assuming unlimited potential to reduce the carbon intensity of global GDP through technological changes, the relationship between poverty reduction and global production and consumption is critical. From this perspective, any scenario which fails to increase the growth rate of the income of the poorest decile relative to that of global GDP appears, quite simply, untenable from the perspective of sustainable poverty eradication.

Poverty eradication based on a $5-a-day poverty line requires an increase in the incomes of the 62.3% of the world population below this income level in 2010 by a total amount equivalent to around $4.57 trillion, or 6.8% of global GDP (at 2005 PPP). In the baseline and high-growth scenarios – which, as noted above, implicitly assume an indefinite continuation of the pro-poor shift in development policy at the 1993-2005 rate – this not only takes between 123 and 209 years, but also requires an increase in global production and consumption by a factor of around 175 from the 2010 level, raising global GDP per capita above $1.3m. Thus closing a poverty gap of $4.5 trillion requires an increase in global GDP in excess of $11,500 trillion.

In a global economy faced by increasingly binding carbon and other environmental constraints, an approach which entails increasing production and consumption by more than 2,000 times the additional income we are seeking to generate, in order to achieve our objective at a glacial pace, can at best be described as grotesquely inefficient.

It also seems highly likely to be unviable. Even allowing for two centuries of technological change (and improvement in global economic and environmental governance), it would seem extremely optimistic to assume that that an increase in global production and consumption on this scale could be achieved within global carbon constraints and without irreparable damage to global ecosystems which are already under

considerable strain even at current levels of economic activity (Millennium Ecosystem Assessment, 2003). Moreover, it is precisely the poorest who stand to suffer first and worst from the consequences of ecosystem failure.

Unless we assume that global economic growth has no adverse environmental effects, or that these do not affect the poorest, such consequences must be off-set against the projected increase in their incomes. For a given trajectory of the carbon intensity of global GDP, faster economic growth will generate additional carbon emissions, accelerating the increase in atmospheric carbon concentrations, and thus the process of climate change. This effect can only be neutralised if a given percentage increase in GDP in itself *induces* an equivalent proportional reduction in carbon intensity – that is, that a 1% acceleration in global growth necessarily brings about a 1% faster reduction in carbon intensity than could have been achieved at the slower growth rate. There is no obvious reason to believe that this is the case, and it would seem difficult to construct a rationale for making an assumption to this effect. Essentially the same argument applies to other ecosystem services.

At best, this means that the acceleration in poverty eradication as a result of faster global growth will be significantly less than indicated above. It also raises a real possibility that faster growth could be counter-productive. Even if we assume (contrary to the general view) that the poorest suffer no more than the average in terms of the impacts of climate change, a positive net effect of faster global growth on the bottom decile requires that the impact of an additional $100 of global production, through its effect on carbon emissions, is less than about $2 (equivalent to the share of the poorest decile in global growth between 1993 and 2008 relative to their *pro rata* share of 10%). This would again seem an extraordinarily optimistic assumption – but without it, we are forced to the conclusion that, from a poverty perspective, it is only such increases in GDP as contribute directly either to poverty reduction or to reducing the carbon intensity of global production and consumption that are justified.

The key issue is thus, not the growth rate of the global economy, but rather the relationship between global GDP and the incomes of poor households, and that between global GDP and global carbon emissions. In effect, this amounts to shifting away from global growth as an objective in itself.

The implications of such a shift can be demonstrated by comparing the high-growth scenario outlined above with scenarios which vary the relationship between global GDP growth and the growth rate of incomes in the poorest decile at the baseline rate of global growth.

If we consider a “proportional” scenario, in which the incomes of the poorest decile rise at the same rate as global GDP per capita, while maintaining global growth at the 1993-2008 rate, the effect on poverty eradication is marginally better than that of “high-growth” scenario: poverty is eradicated in 2068 at $1.25-a-day and 2128 at $5-a-day (respectively four years and eight years earlier). The key difference, however, is the increase in global GDP which is required, which is reduced by a factor of around three in the former case and nine in the latter. This clearly increases the possibility of environmental sustainability considerably.

In the final scenario, we take a step further in this direction, and assume a pro-poor bias in global growth equivalent to the anti-poor bias in 1993-2008, changing the ratio of income growth for the poorest decile to that of global GDP per capita from 0.55 to 1.82. This results in poverty eradication based on the $1.25-a-day poverty line by 2042 (global GDP increasing by a factor of 2.6), and $5-a-day by 2076 (global GDP increasing by a factor of 6.1), respectively 30 and 60 years sooner than in the “high growth” scenario, and 73 and 146 years sooner than in the baseline scenario. The scale of the increase in global GDP required is reduced by a factor of more than 28.

This raises the question of how such a shift could be achieved. As noted above, the relationship between the incomes of the poorest decile and global GDP between 1993 and 2005, unfavourable as it was, itself reflected a marked improvement in the primary instruments for poverty reduction within the current model of development – aid, debt cancellation and the economic policies promoted by the IMF and World Bank in low-income countries. If we are to go substantially further than was achieved by these changes, this implies a much more fundamental shift: a reassessment of the current models of the global economy and of development in terms of their impacts on the poorest, and of their long-term viability and sustainability in a carbon-constrained world, and a real willingness to make fundamental changes in both to maximise their poverty impact and sustainability.

This challenge extends far beyond the scope of the present paper; but, if we genuinely want to eradicate poverty in a meaningful sense within a reasonable timeframe – decades rather than centuries – then it is a challenge we must address as a matter of urgency.

7. Conclusions

Even on the most optimistic assumptions, and even if we sustain the rate of increase in attention to poverty seen since the early 1990s, relying on global economic growth seems an almost certain route to ensuring that the poor are, indeed, "always with us": such benefits as may trickle down from global growth to the poorest will almost inevitably be countered by the adverse effects of climate change and the costs of adaptation. The only viable route towards eradication of poverty, by a meaningful definition and within a remotely reasonable timeframe, is to improve the relationship between global economic growth and poverty – and to do so much faster and more effectively than has been achieved over the last 20 years.

This means shifting our attention from global economic growth itself, and towards improving the distribution of the benefits of global production and consumption. The instruments available within the current global economic model achieved a marked improvement between 1981-93 and 1993-2008; but even this has proved inadequate to make the prospect of poverty eradication a realistic one, at least before the 23rd century, and, in all probability, ever.

If we are to make poverty eradication a real possibility, if not within our own lifetimes, at least within those of our children or our grandchildren, there seems little choice but to move beyond limited "add-on" solutions such as aid and debt relief to reassess fundamentally our whole approach to development and to the global economy, and to make such changes as may be required to make it fit for purpose.

Annex: Data Sources and Methodology

All estimates relating to income bands are the author's estimates based on data from PovcalNet: the on-line tool for poverty measurement developed by the Development Research Group of the World Bank (http://iresearch.worldbank.org/PovcalNet/index.htm?1). Data for GDP and population are from the World Bank's World Databank, Global Development Indicators and Global Development Finance (http://databank.worldbank.org/ddp/home.do). The methodology used to arrive at the estimates in this paper is outlined in Annex I, and detailed results in Annex II.

