sanity, humanity and science probably the world's most read economics journal

real-world economics review

Please click here to support this journal and the WEA

- Subscribers: **26,507** **subscribe** RWER Blog ISSN 1755-9472
- A journal of the World Economics Association (WEA) 14,468 members, **join**
- Sister open access journals: *Economic Thought* and *World Social and Economic Review*

back issues

Issue no. 80

26 June 2017

Inequality, democracy and the ecosystem

Money

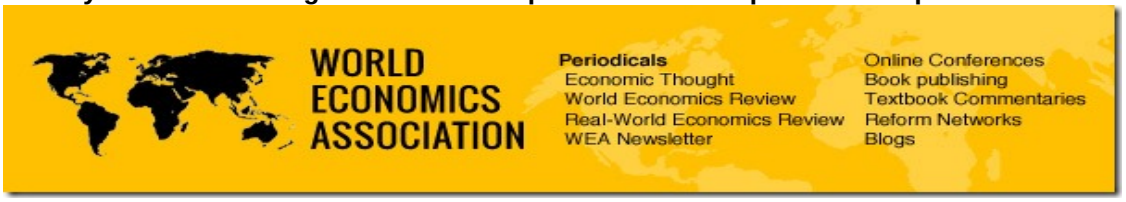
Keynesian issues

Nations

Review essays

A note

By paying a WEA membership fee here, or offering a contribution here, you will be making this and other open-access WEA publications possible.



real-world economics review, issue no. 80
subscribe for free

Politics, preferences, and prices: the political consequences of inequality[1]

Luke Petach [Colorado State University, USA]

Abstract
Drawing on the literature from both political science and economics, I identify four main channels through which democracy may be harmed by economic inequality: an increasing political responsiveness to the wealthy (via lobbying, campaign finance, etc.) simultaneously coupled with a decreasing political responsiveness to the poor and the middle class, increasing political instability, the ability of corporations and financial elites to subvert market reforms enacted by political consensus, and a shift of political preferences toward authoritarian political leaders. I argue that while the first three reasons have received sufficient attention in the literature, not enough attention has been paid to the way in which economic inequality alters the preferences of the electorate. I pull from work on endogenous preferences as well as longstanding arguments from the institutionalist economic perspective to support this argument.

I. Introduction

Political and economic inequalities are inevitably related. Much of the literature addressing the economic consequences of inequality has focused on politics as the mechanism through which economic inequality impinges on growth (Alesina and Rodrik, 1994; Perotti, 1993). The researcher who wishes to develop a complete understanding of inequality must find a way to disentangle the economic effects of inequality from the political effects. The goal of this paper is to isolate the latter. Specifically, I wish to address some of the ways in which economic inequality may be harmful for democracy. If economic inequality inhibits the ability of all members of society to obtain equal representation in government, then – as one researcher put it – such a government is a "democracy in name only" (Gilens, 2005). Drawing on the literature from both political science and economics, I identify four main channels through which democracy may be harmed by economic inequality: an increasing political responsiveness to the wealthy (via lobbying, campaign finance, etc.) simultaneously coupled with a decreasing political responsiveness to the poor and the middle class, increasing political instability, the ability of corporations and financial elites to subvert market reforms enacted by political consensus, and a shift of political preferences toward authoritarian political leaders. In reduced form, economic inequality creates political power, political instability, market power, and shifts in preferences all of which act in ways detrimental to democracy. The first three are well documented in the literature on inequality. I cover these views in Section II. In contrast, I argue not enough attention has been paid to the way in which economic inequality alters the preferences of the electorate. I pull from work on endogenous preferences as well as longstanding arguments from the institutionalist economic perspective to support this argument.

The rest of the paper is organized as follows: after covering the political and market power views and their critiques in Section II, I argue for the importance of the institutionalist /

[1] I would like to thank Steven Pressman for his helpful comments on earlier drafts of this paper.

endogenous preference position in Section III. Section IV considers objections to the endogenous preference view. Section V concludes.

II. Democracy or plutocracy?

There is an extensive literature on both the political (Apel, 2016; Bartels, 2008; Gilens, 2005, 2012; Hacker and Pierson, 2010) and economic (Klasen, 2012; Piketty, 2014; Stiglitz, 2012; Wisman, 2008) causes and consequences of inequality. Rather than addressing each of the works in this literature individually, in this section I turn my attention to two works that advance the two most important points of this literature more broadly: the problem of political responsiveness (Gilens, 2005) and the problem of political instability (Alesina and Perrotti, 1996).

The question of political responsiveness is an empirical one. Gilens (2005) attempts to give an answer using longitudinal data on public opinion and political outcomes over the period 1981–2002. The dataset consists in 1,781 questions indicating the proportion of individuals who "favor" or "oppose" a given policy proposal, with data broken out by categories of income, race, education, sex, etc. In addition to the proportion of respondents, Gilens' (2005) data contains the actual policy response of the government. Thus, breaking up the proportions of "favor" and "oppose" by income level, and relating these to the policy outcome, one has a potential test of the impact of income inequality on policy. If the preferences of the rich tend to outweigh the preferences of the poor and the middle class in determining policy, then income inequality may be harmful to democracy. If a society becomes more unequal over time, one might expect the "voice" of the poor in government to gradually decline if the above holds.

To test this hypothesis Gilens (2005) estimates a logit model in which the policy outcome – taking a value of 1 for a change and 0 otherwise – is regressed on the proportion of respondents within the given category that favor the policy change. Two main results appear. First, it appears that policy has a strong status-quo bias. Even if a majority of the population prefers a new policy, it is still relatively unlikely to pass (the predicted probability of a policy passing if 90% of the population is in favor is only 46% (p.785)). Gilens (2005) argues that this is consistent with the desire of the framers of the constitution to "inhibit the 'tyranny of the majority'" (p.786). The second finding, and the crucial one for the relationship between inequality and democracy, is that the preference-policy link becomes weaker as you move down the income distribution. Gilens (2005) estimates that an increase in the percent of individuals at the 10th percentile of the income distribution who support a given policy from 10% to 90% increases the predicted probability of change from approximately 20% to 40%. In contrast, an increase in the percent of individuals at the 90th percentile of the income distribution who support a given policy from 10% to 90% increases the probability that the policy will be adopted from approximately 12% to almost 50% (note that when the proportion of high-income earners that supports the policy is low, the probability that the policy passes is lower than when the proportion of low-income individuals that supports the policy is low).

This result becomes even starker when divergence in preference between the poor and the wealthy is allowed. Gilens (2005) finds that when preferences diverge between the 90th and 10th percentiles, an increase from 10% support to 90% support at the 10th percentile has zero effect on the predicted probability of a policy change. In contrast, given preference divergence, an increase in support at the 90th percentile from 10% to 90% increases the predicted probability of a policy change from approximately 15% to almost 50%. Thus, in the case of a disagreement between the rich and the poor, the will of the rich wins out.

What Gilens (2005) is not clear on is the mechanism by which the rich assert their influence. Three potential mechanisms suggest themselves however: lobbying, political advertising, and campaign contributions (Hacker and Pierson, 2010). The role of political advertising is of particular interest in the United States since the decision on *Citizens United v. FEC* in 2010. With its decision on *Citizens United* the Supreme Court determined that freedom of speech laws prevent the federal government from prohibiting the creation of, donation of money to, and independent expenditure of money by "non-profit" corporations such as political action committees (PACs) (*Citizens United v. Federal Election Commission,* 558 U.S. 310). These "non-profits" thus present themselves as an opportunity to obtain an essentially limitless source of support for political candidates, insofar as they can acquire adequate donations and purchase expensive advertising. While not directly preventing poor or middle class individuals from expressing their political opinion, insofar as the rich can out-advertise (and out-fund) candidates supported by the rest of the income distribution, they can effectively yell over them. In a ruling that amounts to declaring "money is speech," the Supreme Court determined the poor should have no voice.

In addition to the role played by political advertising, the role of lobbying in shaping the preferences of politicians is well known to economists and political scientists alike. There is a large literature in economics on the role of lobbying in policymaking (Mitra. 1999; Esteban and Ray, 2006), the results of which indicate that the influence of the wealthy increases with the amount of money they contribute to political lobbying.

Finally, the mechanism of simple campaign contributions and contribution of time to political campaigns is suggested by Gilens (2005): "[t]he most obvious source of influence over policy that distinguishes high-income Americans is money and the willingness to donate to parties, candidates, and interest organizations" (p.793). The rich are substantially more likely to donate to an individual candidate, and other research indicates that propensity to donate and the size of donations increase with income level (Gilens, 2005). Additionally, the rich are likely less time-constrained in terms of their income earning ability (i.e. an extra dollar earned by the rich is less likely to go toward purchasing the necessities of life than an extra dollar earned by the poor or the middle class) and as such may be more likely and/or able to spend their time volunteering for political campaigns in the form of canvassing, phone-banking, or otherwise.

The weakness in Gilens' (2005) argument is a lack of control for other variables possibly affecting governmental responsiveness. Indeed, the question of causality is that which plagues Gilens' (2005) findings most. The most obvious omitted variable candidate is a control for voter turnout. Although Gilens' (2005) data measures the level of support for various policies across income percentiles, it is not clear that high levels of support necessarily translate to increased voter turnout at all percentiles. If those at the top end of the income distribution tend to vote more frequently and in higher density, then increasing policy responsiveness to those with top incomes when preferences diverge across the income distribution is explained by the relative amount of votes cast. This differential matters, but it implies the question of interest is not "why is policy more responsive to those at the top?" but "why do those at the top vote more at the bottom?" Voting inequality matters for how economic inequality translates to policy.

Earlier work by Alesina and Perotti (1996) supports the findings of Gilens (2005). These authors find that income inequality tends to fuel social discontent which can increase socio-political instability. Using a sample of 71 countries for the period 1960-1985, the authors estimate a simultaneous equations model looking at the dual impact of economic inequality

on socio-political instability and of socio-political instability on investment. For the question at hand, it is primarily the first relationship I am concerned with. The authors make a compelling case that increasing income inequality is likely to result in increasing demand for political regime change, of which violence, expropriation of property, and illegal seizure of power may be a part. Regime change characterized by the aforementioned characteristics is evidently harmful to democracy. Thus, if the empirical results indicate that income inequality leads to increasing political instability, one would safely conclude that such inequality may undermine democracy.

To measure socio-political instability the authors construct an index which includes various signs of political instability, including: the relative ethnic homogeneity of the population, the average number of deaths in domestic disturbances (per million people), the average number of assassinations, the average number of successful and unsuccessful *coup d'état*, and a dummy for whether or not the current regime is a democracy, semi-democracy, or dictatorship. This composite index is regressed on variables for education, investment, and the inequality measure – an indicator of the share of the third and fourth income quintiles in total output (i.e. a measure of the size of the middle class). The authors find that a decrease in the size of the middle-class increases socio-political instability. The result is statistically significant and robust across a variety of specifications. Insofar as those who take advantage of political instability to gain power are likely to reside in the top portion of the income distribution, Alesina and Perotti's (1996) findings lend support to Gilen's (2005) argument.

Alesina and Perotti's (1996) argument runs into conceptual difficulties when one attempts to translate their findings into policy implications. A comparison with an earlier work, Alesina and Rodrick (1994), will serve to illustrate this point. Alesina and Rodrick (1994) find that economic inequality is potentially harmful for growth because increasing inequality is associated with a tendency to adopt redistributive policies that deviate from Pareto efficiency/The Golden Rule savings rate. Inequality thus creates problems for growth, but it does so without ever approaching the limit of socio-political instability. Redistributive polices may fall entirely within the realm of accepted democratic principles. The question therefore becomes, when is inequality more likely to lead to political instability relative to simple redistribution? The answer will depend on the institutional structure of the country in question. Some institutions may be robust to political instability at relatively high levels of inequality, while others may succumb to a *coup d'etat* for relative trivialities. There is a potential trade-off in policy making between allocating funds to fight inequality, and allocating funds to create institutions that are robust to it. The correct policy bundle will depend, in part, on the preferences of the electorate. Attention to the process through which such preferences are formed is therefore key to understanding how and when inequality can create socio-political instability that is harmful for democracy.

Increasing income inequality thus appears to negatively affect democracy via political channels in two ways: by decreasing the government's political responsiveness to the poor and the middle class and by increasing the likelihood of socio-political instability (through which democratically elected governments may be undermined). However, it is not only through political channels that income inequality may produce harmful political outcomes. Just as it is through political channels that economic inequality has been shown to negatively impact economic growth (Alesina and Rodrick, 1994), so too income inequality may work through economic channels to produce harmful political outcomes.

The main way economic inequality creates negative political outcomes through economic channels is via the increased power to avoid democratically enacted market reform that wealth confers to its owners. In standard economic jargon, wealth allows certain individuals and institutions to be price-makers when they would otherwise be price-takers. Two such examples will serve to illustrate the point: the concentration and magnitude of tax evasion among top income earners and corporations, and the ability of owners and firms to undermine collective bargaining by workers, thus removing Galbraith's (1952) "countervailing power".

One such example of a democratically enacted market reform is an increase in corporate or personal income taxation. If those at the top end of the income/wealth distribution are able to use the power granted them by wealth to move funds around in a way that makes those funds untaxable, then income inequality will tend to undermine democracy – insofar as it makes it difficult to enforce democratically enacted policy provisions. Recent work indicates that the ability of top income earners and wealth holders to engage in such activities is substantial. Zucman (2013; 2015) estimates that the amount of wealth held in tax havens is on the order of 8% of all global financial wealth. Zucman (2013) argues that properly accounting for this wealth would shift the Eurozone's balance-sheet position from net-debtor to net-creditor. Tax evasion on this scale represents a substantial loss – not only in terms of the lost government revenue and the "cost of rent-seeking" paid by those moving wealth, but also of the capacity of the state to enforce laws enacted by democratic means.

The second way inequality creates negative political outcomes through economic channels is via its attenuating effect on collective bargaining. It is no secret that union membership in the United States has seen precipitous declines since the 1980s. At the start of the great recession in 2008, union membership had fallen from its post-World War II peak of around 30% of the employed population to a mere 8% (Galbraith 2016, p.83). While some of this decline can be – and often is (Stiglitz, 2012; Atkinson, 2015; Apel, 2016; Galbraith, 2016) – ascribed to deliberate policy intervention on behalf of the federal government in the 1980s (noting also that a decline in union membership itself has an effect on income inequality), an important contributing mechanism is the rise of inequality itself. By altering the incentives of workers and management, rising inequality weakens the power of unions in both the wage-bargaining process and in the political arena.

To see this process clearly, consider a potentially "exogenous" factor (at least, to the wage-bargaining process) influencing inequality, such as skill-biased technical change. A worker near the top-end of the skill distribution will begin to see her wages rise as the relative demand for workers in her position increases. Workers such as this may also be given promotions (in all likelihood, the aforementioned wage increases may in-fact be tied to promotions). Assuming that such promotions – and the correlative subsequent pay raises – are contingent on conforming to norms of upper-management, skilled workers may be tempted to side with the management in labor disputes. As Atkinson (2015) points out, this dynamic may swing the balance of power in wage negotiations toward management, as employees that would otherwise have supported the union defect in hopes of currying favor. The incentive for such defection is increasing in the extent of inequality, insofar as increasing inequality raises the potential payout to defection. For large levels of inequality the incentive for any one employee to defect in a wage-bargaining game is high, even if the actual probability of receiving the payout is low. Inequality thus exacerbates trends worsening the relative position of workers in the labor market. The second order effect of this decline is the diminution of the political efficacy of unions.

To the extent that effective political organization depends on internal public goods and economies of scale, a decline in union membership resulting from rising inequality will undermine the ability of unions to effectively lobby for policies that further the interest of working class. Furthermore, a decline in union membership has a direct impact on the responsiveness of government to labor vis-à-vis its effect on total union income dues. If the effectiveness of any one lobby is mostly a function of wealth (Esteban and Ray, 2006; Hilber and Robert-Nicoud, 2013) then a decline in union funds will decrease relative lobby effectiveness. Increasing inequality will thus create more uneven political outcomes, and potentially harm democracy, by decreasing the voice of – and political responsiveness of the government to – the working class relative to the wealthy via its impact on unions.

III. Endogenous preferences and political authoritarianism

The impact of inequality on democracy via the political and economic mechanisms discussed above has received much attention elsewhere – if not together, then certainly separately. In this section I wish to turn to a potential consequence for democracy of economic inequality rarely addressed directly in the literature: the ability for changes in market structure to alter the tastes of the electorate. Despite the best laid plans (and models) of economists, preferences (including preference for a particular political candidate) are endogenous. A priori, there is no reason one should expect political preferences to stay constant in the face of rising inequality. I argue that inequality is harmful to democracy because individuals may either A) internalize the notion of a hierarchical society (the "go along to get along" model) or B) be pushed to political extremism. Both outcomes may lead to an increased preference for authoritarian political leaders and thus undermine democracy. In scenario A), an authoritarian leader might be preferred if voters internalize the notion that the wealthiest class is in some way superior. In scenario B), an authoritarian leader may be preferred if s/he promises to put an end to inequality. In either case democracy may be harmed. Ex-post, these two mechanisms may be difficult to differentiate. They need not be mutually exclusive however. Recent events in the politics of at least one major country indicate that these two currents at times overlap.[2] In what follows I discuss both how inequality may induce or alter preference formation, and how the two channels through which such preferences affect democracy overlap.

What is likely the oldest argument for endogenous preferences can be found in Thorstein Veblen's (2009 [1899]) *Theory of the Leisure Class*. Veblen (2009) argues over time there arises a distinction between the "employments proper to several classes" (7). This distinction references the exemption of certain wealthy groups from the need to engage in industrial occupations. Leisure, as opposed to industry, becomes the defining feature for this privileged group of individuals. This class is immediately posed with a problem however. Unlike the industrial occupations, the work of the leisure class has no obvious productive output. To maintain one's status as a member of the privileged class one must come up with creative ways to signal that s/he has not in fact been working. The result of these efforts is the development of the forms of "conspicuous leisure" and later, "conspicuous consumption." Veblen writes:

> "[T]he criteria of past performance of leisure therefore commonly take the form of 'immaterial' goods. Such immaterial evidence of past leisure are

[2] See the Trump phenomena in the United States.

> quasi-scholarly or quasi-artistic accomplishments and a knowledge of processes and incidents which do not conduce directly to the furtherance of human life" (p. 34).

The problem with the development of these patterns of conspicuous leisure and conspicuous consumption is that they are non-neutral with respect to the well-being of those not in the leisure class. Consumption and leisure are honorific, they confer one's social standing. The ability to evidence one's own conspicuous leisure, or to engage in conspicuous consumption, has an effect on an individual's place in the social order and on their social well-being. In Veblen's words, individuals "wish to conform to established usage, to avoid unfavorable notice and comment, to live up to the accepted cannons of decency" (p.78). Around this desire arise "pecuniary canons of taste" set by the leisure class. These function as the "accredited canons of consumption," (p.78) by which the leisure class orders which goods and activities are "honorific" and which are not.

While there may be a gap between the extent to which the wealthiest individuals in a society actually determine the preferences of individuals down the income distribution and Veblen's account of pecuniary canons of taste set by the leisure class, there is no doubt that the relative power of the wealthy to determine trends and standards of taste is greater than that of the general populous. Trends in fashion, politics, sports, and even religion are determined by a rather limited pool of individuals with enough wealth and social standing (be it as a star athlete, fashion designer, long-tenured politician, etc.) to both afford enough advertising and garner enough interest for taste-making campaigns to be effective. There is a reason why Michael Jordan's return from retirement resulted in a nearly 2% increase in stock market returns for endorsed brands (Mathur et. al., 1997) or why Kanye West can sell a plain white T-shirt for $120,[3] erstwhile – despite considering myself quite fashionable – my own endorsements of various products seem to lack the same effect. While perhaps not a one-to-one correlation, the tastes of the wealthy certainly influence the preferences of other classes, and "in some approximation, [these tastes] become incumbent upon all classes lower in the scale" (p. 59).

As the wealthy gain in relative power and social standing as a result of rising economic inequality, social incentives to conform to the canons of taste set by the wealthy become stronger for those further down the income distribution. The stronger the social consequences of non-conformism, the less likely any one individual is to choose that path. Thus individuals may "go along to get along", and the preferences of society as a whole may shift to largely reflect the preferences of the wealthy. The shift of preferences in this direction is harmful for democracy in two ways. First, individuals may be incentivized to vote for policies and representatives that do not have their best interest in mind – thus reducing the overall representativeness of government. Second, individuals will be more likely to vote for authoritarian candidates.