The following methodology was used for the estimation of decile average incomes.

1. Population data from the World Bank's DataBank database (http://databank.worldbank.org/data/home.aspx) were used to estimate the number of people constituting 10% of the world population.
2. Using the World Bank's PovcalNet database (http://iresearch.worldbank.org/PovcalNet/index.htm), an iterative process was used to identify the poverty line (in international dollars per person per month at 2005 purchasing power parity) generating the closest number of people in poverty to 10%, 20%, 30%, 40%, 50% and 60% of the population in the geographical area under consideration. (Note: in the case of the world as a whole and the world excluding China, this is based on an assumption that the number of people in the developed world with incomes below each poverty line is negligible.)
3. Based on world population, the poverty lines thus identified (L) and the corresponding poverty headcount ratio (R) and poverty gap (G), the cumulative income up to each of these poverty lines (Y_{CL}) was estimated as: $Y_{CL} = N*L*(R-G)/100$
4. The average income within each income band (in international dollars per person *per day* at 2005 purchasing power parity) was estimated as the difference between the cumulative income below each band and that below the previous band, divided by the difference between the total population below each band, multiplied by 12/365.

References

Chandy, L., Ledlie, N. and Panciakova, V. (2013) The final countdown: prospects for ending extreme poverty by 2030. Policy Paper 2013-04, Brookings Institution, Wasington D.C., April. http://www.brookings.edu/~/media/research/files/reports/2013/04/ending%20extreme%20poverty%20chandy/the_final_countdown.pdf

DAC (2012) Net ODA from DAC countries from 1950 to 2011 (updated April 2012). Development Assistance Committee, Organisation for Economic Cooperation and Development, Paris. http://www.oecd.org/dac/stats/referencedacstatisticaltables.htm

FAO (2013) Annual real food price indices, 1990-2012. Food and Agriculture Organisation, Rome. http://www.fao.org/fileadmin/templates/worldfood/Reports_and_docs/Food_price_indices_data.xls, accessed April 2013.

Ivanic, M. and Martin, W. (2008) Implications of higher global food prices for poverty in low-income countries. *Agricultural Economics* 39 (s1), pp. 405-16.

Ivanic, M., Martin, W. And Zaman, H. (2011) Estimating the short-run poverty impacts of the 2010-11 surge in food prices. Policy Research Working Paper 5633, World Bank, Washington D.C., April. http://elibrary.worldbank.org/doi/pdf/10.1596/1813-9450-5633

Karver, J., Kenny, C. And Sumner, A. (2012) MDGs 2.0: what goals, targets, and timeframe? Working Paper 297, Center for Global Development, Washington, D.C., June. http://www.cgdev.org/sites/default/files/Kenny_Karver_MDGs-FINAL.pdf

Millennium Ecosystem Assessment (2003) *Ecosystems and human wellbeing: synthesis*. http://www.millenniumassessment.org/en/Products.Synthesis.aspx.

Ravallion, M. (2013) How long will it take to lift one billion people out of poverty? Policy Research Working Paper 6325, World Bank, Washington D.C., January. https://openknowledge.worldbank.org/bitstream/handle/10986/12166/wps6325.pdf?sequence=1

Pogge, T. And Reddy, S. (2006) Unknown: the extent, distribution, and trend of global income poverty. Available at SSRN: http://ssrn.com/abstract=936772

High-Level Panel Of Eminent Persons On The Post-2015 Development Agenda (2013) A new global partnership: eradicate poverty and transform economies through sustainable development – the report of the High-Level Panel of Eminent Persons on the Post-2015 Development Agenda. United Nations, New York, 30 May. http://www.post2015hlp.org/wp-content/uploads/2013/05/UN-Report.pdf

Woodward, D. (2010) How poor is “poor”? Towards a rights-based poverty line. nef (the new economics foundation). http://www.neweconomics.org/publications/how-poor-is-poor

Woodward, D. And Simms, A. (2006) Growth isn’t working: the uneven distribution of costs and benefits from economic growth. nef (the new economics foundation), January. http://www.neweconomics.org/gen/z_sys_publicationdetail.aspx?pid=219

World Bank (2013) A common vision for the World Bank Group: Development Committee Paper Spring 2013. Document SecM2013-0122, World Bank, Washington D.C., 11 March.

A Neoclassical Curmudgeon Looks at Heterodox Criticisms of Microeconomics[1,2]

Donald W. Katzner[3]
University of Massachusetts/Amherst, USA

Abstract

The purpose of this paper is to open a dialogue with heterodox economists about what, from a neoclassical perspective, is valid in heterodox criticisms of neoclassical microeconomics. Many heterodox criticisms of neoclassical microeconomics are valid to a neoclassicist; but some are not. Examples of both are given, mostly taken from E. Fullbrook's anthology, *A Guide to What's Wrong with Economics*, (London: Anthem, 2004). And neoclassical reasons for a lack of validity where that arises are provided. Both criticisms and the judgment of their validity can be subjective or objective.

Key words: heterodox criticisms, neoclassical microeconomics, valid criticisms, invalid criticisms

Neoclassical microeconomics is a well-established intellectual enterprise that sets out to explain economic behavior at the level of individual agents, producing firms, and market complexes. Its development through the nineteenth and twentieth centuries has evidenced increasing formalization and, as is well-known, its pretentions to scientific status have been accompanied by increasing recourse to mathematical formulations. In the course of its historical development, neoclassical microeconomics has properly attracted a healthy, and at times a vigorous and productive, scholarly discourse that has raised questions regarding the methodological character of its investigations, the assumptions it employs, and the interpretation of its results. But caution is necessary in the critique and evaluation of the legitimacy of what it is that neoclassical microeconomic is all about and how it proceeds. Inadequate realization of the objectives of analytical inquiry and the generally accepted methods of its argument can, from the perspective of its practitioners, easily give rise to essentially irrelevant criticisms. The aim of this paper is to throw light on that danger by suggesting what the practitioners of neoclassical microeconomics would regard as both proper and improper grounds and methods of critique, and to present illustrative instances of where the relevant literature has provided forms of criticism that are understood by them as unsustainable. Obviously there is some subjectivism involved in distinguishing proper from improper criticism. The manner in which that subjectivism arises will be taken into account.

There is little doubt that, from the vantage point of the practitioners of neoclassical microeconomics, legitimate forms of criticism of neoclassical microeconomics, together with suggestions for tightening the logic of its content, can be made. But disquisition addressed to either of those ends, instances of which will follow, will miss its target if it remains unaware of the nature and purposes of neoclassical argument itself. That argument is in general responsive to the fact that fully exhaustive description of economic affairs is impossible. Microeconomic argument has therefore necessarily engaged in abstraction from the detailed realities of the phenomenon or state of affairs whose explanation is sought. It is often analytical in nature in that it deals with the logical relations among ideas and frequently involves the development of theories and/or the construction of models. And the explanation it produces has often assumed that economic behavior can be visualized and understood as though it derived from certain well-specified, hypothetical

[1] Martin Shubik (1970) affixed a similar title to his screed on the relevance and appropriateness of microeconomics as of 1970. He began by warning the reader that his article "… is frankly partisan and gives a biased view of what microeconomics is and what is good or bad about microeconomics in its current state" (1970, p. 405). The present paper is of the same spirit and comes with a similar warning. It is intended to open a conversation among neoclassical and heterodox economists about what is valid and invalid when criticizing current microeconomic analysis.