Bowles (1998) offers a more detailed picture of how the endogenous preference channel may operate to induce conformism. Identifying changes in framing and situation construal, intrinsic and extrinsic motivation, effects of the evolution of norms, task performance effects, and effects on the process of cultural transmission as the five main channels through which changes in market structure impact preferences, Bowles (1998) argues that market outcomes are not independent of the social structure(s) that markets create. He writes, "[A]llocation

[3] See the following: http://www.businessinsider.com/kanye-designed-items-sell-out-2013-7

rules also establish relationships among people, based on assignment to distinct positions with corresponding rights, status and obligations" (p.85). For the purpose of the present argument, the evolution of norms and task performance effects will serve best to illustrate the way changes in preferences arising from inequality can be harmful to democracy in the way outlined above.

Consider again the case of the skilled worker. Let us assume for the moment that Atkinson (2015) is correct and "technological change biased towards skilled workers undermines the coalition between them and unskilled workers" (p. 94).[4] In this scenario, the skilled worker has acquired a new norm regarding her behavior toward unskilled workers. If she finds the application of this norm successful at work, she may find it optimal to extend the norm to other spheres of life (Edgerton, 1971; Breer and Locke, 1965; Bowles, 1998). Bowles (1998) writes, "[S]trategies found successful in coping with the tasks defined by one sphere of life are generalized to other realms of life" (p.97). By altering the tasks performed at work (i.e. cooperation vs. non-cooperation among different skill groups in the bargaining process), changes in the structure of the income distribution may alter preferences in other arenas – such as politics – through the process by which successful norms are adapted and generalized from one activity to the next.

This process is harmful for democracy if – as in the case of the skilled worker – the learned behavior is one of non-cooperation. Specifically, antagonism against individuals of lower socio-economic status may show up in a preference for leaders who – at best – are representative of only the interests of the wealthy. Powdthavee and Oswald (2014) find that dramatic changes in one's place in the income distribution – as in the case of the skilled worker – result in a shift of individual preferences toward policies that are more right-wing and inegalitarian. Using exogenous variation in income generated by lottery winnings, the authors find that lottery winners tend to tilt their political preferences to the right following the windfall. The results are found to be robust to different ways of defining "right wing" and conditioning on an individual's having voted for leftist candidates in the past (Powdthavee and Oswald, 2014). One can expect those who benefit from income inequality to shift their vote toward less democratic candidates if the aforementioned benefits – whether they be in the workplace or elsewhere – are contingent on a manifest antagonism toward those lower down the socio-economic ladder. This may be true even if benefits are small, depending on the extent to which a given individual generalizes task-based norms, and can result in individuals further down the income distribution voting for candidates that only have the interest of the wealthy in mind (think for example, of a worker who defects from cooperation when the likelihood of obtaining a promotion is small, but the potential payout is high).

The losers of increasing economic inequality are also subject to changes in preferences. If increasing inequality is – in part – a result of changes in the composition of the labor force due to international competition, then higher unemployment rates and lower wages for formerly well-compensated workers in sectors such as manufacturing will result in shifts in political preference. As labor demand falls for formerly well-paid unskilled labor, the relative bargaining positions of employers and employees are altered as the fallback position of labor if made less attractive. The result of this, Bowles (1998) writes, is that "[t]he surrender of authority to employers which characterizes the labor market appears to support far-reaching psychological effects, some of which undermine the sense of control of one's life" (p. 99). This relative powerlessness of employees *vis-à-vis* employers can negatively affect

[4] See Acemoglu, Aghion, and Violante (2001) and Brandolini (2010) for further support of this point.

democracy in two ways: by increasing individual's incentives to "go along to get along" – and thus increasing voter turnout for authoritarian candidates – or by inciting backlash against authority, increasing the preference for reformist candidates who may be willing to suspend democratic principles to enact economic reform.

In the first case, consider the individual who – as inequality increases – finds unquestioning submission to authority an increasingly successful heuristic in the work place. As this norm gets adopted and extended to other spheres of life, a pattern of preference for authority emerges. This preference is harmful for democracy to the extent that a arge portion of the electorate turn out for authoritarian candidates. If the population whose employment and earnings prospects are negatively affected by increasing inequality is large, this outcome is a distinct possibility. A shift in preferences of this sort seems to be one plausible explanation of the high rates of support for a candidate like Donald Trump in geographic areas of the United States associated with declining employment in manufacturing, mining, and other heavy industry (Irwin and Katz, 2016; Inglehart and Norris, 2016). Additionally, this version of the Bowles-Veblen endogenous preference link provides a behavioral mechanism for explaining the phenomena observed in philosopher Hannah Arendt's (1951; 1963) work on the origins of totalitarianism. Arendt's study of totalitarian regimes in the 20th century relies on a notion of "the banality of evil" as a state of thoughtless in which an individual becomes "rinsed with clichés, norms, ideologies, and national ethos" (Kilkenny, 2016), and uses those norms as justification for acts supporting political authoritarianism. The condition(s) under which the banality of evil may prevail are socio-economic in nature, and may thus be explained through the process of preference formation outlined above.

In the second case, consider an individual who finds submission to authority an unsuccessful heuristic. This may be one of the many long-term unemployed individuals who lost their job as a result of the 2008 financial crisis. Finding application of the "go along to get along" principles of their peers unsuccessful, these individuals may move to demand political reform. There is here a link between the endogenous preference argument and the earlier arguments from Alesina and Rodrick (1994) and Alesina and Perotti (1996). The extent of political reform demanded (and whether a line is crossed between pressure for reform and political instability) will depend on the extent of inequality, as Alesina and Perotti (1996) argue. Furthermore, the robustness of the political system to instability will depend on the extent to which individuals adopt either the "go along to get along" model or the "political extremism" model. A particular country's political institutions may be robust to high levels of inequality if the preferences of the electorate have been shaped in such a way as to prefer submission to authority.

If the candidate both promises reform and satisfies the desire for submission to authority, this candidate may be successful at capturing a large swath of the electorate that has been – in one way or another – alienated by increasing inequality. There is no reason a priori why an individual may not hold both a deference to authority and a desire for strong redistributive reform. Both of the preference formation processes outlined above may operate on a single individual along different dimensions. Depending on the extent of reform promised and the degree of authoritarianism, the result will be detrimental to democracy.

IV. Critiques and discussion

The difficulties for the endogenous preference argument are both conceptual and empirical. In this section I will attempt to outline what I see as the biggest problem for the argument, as well as offer a response from the institutionalist perspective.

The biggest problem for the endogenous preference argument is a lack of formal identification. The exact process through which preferences are altered via economic, social, or political institutions is something of a black box. One can make inferences about an individual's preferences only indirectly through behavioral outcomes. While observed outcomes may indicate that preferences have indeed shifted – the theoretical models of economics are ill-equipped to describe a problem of choice under shifting preferences. Indeed, it is not clear how one would write down a utility maximization problem at all if preferences were not assumed constant. Thus, as one institutionalist wrote, economists "unduly given to difficult theoretical niceties" may find themselves at odds with the Bowles-Veblen approach (Veblen, 2009[1899], p. 31).

Despite the difficulty in developing a formal theoretical model of endogenous preferences, the increasing quality and availability of public opinion and consumer expenditure data makes possible the empirical identification of shifts in preferences. Using an appropriate combination of survey responses to questions about preferences and consumption bundles, it is possible to identify empirically whether a given behavior is a result of a shift in preferences (a change in the utility function) or the result of changes in prices and income along a given utility function. There is no reason to assume – as is often done in empirical work – that preferences are constant and all changes in behavior are the result of shifts in relative prices or income, a point emphasized by Joan Robinson (1962).[5]

Furthermore, a lack of analytical tractability does not imply a lack of fidelity to reality. In fact, as Friedman (1953) and others freely admit, it is often precisely *unrealism* that characterizes the assumptions of neoclassical models. One should fear for the state of the science if mathematical tractability is prized above the ability to depict reality. The picture of the world painted by the institutionalist may be messier than mainstream economists would prefer, but that does not mean that is not in fact how the world is. As Veblen (2009) wrote, "[I]n the organic complex of habits and thought which make up the substance of an individual's conscious life the economic interest does not lie isolated and distinct from all other interests" (p.78).

V. Conclusion

I will conclude by summarizing the arguments made thus far and drawing out some directions for future research. I have shown that there are four main ways economic inequality may potentially harm democracy. Economic inequality may simultaneously increase political responsiveness to the wealthy and decrease responsiveness to the poor and the middle class, it may increase political instability, it may increase the ability of the wealthy to subvert market reforms enacted by political consensus, and it may alter the preferences of the electorate in ways that favor authoritarianism.

[5] And, as usual, ignored by the rest of the profession for the sake of convenience.

I have argued that, while the first three aforementioned consequences are studied elsewhere in the inequality literature, little attention has been paid to the way inequality shapes tastes of the electorate. This channel is important if one is concerned with the type of leaders that may be elected during times of high inequality. If increasing inequality either A) increases individuals' preferences for authority or B) increases individuals' preferences for downward redistribution (rather than assuming – as in Alesina and Rodrick (1994) – a given preference for redistribution), the outcome may be bad for democracy. Furthermore, the way inequality shapes preferences is important for the other channels by which economic inequality affects the political process. Depending on preferences and the institutions that flow from them political institutions may not be robust to high levels of inequality. The extent to which inequality is harmful to democracy depends on the institutional arrangements in the country in question.

Two outlets for future research suggest themselves as profitable. The first, as mentioned earlier, is a general test of the institutionalist hypothesis using joint data on consumption and public opinion. The second is the use of public opinion data to test which of the preference modification channels – the "go along to get along" channel, the "political extremism channel," both, or neither – have tended to operate as a result of increasing economic inequality over time. Survey questions addressing preference for authority and downward redistribution, along with data on inequality and other explanatory variables, may be used in such a study. Answers to these questions can help guide policymaking in the presence of a shifting income distribution by alerting policymakers to the potential trade-offs that occur as preferences in the electorate shift in response to new policy. Such answers are important, the study of inequality is incomplete without understanding the institutional context within which it operates.

References

Acemoglu, D. et. al. (2001). "Deunionization, Technical Change, and Inequality." *Carnegie Rochester Conference Series on Public Policy,* 55: 229-264.

Alesina, A. & Perrotti, R. (1996). "Income Distribution, Political Instability, and Investment." *European Economic Review*, 40(1): 1203-1228.

Alesina, A. & Rodrick, D. (1994). "Distributive Politics and Economic Growth." *Quarterly Journal of Economics*, 109(2): 465-490.

Apel, H. (2015). "Income Inequality in the U.S. from 1950 to 2010 – the Neglect of the Political." *real-world economics review*, 72(1): 2-15.

Arendt, H. (1951). *The Origins of Totalitarianism.* Fort Washington: Harvest Book.

Arendt, H. (1963). *Eichmann in Jerusalem: A Report on the Banality of Evil.* New York: Penguin Books

Atkinson, A. (2015). *Inequality: What Can Be Done?* Cambridge: Harvard University Press.

Bartels (2008). *Unequal Democracy: The Political Economy of the New Gilded Age.* Princeton: Princeton University Press.

Bowles, S. (1998). "Endogenous Preferences: The Cultural Consequences of Markets and other Economic Institutions." *Journal of Economic Literature,* March: 75-111

Brandolini, A. (2010). "Political Economy and the Mechanics of Politics." *Politics and Society,*38: 212-226.

Breer, P. & Locke, E. (1965). *Task Experience as a Source of Attitudes.* Homewood: The DorseyPress.

Citizens United v. Federal Election Commission 558 U.S. 310

Edgerton, R. (1971). *The Individual in Cultural Adaptation.* Berkeley: University of California Press.

Esteban, J. & Ray, D. (2006). "Inequality, Lobbying, and Resource Allocation." *American Economic Review.* 96(1): 257-279.

Friedman, M. (1953). *Essays in Positive Economics*. Chicago: University of Chicago Press.

Galbraith, James. (2016). *Inequality: What Everyone Needs to Know*. Oxford: Oxford UniversityPress.

Galbraith, John. (1952). *American Capitalism*. Piscataway: Transaction Publishers.

Gilens, M. (2005). "Inequality and Democratic Responsiveness." *Public Opinion Quarterly*, 69(5): 778-796.

Hacker, J. & Pierson, P. (2010). *Winner-Take-All Politics.* New York: Simon & Schuster.

Hilber, C. & Robert-Nicoud, F. (2013). "On the Origins of Land Use Regulations: Theory and Evidence from US Metro Areas." *Journal of Urban Economics*, 75: 29-43.

Inglehart, R. & Norris, P. (2016). "Trump, Brexit, and the Rise of Populism: Economic Have-Nots and Cultural Backlash." *HKS Faculty Research Working Paper Series RWP16-026*, August.

Irwin, N. & Katz, J. (2016). "The Geography of Trumpism." *The New York Times.* March 12.

Kilkenny, K. (2016). "Why Hannah Arendt Matters in 2016." *Pacific Standard.*

Klasen, S. (2008). "The Efficiency of Equity." *Review of Political Economy,* 20(2): 257-274.

Malthur, L. et. al. (1997). "The Wealth Effects Associated with a Celebrity Endorser: The Michael Jordan Phenomenon." *Journal of Advertising Research*, 37(3): 67-73

Mitra, D. (1999). "Endogenous Lobby Formation and Endogenous Protection: A Long-Run Model of Trade Policy Determination." *American Economic Review*, 89(5): 1116-1134.

Perrotti, R. (1993). "Political Equilibrium, Income Distribution, and Growth: Evidence from American Data." *Review of Economic Studies,* 60(4): 755-778.

Piketty, T. (2014). *Capital in the Twenty-First Century*. Cambridge: Harvard University Press.

Powdthavee, N. & Oswald, A. (2014). "Does Money Make People Right-Wing and Inegalitarian? A Longitudinal Study of Lottery Winners." *IZA Discussion Paper No. 7934.*

Robinson, J. (1962). *Economic Philosophy*. New York: Penguin Books.

Stiglitz, J. (2012). *The Price of Inequality*. New York: Norton.

Veblen, T. (2009)[1899]. *The Theory of the Leisure Class*. Oxford: Oxford University Press.

Wisman, J. (2013). "Wage Stagnation, Rising Inequality and the Financial Crisis of 2008." *Cambridge Journal of Economics*, 37(4): 921-945.

Zucman, G. (2013). "The Missing Wealth of Nations: Are Europe and the U.S. Net Debtors or Net Creditors?" *Quarterly Journal of Economics*, 128(3): 1321-1364.

Zucman, G. (2015). *The Missing Wealth of Nations*. Chicago: University of Chicago Press.

Author contact: lapetach@rams.colostate.edu

SUGGESTED CITATION:
Luke Petach, "Politics, preferences, and prices: the political consequences of inequality", *real-world economics review*, issue no. 80, 26 June 2017, pp. 2-13, http://www.paecon.net/PAEReview/issue80/Petach80.pdf

real-world economics review, issue no. 80
subscribe for free

The triumph of Pareto*

Gary Flomenhoft [University of Vermont, USA]

1. Introduction

Neo-classical economics (NCE),[1] or mainstream economics, is touted as a positive, "value-free" science. Economic efficiency is defined by the criteria of Pareto optimality as if it were a perfectly objective, value free measure of efficiency. When looked at in depth we find that Pareto optimality is highly charged with underlying normative values, assumptions, and implications. In part one of this article I will explore the underlying meaning of Pareto Optimality (aka efficiency), and conclude that it is actually a normative criteria masquerading as "value-free". Part two explores the philosophies of Vilfredo Pareto, shows that they are entirely consistent with his principle of optimality, demonstrates their relationship to recent economic conditions, and concludes that his views and values have triumphed. Finally in part three, I will propose some solutions. Of particular relevance to Daly's work are the growth imperative and inequalities that result from using Pareto optimality as a measure of efficiency. Daly has long advocated for consideration of fair distribution in his work as one of the key principles of ecological economics, contrary to most of the economics profession, which has abandoned fairness in favor of growth.

1.1 Pillars of neo-classical welfare economics

The three pillars of neo-classical welfare economics are Homo-economicus – self- interested, utility maximizing, rational economic man; Perfect competition – independent actions of firms, no market power, constant returns to scale, perfect information, no uncertainty; and Pareto Optimality or Potential-Pareto Optimality (also known as the Kaldor-Hicks criteria). Behavioral science has convincingly refuted the premise of economic man (e.g. Tversky and Kahneman, 1981; Gintis, 2000; Kahneman and Tversky, 2000; Ariely, 2008; Thaler and Sunstein, 2008), while the "Nobel prize" to Joseph Stiglitz for asymmetrical information demonstrated the rarity of perfect competition. Though neoclassical economists have largely abandoned the assumptions of economic man and perfect competition in their empirical research, "the profession is just beginning to come to grips with the policy implications of abandoning Pareto optimality" (Gowdy and Erickson, 2005).

Pareto optimality is still the fundamental goal of neo-classical welfare economics. It says that a state in which no one can be made better off without making someone else worse off is Pareto optimal or efficient. However, Pareto optimality offers no guidance when confronted with a policy that lifts a million people out of dire poverty but imposes a minor loss on a single billionaire. Any policy alternative that makes anyone worse off in their own estimation, even while other people are made better off in their own estimation, are "Pareto incomparable." Since they fail the Pareto efficiency criteria they are by implication undesirable. Economics

* This paper was originally written for the Festschrift honoring Herman Daly, which explains all the references to Daly.
[1] See Gowdy (chapter 6 in this volume) for a more detailed definition of neoclassical economics.

claims to be a "value-free", objective science. If this is the case, then widely used principles of economics, such as Pareto Optimality, should stand up to scrutiny as being value-free.

1.2 Pareto optimality in the view of economics

Mainstream economists refer to Pareto optimality as Pareto efficiency, or simply efficiency. No one is opposed to "efficiency". If markets guarantee "efficiency", then by implication they must be good: If we can make at least one person better off with no sacrifice, then shouldn't we do so? Mainstream economists obsess on the importance of diminishing marginal utility in the context of individual choice. However, they virtually ignore the implications of diminishing marginal utility in the context of income redistribution, and exalt Pareto optimality as the single criterion for economic desirability. Mainstream economists admit that there are an infinite number of possible Pareto efficient outcomes, with a different one for every possible distribution of wealth and income. How do we choose between them? Daly has argued that sustainability and distribution take priority over any measure of efficiency and therefore a Pareto efficient outcome will only be desirable if it results from sustainable levels of throughput and a just distribution of wealth and resources (Daly, 1992).

Efficiency in any technical field is normally defined by the ratio of output to input. Why choose Pareto optimality as the measure of efficiency? Pareto efficiency seems trivial compared to other measures of efficiency such as GDP per unit of energy, value of output per unit of labor (labor productivity), well being per unit of throughput, Genuine Progress Indicator (GPI) per unit of ecological footprint, and so on.

2. Problems with Pareto optimality

There are several serious problems with the use of Pareto Optimality as the central goal of economics. One of the primary implications of "Pareto Optimality" is that it accepts the current distribution of wealth as a given. If any income distribution can lead to an "optimal" outcome, there is no need to be concerned about just distribution. Mainstream economists generally make no distinction between earned or unearned income, and when they do, it is typically to recommend lower taxes on the latter, a policy that has been adopted in the USA. Mainstream economists rarely question the legitimacy of the source of wealth. Fairness or origin of the current distribution is a problem left to politicians and society. It is not considered a legitimate question for economics. As Steven Hackett points out, "Slavery was widely seen in the North as being unethical from a deontological perspective, but a policy alternative of ending slavery would make slave owners worse off than under the status quo, and thus would have failed the Pareto efficiency criterion" (Hackett, 2001: 26).

While Pareto developed his theory of Optimality, he also described the idea of "indifference curves" in conjunction with Edgeworth, which captures each person's preferences for different combinations of various goods. It established the idea of ordinal rather than cardinal welfare, and eliminated comparisons of utility between people. According to Pareto and current neo-classical orthodoxy, each person can only decide how well off they are "in their own estimation". This avoids any consideration of justice in the current social conditions. In the words of Daly, "The extreme individualism of economics insists that people are so qualitatively different in their hermetical isolation one from another that it makes no sense to say that a leg amputation hurts Smith more than a pin prick hurts Jones" (Daly and Farley, 2010: 306).

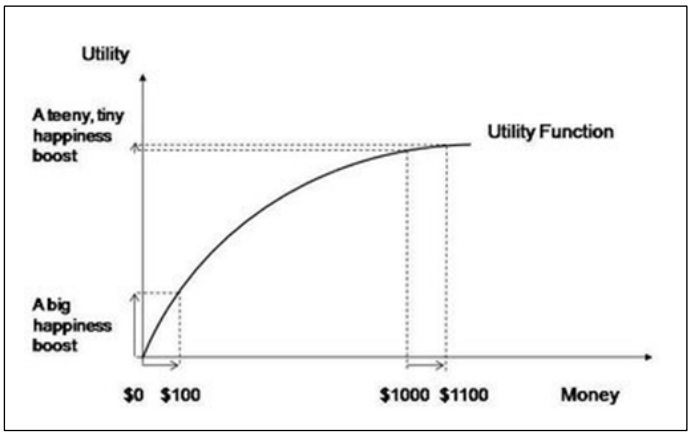
Economics was originally based on classical utilitarianism, which followed the philosophy of "the greatest good for the greatest number", and thus was very concerned with issues of distribution. Maximizing the total utility of society was the goal, and it was well understood that extra income provided more utility for a poor person than for a rich one.

Following Pareto, this was abandoned in favor of ordinal measures of welfare. Interpersonal comparisons of utility are still generally considered outside the bounds of neo-classical welfare economics.

Utility curves of an individual assume diminishing marginal utility: the more of something a person has, the less utility an additional unit provides. Figure 1 below depicts the difference in utility from receiving $100 for a person with $1000 compared to the same person with $0. Is it really so far-fetched to believe that two human beings might have similar utility curves, especially when satisfying basic physiological needs? If this curve represented two people, rather than one person at different times, then the wealthier person obviously gets less utility from $100 than the poorer person. Economists can only accept Pareto efficiency as a central goal of economics by largely rejecting the notion of diminishing marginal utility.

Therefore Pareto Optimality is self-contradictory.

Figure 1 Interpersonal comparison of Utility



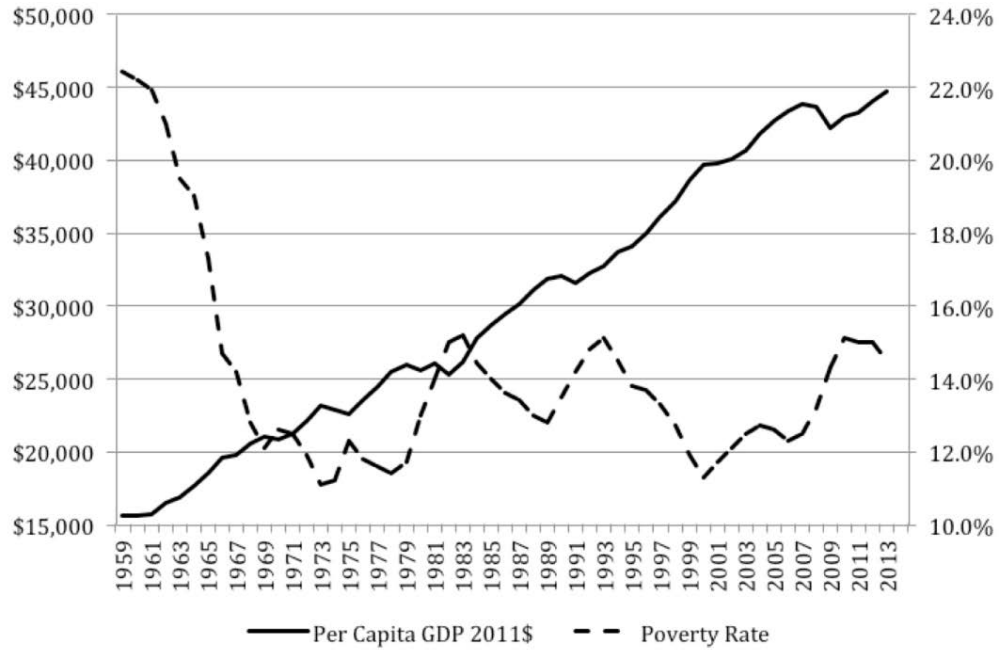
Source: http://mrski-apecon-2008.wikispaces.com/Sun%27s+Page.

One of the primary implications of Pareto optimality is that economists cannot pass judgment on the desirability of different distributions of wealth and income. "Potential Pareto optimality" relaxes this criterion by declaring one option superior to another if the winners could potentially compensate the losers through transfer payments, even if no compensation actually takes place (Kaldor-Hicks criteria). Actual compensation is left to society to take care of and is out of the realm of economics.

Economists' obsession with Pareto optimality has led them to virtually ignore the welfare implications of upward and downward redistribution. If redistribution of the pie is considered off-limits, then the only option left for improving welfare is a bigger pie, typically measured by higher GDP. Also, since the measure of welfare is entirely subjective, how could one determine who feels better or worse? It's much safer to assume that a rising tide lifts all boats, and don't worry about the people with no boats. It is universally believed that we can grow our way out of poverty. Does reality support the myth? Prior to 1967 the poverty rate appeared to

decline as GDP increased. However, as shown in Figure 2, since 1968 there has been no relationship between long term GDP growth and poverty alleviation as is commonly believed.

Figure 2 Poverty rate vs GDP



Source: Author from US Census Bureau data, 2014.

Even if growth did contribute to poverty alleviation it is questionable if it is a viable option any longer due to the limits of planetary growth. It would require an estimated 4.5 planets to extend current levels of US consumption of resources and emission of pollutants to all 7 billion people on the earth (Ewing et al., 2010).

3. Pareto's other philosophies, their connection to Pareto optimality, and to current social conditions

In order to understand the genesis of Pareto's principle of optimality it is important to realize that this principle did not emerge whole cloth in isolation, but was a reflection of his deeply social Darwinist views. The overwhelming dominance of neoclassical market economics in modern society, particularly in the US, therefore affected other attitudes and policies in our society, and getting economists to accept Pareto optimality paved the way for the triumph of Pareto's ethical values. In adopting Pareto Optimality, economists unleashed a hidden principle driving society toward ever-increasing inequality. Since the official end of the recession caused by the financial crisis, the top 1% of income earners captured 95% of the income gains and now earn over 20% of the total income, with the top 10% receiving over 50% of total income (Pikkety and Saez, 2006).

This section of the paper therefore examines Pareto's moral philosophy, and assesses the extent to which it is reflected in the American economy. In 1916 Pareto wrote what he considered his greatest work, *Mind and Society*. In this work he elaborated his sociological theory of the "Circulation of the Elites". Persons of superior ability actively seek to confirm and

aggrandize their social position. The best-equipped members of the lower classes rise to challenge the position of the upper-class elites. This theory best captures Pareto's social Darwinist beliefs, and also corresponds with recent economic history.

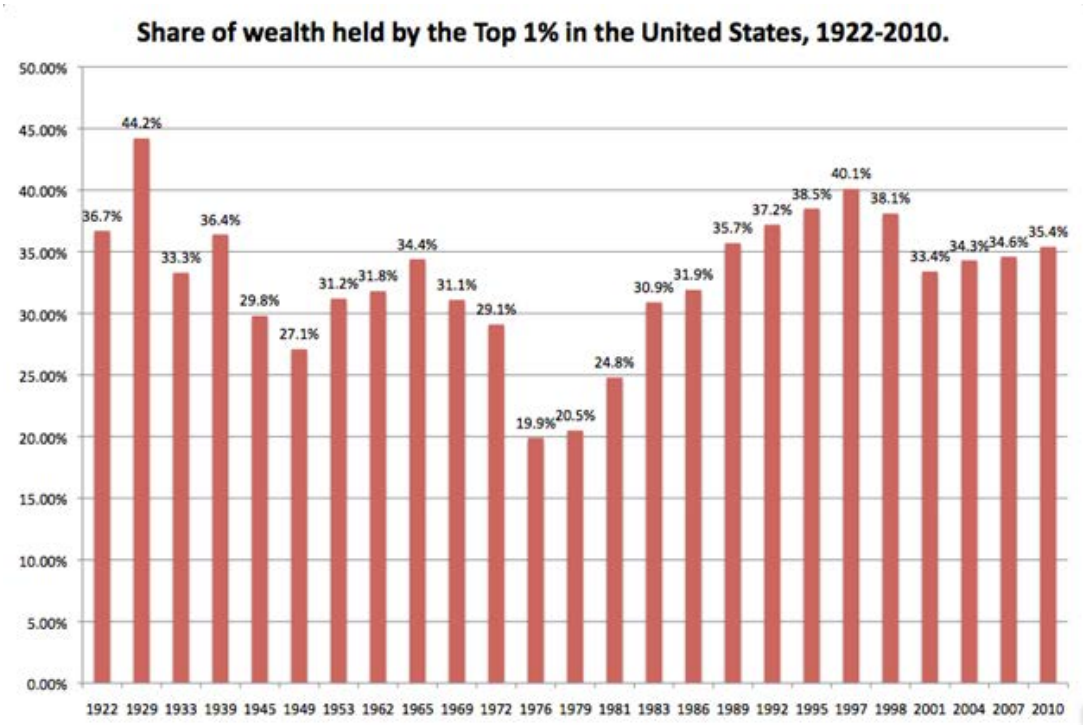
The period of 1980-2008 was one of the longest sustained periods of economic growth in US history. It included a period of neo-liberal supremacy after the demise of the Soviet Union. A huge increase in economic and social inequality also occurred during this time. A combination of factors led to this result. Because of the demise of the communist system, the promoters of capitalism now had "proof" that their system was better and celebrated the "End of History" with the emergence of free market democracies as supposedly the best possible system (Fukuyama, 1989). Free from a countervailing adversary, neo-classical economists, in conjunction with neo-liberal business and political interests, were given free reign to promote their agenda in advocating free-trade, globalization, structural adjustment, corporate downsizing, elimination of the social safety net, destruction of unions, privatization, etc. How did it work out? Pareto's other principles offer a lens through which to evaluate the results of the neo- classical economic (neo-liberal political) resurgence from 1980 to present.

3.1 Pareto on social class

> "…no social class can for long hold its property or its power if it does not have the strength and vigor necessary to defend them. In the long run only power determines the social forms; the great error of the 19th century will be to have forgotten this principle" (Pareto, 1906: 361, cited in Gaffney and Harrison, 1994).

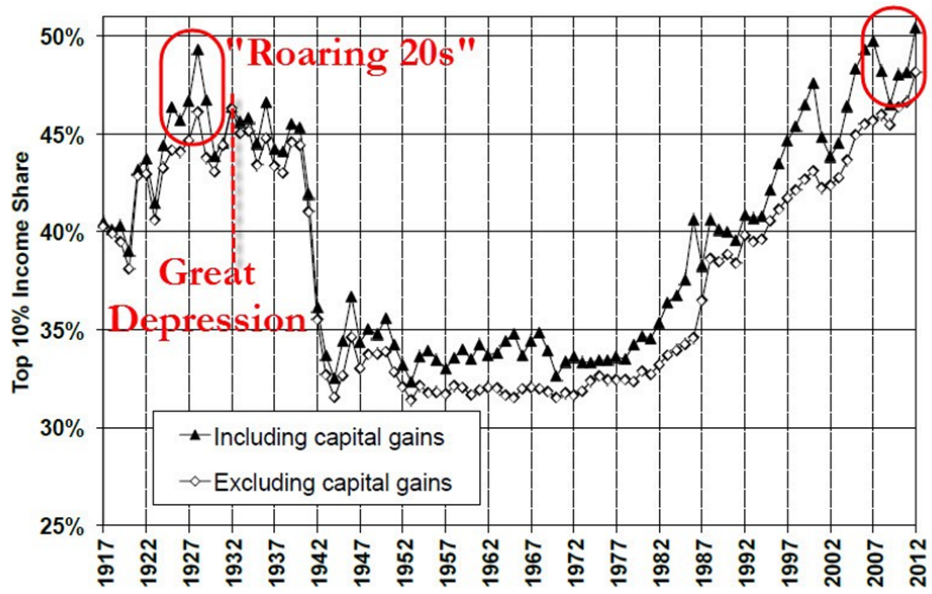
Pareto was certainly a supporter of class warfare, the right of elites to control property, and the primacy of power over other social considerations such as justice or equality that are espoused by democratic societies. Let's examine the evidence of income and wealth distribution by economic class, to see if it is consistent with this principle. The first factor we can observe since the 1980s is the concentration of US wealth held by the top 1% of the population. It reached 40.1% in 1997, its greatest level since just before the stock market crash of 1929, when it reached 44.2%. Concentration dropped slightly due to the dot.com crash around 2000, but trended upward to 34.6% until the financial crisis of 2008. After the crash the trend continued up to 35.4% in 2010 (Figure 3). This was due in large part to stock market gains going mainly to the wealthy, but also correlated with increasing disparities in income and benefits, tax reductions, and other factors. Certainly, control over approximately 40% of all wealth would appear to give the top 1% the strength to defend its property and power.

Figure 3 Top 1% share of household wealth (Wolff, 2001; 2010; 2012)



In terms of income we see a similar story. The top 10% of society was able to exceed 50% of the total income share in 2012 (Figure 4).

Figure 4 The top decile income share, 1917-2012



Source: (Piketty and Saez, 2006; Saez, 2013). For more information on trends in the distribution of wealth and income, see Atkinson et al. (2011); Alvaredo et al. (2013).

Pareto further claimed that:

> "the dominant class A has an alpha part with still enough strength and energy to defend its share of authority; and a beta part made up of degenerated individuals, with feeble intelligence and will, humanitarians, as is said today…Objectively, the struggle consists solely in the B-alpha trying to take the place of the A-alpha; everything else is subordinate and incidental" (Pareto, 1927: 91, cited in Gaffney and Harrison, 1994).

Pareto explicitly describes a patriarchal battle for dominance as the existing and desirable state of society. While admiring the elite, Pareto also somewhat admired the self-seeking B-alphas, leaders of the lower classes who seek to overthrow the upper classes. Pareto's "Circulation of the Elites", to anyone who has studied primate behavior, is nothing but a description of the dominance hierarchy of chimpanzee society. Alpha males (A-alphas) rule, and are constantly challenged by beta males (B-alphas) for dominance. Despite the wide variety of primate and human societies throughout evolutionary history, including matriarchal and matrilineal societies, and egalitarian Neolithic and tribal societies, Pareto holds Chimpanzee society up as the human ideal. Even among apes, chimpanzees are more aggressive and hierarchical compared to others such as Bonobos or Orangutans. Is chimpanzee society what we aspire to as human beings?

3.2 Distribution of wealth 1947-2013

Additional evidence on income distribution reveals that the "Circulation of the Elites" is consistent with recent history. While real family income grew fairly equally across the board from 1947-1979 (Figure 5), in the period from 1982-2006 inequality vastly increased (Figure 6). The gains increased with increasing income, and those at the top increased their already much higher income by 16 times as much as those at the bottom. This is considered Pareto "efficient" since no one received less income. Yet do we consider this outcome desirable? It appears to be exactly what Pareto had in mind. In household net worth, the situation was even worse. People at the top got better off, while people at the bottom got much worse off, therefore failing the Pareto efficiency criterion (Figure 7). During these years we heard much about "irrational exuberance", and little or nothing from economists about wealth or income inequality. We certainly heard nothing about the decline in household net worth of the bottom 40% of society as a violation of Pareto efficiency. Evidence suggests economic growth in recent decades has made the rich richer and poor poorer. If Pareto efficiency is our sole criterion for judgment, we cannot say whether such growth is good or bad, while the Kaldor-Hicks criterion of potential Pareto efficiency says that aggregate growth is good no matter how unequal the distribution.

Figure 5 1947-1979 Real family income growth by quintile for top 5% (Collins, 2000)

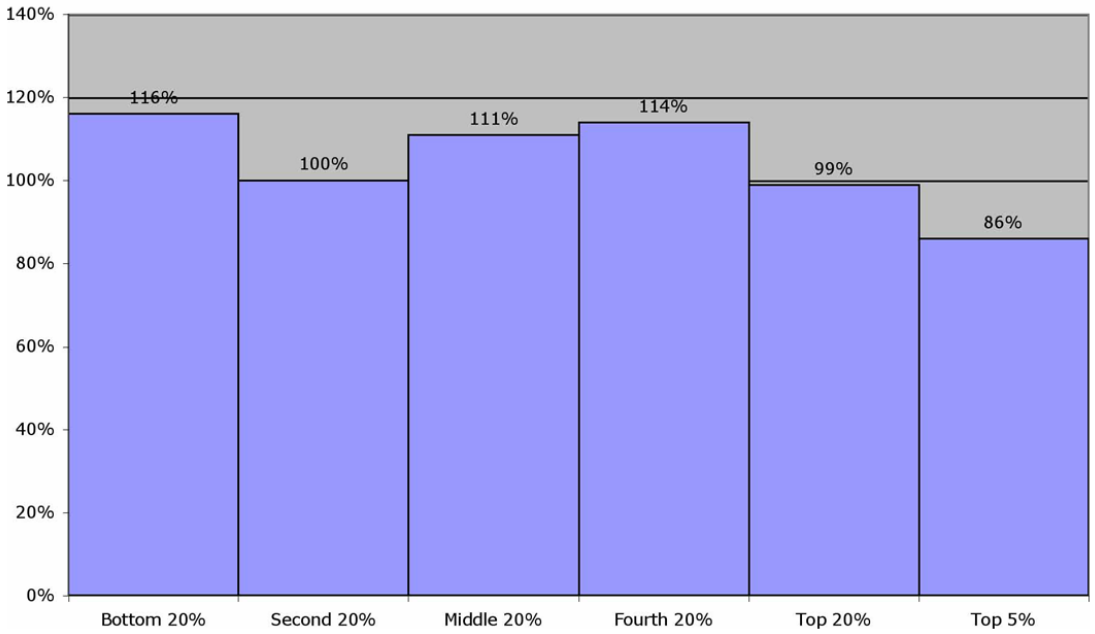


Figure 6 1982-2006 Real household income growth (Wolff, 2010)

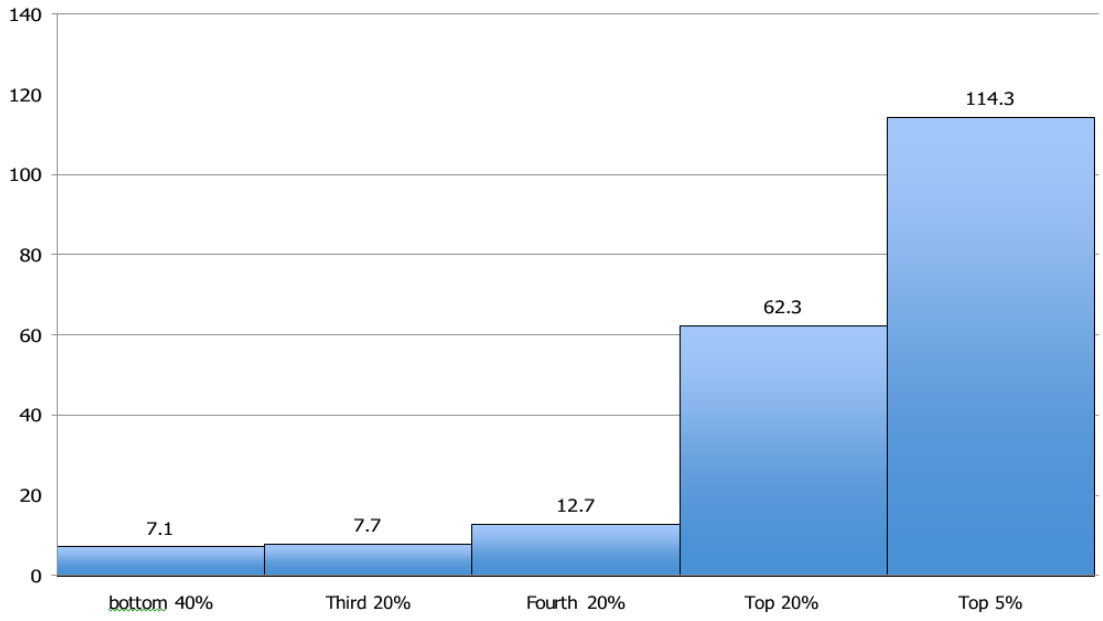
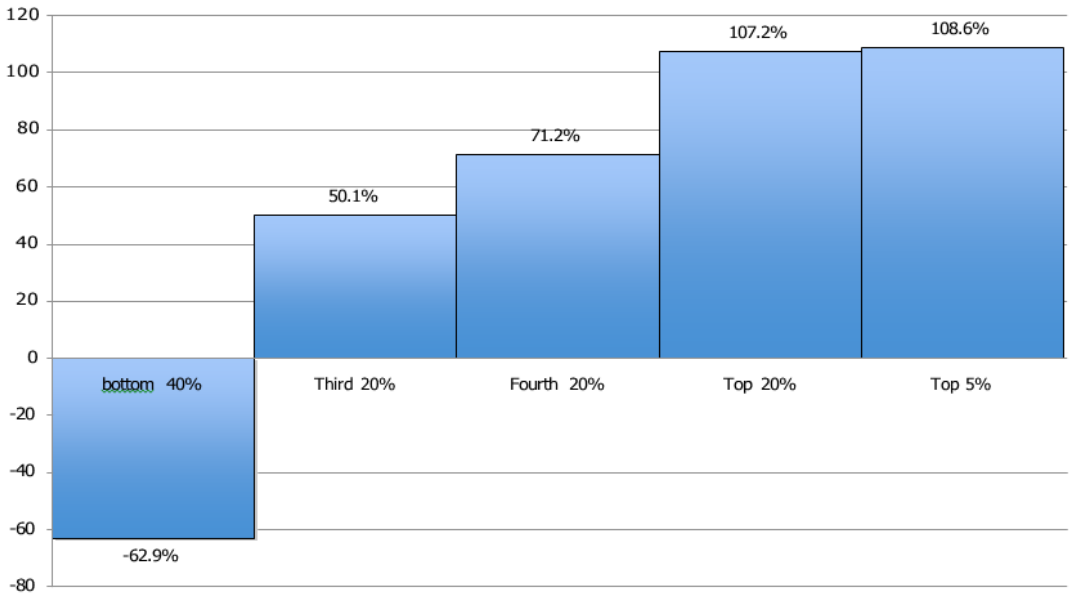


Figure 7 Change in average household net worth (Wolff, 2010)



3.3 Pareto on genetic selection

> "The struggle to appropriate the goods of others may be favorable to (genetic) selection" (Pareto, 1906: 341).