[2] The author would like to thank Roberto Veneziani for suggesting this paper and, along with Douglas Vickers and John Harvey, for helping with its execution.

[3] Department of Economics, University of Massachusetts/Amherst, Amherst, MA 01003, USA.

functions. Such an approach has a rationalist or *a priori* flavor[4] and clearly gives rise to what is essentially an 'as if,' or 'something like,' explanation. For example, it might be contemplated analytically that consumers make purchasing decisions as though they were optimizing an appropriately constrained utility function, or that a business firm chooses its output levels and factor combinations as though its decisions were taken against a suitably defined production function. But assuredly, real-world consumers and real-world firms do not and cannot actually have at their disposal such analytically specified functions.

Within this frame of reference economists have always thought that there was much wrong with neoclassical microeconomics. Any perusal of the history of microeconomic thought (or, for that matter, the history of economic thought in general[5]) will find numerous instances in which one theory, analysis, model, or argument was replaced by another because the one that was replaced was denigrated as incorrect or found to be inadequate to explain the phenomenon at issue. In perhaps the most startling rebuke of received 20th century microeconomic analysis, one of its founding fathers, none other than Nobel laureate J.R. Hicks repudiated his own work (the same work for which he was cited in his Nobel award) by altering his identity: "Clearly I need to change my name," he wrote in 1975. **"**Let it be understood that *Value and Capital* (1939) was the work of J.R. Hicks, a 'neoclassical' economist now deceased; *Capital and Time* (1973) – and *A Theory of Economic History* (1969) – are works of John Hicks, a non-neoclassic who is quite disrespectful toward his 'uncle'" (1975, p. 365**).** Much of Hicks' subsequently published work appeared under the newly assumed identity of its author. More recently, in a collection of essays edited by E. Fullbrook entitled *A Guide to What's Wrong with Economics*, (London: Anthem, 2004) twenty-seven economists detail their criticisms of economics in general and microeconomics in particular.

Criticisms, of course, can be valid or invalid. Valid criticisms can inspire. They can lead to changes in analytical or theoretical constructions, to their extensions, or to entirely fresh ones. Modifications and extensions of old analyses and theories and the development of new analyses and theories enhance explanation and the ability to predict. They can provide the means for communication among scholars and the basis for policy decisions. But invalid criticisms can have detrimental effects. If accepted by many, they can lead to the rejection of significant and useful analyses and theories and possibly to the development of irrelevant or 'false' ones. They can discourage communication among scholars because those who subscribe to the analysis or theory being criticized may think it a waste of time to respond to the individuals who appear to the subscribers not to understand what they are criticizing. And invalid criticisms can have serious consequences if damaging policy decisions eventually emerge from them.

1. Types of criticism

Criticisms, of course, may appear in different forms. Certain types, which might be referred to as objective in their intentions, may recognize and take full account of the methods and stated aims and claims of neoclassical microeconomic theory. But in doing so, they may call attention to imagined logical flaws or internal inconsistencies in parts of its argument or methods while containing no subjective presuppositions or preferences as to assumption content or methods of analysis. These criticisms remain consistent with the principle that a critique of a body of thought, quite apart from whether or not it is judged to be valid, should evidence an understanding of, and fully recognize the real nature, purpose, and intention of that which is being criticized. If, in such an instance, a charge of logical flaw or internal inconsistency was shown to be in error, or if that charge demonstrated a lack of understanding of the argument being criticized, its purpose, or its intention, or if acceptance of the criticism would vitiate the possibility of doing analysis as that activity is generally understood, a neoclassical's response to the criticism may be quiet straightforward in its reassertion of the validity of the criticized argument. In that way the neoclassicist would be making a clearly objective response that, as he sees things, vindicates the original construction. But it is always possible that a criticism that carries the charge of logical flaw or internal inconsistency may, in fact, be shown to be valid. In that case, the objective response of the neoclassicist would be to abandon or adjust

[4] See Prasch (1996).
[5] Indeed, much of the following through Section 1 may also be applied more generally to portions of economic analysis well beyond microeconomics.

the argument accordingly.

On the other hand, instances of criticism may arise (and the literature of criticism, as will be explored in what follows, makes their occurrence abundantly clear) where the critical attack emanates from more clearly subjective claims relating to assumption content or analytical methods. There are at least three kinds of subjective criticisms. First, a neoclassical argument could be criticized for its methodology. For example, the critic might be unwilling to accept the rationalist or *a priori* method employed by neoclassicals, asserting that an empiricist approach is more appropriate for understanding the microeconomic phenomena at issue. But a criticism that directs attention in the direction of empiricism is in danger of throwing out the objective or purpose of the analysis and lapsing into a plea for what is substantially a mere description of the phenomenon under investigation.

Second, it may be claimed that certain neoclassical assumptions are unwarranted. Such a critique may arise either as an intra-mural disagreement within the neoclassical camp, or from a heterodox or other perspective that understands the phenomenon in question or what is significant about it differently.

And third, a neoclassical analysis could be faulted for being unable to achieve its own stated goals. By that is meant that the analysis is criticized for not doing what it claims it sets out to do. For example, an analysis that purports to explain a certain economic behavior may be censured for not, in the critic's opinion, providing a meaningful or pertinent explanation of that behavior.

As a combination of the second and third types of subjective criticism, it might be held that the particular assumptions and/or construction in view remove the analysis sufficiently far from practical affairs and empirical reality as to render it irrelevant to the real microeconomic world. To illustrate, it may be contended that the widely-accepted assumption of perfect certainty takes neoclassical microeconomic theory quite outside the bounds of possible empirical actuality. Such a critique carries the potential of propelling analytical attention to a more robust development of other competitive forms, as well as to stimulating discussion across the neoclassical-heterodox divide. For, in a fundamental sense, that kind of criticism drives to the heart of neoclassical assumptions, analytical methods, and purported links to the real economic world. The neoclassical response to subjective criticisms may itself be subjective in the sense that it acknowledges that differences exist between the neoclassicist and his critic as to the realism, degree of potential usefulness, and justifiability of the scheme under critique. Clearly, then, neoclassical and heterodox economists may arrive at different judgments regarding the validity of subjective criticisms. And it is here that a dialogue between neoclassical and heterodox economists could benefit both sides and the profession at large by leading to a deeper understanding and more significant analyses of the issues involved.

Thus criticisms of neoclassical microeconomic theory as well as the neoclassical response to them can be either objective or subjective. To throw further light on the criticisms, responses, and the issues involved, this paper will examine in what follows a number of criticisms of neoclassical microeconomic theory that have been advanced by heterodox economists against various aspects of the neoclassical tradition. The specific criticisms selected for consideration are presented as a convenient means of drawing attention to more prominent and popular criticisms that are frequently offered from heterodox economic standpoints. The views represented are broadly shared. Because they are typical of the elements of a wider landscape, there is little harm in focusing on several of the essays in Fullbrook's anthology in which they appear. There are many valid criticisms of microeconomics in this anthology, and several examples of them are presented here without much evaluative comment. Unfortunately, the anthology also contains, from a neoclassical vantage point, examples of invalid criticisms of microeconomic analyses and theories. Four examples of the latter will be considered. In each case the reasons for the lack of validity are described. Additional criticisms, mostly invalid, not present in Fullbrook's book appear, with considerably less discussion, in footnotes. Understanding why the criticisms singled out for discussion here are invalid will hopefully pave the way to more solid and convincing criticisms of neoclassical microeconomics in the future.