CEO pay reflects this principle as corporate management in the US drastically increased its share of corporate income from 1980-present. We have heard the justification from Wall Street that CEOs are the "Job creators", and that "they need high salaries to recruit good people", and that CEO pay reflects performance. But this is all very self-serving as CEOs received huge bonuses even while their companies stock prices were dropping drastically in 2008, or going bankrupt. From 1980 to 2000 the ratio of CEO pay to the average worker increased from 42:1 to 411:1 (Figure 8). Though the ratio of CEO to worker pay in the US dropped to 209:1 in 2009, it has since rebounded to 354:1 according to the AFL-CIO (2013). Many other countries such as Germany and Japan are also experiencing dramatic increases in the ratio of CEO to worker pay, but starting from a much more equal base (Khazan, 2013).

Labor productivity increased eight times faster than pay during this same period of time, 1979-2013 according to the Economic Policy Institute (Figure 9). According to neo-classical economic doctrine, wages always rise in proportion to labor productivity. From 1972 onward, there was basically no corresponding increase in wages. Real wages remained virtually unchanged. CEOs appropriated huge increases in compensation from increased productivity while giving their employees nothing, in keeping with Pareto.

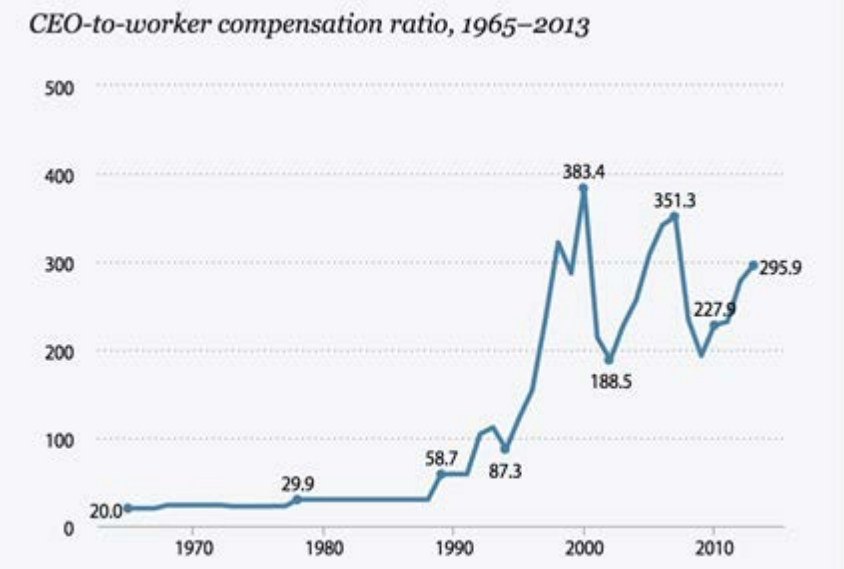
Much of Pareto's writing suggests that he would exalt the CEOs position and be contemptuous of the passive masses who accepted this condition. Several important factors led to the stagnation of real wages in the US:

- The decline in unionization from 35% in 1948 to 14% in 1999.
- The export of jobs due to globalization and a "race to the bottom" for wages (Daly and

Goodland, 1994). (Free-trade zones, NAFTA, WTO, etc.)

- Increased temporary and contract workers with no benefits
- Layoffs resulting from corporate downsizing and mergers.
- The loss of manufacturing jobs.

Figure 8 Ratio of US CEO pay to average worker



Source: Economic Policy Institute (http://www.epi.org/publication/the-top-10-charts-of-2014/)

3.4 Pareto on suffrage

> "When suffrage has been given to all men, including madmen and criminals, when it has been extended to women, and if you like, to children, it will have to stop. One cannot go any lower, unless the suffrage is extended to animals" (Pareto, 1906: 100, cited in Gaffney and Harrison, 1994).

Aside from being a social Darwinist, Pareto was also a misogynist. Is it surprising that economic statistics show that women suffer the most from economic deprivation in a culture where economists follow Pareto's principles? Women, children, and single parent families are most often the victims of hunger. In 2001, 60% of adults seeking food assistance were women, and more than 40% of the total were children (Gazette Community News, 2001). Poverty has risen since 2001. In 2010, 58.7% of single mothers with children under 6 were poor, and 34.2% of all women-headed households were poor (Figure 10). Pareto's goal of disenfranchising women and children has been achieved.

Figure 10 Poverty among families and children

2010	All families	Children <18	Children <6
All families	13.2%	21.5%	25.4%
Married couple	7.6%	11.6%	13.3%
Male householder, no wife	17.3%	28.1%	32.2%
Female householder, no husband	34.2%	46.9%	58.7%

Source: Author from US Census Bureau data 2012.

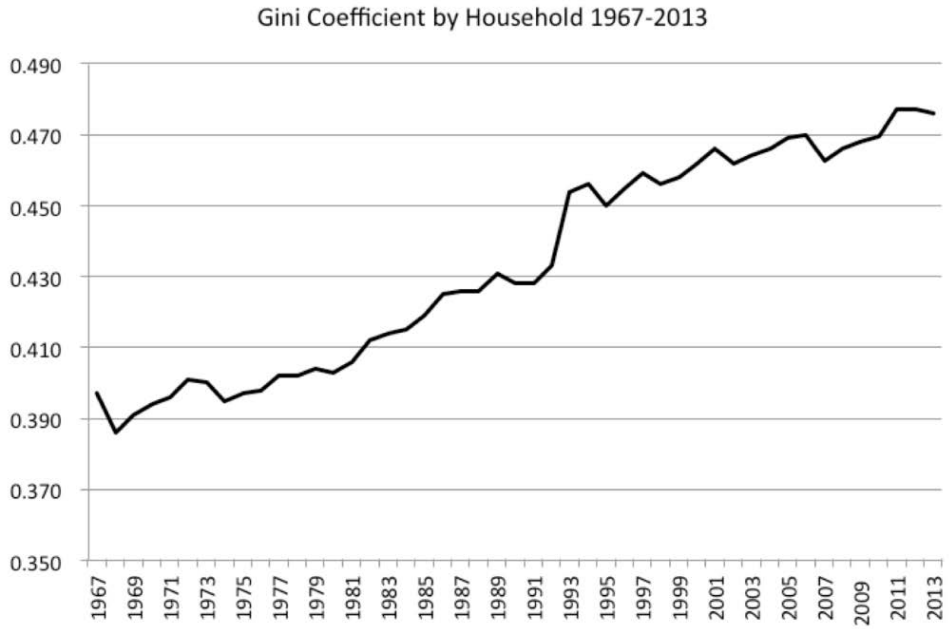
3.5 Pareto on equality

> "Equality before the law is a dogma for many people…it is not at all evident a priori that such equality is advantageous to society;…the contrary appears more probable" (Pareto, 1906: 95, cited in Gaffney and Harrison, 1994).

Pareto's belief that inequality is good for society seems to be shared by neo-classical economists. Mainstream economics over the last several decades have espoused the belief that inequality is necessary because it provides incentives for economic growth. Furthermore, they claim that inequality is the natural state of society and that to oppose it would be against natural law. On that basis they have opposed policies such as increasing the minimum wage, and ignored the fact that lower income people have a higher propensity to spend and hence to stimulate economic activity. Therefore, economists have fostered a tolerance for the extreme disparities in wealth that we see today. There has been extensive recent experience with inequality as it has rapidly increased over the last several decades. So we can evaluate the results empirically to test Pareto's theory that inequality is good for society.

The Gini coefficient is a commonly used measure of income inequality. While declining slightly prior to 1968 in the US, it has been rising dramatically since then (Figure 11).

Figure 11 US Gini coefficient



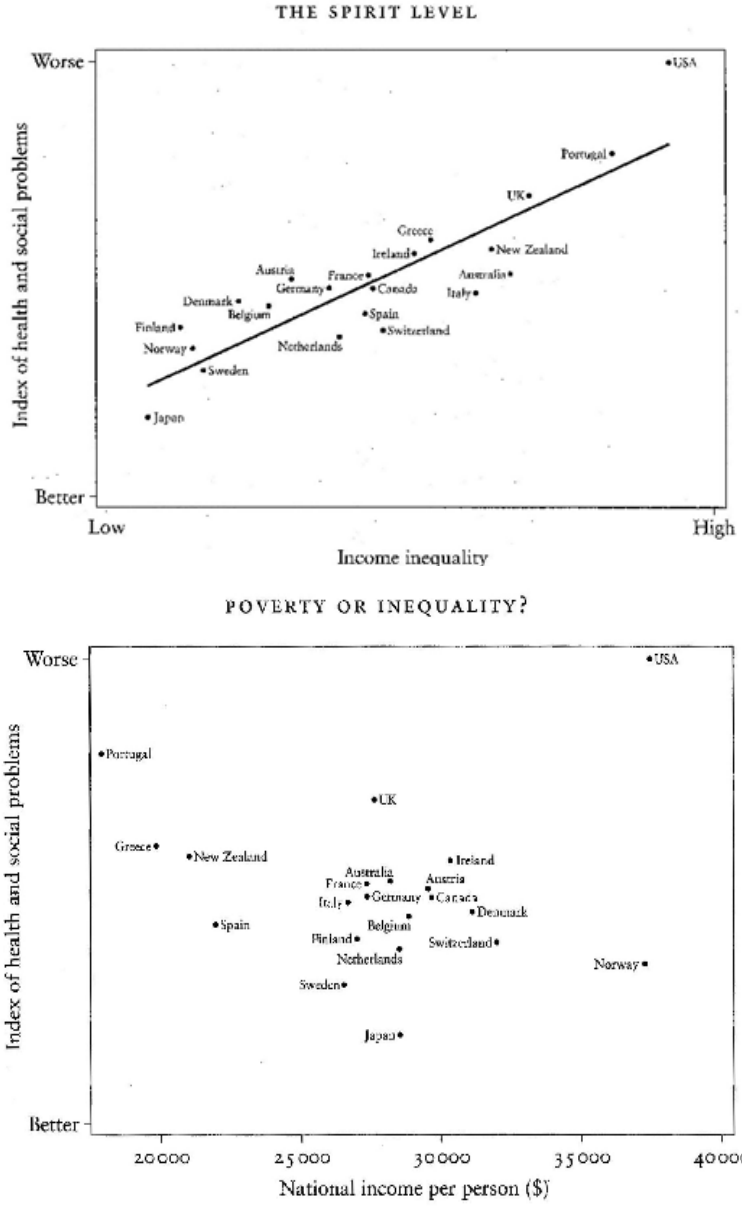
Source: Author from US Census Bureau data 2014.

Is inequality good for society?

The notion that inequality is good for society, although espoused by neo-classical economists and neo-liberal (US conservative) politicians following Pareto, has been utterly refuted most recently by the book "The Spirit Level". Authors Wilkinson and Pickett (2009), found that numerous indicators of health and social problems, including trust, mental illness, life expectancy, infant mortality, obesity, children's educational performance, teenage births, homicides, imprisonment rates, and social mobility are all made worse by inequality. They found this correlation to hold both between countries and between the 50 US states. Furthermore, the greater the degree of inequality, the worse the outcome as shown in the following charts.

Surprisingly, health and social problems show a weak inverse correlation to average income among similar countries, but strong positive correlation with inequality:

Figure 12 and Figure 13 Source (Wilkinson and Pickett 2009)



3.6 The triumph of Pareto, economic growth, and welfare

In the last 30 years we have seen the implementation of Pareto's values through laissez-faire, free-market capitalism as promoted by neo-classical economists and neo-liberal advocates. We have seen a huge redistribution of wealth and income from the lower income groups to the upper income groups. The greatest poverty is among single women and children. Nonetheless, there seems to be a growing belief that downward redistribution of any kind is considered morally repugnant.

Remember, Pareto considered humanitarians who want to share wealth equitably to be "degenerated individuals". The economy grew, but inequality grew much greater. It is truly the triumph of Pareto.

The response of mainstream economists to this growing inequality has been to call for yet more growth to reduce poverty and inequality. As this article illustrates, economic growth and poverty reduction, in the US at least, are completely unrelated. Furthermore, it is debatable that growth in material throughput is sustainable for much longer. The most effective way to increase the total utility of society is to provide for the material needs of those with the greatest marginal utility from increases in income, namely those at the bottom of the economic scale.

How to go about this is an essential question.

4. Policies to address poverty and inequality

Given the political difficulty of a Robin Hood tax-the-rich and redistributed downward scheme, perhaps the best way accomplish to reduce poverty and inequality is through the sharing of rent from the commons[2] Many policy recommendations have been made to address poverty that do not necessarily require additional economic growth. They include higher and more progressive taxation, eliminating the cap on payroll taxes, negative income tax or basic income, refundable care-giver tax credit, affordable housing, universal health care, income deductible rent, and so on. Many recommendations addressing inequality are based on a Robin Hood approach, transferring wealth from the rich to the poor. Debates usually deteriorate into unproductive and divisive arguments between supporters of downward redistribution and opponents. Advocates seek to remedy the injustice of the current inequality of wealth, while opponents consider it unfair to confiscate and redistribute what they consider private wealth, and consider it a deterrent to initiative and productivity.

4.1 Addressing inequality

A better approach to inequality would be to make a distinction between those assets that are created by private effort and those which are inherently common property, namely those assets produced neither by individual labor nor by capital. Every person shares a common inheritance of natural and cultural assets. These assets contribute actual or potential dividends to everyone on Earth. Justice requires that contributions to welfare from natural capital and the shared cultural heritage of humankind should be equitably distributed among all, while the value added to these assets by individual or collective effort (labor or capital)

[2] See Cobb (chapter 10 in this volume) for a more detailed discussion of economic rent.

should belong to the individuals who contributed that value. Daly has repeatedly argued that taxing low entropy matter and energy inputs to the economy is preferable to taxing value added by labor and capital (Daly, 2007). Actual and potential dividends can be computed on a per capita basis for ecosystem services, human and social services, and rent on natural and social assets. For the purposes of redressing inequality I will focus on rent from natural common assets.

4.2 Equity from common assets

Many resources cannot be assigned individual property rights and are inherently public, such as climate stability and the ozone layer. It is however possible for individuals to destroy them. A number of other resources created by nature can be privately owned, but were traditionally shared by communities, often with explicit rules or norms that limited the amount of resource that any individual could use. Examples include grazing rights, fishing rights, and the right to collect firewood from public lands. The enclosure movement, beginning in England in the 17th century, converted many of these common property rights into private rights (Bollier, 2002). Mainstream economists frequently claim that common property resources are poorly managed, resulting in overuse and a "tragedy of the commons", and therefore should be privatized, though abundant evidence shows that common property institutions can sustainably manage such resources for millennia. Lack of property rights is the problem, not common property rights (Bromley, 1992; Ostrom, 2002). New technologies frequently create whole new types of resources, such as the electromagnetic broadcast spectrum, aircraft landing rights, orbital satellite slots, the Internet and other products of government research, and genetic information. New institutions can also establish ownership rights to previously unowned resources, such as cap and trade permits for emissions of pollutants such as sulfur dioxide, NO_2, or carbon dioxide, CO_2. Both new resources and newly created property rights raise the important question of to whom these resources should belong. Following the advice of mainstream economists, most of these resources have been turned over to society's wealthiest individuals and corporations, even though when public enforcement of private property rights is very expensive. A more just alternative would be to retain rental values on these assets, as well as charges for harm done to them, as citizen equity.

4.3 Equity from natural capital

UN Resolution 1803 (XVII) of 14 December, 1962/ Declaration of Permanent Sovereignty over Natural Resources states:

> "Violation of the rights of peoples and nations to sovereignty over their natural wealth and resources is contrary to the spirit and principles of the Charter of the UN, and hinders the development of international co-operation and the maintenance of peace."

Another view:

> "The meek shall inherit the earth... Except for the mineral rights" J. Paul Getty.

An existing model of citizen equity in natural capital, combined with weak sustainability[3] currently exists in the US state of Alaska. Oil resources in Alaska belong to the people of the state. The severance tax rate on oil is 12.25%-15% of extraction value depending on the age of the oil field, and 10% on natural gas. Royalties paid by oil companies drilling in Alaska are partly used for state revenue, but a large portion is placed in a permanent fund (APF), which is invested for the benefit of the citizens of Alaska.

Without depleting the capital fund, interest is paid as an annual dividend to every resident of Alaska who has lived in the state for more than one year. Payments have averaged over $1000 per year in recent years (Figure 14). This illustrates the transformation of natural capital into a sustainable stream of financial capital. As state oil resources are used up, the citizens of Alaska will still have a large and growing capital fund earning interest for them (Figure 15). This will continue indefinitely as long as the fund is managed well, and the state government is prevented from spending the fund.

The APF is one of the few cases in the world where property rights to natural capital have in essence been equally distributed among all citizens. Governments usually retain rights to these resources, and revenues are used for general government expenditures, or quite often end up in the bank accounts of government officials, especially in authoritarian regimes. Under the Bush administration oil companies in the Gulf of Mexico were given a waiver on paying royalties to the government (Andrews 2006). These tax breaks were upheld in Congress as of April 4, 2012. By establishing citizen equity to common assets and natural capital, these benefits will accrue to the population at large, rather than to government officials and their associates, or to corporate owners. What does the Alaska Oil dividend have to do with equity? It turns out that Alaska has the lowest Gini ratio of any state in the US at .40217 (2000 census figure).