It is evident that some valid criticisms will be more consequential in terms of the acceptability or usefulness or relevance of the criticized object than others. There could also be disagreement within the neoclassical camp on the importance of some valid criticisms or their relevance to the purpose of the criticized object. Nevertheless many analyses, theories, and models may continue to be embraced in full recognition and acceptance of the valid criticisms leveled against them because there are no reasonable

alternatives and because, in spite of the criticisms, the assumption content they contain seems to provide plausible approximate explanations of the phenomena under consideration.

2. Examples of criticisms likely to be judged valid by neoclassicals

Now many criticisms of neoclassical microeconomics, including, as previously suggested, several of those in Fullbrook's anthology, would be judged from a neoclassical perspective to meet standards of validity. For example, there would be little quarrel with Hicks' complaint (1976) that much of economic analysis is "out of time" because time appears in those analyses as reversible when, in fact, real historical time is irreversible.[6] That is, the appearance of time in the analyses Hicks refers to is far removed from the time that is actually experienced in human life.[7] Nor would a neoclassicist likely disagree with G.M. Hodgson who asserts (2004, p. 58) that the isolated individual who often appears alone without reference to an institutional structure is not always viable as an analytical starting point for economic model building. Institutions as well as individuals have to be accounted for from the outset. In particular, since all individual behavior takes place in an institutional context (Hodgson, 2004, p. 65), the institutional background in which that behavior occurs has to be a part of the assumption content of any model that purports to explain that behavior.

Three additional examples of similarly valid criticisms appearing in Fullbrook's collection of essays are as follows: First is H. Stretton's (2004, p. 20) clear statement of the inappropriate limits of some economic analyses: "Economic activity depends, in varying degrees, on its surroundings: on natural resources, law, culture, experience, know-how, mutual trust or distrust, and so on. It is part of life as a whole. It can be affected in many ways by the political, intellectual, social, and global conditions around it." The boundaries of economics need to be expanded to account for instances in which such interdependence can be judged to be significant.

Second, P. Ormerod argues that, "The single most restrictive assumption of conventional economics is that the tastes and preferences of individual agents – whether people or firms – are fixed" (Omerod, 2004, p. 41). In many cases, this assumption is not reasonable since "… the behaviour of an agent can be … altered by the behaviour of other agents" (Omerod, 2004, p. 41). "In models in which tastes and preference can vary according to the actions of others, there is no point at all in looking for optimal rules of behaviour. Instead, we are seeking plausible rules of thumb for agents to follow" (Omerod, 2004, p. 45). Certainly, Ormerod is not suggesting that an assumption of fixed preferences would not be appropriate under certain circumstances such as the purchase of commodities from a supermarket when the behavior of others has minimal influence.

Third, consider the comments of M.A. Bernstein: In mainstream economics, competition (whether perfect or imperfect) is in most instances a state rather than a process.[8] That state, with its unrealistic assumptions, is considered a first approximation of reality that serves as a starting point for further more realistic investigation. But the later investigation is seldom undertaken (Bernstein, 2004, p. 34). And because mainstream economics characterizes competition as a state rather than a process of economic change, "… it rather frames a way of thinking about competitive behavior that forever and completely prevents a full understanding of the [process] phenomenon itself" (Bernstein, 2004, p. 35). This does not say, nor does Bernstein imply, that the concept of competition as a state does not have its uses.

The first four of the above criticisms are focused on an absence of realism in the objects or forms of analysis being criticized. The last draws attention to an inadequate methodology. Each is directed toward specific assumptions or specific aspects of the methodology employed. None can be generalized to render analysis impossible. Criticisms such as these are valid and push economists to explore issues and modify or create analyses that address the problems to which the criticisms point.[9] But from the neoclassical perspective, certain other criticisms are based on misunderstandings of the object of criticism, or purport to

[6] Here Hicks is writing under his altered identity.
[7] See Katzner (1998, pp. 5-8).
[8] An example of an exception is the competition analyzed as a Cournot duopoly.
[9] For example, Hicks' valid criticism cited above that economic analysis is largely out of time has resulted in the development of economic models that are 'in time' (*e.g.*, Katzner, 1998). And the valid criticisms of the neoclassical approach to capital and the marginal product theory relating to it (*e.g.*, Moseley, 2012) have led to models based on money capital that arrive at parallel conclusions and avoid the difficulties raised in those criticisms (*e.g.*, Vickers, 1968, and Katzner 2006, Chs. 13, 14).

expose defective elements in that object that do not actually exist. These purported defects include claims of logical or methodological weaknesses, and lack of reflection of, or contradictions to reality. Such criticisms would be judged by most neoclassicals to be invalid. The remainder of this paper provides the four examples of invalid criticisms of microeconomic analysis alluded to earlier that appear in Fullbrook's book. The issues raised by the criticisms are analytically, methodologically, or empirically important. In each case the source of the invalidity will be described. As illustrated by the previous examples, even from a neoclassical point of view there is plenty to criticize in neoclassical microeconomics, but such criticisms should be (as are those presented above) appropriately focused and well grounded.

3. Lawson's criticism

Consider first the following methodological criticism of T. Lawson:[10] Economic analysis, he argues, should not employ mathematical-type functions because they imply a stable and isolatable relationship between the variables in the function's domain and its range. And such relationships do not exist in economic reality (Lawson, 2004, pp. 22-24).[11] This criticism is subjective in nature, leaving intact other possible analytical forms such as storytelling and argument by example. The following discussion of its invalidity is also subjective in the sense that different judgments are involved as to the meaning and operative usefulness in scientific argument of possible functional forms.

Lawson's argument does not convincingly demonstrate that the functional methodology employed in much of microeconomics is inappropriate. Of course such functions do not exist in reality. No functions can. They are all creations of the human mind – abstractions, figments of the imagination that serve as an aid in formulating an understanding of what is going on. The same is true of all functions in all scientific endeavor including the hard sciences such as physics and chemistry, *e.g*., the equations describing Newton's law of gravity or Boyles law relating the volume of a gas to its temperature and pressure.

Analyses (and theories and models) often entail the isolation of what are thought to be the important parts of the phenomenon under investigation in order to be able to study their properties and how they combine to approximate the whole. As noted at the outset, each element in the analysis is an abstraction, necessarily somewhat distant from reality both in its form and in its isolation from other real parts, and the construction of such abstractions is often required if the analysis is to achieve its goal of explanation. That the use of functions abets the process of construction and isolation is one of the reasons for their frequent employment.