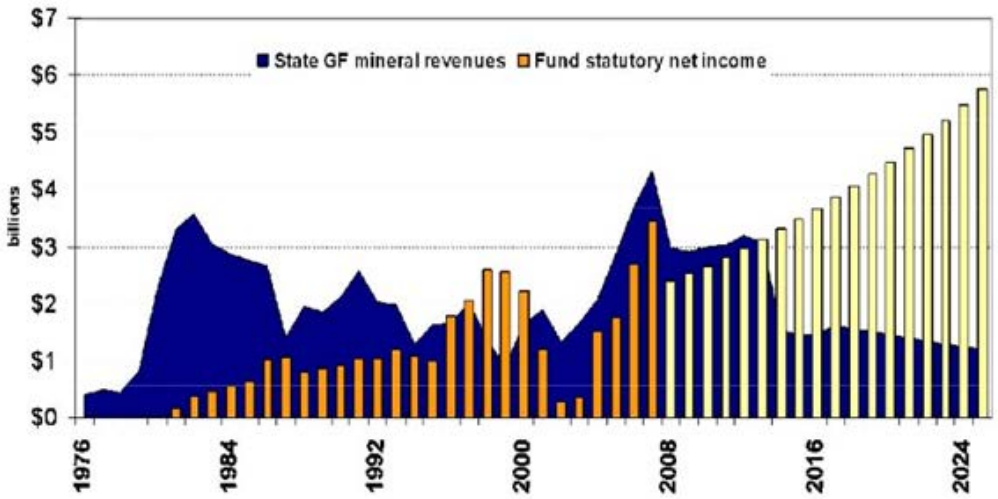
Figure 14 Permanent fund dividend 1982-2014

		1990	$952.63	2000	$1,963.86	2010	$1,281.00
		1991	$931.34	2001	$1,850.28	2011	$1,174.00
1982	$1,000.00	1992	$915.84	2002	$1,540.76	2012	$878.00
1983	$386.15	1993	$949.46	2003	$1,107.56	2013	$900.00
1984	$331.29	1994	$983.90	2004	$919.84	2014	$1884.00
1985	$404.00	1995	$990.30	2005	$845.76		
1986	$556.26	1996	$1,130.68	2006	$1,106.96		
1987	$708.19	1997	$1,296.54	2007	$1,654.00		
1988	$826.93	1998	$1,540.88	2008	$2,069.00		
1989	$873.16	1999	$1,769.84	2009	$1,305.00		

Source: http://www.apfc.org/home/Content/dividend/dividend.cfm

[3] Weak sustainability is the belief that natural and human made capital and completely substitutable, and sustainability requires only non-diminishing quantities of total capital. In contrast, strong sustainability requires non- diminishing amounts of natural capital.

Figure 15 The relationship of fund income to state oil revenues: past and future. Alaska Permanent Fund Corporation "Fund Works", 2007



Source: http://www.housemajority.org/coms/hfsp/pdfs/CWN_PF_Study.pdf

5. Conclusion

Pareto Optimality is consistent with Pareto's other writings that are explicitly social Darwinist, advocating "survival of the fittest" in the human economy. Use of this principle in neo-classical economics and calling it "value-free" is inconsistent with reality if not deceitful. It contributes to inequality in society and the constant call for economic growth. Ecological economists such as Herman Daly are correct in questioning Pareto Optimality as a measure of efficiency. "Efficiency" could just as legitimately be measured by reductions in the Gini coefficient, which with all else equal would be likely to increase the welfare produced per unit of output.

Redressing the vast inequities in wealth created by the triumph of Pareto does not require a transfer of "hard earned" wealth from the rich to the poor. A better way would be to distribute common assets, which do not involve confiscating anyone's rightfully earned property.

Distributing dividends universally from natural and social assets such as minerals, land, and the electromagnetic spectrum might be relatively trivial to those at the top of the income scale. To those at the bottom who receive much greater marginal utility from small increases in income, these dividends could make a significant difference. The reduction in inequality that results, as demonstrated by Alaska, could reduce the perceived need for economic growth. It might also achieve the common good for the many as advocated by Daly, rather than the private good for a few as advocated by neo-classical economists following Pareto.

References

AFL-CIO (2013), "Executive paywatch: CEO-to-worker pay gap in the United States".

Alvaredo, F., A. B. Atkinson, T. Piketty and E. Saez (2013), "The top 1 percent in international and historical perspective", *Journal of Economic Perspectives*, **27** (3), 3–20.

Andrews, E. (2006), "U.S. has royalty plan to give windfall to oil companies", *New York Times*, New York.

Ariely, D. (2008), "*Predictably Irrationl"*, New York: Harper Collins.

Atkinson, T., T. Piketty and E. Saez (2011), "Top incomes in the long run of history", *Journal of Economic Literature*, **49** (1), 3–71.

Bollier, D. (2002), *Silent Theft: The Private Plunder of Our Common Wealth*, New York: Routledge.

Bromley, D. W. (ed.) (1992), *Making the Commons Work: Theory, Practice, and Policy*, San Francisco: ICS Press.

Collins, Y. (2000), *Economic Apartheid in America*, New York: New Press.

Daly, H. E. (1992), "Allocation, distribution, and scale: Towards an economics that is efficient, just, and sustainable", *Ecological Economics*, **6** (3), 185–193.

Daly, H. E. (2007), *Ecological Economics and Sustainable Development: Selected Essays of Herman Daly*, Northampton, MA: Edward Elgar.

Daly, H. E. and J. Farley (2010), *Ecological Economics: Principles and Applications*, Washington, DC: Island Press.

Daly, H. E. and R. Goodland (1994), "An ecological-economic assessment of deregulation of international commerce under GATT", *Ecological Economics*, **9** (1), 73–92.

Economic Policy Institute, "*The Top 10 Charts of 2014"*, http://www.epi.org/publication/the_top- 10-charts-of-2014/ (retrieved 12/21/14)

Ewing, B., D. Moore, S. Goldfinger, A. Oursler, A. Reed and M. Wackernagel (2010), *Ecological Footprint Atlas 2010*, Global Footprint Network.

Fukuyama, F. (1989), "The end of history?", *The National Interest* (Summer).

Gaffney, M. and F. Harrison (1994), *The Corruption of Economics*, London: Shepheard-Walwyn in association with Centre for Incentive Taxation.

Gintis, H. (2000), "Beyond *Homo economicus*: Evidence from experimental economics", *Ecological Economics*, **35** (3), 311–322.

Gowdy, J. and J. Erickson (2005), "The approach of ecological economics", *Cambridge Journal of Economics*, **29** (2), 207–222.

Hackett, S. C. (2001), *Environmental and Natural Resource Economics*, NY: ME Sharpe. Heintz, J. and N. Folbre (2000), *The Ultimate Field Guide to the US. Economy*, New York: New Press.

Kahneman, D. and A. Tversky (2000), *Choices, Values, and Frames*, New York: Cambridge University Press, Russell Sage Foundation.

Khazan, O. (2013), "U.S. corporate executives aren't the only ones making tons of money", *The Atlantic* (March 19).

Misheland, L. and N. Sabadish (2012), "CEO pay and the top 1%", *Issue Brief #331*, Washington, DC: Economic Policy Institute.

Ostrom, E. (ed.) (2002), *The Drama of the Commons*, Washington DC: National Academy Press. Piketty, T. and E. Saez (2006), "The evolution of top incomes: A historical and international perspective", *American Economic Review*, **96** (2), 200–205.

Saez, E. (2013), "Striking it richer: The evolution of top incomes in the United States (Updated with 2011 estimates)", *University of California-Berkeley Working Paper*, University of California, Berkeley, available at http://emlab.berkeley.edu/~saez/saez-UStopincomes-2011.pdf (accessed 27 July 2013).

Thaler, R. H. and C. R. Sunstein (2008), *Nudge: Improving Decisions About Health, Wealth, and Happiness*, New Haven: Yale University Press.

Tversky, A. and D. Kahneman (1981), "The framing of decisions and the psychology of choice", *Science*, **211** (4481), 453–458.

US Census Bureau (2011), *Current Population Survey: 2011*, US Government. US Census Bureau (2012), *Historical Poverty Tables*, US Government.

Wilkinson, R. and K. Pickett (2009), *The Spirit Level: Why Greater Equality Makes Societies Stronger*, New York: Bloomsbury Press.

Wolff, E. N. (2001), *Top Heavy: The Increasing Inequality of Wealth in America and What Can Be Done About It*, New York: The New Press.

Wolff, E. N. (2010), "Recent trends in household wealth in the United States: Rising debt and the middle-class squeeze—An update to 2007", *Working Paper No. 589*, Levy Economics Institute of Bard College.

Wolff, E. N. (2012), "The asset price meltdown and the wealth of the middle class", *NBER Working Paper Series No. 18559*, National Bureau of Economic Research.

Author contact: Gary.Flo@uvm.edu

SUGGESTED CITATION:
Gary Flomenhoft, "The triumph of Pareto", *real-world economics review*, issue no. 80, 26 June 2017, pp. 14-31,
http://www.paecon.net/PAEReview/issue80/Flomenhoft80.pdf

Do we need a new economics for sustainable development?

Peter Söderbaum [Mälardalen University, Sweden]

When asked about the "economic" aspect of an activity or project people tend to think about "money". Focus on this monetary or financial aspect is broadly in line with neoclassical economic theory which is the conceptual framework and theory taught at university department of economics in different parts of the world. Economics is defined as allocation of, or "management of scarce resources" (Mankiw, 2011, p. 2) where optimal solutions are normally explored in monetary terms. As part of this neoclassical theory two actor categories are considered, i.e. consumers (or households) and firms. Consumers and firms interact in three kinds of markets: for commodities, for labor and for financial capital. It is assumed that consumers maximize utility within the scope of their monetary budget while firms maximize monetary profits. At the macro or national level Gross Domestic Product, a specific summary of all market transactions, is the main progress indicator. When preparing investments in roads, airports, energy systems and other infrastructure projects Cost-Benefit Analysis (CBA) in monetary terms is the approach advocated.

Neoclassical theory with its assumptions has certainly had an impact upon the cognitive habits of individuals as consumers, of business leaders, politicians, and governmental administrators. In some sense consumers may feel that they are successful in maximizing utility. Business leaders may be successful in terms of monetary profits and so on. At issue, however, is if such thinking patterns are appropriate in the present situation. Sustainable development or sustainability is one of the present challenges. It is true that neoclassical theory has something to offer even in relation to sustainable development. It is the limitation to one theoretical framework that is questioned here. Rather than the present neoclassical monopoly we need competition. Rather than thinking in terms of "paradigm-shift" (Kuhn, 1970) we need to think in terms of "paradigm coexistence" (Söderbaum, 2000, pp. 29-30) and pluralism.

Neoclassical theory at the micro and macro levels has influenced thinking patterns in a period when development has become unsustainable in many ways. Climate change is one example, loss of biological diversity and pollution of air, water and soil are other examples. While not excluding ideas coming from neoclassical theory it is a bit strange to argue that we in the present situation should exclusively listen to advocates of a theory that has failed in relation to sustainability. We also need to "test" other theoretical perspectives in economics. Can they improve our chances of dealing with sustainability and other issues?

Redefining economics in terms of multidimensional analysis and democracy

A proposed new theoretical perspective starts with a partly different definition of economics:

> "Economics is multidimensional management of (limited) resources in a democratic society"

Why "multidimensional" management? Multidimensional goes against the one-dimensional analysis of neoclassical theory and method. "Monetary reductionism" is no longer accepted. The idea that we should put a monetary price on all impacts, ecosystem services included, to make them commensurable and tradeable, is abandoned. Instead impacts of different kinds are kept separate throughout analysis. And non-monetary impacts are viewed as being as "economic" as monetary ones. This may make analysis more complex but also more relevant. In relation to environmental issues, Sweden has proposed 16 environmental objectives to monitor the quality and state of the environment as well as possible (Swedish Environmental Agency, 2016). At the UN level, world leaders agreed about 8 Millennium Development Goals in the year 2000. This process of broadening the agenda has continued with 17 Sustainable Development Goals (SDGs) sanctioned at the UN-level in November 2015. They are:

- No poverty
- Zero hunger
- Good health and well-being
- Quality education
- Gender equality
- Clean water and sanitation
- Affordable and clean energy
- Decent work and economic growth
- Industry, innovation and infrastructure
- Reduced inequalities
- Sustainable cities and communities
- Responsible consumption and production
- Climate action
- Life below water
- Life on land
- Peace, justice and strong institutions
- Partnerships for the goals

Each set of goals is further elaborated including a number of sub-goals and targets to be achieved in the next 15 years. "For the goals to be reached, everyone needs to do their part: government, the private sector, civil society and people like you", it is argued (United Nations, Sustainable Development Knowledge Platform, 2016).

It is clear that the above list of objectives points to a disaggregated and multidimensional view of sustainable development. One dimensional aggregation in monetary or other terms does not appear meaningful. It can furthermore be observed that economic growth in traditional terms is part of the list but only one of the 17 SDGs. Performance in relation to the goals is finally made an issue for all individuals in all potentially relevant roles. This suggests that we need to abandon the neoclassical mechanistic idea of individuals and organizations in favor of a view where actors are responsible and accountable. To this we will return.

Why reference to a democratic society? When reading neoclassical introductory textbooks in economics it becomes clear that "democracy" is not a theme taken seriously. These texts rather reflect an emphasis on economists as experts, i.e. a kind of technocracy. In the

mentioned textbook by Mankiw and Taylor "democracy" is not even part of the index or glossary. As I see it, democracy should play a crucial role at two levels:

- Regarding the discipline of economics itself. We should not accept neoclassical monopoly and dictatorship
- Regarding our understanding of individuals as actors in the economy. We should respect that individuals differ with respect to value or ideological orientation.

Economics as a discipline is in need of being democratized (Söderbaum and Brown, 2010). As scholars we should admit that a specific paradigm (theoretical perspective) in economics is not only theory; it is also specific in value or ideological terms. Gunnar Myrdal, a well-known institutional economist, argues that "values are always with us" (Myrdal, 1973; 1978) in the different stages of the research process from formulation of problems to presentation of results. Neoclassical theory is specific in ideological terms with its focus on markets and the monetary dimension. Applying Economic Man assumptions and looking upon organizations in profit-maximizing terms are other examples. The reasons for some of us economists to look for other theoretical perspectives are not only scientific (in some narrow sense) but also ideological. Those of us who claim to take environmental issues seriously and advocate institutional environmental (or ecological) economics as alternative perspective do so for scientific as well as ideological reasons.

If there are options at the level of ideology then democracy becomes important also when understanding individuals in the economy and respecting differences in their ideological orientations. Neoclassical economists with their CBA claim expertise in "correct values" in market terms for purposes of assessing impacts of investments in infrastructure. They claim ability to identify the best or optimal solution for society as a whole among competing investment projects. Applying Positional Analysis (PA) as an alternative method means that we in many situations have to refer to conditional conclusions rather than one single optimal solution. Alternative investment projects will be ranked differently depending upon the ideological orientation considered. And in many decision situations it is not realistic to assume that all stakeholders and concerned actors share the same idea about what is progress in society. Such ideas may differ between political parties in a democratic society for example.

Overview of differences between neoclassical and institutional economics

What are considered as essential elements of neoclassical economic theory are indicated in the left-hand column of Table 1. I have chosen to point to the view of individuals, of organizations and of markets, as well as to ideas about institutional change, progress in society and, finally, an approach to sustainability assessment. In the right hand column I have identified corresponding views and concepts of institutional theory as parts of an alternative perspective. Political Economic Person assumptions are presented as an alternative to Economic Man assumptions, and Political Economic Organization as alternative to assumptions of profit-maximizing firms, etc.

Table 1 Differences between neoclassical and institutional economics

	Neoclassical economics	**Institutional economics**
Individual	Economic Man	Political Economic Person (PEP)
Organization	Profit-maximizing firm	Political Economic Organization (PEO)
Market	Mechanistic ideas of supply and demand	PEPs and PEOs as market actors, relationships and networks
Institutional change	Change in regulation of market incentives	Also actor-network approach
Progress indicators	Gross Domestic Product (GDP) (side concern: employment)	Open issue where Sustainable Development Goals (SDGs) are part
Approach to sustainability assessment	Cost-Benefit Analysis (CBA)	Positional Analysis (PA)

The role of values and ideology in economics suggests that economics is always "political economics". The individual as political economic person (PEP) is "an actor guided by her ideological orientation". Ideological orientation is something that differs between individuals; some bother more than others about threats of climate change for example. Such differences are of relevance in economic analysis.

A political economic organization (PEO) is similarly an actor guided by its 'ideological orientation' or 'mission'. Business corporations or firms is only one category of organizations and the role of monetary profits or other financial considerations is partly an open issue. The debate about Corporate Social Responsibility (CSR) is an example of this. Differences between business leaders and other actors in organizations in relation to sustainable development are worth studying.

Rather than the mechanistic view of markets in neoclassical theory, a focus will be on the behavior of individuals as market actors. The actors are understood in terms of social psychology and reference is made to an actor's ideological orientation which is not limited to self-interest. Each market actor is related to other market actors in supply chains and networks.

When institutional change is concerned, the tendency of neoclassical economists is to point to incentive systems in monetary market terms. Taxes and environmental charges are examples. As an institutional economist I am a bit sceptical about such proposals because inherent in the neoclassical ideology is a fundamental belief that that interventions in existing markets are generally dysfunctional. As an institutionalist I see changes in institutional arrangements more as a result of power games with the different actors involved. State intervention is an option where certainly taxes and monetary incentives can play important roles but where also voluntary action is possible.

GDP as an indicator of progress has already been touched upon. This indicator suffers from the mentioned "monetary reductionism" and the recent emphasis at the international level on the 17 SDGs is here regarded as an important step forward. When approach to sustainability assessment is concerned neoclassical economists recommend CBA, i.e. analysis where all

kinds of impacts are reduced to their alleged monetary equivalents. An ambitious study of ecosystem services and biodiversity where United Nations Environmental Program (UNEP) played an important role, the so called TEEB-study, is an example of this (Kumar, ed., 2010). I have elsewhere discussed this study (2013, 2015). There are alternatives to CBA and my preference is Positional Analysis (PA), an approach based on the multidimensional and disaggregated idea of economics previously discussed.

The concepts of democracy and ideological orientation

In a democratic society there are normally different political parties, each with its ideology or ideological orientation. Actors belonging to specific political parties turn to citizens (or actors in other roles) in attempts to convince them with their ideas. And citizens respond in some way suggesting that they too are guided by a kind of "ideological orientation".

Transforming societies from present unsustainable tendencies in a Green or more sustainable direction is clearly not exclusively a matter of science and management but also of ideology. And just as there are varieties of ideological orientations that downplay climate change and other sustainability concerns, there are also competing views about how to best transform society in a "Green" direction.

Green issues compete with other issues in the brains and thinking patterns of individuals as actors and there are capacity limits for dialogue in different arenas, for example in specific newspapers. Climate change appears as an urgent issue for many of us but does it get the priority it deserves when compared with other issues? These days public debate is focused on war in Syria and to my knowledge there are not many attempts to systematically relate these destructive activities to climate change or the 17 SDGs previously listed.

Democracy stands for a political system where it is possible to raise various concerns or issues to make them part of the agenda of actors in various arenas. Democracy is not only a system of voting rules in parliamentary elections. It is also about human rights such as freedom of speech, freedom of association and generally a willingness to listen to actors with an ideological orientation that differs from your own (as long as these voices do not go against democracy itself). Actors who differ in views can learn from each other, suggesting that democracy can contribute to creative outcomes of public dialogue. In this sense democracy can be part of the security system of a society. Moving from democracy in the direction of dictatorship would then reduce security.

Democracy versus dictatorship is not a simple matter of either-or. Democratic elements can be identified in a society described as dictatorship and democracy can always be strengthened in countries such as Canada or Sweden. In any country and globally there are institutions protecting mainstream interests. The monopoly of neoclassical economic theory at university Departments of Economics is a case in point. Even institutions such as the Bank of Sweden Prize in Economic Sciences in Memory of Alfred Nobel play, as I see it, a role in making a transformation to sustainability more difficult by systematically protecting the neoclassical paradigm. This neoclassical perspective contributes to make the present political system legitimate in a situation where institutional change is needed.

There are different theories of democracy. Some are built on the assumption that democracy aims at consensus and that consensus and agreement can be obtained through dialogue.

There is certainly a need for many kinds of agreement at local, national, regional and global levels but with Chantal Mouffe (2005; 2013) I tend to believe that democracy is also and perhaps primarily about how to live with disagreement. Aiming always at consensus involves a risk of neglecting or even eliminating those with a different opinion. This is what happens for example in the international dialogue (or lack of it) about economics where the mentioned Economics Prize plays a role. Also when looking at some nations today, Turkey being a possible example, the attempts to eliminate opposition can hardly be regarded as compatible with democracy.

"Ideology" is a "contested concept" (Connolly, 1974) in the sense that it can be understood in more ways than one. "Power", "institution", "democracy", "value" are other contested concepts relevant in social science. When using concepts of this kind, one has to clarify their meaning as well as possible. Ideology and ideological orientation are here understood in means-ends terms. It is about where you are (original position), where you want to go (future positions) and how to get there (strategy). The ideological orientation of an actor varies more or less over time with situation and context. While neoclassical economists refer to mathematical objective functions and optimal solutions in precise terms, an ideological orientation is often fragmentary and uncertain. An ideological orientation can furthermore be narrow or broad. It can reflect self-interest but also ideas about common interests, concerns for future generations, etc. Ideological orientation in this sense points in the same direction as Herman Daly's reference to "the common good" (Daly and Cobb, 1989). Reference to "the" common good may suggest that there is one single idea of common interests. In a democracy there are – as we have argued – competing varieties of ideological orientation and thereby competing ideas of the meaning of "common good". Whether expressed in quantitative, qualitative, visual terms or some combination thereof, an actor's ideological orientation can guide behavior and investment decisions.