Moreover, for an analysis, theory, or model to provide an explanation, its formal structure cannot change in response to changes in the phenomenon over the time interval that the analysis, theory, or model purports to cover. The structure required to explain a rapidly changing phenomenon may be different from that needed for a less rapidly changing phenomenon. If, over the period during which the phenomenon is studied, no significant change is thought to have occurred, or if whatever change that has occurred is thought to be reasonable to ignore, then a static analysis, theory, or model may be appropriate. Otherwise, a dynamic structure would be needed. And each time an explanation of the phenomenon is undertaken (even if it is to be for the same time interval) a different formal structure may be necessary.

These last two points concerning the facts that functions isolate relationships and impose stability on them apply, as does the earlier point about their non-existence in reality, to all scientific inquiry. The fact that stable and isolated components might not exist in the real phenomenon at issue is of no consequence.

In neoclassical microeconomics, for example, consider the supply-demand model explaining how the observed price of a single commodity is determined and how it rises and falls during a period when rapid inflation is not present. The commodity at issue is assumed to be homogeneous and bought and sold in a

[10] A more extensive discussion of these and related issues emerging from Lawson's criticisms is contained in Mohun and Veneziani's (2012) review of Lawson's *Reorienting Economics* (London: Routledge, 2003).

[11] Another subjective-methodological-type criticism that is subjectively invalid and has a different focus is Fleetwood's (2002, p. 1) assertion that "… the presence of known falsehoods [necessarily arising through the use of deductive reasoning in model construction] removes all explanatory power from neoclassical economics." Of course, due to their abstract nature, all neoclassical models contain known falsehoods and omissions of aspects of the reality under consideration (*e.g.*, Wimsatt, 1987, pp. 28-29). But this does not necessarily destroy their explanatory power. A case in point is the supply-demand model of an isolated market described momentarily.

single isolated market. This is already an abstraction for which, necessarily, there is no counterpart in reality. Gasoline, say, even within a single country, is not homogeneous and is sold in many places and under varying circumstances. The same is true of the 'observed price' which, at the empirical level, is often taken to be an average of observed prices over a collection of locations and similar products. Regardless, demand and supply functions (possibly represented by curves) are postulated that relate single quantity values to single price values and that are assumed to be stable for the period or moment of time in question. Clearly, such functions are further abstractions that are not reflected in actuality. When represented by curves and an observed price-quantity point is to be explained, both demand and supply curves are assumed to pass through that same observed point. The observed price is then viewed as the equilibrium outcome of the interaction of these representations of demand and supply. Furthermore, an observed rise in price, say, is said to be a consequence of either an increase in demand or a reduction in supply.

This explanation is so useful and resonates so strongly as an understanding of what is going on that it has passed from the realm of academic economics into the public domain.[12] But it is nevertheless all based on a fiction. Admittedly, when the media write or speak of demand and supply, they do not explicitly refer to functions or curves. But without having a demand function or its geometric representation as a curve explicitly stated or implicitly present in the background, it is not possible to give clear and full meaning to such words as 'demand' or such phrases as 'increase in demand.' That is because an increase in demand, say, may be taken to mean a change in the form of the relevant part of the demand function or a movement of the curve. Without functions or curves, there is no way to distinguish between changes in demand (shifts in the curve) and changes in quantity demanded (movements along the curve). And rigorous analysis falters in the absence of such meanings. None of this is to say that the demand and supply model gives a complete picture of all influences that determine the short-run price of the good. But in spite of its explicit or implicit use of a fictitious market, a fictitious commodity and price, and fictitious functions, it certainly is a reasonable start. Marshall (1948, p. 461) put it this way:

> "The theory of stable equilibrium of normal demand and supply helps indeed to give definiteness to our ideas; and in its elementary stages it does not diverge from the actual facts of life, so far as to prevent its giving a fairly trustworthy picture of the chief methods of action of the strongest and most persistent group of economic forces. But when pushed to its more remote and intricate logical consequences, it slips away from the conditions of real life."

Lawson continues his argument by suggesting an alternative to the use of functions. 'Contrast explanation,' he says, should replace the use of functions:

> "All we need for this method to work is a situation in which … two outcomes are different … in conditions where it was expected that they would have been the same, resting on an assessment that they shared the same, or sufficiently similar, causal history. Alternatively put, in contrast explanation we seek *not* to explain some X, but to explain why some 'X rather than Y' occurred in a situation where Y was expected (given our understanding of the causal history of the relevant phenomenon). In such a situation we do not seek all the causes of X but the one that made it different from the Y that was anticipated" (Lawson, 2004, p. 27).

One weakness of this part of Lawson's criticism is that he does not suggest any way of forming the initial expectation. To assert only that it emerges from past experience or understandings of that experience is not

[12] Nevertheless, and contrary to its popularity as explanation, Lee and Keen (2004, p. 190) claim that the supply-demand model has "no theoretical or explanatory substance whatsoever." Their subjective conclusion is derived from their critical view of the theoretical structure (*i.e.*, traditional models of the consumer and firm) underlying market demand and supply curves. Now apart from whether their criticisms of that underlying structure are valid or not, it is clear from the text above that the supply-demand model can stand on its own and have explanatory significance independently of reliance on any underlying theoretical structure. Hence, from this perspective, the criticism that the model has no theoretical or explanatory substance is subjectively invalid.

enough. For in order to be able to conclude that an initial expectation has been falsified by an actual occurrence, it is first necessary to establish clearly the grounds upon which that expectation was formulated. That is, the nature of the mental process that must have operated on the experience in question to produce the expectation needs to be described. And that may implicitly involve the construction of an analysis, theory, or model with functions.

Furthermore, to expect an outcome such as Y from conditions like A and B can only mean that Y is associated with both A and B. Now a function consists of three parts: a set called its domain, a set called its range, and a rule that identifies to each element in its domain a unique element in its range. Thus, to expect an outcome Y from conditions A and B is to implicitly specify a single-element domain containing the vector whose components are A and B, a single-element range containing Y, and a rule that relates Y to both A and B, that is the vector (*A*,*B*). Similarly, to assert that A, for example, has a causal history is to claim that certain elements that have arisen in the past and that implicitly constitute a single-vector-element domain are associated by a rule to A in a single-element range. Furthermore, these two rules or functions have to be stable for Lawson's approach to work. Therefore, although Lawson does not say so, contrast explanation is implicitly based on the same stable-function methodology that Lawson states should not be employed in economic analysis.

4. Fullbrook's criticism

Focus next on E. Fullbrook's criticism of the standard analysis of consumer demand behavior. In this essay, Fullbrook describes eight categories of consumer behavior that do not fit into the mold of that analysis. He then asserts that:

> "Projects to understand the logic of economic choice that do not engage with real-world situations of the same are doomed to epistemological failure and axiomatic delusions for the same reasons as are also attempts to theorize about the natural world without observing it. Economists' interest in choice behaviour has in the main been and continues to be far removed from the spirit of empirical, let alone scientific, inquiry" (Fullbrook, 2004, p. 83).

One interpretation of this statement (not necessarily that which Fullbrook intended) is the criticism that the models of the standard analysis of consumer behavior based on the constrained maximization of utility or preferences are irrelevant to economic explanation because, as Fullbrook understands it, they have no connection to real consumer activity.[13] Under that interpretation, the criticism is subjective and if one accounts for the facts that individuals do have limits on what they are financially able to buy and preferences among the possible objects of choice, and, at least in Western economies, they often do attempt to buy what they tend to prefer, then, as will be argued from a neoclassical vantage point below, the criticism is subjectively invalid.[14] Moreover, Fullbrook's criticism would appear to be subjectively defective on two additional counts. First, his assertion regarding distance of choice behavior from the "spirit of empirical … inquiry" is in danger of confusing descriptive with analytical economics. And second, his dismissal of economic analysis as scientifically deficient misunderstands the proper and legitimate place of economic analysis as characterizing a unique scientific methodology. The latter point will also be considered momentarily.

With respect to the above claim of Fullbrook's paper regarding the relevance of the standard

[13] Other invalid subjective criticisms of this type include the Benicourt and Guerrien (2008a, pp. 318, 319, and, respectively, 2008b, p. 336) statements that (i) general equilibrium analysis "… is irrelevant for understanding market economics" and that (ii) the supply-demand model of an isolated market contains logical flaws relating to aggregation and price determination that render it irrelevant to understanding how a real market operates. The former criticism is objectively invalid because it is based on a misunderstanding of what models do and how they relate to real phenomena; the latter has been shown to be objectively invalid in Section III above in which the supply-demand model presented there does not suffer from the logical flaws described by Benicourt and Guerrien.

[14] Martinás' (2002) absence-of-connection-to-reality rejection of consumer utility maximization (arising from the timelessness of the context in which the maximization takes place) is subjectively invalid for a similar reason.

analysis of consumer behavior to real consumer behavior, it should be acknowledged first that empirical testing of consumer behavior derived from that analysis may have little power in determining its accuracy or fitness to explain real behavior. For to refute any particular analysis it is necessary to demonstrate that the theoretical properties of demand functions derived from the assumption content of that analysis are inconsistent with observed demand behavior. But it is a well-known fact that one can never in practice conclusively determine inconsistency. The problem is that observations are necessarily taken over time and there is no guarantee that consumer preferences are constant from one observation to another. Were preferences to change from one observation to the next, then the analytical demand functions relevant to the first observation of demand behavior would be different from those relevant to the second, and the change in demand behavior necessary for comparison to the theoretical properties of a single set of demand functions would appear to be inconsistent with both sets of demand functions. A test of the consistency of the properties of one set of demand functions with observed demand behavior, although possible in principle, is rendered impossible in practice. Thus any disconfirming instance observed in demand behavior can always be explained away by asserting that preferences modified between observations.[15] Moreover, even if there were no disconfirming instances, the analysis still cannot be thought to provide the only definitive explanation of the observed demand behavior because there is no guarantee that some other analysis, which may actually yield the 'true' or 'correct' understanding, does not explain the same behavior. Therefore a negative outcome from empirical testing of the theoretical properties of demand behavior implied by the analysis that suggests rejection is irrelevant to a judgment of whether that analysis possesses the ability to explain real consumer behavior. This is not to suggest that empirical analysis of demand behavior (such as the estimation of demand functions, *e.g.*, Theil [1975, 1976], and the prediction of consumer action) is neither significant nor useful – only that it is unable to contribute to a determination of the model's viability.

But that in no sense provides a sustainable criticism of the received theory of consumer behavior. For it follows from the preceding argument that the question of relevance to reality in this case has to be based on the appropriateness of the assumptions[16] that gave rise to the theory. Evidently, some assumptions are imposed for convenience: If constrained utility maximization is to be invoked, enough properties have to be assumed on the utility function so that the function can, in fact, be maximized. Nevertheless it is a widely recognized and observable fact that in Western economies individuals, as noted above, have preferences and limited budgets, and often buy what they prefer. Thus, as an approximation, the standard theory of consumer demand can usefully be applied to explain consumer behavior under situations where 'outside' influences such as altruism or savings are thought not to play a significant role.

As already observed, to claim that the use of such theories results in an analysis that is not scientific is to misunderstand the nature of science in regard to economics and to ignore the fact that economics cannot be a science in the same sense that, say, physics or chemistry is a science.[17] That is so for at least three reasons: First the objects of study in economics are human activities and not the result of action by non-sentient, unmotivated physical entities. Human activities can change in ways and for reasons that physical objects cannot. Economic phenomena are therefore more unstable and less predictable than physical phenomena. Second, economists do not have as much control over human subjects in laboratories as physical scientists have over physical subjects. Unlike temperature, weight, and pressure that could be held fixed during a physics experiment, attitudes, beliefs, emotions, and desires cannot be counted on to remain constant over the course of an economics experiment. The modifications in preferences that might, as mentioned above, vary as observations of demand behavior are taken is one of many examples. And third, it is often easier to measure the significant features of physical phenomena than it is to quantify the human features of economic phenomena – although, it should be noted that rigorous analysis, both theoretical and empirical, are entirely possible in the absence of measurement.[18] In any case, the unavailability of measures can impact both the content and methodology of an analysis. Taking these considerations into account, there is still no problem in identifying economics as a science, even as an

[15] See Caldwell (1982, pp. 156-157).
[16] Appropriateness of assumptions may be defined in such a way as to be consistent with Kincaid's criteria for good science set out below.
[17] A more detailed discussion of this point may be found in Katzner (2001a, pp. 2-5).
[18] Katzner (1983, 2001b).

empirical science. But its science is distinct from that of the physical sciences.

With these departures from the physical sciences in mind, the traditional theory of consumer demand may, according to Kincaid's criteria (1996, pp. 50-55), be regarded as good economic science: it coheres within itself and with knowledge outside its purview (*e.g.*, known cultural values); it provides information relating to the actual buying behavior it addresses; it is objective given the value context within which it was created; it can be accurately fitted to demand data (*e.g.*, Theil *op. cit.*) – although, as indicated above, rejection based on such data is not possible; it can be falsified in terms of the appropriateness of its assumption content in relation to the characteristics thought to be present in the reality it reflects; and it is causal in nature.

5. Benicourt's criticism

Also invalid because its assertion is based on a misunderstanding of the use of concepts in analysis and communication is the methodological criticism of E. Benicourt. According to her, words like 'goods,' 'household,' 'firm,' 'market,' and 'price,' used by economists have nothing to do with the meaning these words have in reference to reality. "Instead they refer to fictitious entities, and not approximations of something that really exists." This renders explanation of real economic phenomena that employs these words useless (Benicourt, 2004, p. 85).

It will now be argued that this subjective criticism is objectively invalid because to accept it would eliminate the possibility of engaging in any analytical endeavor in any science – not only economics. As previously indicated, this judgment of invalidity would be classified as relatively objective.

In general, to understand and communicate ideas, human beings have learned rules that tell them to identify repeatedly specific things with specific words.[19] A chair is called a chair because a rule has been accepted asserting that a physical structure with certain characteristics and functions is called a chair. Similarly, behavior can be meaningful only if it involves the application of a rule. To cast a vote is to follow a rule that identifies a particular action under particular circumstances with the behavior voting. The learning and application of rules necessarily takes place in the mind.