In what way can ideological orientation guide an actor's behavior and decision making? This can be understood as a 'matching' process where the multidimensional and multi-facetted ideological orientation of an actor is matched against the multidimensional impact profile of each alternative of choice considered. Our actor and decision-maker looks for an alternative that with respect to expected impacts matches her ideological orientation as well as possible. Reference can also be made to "compatibility", "appropriateness" or "pattern recognition". E.F. Schumacher pointed to the need for "appropriate technology" at an early stage in his book "Small is Beautiful" (1973) and pattern recognition is a term closer to computer language. Ideological orientation as well as expected impacts connected with specific alternatives can be understood in terms of patterns.

Reconsidering "costs" and "benefits"

As already indicated costs and benefits in neoclassical analysis tend to be interpreted in monetary or financial terms. A case in point is neoclassical Cost-Benefit Analysis (CBA). The idea is that only that counts which can be quantified and a way of achieving this is to put price tags on all kinds of impacts. A problem with this is that actors may assume that an impact has been fully considered when a monetary price has been chosen. Non-monetary impacts somehow disappear from analysis to become invisible.

A way of dealing with this is to keep monetary and non-monetary impacts separate and use the terms of "cost" and "benefit" both on the monetary and non-monetary side (Table 2). Non-

monetary costs and benefits ("b" and "d" in Table 2) are regarded as important as such and not reducible to monetary impacts. The importance of specific impacts in a decision situation becomes an issue of the knowledge available to an actor, her specific situation and ideological orientation.

Table 2 A classification of costs and benefits in economic analysis

	Monetary	**Non-monetary**
Cost	"a"	"b"
Benefit	"c"	"d"

In the case of constructing a motorway where there is agricultural land and forest with connected ecosystems, the CBA-analyst tends to focus on construction costs in financial terms. Buying the land needed is one part of the financial outlays and paying the construction company another part. But at issue is whether this covers the whole "cost". Agricultural land and forests are turned into asphalt surface suggesting that there are more or less irreversible losses in ecosystem function. Irreversible losses are hardly measurable in monetary terms. And where does the asphalt come from?

One way of understanding what happens in non-monetary terms is to think in terms of how options for future decisions are affected. A new road will increase opportunities for those using the road for transportation purposes but the transformation from agricultural land to hard cover or asphalt surface is a largely irreversible process. Some farms may be closed down and the farmers look upon monetary and non-monetary impacts in their own way. The important thing for the decision-maker is to know what she is doing in multidimensional terms and for whom.

When dealing with inertia (in its different forms) and future options, it becomes relevant to distinguish between impacts in terms of "flows" and "positions" (Table 3). On the monetary side, the turnover of a company exemplifies a monetary flow ("e" in Table 3) while financial assets and debts at a point in time are examples of monetary positions. When non-monetary impacts are concerned, CO_2 emissions from transportation activities per year exemplify a flow whereas the size of land used for transportation purposes in a city at a point in time exemplifies a non-monetary position ("h" in Table 3).

Table 3 A classification of impacts in terms of flows and positions

	Flow (referring to a period of time)	**Position (referring to a point in time)**
Monetary	"e"	"f"
Non-monetary	"g"	"h"

Positional Analysis is an approach to sustainability assessment that considers monetary as well as non-monetary impact but with an emphasis on non-monetary flows and positions. It is argued that impact studies, monitoring and accounting systems in relation to sustainability should focus much more (than in CBA) on non-monetary impacts and non-monetary positions (or states) in particular.

Positional Analysis (PA) is described elsewhere (Söderbaum, 2000; Brown et al., 2017). The purpose is to illuminate an issue in a many-sided way with respect to relevant ideological orientations, alternatives of choice and impacts. An attempt is furthermore made to identify conflicts of interest and make them visible (rather than assume that all stakeholders and concerned actors agree about one single objective function as in CBA). Conclusions in terms of ranking the alternatives considered are conditional in relation to each ideological orientation considered. PA then claims to be an approach to decision-making that is compatible with essential features of democracy. At an early stage the analyst also enters into a problem-solving dialogue with stakeholders, others concerned and decision-makers.

Individuals and organizations as political actors

The neoclassical approach to environmental and sustainability problems can be described as follows. It is believed that markets on their own automatically can handle many kinds of problems in the economy and that they generally stand for efficiency. A first recommendation is therefore to refrain from "unnecessary government intervention". But neoclassical economists understand that there are what they see as exceptions where markets can fail and produce environmental impacts of a negative kind. They then refer to externalities, i.e. single impacts on third parties (being outside the market transaction). The third party suffers but can be compensated for in monetary terms according to the 'polluter-pays principle'. There is also an idea of how the correct compensation can be estimated. Neoclassical economists (with their neoliberal agenda) are eager to point out that also governments may also fail by subsidizing activities in the economy that are harmful to the environment. Such subsidies should be removed, it is argued.

Institutional economists such as William Kapp (1970, Berger and Steppacher, eds 2011) do not see environmental impacts as single disturbances in a system that otherwise is functioning well. Environmental impacts connected with production, trade and transportation are rather ubiquitous and there is a systematic tendency for business corporations to eliminate monetary costs by carrying over negative environmental impacts on society at large (just as many corporations do what they can to avoid paying taxes that would, when paid, make it easier for local and national governments to cover outlays when dealing with environmental and other problems). Or, in the words of William Kapp himself:

> "Thus, a system of decision-making, operating in accordance with the principle of investment for profit, cannot be expected to proceed in any other way but to try to reduce its costs whenever possible by shifting them to the shoulders of others or to society at large" (Kapp, 1970, p. 18)

According to a neoclassical perspective, policy and politics is in the hands of national and local governments. My view of policy and politics as an institutionalist is much broader. Each individual and each organization is a potential and actual policy-maker, as suggested by the concepts "political economic person" and "political economic organization". All actors contribute to aggravate or improve the state of the environment (and other sustainability indicators in positional terms). Each actor has her rights and responsibilities in a democratic society.

Individuals as actors relate to each other in networks. Based on their ideological orientations or missions they interact – cooperate or compete – with other actors. A problem is that our

present political economic system with all its institutions is not easily changed. Inertia and path dependence makes it difficult to transform the economy and society from its present unsustainable path to one which is more sustainable. Some powerful individuals and organizations as actors, for example transnational corporations furthermore protect present institutional arrangements (and may work for institutional change in an even more unsustainable direction).

In this situation where a transformation is needed but where there are many kinds of inertia, all actors are important but some actors are perhaps more important than others. Eva Kras (2007), former president of the Canadian Society for Ecological Economics, has repeatedly argued that we need to listen to "visionaries", i.e. actors who tend to face fundamental scientific and ideological issues rather than avoid them. The visionaries play an important role but so do also all those who are open-minded enough to listen to voices outside the mainstream. There are visionaries within science but also among other actors in the broad public dialogue about progress in society.

While many individuals qualify as visionaries I will here only single out two early and two more recent visionaries in the debate about the future of our societies. E.F. Schumacher has already been mentioned. In his book he reminds us that *Small is Beautiful* and that there is a need for an economics "as if people mattered". Smallholders may be successful in many ways and dimensions, and neoclassical ideas about "economies of scale" are too simplistic. A second early writer is Hazel Henderson whose collective writings published in *Creating Alternative Futures. The End of Economics* (1980) in my judgment is as relevant today. I understand however her subtitle "The End of Economics" as directed to the neoclassical mainstream rather than all kinds of economics. Henderson is worried about "creeping gigantism" (p. 163) and the power of transnational corporations when compared with smallholders and other small companies. In her view "the reliance on global price mechanisms blinded economists" to mounting environmental and ecological problems (p.3). She also refers to the "bankruptcy of economics" (p. 4). In more positive terms she points to a large number of examples of an "emerging counter-economy" (pp. 381-389), from UN conferences dealing with different issues to local initiatives.

More recent writings by visionaries can be exemplified by Naomi Klein's book *This Changes Everything. Capitalism vs. the Climate* (2014) and George Monbiot's *How did we get into this mess? Politics, Equality, Nature* (2016). Naomi Klein points to the abuse of power of oil companies in relation to other interests for example those of indigenous people and Monbiot, columnist in *The Guardian*, is attempting to bring issues into the political agenda where establishment actors tend to be silent.

This reference to some visionaries is of course not meant to do justice to their writings (and even less to the writings of all other visionaries) but should mainly serve to remind us that there is a danger in being limited to mainstream arguments and mainstream literature.

The case of "free trade" agreements: Ceta and TTIP

Individuals as authors of books and in other roles can act as precursors. But so also can business companies, local governments, cities, universities, civil society organizations and even nations and groups of nation states, such as the European Union. The opposite is also possible: these entities can play a role in delaying necessary change. Let us look at proposed

free trade agreements, such as Ceta (Comprehensive Economic and Trade Agreement) between Canada and the European Union (EU) and TTIP (Transatlantic Trade and Investment Partnership) between USA and EU. A long period of negotiations is now followed by a period when contracts are supposed to be signed.

In relation to international trade there is a mainstream in economics telling us that free trade is good and attempts to protect home industry from competition are bad. "Protectionism" through tariffs and quotas is bad and both nations benefit if such obstacles are reduced or eliminated. These recommendations are built upon a simplistic neoclassical economic theory which has little to do with the complexities of the real world.

Neoclassical theory about international trade is at the same time ideology and this ideology has been internalized through economics education and in other ways by a large number of actors in society. The theory more precisely plays a role in making neoliberalism legitimate. It is not possible here to discuss all the assumptions behind the theory of international trade. There is an assumption however about homogenous commodities, i.e. that the same goods or services can be produced in both of the trading countries which is seldom the case in reality. There is also an assumption that the interests of each country can be dealt with in one-dimensional terms. In reality there are normally conflicts of interest in each of the trading countries, for example between labor interests and shareholder interests or between consumer interests and producer interests. Reducing conflicts of interest to one interest can only be done by choosing ideological position. Environmental impacts and other externalities are not part of, nor play a peripheral role in, international trade theory. International trade in goods necessarily involves transportation which leads to CO_2 emissions. How can such impacts be assumed away today when we claim to understand what we are doing to the climate?

A more realistic perspective on international trade in goods and services would suggest that "protectionism" often plays a positive role. In each country we have an Environmental Protection Agency with the objective to protect ecosystems and natural resources. Today, reference is increasingly made to a need to "protect the planet" and why should those who think in such terms apologize? Similarly, there are national and international institutions with the purpose to protect human health. Transnational corporations prefer one system of health standards or environmental standards globally (which is stable over time) while some actors in a nation may prefer to control its own standards and make changes in standards when needed.

Included in the present negotiations is a kind of conflict resolution mechanism, ISDS (Investor-State Dispute Settlement), which attributes a right to corporations to sue nations if their monetary profits are reduced by changes in national policy, for example environmental policy or health policy. This is not so much about so called free trade but about protectionism of business interests. From a democracy point of view this is a grotesque proposal as everybody can understand. Should similarly civil society organizations (CSOs) get a right to sue transnational corporations in a Court of Justice (built on principles at their own choice), if they feel that their interests are threatened?

This is where we return to the possible role of nations, or part of nations, as leaders in sustainability policy making. There is a neoliberal mainstream of actors on both sides of the Atlantic but also numerous actors with a critical attitude. Among regions within EU, Wallonia, an independent part of Belgium in some respects, protested and wanted to use their veto but

finally were convinced by Cecilia Malmström, Commissioner for International Trade Agreements in EU, and others to accept. Malmström is a Liberal politician and it is clear that the ideological orientation of the EU mainstream is neoliberal. The way this trade issue has been handled is an indicator of the strength of the mainstream and the ignorance by mainstream actors of other ways of thinking. Things are not finally settled yet and my point here is that Wallonia with its way of arguing has become a bit of a "hero region" for me. Hopefully, other regions and entities will follow the example.

International trade theory does not deal with power issues and the ideological orientation of various actors. The role of transnational corporations in the global economy is hardly mentioned, if at all, in mainstream economics textbooks. Free trade in the sense of tariffs and quotas already exists for many commodities suggesting that cheap products in monetary terms are already available. Some of us may argue that some products are too cheap when considering the ways cheap labor is exploited and other ways of lowering monetary costs. Further moves toward free trade in the mentioned sense are mainly a financial interest of big transnational corporations. They want to dictate the rules of the game. But again joint stock companies are defined in reductionist, monetary, financial terms while present challenges are of a non-monetary kind, for example the 17 UN sanctioned SDGs previously discussed. Does this mean that joint stock companies are miss-constructed in relation to present needs? I tend to think so. We need a debate about different forms of organization and their possible roles in society.

Conclusions for sustainability politics

When asked about present unsustainable trends and how they can be transformed in a more sustainable direction, many persons would probably point to what is perceived as concrete issues, such as climate change and then list various ways of reducing CO_2 emissions. In the present essay I have argued that such efforts are certainly desirable but that we also should focus on perspectives, such as the close to monopoly position in economics of neoclassical theory and the dominance of neoliberalism in politics. In fact I regard these issues of perspectives as fundamental in dealing with present problems. It is an illusion that there can be a value-free or value-neutral economics (von Egan-Krieger, 2014) and therefore we need to study and acknowledge the values or ideological orientation of each kind of economics. Just as neoclassical theory is close to neoliberalism, the kind of theory advocated in this essay claims to be closer to a Green or sustainability ideology where for example the 17 SDGs play an essential role.

Rather than exclusively connecting sustainability policy and politics with government action and intervention in the economy I have suggested that each individual in her different roles is a policy maker. The individual as actor is invited to participate in a dialogue about the future by listening to various and often contradictory voices. It is no longer enough to listen to those who claim expertise and act accordingly since experts may – as in the case of economics – be 'disciplinary monopolists' whose primary interest appears to be to protect the monopoly. This is a case of negative protectionism while there are also positive examples of protectionism as we have argued. It should be added that it is not even enough to listen to various scientists. Sometimes it is important to listen to visionaries and to encourage scientists to do the same. And there are many other influential actors to bring into this dialogue, such as journalists, business leaders, actors from Civil Society Organizations and politicians.

An alternative to neoclassical theory has to be built from the very beginning with a new or complementary definition of the subject. I have chosen to emphasize multidimensionality and democracy. Alternative views of individuals, organizations, markets, institutional change, sustainability assessment then follows as indicated in this essay and at other places (Söderbaum, 2016, 2017). But it should be made clear that there are ecological economists who hold a partly different view (Costanza et al. 2017). Neoclassical economists resist radical change but may modify their arguments. "Natural capital" has become a key concept for some, comparable to social capital, human capital and built capital (Helm, 2015). But these concepts are met with criticism for example by Richard Smith in the book *Green Capitalism. The God that failed* (2016).

Economists, in particular the neoclassical ones, claim expertise in forecasting while often failing to predict in practice. But other kinds of forecasting may be easier. Let me therefore predict that the debate about economics, ideology and sustainability will continue in the future. At some stage there will hopefully be a break-through with a more pluralist economics education and research at university Departments of Economics. Ideas about how such a change can take place exist in books, (e.g. Madi and Reardon, 2014) as well as journals (e.g. International Journal of Pluralism and Economics Education).

References

Berger, Sebastian and Rolf Steppacher, 2011. *The Foundations of Institutional Economics*. K. William Kapp. Routledge, London.

Brown, Judy, Peter Söderbaum, Malgorzata Dereniowska, 2017. *Positional Analysis for Sustainable Development: Reconsidering policy, economics and accounting.* Routledge, London.

Connolly, William E., 1974, 1983. *The Terms of Political Discourse*. Blackwell, Oxford.

Costanza, Robert, Gar Alperovitz, Herman Daly, Joshua Farley, Carol Franco, Tim Jackson, Ida Kubiszewski, Juliet Schor and Peter Victor, 2017. Chapter 16, pp. 367-454 in Stanislav Shmelev, Ed. *Green Economy Reader. Lectures in Ecological Economics and Sustainability*. Springer International Publisher Switzerland AG.

Daly, Herman E. and John E. Cobb, 1989. *For the Common Good. Redirecting the Economy Toward Community, the Environment and a Sustainable Future*. Beacon Press, Boston.

von Egan-Krieger, Tanja, 2014. *Die Illusion wertfreier Ökonomie*. Eine Untersuchung der Normativität heterodoxer Theorien. Campus Verlag, Frankfurt.

Helm, Dieter, 2015. *Natural Capital. Valuing the planet*. Yale University Press, London.

Henderson, Hazel, 1980. *Creating Alternative Futures. The End of Economics*. Perigee Books, New York.

Kapp, K. William, 1970. "Environmental Disruption: General Issues and Methodological Problems," *Social Science Information* (International Social Council), Vol. 4, No. 9, pp. 15-32.

Klein, Naomi. 2014. *This Changes Everything. Capitalism vs. the Climate*. Allen Lane, London.

Kras, Eva, 2007. *The Blockage. Rethinking Organizational Principles for the 21st Century*. Americaan Literary Press, Chicago Ill.

Kuhn, Thomas S., 1970. *The structure of scientific revolutions* (Second edition). Chicago University Press, Chicago.

Kumar, Pushpam, ed. 2010. *The Economics of Ecosystems and Biodiversity. Ecological and Economic Foundations* (the TEEB-study). Earthscan, London.

Madi, Maria Alejandra and Jack Reardon, 2014. *The Economics Curriculum: Towards a Radical Reformulation*. World Economics Association (WEA) Book Series Volume 1. College Publications (on behalf of WEA). Available at: www.worldeconomicsassociation.org

Mankiw, N. Gregory and Mark P. Taylor, 2011. *Economics*. Gengage Learning, Andover.

Monbiot, George, 2016. *How did we get into this mess? Politics, equality, nature*. Verso, London.

Mouffe, Chantal, 2005. *On the Political*. Routledge, London.

Mouffe, Chantal, 2013. *Agonistics. Thinking the World Politically*. Verso, London.

Myrdal, Gunnar, 1973. *Against the Stream. Critical Essays on Economics*. Random House, New York.

Myrdal, Gunnar, 1978. "Institutional Economics". *Journal of Economic Issues*, Vol. 12, No 4, pp. 771-783.

Schumacher, E. F., 1973. *Small is Beautiful. Economics as if People Mattered*. Harper & Row, New York.

Smith, Richard, 2016. *Green Capitalism. The God that failed*. World Economics Association (WEA), Book Series, Volume 5. E-book and print (WEA with College Publications). Available at www.worldeconomicsassociation.org

Swedish Environmental Agency, 2016. Environmental Objectives www.miljomal.se/sv/Environmental-Objectives-Portal/ (Accessed October 16, 2016)

Söderbaum, Peter, 2000. *Ecological Economics. A political economics approach to environment and development*. Earthscan, London.

Söderbaum, Peter and Judy Brown, 2010. Democratizing economics. Pluralism as a path toward sustainability, *Annals of the New York Academy of Sciences* 1185 (Ecological Economics Reviews, edited by Robert Costanza and Karin Limburg), pp. 179-195.

Söderbaum, Peter, 2013. "Ecological economics in relation to democracy, ideology and politics," *Ecological Economics* Vol 95, pp. 221-225.

Söderbaum, Peter, 2015. "Varieties of ecological economics: Do we need a more open and radical version of ecological economics?" *Ecological Economics* Vol. 119, pp. 420-423.

Söderbaum, Peter, 2016. *Economics, ideological orientation and democracy for sustainable development*. World Economics Association (WEA) e-Books. Available at www.worldeconomicsassociation.org

Söderbaum, Peter, 2017. "How Economics Can Become Compatible with Democracy", Chapter 2, pp. 25-38 in Stanislav Shmelev Ed. *Green Economy Reader. Lectures in Ecological Economics and Sustainability.* Springer International Publishing AG, Switzerland.