A concept is a mental image obtained by applying a rule to a collection of different objects having the same characteristics as specified by the rule. Thus concepts take their places in a microeconomic analysis, theory, or model as the referents of things that exist in the real world and are often identified by words. That is, words necessarily refer to concepts, figments of the human imagination, that are themselves abstractions from what actually exists. For example, the word 'chair' refers to a concept, an abstraction from a large class of real objects that have certain seemingly identical properties that according to the relevant rule characterize what is known as a chair. That is how human beings formulate and communicate ideas in reference to reality.

Now in any scientific investigation the object of study must be observed, relevant facts about it noticed, and then, as in the case of the analysis of market behavior discussed earlier, abstract constructions developed. But to be able to notice relevant facts means that the researcher is able to identify certain characteristics; and to do this he or she must have some concept of what the characteristics are. And the latter is possible only with appropriate rules identifying communicable symbols like words with the characteristics, and these, in turn, will depend on what has been socially acceptable to others working in the same area.

This is exactly what is done in economics. "Words like 'goods,' 'household,' 'firm,' 'market,' and 'price,'" even when Benicourt uses them, necessarily "refer to fictitious entities." And in analyzing the behavior of, say, an individual or a firm, economists have to abstract from reality to focus on the conceptual forces that seem to be most important in understanding the economic events that have transpired. (Recall the Marshall quotation above.) There is nothing wrong with this. It is the way in which all analysis and explanation, economic or otherwise, proceeds. Moreover, it does not matter that the firm, for example, might be defined in terms of production and cost functions in one instance or with any collection of

[19] The idea is due to Wittgenstein. See, for example, Winch (1958).

nonmathematical sentences seemingly unrelated to production and cost functions in another. Both are abstractions based on fictions, and parts of each may be implicit in parts of the other. To assert that one conceptualization is closer to reality than the other is misleading since there can be no objective or meaningful measure of distance from any concept to the reality it represents.

It is interesting that Benicourt's view of the way economic analysis should proceed is not unlike that described here:

> "… observe what happens (or has happened) in known societies, look at how prices are actually formed and then try to find rules or tendencies on which a theory can be built (or developed)" (Benicourt, 2004, p. 93).

But in this statement, 'known societies,' 'price,' 'rule,' and 'tendency' are all abstract concepts – fictitious entities that do not really exist. To illustrate, the price of a particular good varies from time to time and place to place. The implicit notion that the good is homogeneous, and the explicit notion that a good has a single price that can be designated as 'the price' of that good, are both abstractions. (As noted previously, the average of observed prices is an abstraction at the empirical level.) Thus Benicourt's method of analysis has to employ fictitious entities while, at the same time, she criticizes economists for doing so.

6. Two criticisms of Keen

The last example of invalid criticism presented here is that of Keen who argues that microeconomic analysis, theory, and models often contain logical mistakes:[20]

> "Neoclassical economics is a mathematical science? Give me a break! … its so-called mathematical credentials are a myth. Far from using mathematics to elucidate economic issues, neoclassical economics has twisted and distorted mathematics to maintain an ideological vision of the market which proper mathematics shows is unsustainable" (Keen, 2004, p. 221).

Quite the contrary. This objective criticism is judged invalid objectively because, at least with regard to microeconomics, it is not based on an accurate representation of microeconomic analysis or microeconomic models. Neoclassical microeconomic constructions are transformed into something they are not, and that something is used as evidence against neoclassical microeconomics. Here are two examples:

First, Keen argues that, in the neoclassical analysis of the firm, the "… slope of the individual firm's demand curve is exactly the same as the slope of the market demand curve." His argument is based on the equation

$$\frac{dp}{dq} = \frac{dp}{dQ}\frac{dQ}{dq} = \frac{dp}{dQ},$$

where p represents market price, q denotes quantity of a particular firm's output demanded by all buyers in the market (*i.e.*, demand quantity facing the firm), and Q indicates market quantity demanded. The right-hand equality obtains because Keen assumes that $dQ/dq = 1$.[21] Thus Keen concludes that under perfectly competitive conditions, say, the demand curve facing the firm cannot be a line parallel to the q axis in the p-q plane as neoclassical microeconomic theory says it is because the market demand curve has a negative slope (Keen, 2004, p. 210).

This distorts the neoclassical microeconomic analysis of the firm by ignoring the fact that the market

[20] Barzilai (2013, 2014) makes a similar objectively invalid objective claim with respect to the utility function in the traditional model of consumer demand. His argument that, due to its ordinality, the utility function of that model (whether actually present or only implicit in preferences) introduces logical flaws that render the model useless in explaining consumer behavior is based on a misunderstanding of the ordinal nature of the utility function. See Katzner (2014).

[21] This assumption would follow from, say, requiring that market demand be the sum of the demands that face all firms in the market.

demand curve and the demand curve facing the individual firm are two independent and, at the conceptual level, unrelated constructions. Neither is defined in terms of, or in relation to, the other, and the former is not the sum of the latter over all firms in the market. Rather, the market demand curve indicates the quantity that would be demanded by all buyers in the market at various prices (without reference to the particular firm from which each unit of the good might be purchased); the demand curve facing the firm describes the quantity that the firm expects to be able to sell at those prices.[22] Except in the case of monopoly where they become identical, the relationship between them is rather complex[23] and cannot be simplified to imply that $dQ/dq = 1$. With regard to the perfectly competitive firm, because it is such a small part of a large market in which buying and selling a homogeneous product takes place, the firm believes that it can sell all it wants at the prevailing price. Based on this belief, the demand curve it faces is a line parallel to the q axis at the prevailing market price.[24]

Second, Keen contends that, in a neoclassical Walrasian model, if market prices and market quantities begin to adjust from an out-of-general equilibrium position, convergence to a new general equilibrium position is not possible. He does this by postulating independent dynamic movements for market prices and market quantities over time, and obtains the result that if one of market prices or market quantities converges to their equilibrium values, then the other must diverge from their equilibrium values (Keen, 2004, pp. 220-221).

This distorts neoclassical analysis because Walrasian general equilibrium models do not postulate independent dynamics for market prices and market quantities. Rather they focus attention on what are called 'excess demand' functions. These functions express quantities demanded minus quantities supplied, *i.e.*, excess demands, as dependent on market prices. All markets in the model are in equilibrium together when all excess demands equal zero, and in that instance, supply equals demand in all markets. To handle out-of-equilibrium positions, a dynamic is expressed in terms of changing prices over time and this implies, through the excess demand functions, movements in excess demand quantities over time. Convergence in a Walrasian model to equilibrium price values everywhere then points to convergence of excess demand quantities to zero and hence to equilibrium quantities throughout the system. Alternatively, the dynamic could be expressed with respect to changing excess demand quantities and this would imply modifications in market prices across time. In this case, convergence to equilibrium quantities would force convergence to equilibrium prices.[25]

7. Conclusion

Each of the four invalid criticisms described above is, from a neoclassical perspective, based on a misunderstanding of important aspects of microeconomics – its methodology, its relation to the real world, or its logical structure. This is not to say that various aspects of microeconomics in general cannot be validly criticized on these grounds; validly, that is, in the sense of the definition provided in this paper. But the particular criticisms of Lawson, Fullbrook, Benicourt, and Keen presented here decidedly miss their mark.