United Nations Sustainable Development Platform, 2016. http://www.un.org/sustainabledevelopment/sustainable-development-goals/ (Accessed 2016-10-14)

Author contact: peter.soderbaum@mdh.se

SUGGESTED CITATION:
Peter Söderbaum, "Do we need a new economics for sustainable development?", *real-world economics review*, issue no. 80, 26 June 2017, pp. 32-44, http://www.paecon.net/PAEReview/issue80/Soderbaum80.pdf

real-world economics review, issue no. 80
subscribe for free

From green growth towards a sustainable real economy[1]

The myth of decoupling, rebound effects, and the $I = P \cdot A \cdot T$ equation

Jørgen Nørgård[2] & Jin Xue[3] [Technical University of Denmark; Norwegian University of Life Sciences]

Abstract

Taking the simple static equation: *I* (impact) = *P* (population) ·*A* (affluence) ·*T* (technology) as the point of departure, this paper discusses the delusion of decoupling economic activities from environmental impacts (*I*) by resorting to simply reducing eco-intensities (*T*), (i.e. increasing efficiencies) through technological advancement. It is argued that the rebound effect reflects some dynamic interdependence between the factors. For instance a higher efficiency in the use of resources will partly be turned into growth in the economy i.e. (*A*) and (*P*) rather than into lower environmental impact (*I*). This is both a natural consequence of the growth dedicated society, and a driver of further economic growth. Through rebound effects, eco-efficiency efforts in the growth society tend to contradict and counteract the goal of environmental sustainability. To address the global environmental problems properly, our critique should therefore be redirected towards the growth ideology and growth policy itself. Drawing on the global inequity and emerging degrowth debates in the affluent countries, the paper proposes pathways towards a degrowth strategy by discussing the respective roles of population *P*, affluence *A*, and technology *T*. Overall, it is suggested that given an analysis not confined to monetary terms, but with *real cost and real benefits* represented by *environmental damage* and *human satisfaction,* respectively, sustainability in today's affluent countries might be achievable at no net cost.

Keywords delusion of decoupling, rebound effect, degrowth sustainability

The equation $I = P \cdot A \cdot T$, which combines population (*P*), affluence level (*A*), and technological eco-intensity (*T*) in the consideration of environmental impacts (*I*), has been well-known for a long time. Since the equation's development by Ehrlich and Holdren in 1971, the relative focus on the three factors has shifted. Although Ehrlich and Holdren at the initial stage emphasized the impacts of population (*P*) on the environment (Ehrlich & Holdren, 1971), today, the factor (*T*) almost totally dominates debates about solutions to resolve environmental deterioration (e.g. Weizsäcker et al., 1998; WCED, 1987; OECD, 2011). Population (*P*) is often tabooed and rarely included as a variable in the analyses and debates, although it obviously is still a key factor. In more recent debates, particularly in affluent regions, the factor (*A*) is sporadically being highlighted as a key to solve the ecological crisis

[1] This paper is a modified version of: Nørgård, J. and Xue J. 2016: *Between Green Growth and Degrowth, Decoupling, Rebound Effects and the Politics for Long-Term Sustainability*, Chapter 15 in Book: Rethinking Climate and Energy Policies (Santarious, Walnum and Aall, eds,) Springer, Switzerland.

[2] Technical University of Denmark, Department of Civil Engineering, Building 118, DK 2800 Lyngby, Denmark.
[3] Norwegian University of Life Sciences, Department of Landscape Architecture and Spatial Planning.

(Martinez-Alier et al., 2010). For instance, a call for reducing affluence levels is *well captured* in current discussions on "degrowth".

The quest for reducing the affluence (*A*) in rich nations, measured as per capita level of consumption of goods and services, is partly based on the impossibility of reducing resource consumption and pollution (*I*) to a level necessary for environmental sustainability by resorting to technological advancement (*T*) alone. This failure is to some extent attributed to the ignorance of the rebound effects from increasing resource use efficiency, which pushes up (*P*) and (*A*). In other words, the right-side factors in the equation are not constants, but mutually interdependent and dynamic (Alcott, 2008; 2010). This paper goes beyond this explanation, and points to the constant political quest for economic growth, i.e. growth in (*P*·*A*), as the fundamental problem. This suggests that it is difficult to eliminate rebound effects and sufficiently reduce environmental impacts without addressing directly the (*A*) and (*P*) factors, in addition to lowering the eco-intensities (*T*). Invoking the term "degrowth", we also propose for affluent, unsustainable regions alternative economics beyond the ideology of growth, and indicate desirable pathways down towards a sustainable level. In this paper, environmental impact (*I*) is mostly exemplified by energy consumption and carbon emissions, but (*I*) does refer as well to all other degradation of nature, such as biodiversity loss, resource depletion, and pollution of air, water and soil. By definition, a degrowth society "challenges the hegemony of growth and calls for a democratically led redistributive downscaling of production and consumption in industrialised countries as a means to achieve environmental sustainability, social justice and well-being" (Demaria et al., 2013). Degrowth therefore calls for strategies to reduce the aggregate impact from *P*·*A* in addition to the efficiency improvements (*T*). A degrowth society cannot be interpreted merely as a downscaled economy in the quantitative sense. It implies a qualitatively different society with different socio-economic structures and institutional settings (Asara et al., 2015). In addition, degrowth carries the ethical premise of distributive justice and intergenerational equity. Although the *P* factor is not given sufficient attention in the degrowth debates today, and propositions on population development among degrowth proponents are inconsistent (Kerschner, 2010; Latouche, 2009; Martinez-Alier, 2009), we believe that reducing the global population is essential for bringing human economic activity down to a sustainable level and thus should be advocated as a strong part of the solutions.

The paper will proceed as follows. In section 1, the 'growth and decoupling' approach for environmental sustainability is criticised as a delusion. Section 2 analyses how rebound effects are associated with the growth economy, and proceeds by arguing that attempts at enhancing labour and resource efficiencies in a growth dedicated society tend to contradict the goal of environmental sustainability. We therefore call for shifting the focus of critique from rebound effects to the growth ideology and policy in order to resolve environmental problems. This is followed by proposing a degrowth society. Drawing on the equation $I = P \cdot A \cdot T$ as an analytical framework, in sections 3, 4, and 5 we discuss options for degrowth of all three right-side factors, as well as some of their dynamics. Finally, the concluding section provides some reflections on the need for a concerted degrowth strategy taking into account capping the left-side factor *I,* and emphasizing the importance of addressing the deep socio-economic structures as part of the degrowth transformation, apart from the factors in the $I = P \cdot A \cdot T$ equation.

1. The delusion of decoupling economic activities from environmental impacts

During the 1960s and 70s ecological crises attributed to an exponential economic growth triggered a critical discussion on the environmental and social consequences of growth. This culminated with the publication of the report from the Club of Rome, *The Limits to Growth* (Meadows et al., 1972), together with other reports and books presenting similar growth critique (e.g. Daly, 1973; Goldsmith & Allen, 1972; Schumacher, 1973). During the 1980s, the growth critique was played down as the economy regained momentum, and was gradually replaced by the view of 'decoupling' economic growth from environmental deterioration. This 'decoupling' view was emphasized, for instance, by the World Commission on Environment and Development as a key strategy of sustainable development in their report *Our Common Future* (WCED, 1987) as well as in a number of publications that developed the concept of 'Ecological Modernization' (Huber, 1985; Spaargaren & Mol, 1992; Hajer, 1995). More recently, however, the possibility of maintaining environmentally sustainable economic growth through decoupling has been questioned by critics. Together with multiple socio-economic-political crises, this has revitalized the criticisms of economic growth, manifested in the increasingly heated debates on degrowth (Asara et al., 2015; Jackson, 2009; Martinez-Alier et al. 2010).

According to the decoupling view, economic growth and environmental sustainability are *not* incompatible, but can be combined. To illustrate the decoupling notion with the $I = P \cdot A \cdot T$ equation, it means that ecological impact (*I*) can grow at a lower rate than the growth in economic affluence level of a whole population, i.e. $P \cdot A$. In order to materialize such decoupling, the *T* factor is the key. The belief in decoupling was based on good efficiency progress in the wake of the 1970s' oil crises. Many analyses then showed remarkable potentials for increasing the efficiencies of energy use (e.g. Goldemberg et al., 1985; Lovins, 1977; Nørgård, 1979 a). In the 1990s, the concepts of Factor 4 (Weizsäcker et al., 1998), and Factor 10 (Schmidt-Bleek, 2001) emerged and provided a broad basis for now widely agreed national policies. Factor 4 means that the same amount of commodity being produced with only a quarter of the previous resource consumption (thus, factor 10 means using one tenth of the previous resource consumption). At an aggregate level, Factor 4 could imply "doubling wealth while halving resource use".

These large potentials in reducing eco-intensities were mainly low-hanging fruits resulting from the neglect of efficiency options during the post-war period of almost free oil. The efficiency boom, however, then resulted in a strong faith in the possibility of decoupling economic growth from environmental damage. "Reviving growth" was pointed out as an essential objective, the suggestion being that only the quality of growth should be changed (WCED, 1987). The Western faith in technology as the overriding solution has now shaped environmental policies for half a century.

This decoupling notion can be challenged in several different ways and is subject to serious criticisms. The conventional use of the decoupling term distinguishes "relative decoupling" from "absolute decoupling". If the ecological impact (*I*) grows at a lower rate than economic growth measured as GDP or *P·A*, this is expressed as relative decoupling. The term absolute decoupling has been used when growth in GDP - or $P \cdot A$ - does not result in increase in the overall ecological impact, i.e. (*I*) is kept stable or even declines (Jackson, 2009). What matters for ecological sustainability is whether the absolute environmental impacts, (*I*), increase or decrease. From this perspective, absolute decoupling is of fundamental concern.

However, a broad range of empirical evidence indicates a low achievement of absolute decoupling. At the aggregate economy level, the total emissions of CO_2 in OECD countries showed relative decoupling from economic growth during the 1990s (OECD, 2002), and similar modest relative decoupling has been observed in traffic volume versus CO_2 emissions (Tapio, 2005), as well as in housing sector's growth versus growth in residential energy consumption (Xue, 2014). However, the overall picture is that such reductions in environmental impacts *per unit* of product are cancelled out by the increases in volumes within any growing economy. Only if the speed of *T* going down equals the growth rate of *P·A*, can environmental impacts *I* be stabilized.

But stabilisation will not suffice.

Among environmentally concerned scholars it is generally agreed that the present global environmental impact is not sustainable. Taking Ecological Footprint as the indicator, we are presently overloading the Earth by a factor of around 1.5 (WWF, 2014), implying a need to reduce the global ecological footprint by 35 percent. Another study shows that four out of nine planetary boundaries have already been crossed by human activities, including climate change, biosphere integrity, biogeochemical flows and land system change, which might push the Earth system into a new state (Steffen et al., 2015). This means that for some specific environmental damages, a more drastic reduction is required. For the emissions of CO_2 and other greenhouse gases, the reduction has to be "net 100 %" by 2100 if we are going to achieve the goal of keeping global warming below 2 degrees (IPCC, 2014).

By acknowledging the enormous inequity in the world and ascribing all humanity equal environmental rights, it is argued that people in affluent countries, such as USA, EU, and Japan, would need to reduce their impacts (*I*) down to only around one tenth of present levels to ensure world-wide sustainability (Schmidt-Bleek, 2001). To achieve this in the course of 50 years given an annual GDP growth of 3 % would require eco-intensity (*T*) to be reduced by a *factor of around 40* (Nørgård, 2009; Jackson, 2009). For comparison, the much praised efficiency efforts in Denmark's energy system have over the past 25 years managed only to reduce (*T)* by a *factor of 1.4 !* (Danish Energy Agency, 2015). And this has been achieved by implementing the easiest and most cost-effective options. Could all technologies, in all sectors, possibly over the next 50 years become 40 times more efficient? Hardly!

Similar conclusions are reached by several other studies and examples, summarized by Trainer (2016). It therefore seems even theoretically implausible to reduce the environmental impacts *I* to reach and maintain a sustainable level by relying on reducing the *T* factor alone while simultaneously maintaining growth in GDP, i.e. $P \cdot A$.

Linguistically, decoupling implies that the two parameters – economic activity and environmental impact – are separated (Webster, 1986) with no coupling at all. OECD (2002) in their energy analyses define "decoupling" as breaking the links between "environmental bads" and "economic goods". Physically, however, there will always be some amount of 'coupling', since every economic activity is – directly or indirectly – reliant on a minimum of resource supply from nature and emission of wastes back into nature. And *vice versa*, all eco-impacts have their roots in economic activity. The fact that economic activity and the eco-impacts grow at different rates does not imply that the two parameters are not coupled.

The misleading term "decoupling" should therefore *not* be used in analyses and debates about economic growth and the environment. Instead, the term "relative decoupling" should be referred to as a *change in eco-intensity* (*T*), while the term "absolute decoupling" should be referred to simply as a *change in overall eco-impact* (*I*). These are not just linguistic trifles. The real problem is that the very use of the term "decoupling" might – probably sometimes intentionally – leave the false perception that we can let economic activities grow forever, without having to worry about any ecological constraints. The use of the term "decoupling" can be seen as a false "peacemaker" between environmentalists and growth-dedicated politicians, and thereby contributes to the maintenance of growth far beyond the economy's optimal size (Nørgård, 2009).

2. Rebound effects in a growth dedicated society

Normally the concept "rebound effect" depicts the phenomenon that eco-efficiency improvements through technological advancement do not reduce the adverse environmental impacts as expected from simple calculations, because the efficiency induced increases in production and consumption. For example, consumers who make their house technically more energy efficient and hence save on the energy bill are often tempted to take out that saving by more energy use in other ways, e.g. higher indoor temperature or an overseas flight for a holiday. The effect was first observed in the 19th century by British economist Jevons, who noticed that increasing efficiency in the use of coal was not accompanied by corresponding reduction in the use of that resource at the aggregate level, – rather the opposite (Alcott, 2005).

Rebound effects from using energy and other natural resources more efficiently has since the 1980s been a key dilemma of the energy efficiency debate, (Herring and Sorrell, 2009; Weizsäcker et al., 2009).

Here, we extend efficiency improvements to embrace other production factors, which we merge into one factor, *labour input*, by considering capital as accumulated stored labour input. Throughout industrialisation, technology has increased labour efficiency (productivity) in the sense of less work being needed per unit of output. A substantial part of the labour efficiency gains during early industrialization were utilized to reduce the more than 80 hours weekly work time. However, later on, more and more of the labour efficiency gains were turned into growth in overall production and consumption $A \cdot P$. In recent decades this rebound effect has in some affluent regions approached 100 percent, as illustrated by the average work time in the USA, which has been frozen since the 1930s at around 40 hours per week despite large gains in labour productivity (Schor, 2005). Almost all labour productivity gains are presently used to increase GDP – and consumption in general, – rather than to relieve the environmental impacts by lowering work input and consumption. Instead of reducing the input of labour, during the past 50 years the global *workforce* has enlarged substantially, partly by general population growth, and partly by absorbing ever more men and (in particular for the case of affluent countries) women into the economic (monetary) production sectors.

The direct and micro-level causes of rebound effects from eco-efficiency technologies can be largely ascribed to facts, that the estimates for savings were overlooking the socio-cultural elements and individual subjectivity in consumption behaviour (Santarius, 2016a; Peters and Dütschke, 2016; Santarius 2016b). In addition, increasing productivity through technological advancements involves a general trend of social acceleration, where the speed of production,

consumption, and mobility increases, leading to more consumption of resources (Suffolk and Poortinga, 2016). Nevertheless, there is nothing deterministic about the growth impact of improving resource and labour efficiency through technologies. As shown above, labour productivity gains could instead be employed to shorten work time instead of increasing production and consumption. We do, in fact, have choices as to how we utilize the benefits of efficiency improvements.

Arguably, the conversion of efficiency gains predominantly into more production and consumption is due to the ideology of economic growth and the structural growth imperative of a market-dominated socio-economic system. In the growth society, "quality of life" and 'well-being' are still very broadly interpreted as possession of material wealth, and hence, consumerism is a dominant value entrenched in society. Continuously enhancing material living standards becomes a widely accepted social norm without being questioned. When basic needs are satisfied, as in affluent societies, positional goods and conspicuous consumption are promoted as new momentums of growth through advertising and consumption-stimulating policies (Hirsch, 1976). This growth path was (with pressure from business) "deliberately" chosen in 1933 by US President F.D. Roosevelt as a way out of the economic depression (Hunnicutt, 1988; Cross, 1993). For consumers under the hegemony of the growth discourse, it is very likely that reduced costs due to lower *resource intensity* per unit of product is used to secure higher material standards, just as the case with rebound from higher *labour efficiency*. In other words, the growing purchasing power derived from either of the two efficiency gains has to be channelled to somewhere; often to higher levels of consumption (Schneider, 2008).

Furthermore, the fact that most efficiency improvements are turned into drivers of growth is highly associated with the market economy with its structural necessity of growth. Several authors have pointed out that the growth imperative is intrinsic to the market-dominated socio-economic system (Gordon & Rosenthal 2003; Griethuysen 2010; Harvey 2010). Fierce competition in the market economy sets a "grow or die" dynamic in motion. By enhancing both resource and labour efficiency, businesses are able to reduce the costs of products so as to earn more profits than their competitors and increase their market shares (Buhl and Acosta, 2016). Therefore, the rebound effect on the production side is an intentional pursuit by producers towards higher profitability. Not only business, but also governments seek high rebound effects. The Danish government has directly *required* that *"Energy savings should contribute to growth and commercial development"* (Danish Energy Agency, 2004).

Based on the discussion above, it can be argued that the rebound effect is both a natural *consequence* of a growth society and an important *contributor* to further economic growth. It is welcomed in current growth dedicated society and cannot be understood as a problem from a perspective of economic growth. Only when examined from an environmental perspective is rebound regarded as a problematic side effect, since it increases the level of production and consumption thus offsetting intended environmental gains from efficiency strategies. For societies, committed predominantly to perpetual growth in output, the rebound effect is not bad, rather the opposite. What remains as the fundamental problem is this commitment to economic growth and its contradiction with environmental sustainability. Both labour- and eco-efficiency strategies tend to be "co-opted" by the growth ideology and serve the purpose of maintaining growth. The more we reduce the eco-intensity (*T*), the more difficult it will be to decrease the aggregate impact of $A \cdot P$, because implementing technologies for resource efficiency and productivity are key drivers of economic growth. Attempts at enhancing labour

and resource efficiencies *in a growth society* tend to contradict the requirements for environmental sustainability. Hence, it appears patently impossible to reduce environmental impacts as much as needed by resorting only to technological eco-efficiency strategies in a growth society. It is imperative to address directly the growth issue in order to achieve long-term environmental sustainability (Nørgård, 2009).

This suggests the adoption of a degrowth strategy that seeks to stabilize or even lower the affluence level (*A*), and the population size (*P*). In addition to lower eco-intensities, the technology factor (*T)* will be redirected towards prolonging the durability of products. How would such a degrowth society look like? What policies would be needed in order to avoid problems like unemployment, poverty and inequality, during the shrinkage towards a sustainable steady state economy? The following sections sketch some suggestions for achieving a prosperous degrowth society by exploring the role of each of the right side factors in the $I = P \cdot A \cdot T$ equation.

3. Population development in a degrowth society

Global population has over recent decades moved from exponential growth into what appears more like linear growth. But this is still adding a staggering 80 million people annually to a limited planet. With continued population growth, a desirably low level of environmental impact can hardly be reached or maintained. Despite the environmental impact of the world's population, there is a taboo about including population as a variable when analysing future options for sustainability (Nicholson-Lord, 2008). Analyses and future scenarios typically start by just referring to the latest UN medium estimate of future population development (see Figure 1), and accept this one scenario as a fact. This lack of scientific and political vision or courage amongst experts to include a reduction of population as one possible, if not essential, contribution to solving this *man-made* impact, is one of the most glaring flaws in present environmental debates.

When European politicians and scientists in rare cases do touch on the issue of population, it is usually from a growth *promoting* viewpoint; for instance by encouraging higher birth rates to secure sufficient workers and consumers – again, to maintain GDP growth. As was historically hinted at by philosophers as long ago as the 1700s (Lütken, 1760; Malthus, 1798), this strategy must sooner or later result in ecological and human misery and starvation for parts of the world population.

Most demographers and politicians still contend that as standards of living in developing countries approach a Western level, birth rates will drop and thus prevent global overpopulation. One problem has been that in some poor regions, economic growth could hardly keep pace with population growth, which has resulted in stagnating or even declining standards of living, blocking the way for that "automatic" decline in birth rates. Furthermore, it is often forgotten that the European population "explosion" in the 1800s and 1900s was partly "resolved" by millions of Europeans migrating and taking control of "empty" continents: North and South America, Australia and parts of Asia and Africa. In recent decades similar population "explosions" have crowded other world regions, not least in Africa, but there are no more "empty" continents for these people to migrate to.