The discussion of analytical critiques in this paper implicitly points, it will be clear, to significant issues that lie beyond its scope. Questions of the status of economics as a scientific discipline, of matters that probe to deeper epistemological foundations, and conceivably of ethical relevance if the claim is made that economic argument is not, and cannot be, a value-free inquiry, warrant careful reflection. But within the frame of argument adopted here, it is equally clear that the room that economic analysis provides for discussion, and for agreement or disagreement on methods, assumptions, and results, not only furnishes channels for beneficial interaction among opposing camps, but also contributes to the intellectual pluralism

[22] Similar wording appears in most undergraduate microeconomics textbooks. See, for example, Pindyck and Rubinfeld (2009, p. 278).
[23] This relationship would depend on how much each firm expects each buyer to purchase from it at each price – characteristics of buying behavior that are not normally specified in analyses of the firm. Moreover, except possibly at overall market equilibrium, at each price the sum of all expected purchases over all firms generally would not, and could not be expected to add up to actual market demand.
[24] Keen and Standish compound the error by using the incorrect statement that "the demand curve facing the perfectly competitive firm is downward sloping" to deduce that the perfectly competitive firm does not maximize profit where price equals marginal cost (Keen and Standish, 2010, p. 61) and that for such firms price-taking behavior is irrational (Keen and Standish, 2010, p. 69).
[25] See, for example, Katzner (2006, Sect. 9.1).

that both neoclassical orthodoxy (in its varied expressions) and modern (or postmodern) forms of heterodoxy address.

References

Barzilai, J. (2013) Inapplicable operations on ordinal, cardinal, and expected utility, *Real-World Economics Review*, issue 63, pp. 98-102.

Barzilai, J. (2014) Demand theory is founded on errors, *Real-World Economics Review*, issue 68, pp. 62-65.

Benicourt, E., (2004) Five pieces of advice for students studying microeconomics. In *A guide to what's wrong with economics*, Fullbrook, E. (ed.) London: Anthem, pp.84-94.

Benicourt, E. and Guerrien, B. (2008a) Is anything worth keeping in microeconomics? *Review of Radical Political Economics* 40 (3), pp. 317-323.

Benicourt, E. and Guerrien, B. (2008b) A reply to Katzner and Case, *Review of Radical Political Economics* 40 (3), pp. 336-340.

Bernstein, M.A. (2004) The pitfalls of mainstream economic reasoning (and teaching). In *A guide to what's wrong with economics*, Fullbrook, E. (ed.) London: Anthem, pp. 33-40.

Caldwell, B.J. (1982) *Beyond positivism*: *Economic methodology in the twentieth century*, London: Allen & Unwin.

Fleetwood, S. (2002) Why neoclassical economics explains nothing at all, *Post-Autistic Economics Review*, issue no. 17, article 2.

Fullbrook, E. (2004) Are you rational? In *A guide to what's wrong with economics*, Fullbrook, E. (ed.) London: Anthem, pp. 71-83.

Hicks, Sir J. (1975) The revival of political economy: The old and the new," *Economic Record* 51, pp. 365-367.

Hicks, Sir J. (1976) Some questions of time in economics. In *Evolution, welfare, and time in economics*, Tang, A.M., Westfield, F.M. and Worley, J.S. (eds.) Lexington: D.C. Heath, pp.135-151.

Hodgson, G.M. (2004) Can economics start from the individual alone? In *A guide to what's wrong with economics*, Fullbrook, E. (ed.) London: Anthem, pp. 57-67.

Katzner, D.W. (1983) *Analysis without measurement*, Cambridge: Cambridge Univesity Press.

Katzner, D.W. (1998) *Time*, *ignorance*, *and uncertainty in economic models*, Ann Arbor: University of Michigan Press.

Katzner, D.W. (2001a) Introduction: Science, social science and measurement. In *Unmeasured information and the methodology of social scientific inquiry*, Boston: Kluwer, pp. 1-24.

Katzner, D.W. (2001b) *Unmeasured information and the methodology of social scientific inquiry*, Boston: Kluwer.

Katzner, D.W. (2006) *An introduction to the economic theory of market behavior*: *Microeconomics from a Walrasian perspective*, Cheltenham: Elgar.

Katzner, D.W. (2014) Ordinal utility and the traditional theory of consumer demand," *Real World Economic Review*, issue 67, pp. 130-136.

Keen, S. (2004), "Improbable, incorrect or impossible: The persuasive but flawed mathematics of microeconomics. In *A guide to what's wrong with economics*, Fullbrook, E. (ed.) London: Anthem, pp. 209-222.

Keen, S. and R. Standish, R. (2010) Debunking the theory of the firm – A chronology," *Real-World Economics Review*, issue 53, pp. 56-94.

Kincaid, H. (1996) *Philosophical foundations of the social sciences*, Cambridge: Cambridge University Press.

Lawson, T. (2004) Modern economics: The problem and a solution. In *A guide to what's wrong with economics*, Fullbrook, E. (ed.) London: Anthem, pp. 21-32.

Lee, F.S. and S. Keen (2004) The incoherent emperor: A heterodox critique of neoclassical microeconomic theory, *Review of Social Economy* 62, pp. 169-199.

Marshall, A. (1948) *Principles of economics*, 8th ed. New York: Macmillan.

Martinás, K. (2002) Is the utility maximization principle necessary? *Post-Autistic Economics Review*, issue 12, article 4.

Mohun, S. and R. Veneziani, R. (2012) Reorienting economics? *Philosophy of the Social Sciences* 42, pp. 126-145.

Moseley, F. (2012) A critique of the marginal productivity theory of the price of capital, *Real-World Economics Review*, issue 59, pp. 131-137.

Ormerod, P. (2004) Neoclassical economic theory: A special and not a general case. In *A guide to what's wrong with economics*, Fullbrook, E. (ed.) London: Anthem, pp. 41-46.

Pindyck, R.S. and Rubinfeld, D.L. *Microeconomics*, 7th ed. Upper Saddle River: Pearson Prentice Hall.

Prasch, R.E. (1996) The origins of the a priori method in classical political economy: A reinterpretation, *Journal of Economic Issues* 30, pp. 1105-1125.

Shubik, M. (1970) A curmudgeon's guide to microeconomics, *Journal of Economic Literature* 8, pp. 405-434.

Stretton, H. (2004) The quarrelsome boundaries of economics. In *A guide to what's wrong with economics*, Fullbrook, E. (ed.) London: Anthem, pp. 9-20.

Theil, H. (1975, 1976) *Theory and measurement of consumer demand*, v. 1, 2 Amsterdam: North Holland.

Vickers, D. (1968) *The theory of the firm: Production, capital, and finance*, New York: McGraw-Hill.

Wimsatt, W.C. (1987) False models as means to truer theories. In *Neutral models in biology*, Nitecki, M.H. and Hoffman, A. (eds.) New York: Oxfored University Press, pp. 23-55.

Winch, P. (1958) *The idea of a social Science*, London: Routledge and Kegan Paul.