What then is an *optimal* number of people that could live comfortably on the planet with a sustainable natural environment? It seems evident that such optimisation must be a trade-off between the two aspects, since more people (*P*) imply reduced options for a good life.

A number of studies have under certain assumptions estimated optimal world population sizes. One such analysis from the 1990s was based on the energy needed for high quality of life versus the environmental problems from using fossil fuels. Given the average energy consumption of 7.5 and 1 kW per capita in respectively industrialized and developing nations at that time, researchers suggested 3 kW to suffice, and found an optimal world population of 1.5 - 2 billion people (Daily et al., 1994). More recent analysis, based on ecological footprint versus the earth's bio-capacity found that future optimal population levels range from 2.7 to 5.1 billion people (Desveaux, 2008), depending on average footprints, maintenance of bio-capacity and allowances for biodiversity. The lowest figure, 2.7 billion, allows for a 20 % margin for biodiversity. Obviously estimates of an *optimal* population are highly subjective, but it is worth noticing that they all indicate an optimal population size far below the present 7.5 billion.

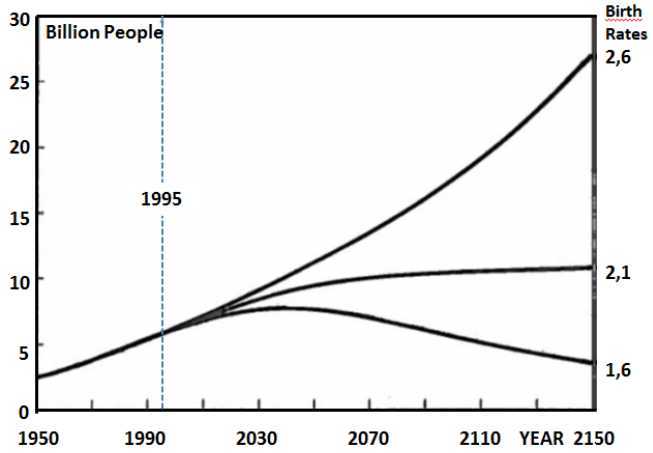
One recent estimate of how world population will actually develop if no extra actions are taken, suggests it will peak around 2040 at above 8 billion, followed by a very slow decline (Maxton and Randers, 2016). This seems to imply an overpopulated planet for a long time to come and underlines the urgency of aiming at a more optimal path by supporting a gentle reduction of the number of people on the planet.

The taboo on the population issue is widespread among decision makers, who wish to appear neutral on the sensitive issue of people's choice of family size. But no policy can be neutral on population development. All political decisions have indirect effects on fertility rates through tax systems, education, health care, social security, etc. Some decision makers defend their silence on active population limitation policies by the fact that such policies mainly have long-term effects. This is an odd way to justify *postponing* actions to secure the future of our descendants.

Imagine mankind decides to aim for half as many people on the planet as today, which is here suggested as the lowest UN scenario for world population development (Figure 1). If we act now, we could reach this goal around 2150 by convincing all women to have on average 1.6 children, rather than the present average of 2.6. Reaching a worldwide birth rate of just 1.6 should not be ruled out, considering that it is in fact the present average birth rates in two politically quite different regions of the world, Europe and China. In Europe the low birth rate and the resulting future contraction of population has been reached unintentionally (politically speaking) and voluntarily as a consequence of the general economic and welfare policy. In China it has been reached by a more conscious and direct active family planning policy.

Although China's family planning policy is effective in slowing population growth, it has been criticised for its authoritarian and coercive approach (Dietz & O'Neill, 2013). Later optimistic experiences from a number of developing countries, mainly in Asia, have shown how similar effects as those in China has been achieved based on non-coercive means, including education and empowerment of women (Kingholz & Töpfer, 2012).

Figure 1 UN scenarios for future world population development as a consequence of three average fertility rates (Source: The Population Division of the UNs Secretariat, 1998)



It is hard to see *disadvantages* of living in a future world with say half as many people as today. On the contrary, the basic problems mankind is facing today would be easier to solve. Even in monetary terms, reducing population is the most cost-effective strategy for mitigating climate change. This also applies to biodiversity loss and other resource and pollution problems. International conflicts too are often linked to shortages of land, food and resources. Lowering fertility rates to below the replacement rate of 2.1 would facilitate improved standards of living in general and provide environmental benefit through the reduced ecological footprint of fewer children and all their future descendants.

Mainstream politicians mainly associate low birth rates with the *transition problems, when* an ageing population will need more care to be provided by a shrinking productive workforce. Given a gentle pace of transition, these problems are, however, manageable, especially when remembering that a shrinking population will require less need for child care and educational services and that infrastructures like highways, buildings, power systems, libraries, schools, etc. inherited from the earlier, larger generations will be more than enough.. This inherited "overcapacity" will need maintenance and replacement, but the huge investments for *growth-induced expansion* are avoided, reducing also environmental impacts.

4. Affluence and work in a degrowth society

In the Western economies, average consumption per capita (*A*) has reached a level which qualifies as a dominant, very obvious factor in the environmental impact $I = P \cdot A \cdot T$. According to the *Living Planet Report 2014* (WWF, 2014), the ecological footprint per capita of high-income countries is about five times more than that of low-income countries. Furthermore, the high income countries often rely on the bio-capacity of other nations or the global commons to meet their consumption demands. Growth in such affluence does not primarily serve to satisfy human basic needs or even deeper needs, but rather to satisfy the "basic needs" of a debt-based financial economy designed for unlimited GDP expansion. This explains why not only financial actors in private business, but also governments insist on encouraging people to consume still more. The tools include massive advertising, key trade and tax policies, and also the goals of education and research systems. It is not hard to

imagine, that consumer preferences would shift somewhat, if these growth biased measures were slowed down or even reversed towards more equity and sustainability.

In general, when aiming for a degrowth economy many goals will be contrary to those of a growth economy. Fortunately, curbing excessive consumption offers rewards in return, mainly in the form of more free time, less stress, better health, more options for meaningful life, in addition to a better rather than worse environment.

It is notable that reducing affluence levels whilst improving life quality is a key focus in many successful sustainability initiatives at the community level (Butters, 2010).

The affluence level (*A*) is not only coupled to environmental impact but also affects our health. Economic growth in wealthy countries might still bring new health improvements through better technology and medicine, but excessive consumption is in many regions also causing huge *negative* health impacts in the form of lifestyle diseases; such as obesity, caused by overconsumption of food and sedentary, motorized lifestyles. Others are smoking, alcohol and drug abuse. Around the year 2000, overconsumption alone in the USA was found to result in more than one million premature deaths every year. With “a slip of the tongue”, the USA’s Secretary of Health stated that these and other “social problems and complaints stem from our affluence, not our poverty” (Samuelson, 2004), exposing such severe and rising human cost of the growth policy.

If we assume that consumption can be expressed by people’s annual income, studies comparing different nations indicate that increase in income gives a diminishing return in the form of well-being or happiness. This is particularly clear when average annual income exceeds $10,000 per person, and beyond $ 15,000 the extra benefit seems negligibly (Jackson, 2009). Moreover, the same study shows that countries with the highest score of happiness, such as Iceland, the Netherlands, Denmark and Sweden have lower incomes than USA. When observing the historical relationship between economic growth and happiness in USA, it is found that the percentage of people who report being “very happy” stabilized at around 30 % during the years 1945-2005, although income more than tripled (Dietz & O’Neill, 2013). This indicates that other aspects of life are more important for people’s wellbeing than their level of consumption or income. Some of these are equity, education, job guarantee, etc. (Wilkinson & Pickett, 2010). An increasing number of such studies show that further economic growth in the developed countries is *not a necessary condition* for progress in human well-being.

Apart from arguing for continuously increasing affluence levels in terms of social benefits, a key political argument for increasing consumption in affluent nations is to avoid increasing unemployment resulting from productivity increases. In general, there are three ways to accommodate this: 1) increase public and private investments, 2) increase consumption, and 3) reduce labour input in terms of lowering annual work time and/or labour productivity to fit the production wanted. The simplest long term solution to avoid unwanted unemployment without growth in investment and consumption is to share the work to be done annually by lowering the work time.

Annual working time in various nations is quite different, with people in USA, Russia, South Korea and Japan working about 20 % more than Europeans. This suggests that Europeans, as in the population issue, are on a positive track towards degrowth and sustainability. In addition, in a future aimed at sustainability, pressure for continuously increasing labour

productivity can be relieved and even reversed as a means for both adapting production to declining consumption and simultaneously making working conditions better and more meaningful in various ways.

Lowering affluence can appear an impossible task, given the dominance of the growth ideology. After lifelong exposure to intensive commercial advertising and political encouragement to buy ever more, plus the peer pressures of neighbours' new cars and bigger house, it is understandable that people may be reluctant to reduce their consumption. On the other hand, surveys on work time preferences have actually indicated an increasing wish among people for less work time (Gorz, 1983; Hayden, 2000; Sanne, 2000; Schor, 1991). A series of surveys conducted in Denmark over some decades showed that the fraction preferring less work over more income increased from 44% in 1964 to 73% by 2007 (Nørgård, 2009).

People's preferences for more leisure over more income, as illustrated above, might well be based on personal concerns, not our common environmental benefits of their choice (Hayden, 2000). With this argument added, preferences would probably be higher. However, increase in leisure activities cannot be anticipated to necessarily lead to fewer environmental problems, due to the possibility of time-use rebound in terms of resource use (Santarius 2016). Not all spare-time activities are environment friendly (Aall, 2011). However, stabilization or even decline in income due to reduction in working hours constitutes one of the mechanisms counteracting consumption. In addition, tax policies can be shaped to encourage people to engage in leisure activities that are relatively less resource intensive or environmentally harmful.

The fact that most people in affluent Western nations express a wish to use productivity gains to get more free time rather than more income, if given the choice, should be seen as a welcome opportunity for politicians to gently change economic paths away from the money dominated growth economy towards a degrowth economy. In a degrowth society, the environmental impact from the affluence level could decline in combination with an improvement in quality of life in the form of better health, more freedom and non-material sources of happiness. (Nørgård 2013).

5. Technology in a degrowth society

Having considered population (*P)* and affluence (*A)* above, we now return to the third factor (*T)*. The rebound effects of eco-efficiency in technology should not lead us to dismiss technological advancement as an obvious part of strategies for environmental sustainability in a degrowth society. A key problem with technical solutions is that they often overshadow many more effective "soft" solutions, including political instruments and social innovations. Arguably, technological advances in a degrowth society with a cap on affluence level (*A*) and population size (*P*) will not lead to net rebound effects and thus will at last contribute to *reducing* environmental impacts.

Besides seeking higher efficiency in *direct use* of resources, this section will also address how technologies on the consumer side can be utilised in interplay with lifestyle and behavioural issues to contribute substantially to reducing also the *indirect* resource use and pollution. The potential of technologies in these aspects include enhancing consumer efficiency by sharing

and prolonging the useful life expectancy of consumer goods, policies which have been neglected or counteracted in economies dedicated to growth.

People's material affluence (*A*) can be expressed by the consumption of three types of goods and services: 1) flows of non-durable goods, defined as consumption of goods, the value of which lies in actually being consumed, such as food, water, electricity, heat etc., 2) stocks of durable goods, defined as physical goods including houses, clothes, appliances and cars, the value of which lies in having a stock of them at one's disposal, and 3) services, such as trade, entertainment, education, administration, health care, administration, which are provided to people by durable and non-durable goods outside their daily sphere (Nørgård, 2006).

Most focus on energy saving options has been devoted to the non-durable flow of *direct energy* used for providing services like transport, light, comfort, meals, etc. by operating energy consuming durables like cars, lamps, houses, refrigerators, TVs, etc. In these fields, substantial room for energy efficiency improvements has been identified and to some extent implemented. These efficiency gains hold many examples of rebound effects, by increasing the sale of durables.

However, investigating *indirect* energy consumption, defined as the energy used to produce and provide the durable goods, opens up more room for reduction in environmental impacts, in particular when technological improvements are integrated with behaviour and lifestyle changes. The potentials for these savings lie in 1) improving energy efficiency in the whole chain of the system providing the durables; 2) reducing the number of durable goods people possess, e.g. by more sharing of goods; 3) extending the useful lifetime, and finally, when scrapped, 4) recycling components or materials in a circular economy system. In the following the focus is on the product lifetime. These considerations apply not only to energy but to resource use in general.

The useful lifetime of durable goods is determined by three factors (Nørgård, 1979b): *technological obsolescence,* meaning the physical wear and tear and inability to fulfil the basic purposes of the products; *functional obsolescence,* in the sense that new products can fulfil the purpose in a better way, for instance by being more energy efficient or providing better service options; and *psychological obsolescence,* e.g. by becoming out of fashion compared to novel designs on the market. The most striking example of fashion driven purchase is clothes. But today sale of most items, including cars and houses, is to a large extent promoted by changing fashion. Obviously, the first occurring obsolescence of a product will determine the factual useful lifetime of the product.

In the growth economy, *planned obsolescence* that deliberately makes products obsolete faster in any or all of the three obsolescence factors is a business strategy to accelerate capital accumulation and at the macro-economy level boosting growth in GDP (Slade, 2006). There is therefore a basic conflict between increase in the consumption of durable goods and preservation of the environment. In a growth dedicated economy, public campaigns aimed at saving energy or the environment have been lukewarm in also emphasizing the *indirect* use of energy, because this would imply a general curb on economic activities. This argument can obviously not hold if sustainability is given higher priority.

In contrast to the call for speeding up the flow of durable goods in the growth society, a degrowth society aims at slowing down this flow and reducing the total amount of durable goods people possess.

Extending the useful lifetime of durable goods might be the most fruitful way of lowering environmental impacts, through combining behavioural and technical changes. This could apply to, e.g. electronic products, clothes, buildings, plastics. Manufacturers could use their technical expertise to design more durable and repairable products with longer intervals between functional and fashion changes. Sharing various goods also constitutes a significant potential for saving energy and other environmental impacts, since this reduces the size of the stock of durable goods. Besides examples like cars, tools, and clothes, this thinking also include architectural design to facilitate flexibility and co-housing (Lietaert, 2010).

The main obstacle on the path towards such indirect resource saving is not the technology, which is almost readily available. We do not have to wait for new invention before starting a transition. As an example, electronic devices like mobile phones now scrapped after a year or two can easily last for ten or more years. Similarly with clothes. In certain areas, e.g. urban sustainable development, it is also a matter of designing for, and reinvigorating, old environmentally friendly technologies, such as bicycles, buses and communal building facilities to replace the relatively new but more environmentally harmful solutions, including the universal goal of private cars (Næss &Vogel, 2012). What seems to be more important is the necessary change in economic and financial targets, in work patterns as discussed in section 4, and in culture and lifestyle. Fashion and advertising can, as demonstrated in recent decades, be quite effective in changing people's lifestyles with satisfaction through faster obsolescence replacement. Alternatively one could use the same advertising expertise to convey to consumers, little by little, the benefits of focusing more on the physical services or use values provided by the car, the clothes, and the other durable goods, and less on fashions and novelty.

To summarize, attempts at enhancing eco-efficiencies through technological advancement should not be abandoned in a degrowth society. However, technological innovation should to a higher degree be reoriented in the direction of focusing on use values and longevity of durable products, along with far more emphasis on cultural and lifestyle change, which, where achievable, are in addition entirely free.

6. Concluding remarks

In this paper we have argued that since the phenomenon of rebound effect constitutes a complicated indirect barrier to achieving environmental sustainability in affluent, growth dedicated economies, we should instead devote our critical attention directly to the growth paradigm itself, which is both a fundamental causal mechanism of the rebound effect and partly a consequence of it.

Throughout the paper, we have employed the $I = P \cdot A \cdot T$ equation to illustrate and develop our argument. We first criticised the belief in decoupling economic growth from environmental impacts and the misleading use of the very term "decoupling" that seems to suggest the material independence of economic activities. We then argued that the options for utilizing efficiency improvements in the use of resources and labour hold much larger potentials than just being rebounded into increased levels of production and consumption. It is the growth ideology and the structural necessity of growth of the market economy that converts efficiency gains into drivers of further economic growth. Hence, rebound effects in a society dedicated to growth in the case of poor, developing countries, are more than welcomed, whilst efficiency

improvements in affluent growth economies are likely to contradict the goal of environmental sustainability. In the light of this argument, we further proposed a degrowth society which addresses simultaneously decreasing population size, reducing affluence level through work sharing, and redirecting technology towards prolonging the functional longevity of goods in addition to increasing resource use efficiency, as pathways towards reducing environmental impacts to a sustainable level. Such an economy geared towards a sustainable steady state is not only beneficial for environmental and resource problems, but may also contribute to a happier and more meaningful life by e.g. more sharing of the fewer work hours and more relaxed work conditions.

Apart from addressing the right-side factors in the equation, it is also an essential pathway towards the degrowth society to combine this with policies of directly capping the resource use and environmental impacts (*I*) on the left side of the equation. As Alcott (2010) suggested, the cap strategy can take the form of 1) production caps where limits are imposed on the input of raw materials to production, 2) consumption caps restricting the end-use of energy and other resources, and 3) pollution/emission impact caps. A multi-scalar cap system can be developed where individual and municipal caps are deduced from the national and global maxima. A capping strategy should be adopted in a concerted and coordinated way with the right-side factors. This will prevent the potential rebound effects which are generated by focusing separately on the factors regardless of the dynamics between them.

To build a degrowth society also requires a profound socio-economic transformation apart from adopting the strategies targeting the four factors in the equation. As discussed, the growth commitment and the consumerism culture emanate from the "grow or die" dynamic in the market economy. Without confronting the hegemony of this economic structure, it is hard to eradicate the growth imperative. The necessary measures like implementing policies aimed at slowing down labour productivity, curbing or discouraging demand for consumption, redirecting technologies towards use value and durability, will all meet resistance from conventional business and finance. The current propagation of a neoliberal agenda all over the world is at odds with policy suggestions for a sustainable society. The weaknesses of the present systems have been increasingly manifested through its failures in tackling the social, ecological, political and economic crises it has generated. It is urgent to transform the economy and society not only for a better environment but also for long-term human prosperity.

A degrowth transformation should be first pursued in the developed, affluent countries where the current economic volume can be qualified as "uneconomic growth" with negative interest rates (Daly, 1999; Daly, 2016; Trainer, 2016). For the least developed countries where economic growth still plays an important role in enhancing people's wellbeing, increase in consumption levels is thus still needed, but only temporarily. After a period of growth in these developing countries to a point placed safely within the planet's ecological capabilities, these countries should also aim for a long term development with a steady state economy. Whether a steady state economy can be run as a conventional capitalist economy, is a question still discussed (Trainer, 2016).

The cost for a degrowth transition can be very low or negative if analysed in *real economy terms,* i.e. not confined to what happens to be measured in money. In the real economy, the *real benefits* are measured in people's satisfaction and the *real cost* in the destruction of the natural environment. In that case, most of the actions needed in the affluent nations to take towards humane and environmentally sustainable societies proposed in this chapter are

available at no *real net cost.* If people prefer to have no more than two children, it makes no sense to ascribe a real human *cost* to this essential ecological *benefit*. Similarly, if people at a certain affluence level are convinced to prefer more relaxed work conditions over more materiel consumption, a degrowth economy can give them more of what they really want, again at no real *cost*, and with *benefit to* the environment and quality of life. And if technological development is directed towards longevity and eco-efficiency in general, it is possible to provide decent and comfortable lives to all humanity and preserve or restore the natural ecology. Only when mankind insists on monetary cost-benefit analyses do we seem to have a real problem.

Acknowledgment

We wish to thank Chris Butters, Oslo University/GAIA architects, for his tireless and most useful suggestions towards both the content and the language of this paper.

References

Aall, C. (2011) "Energy use and leisure consumption in Norway: an analysis and reduction strategy". *Journal of Sustainable Tourism*, 19(6), 729-745.

Alcott, B. (2005) "Jevons' paradox". *Ecological Economics*, 54(1), 9-21.

Alcott, B. (2008) "The sufficiency strategy: Would rich-world frugality lower environmental impact?" *Ecological Economics*, 64(4), 770-786.

Alcott, B. (2010) "Impact caps: Why population, affluence and technology strategies should be abandoned". *Journal of Cleaner Production*, 18(6), 552-560.

Asara, V., Otero, I., Demaria, F., Corbera, E. (2015) "Socially sustainable degrowth as a social-ecological transformation: repoliticizing sustainability". *Sustainability Science*, 10(3), 376-384.

Buhl, J. and Acosta, J. (2016). "Indirect Effects from Resource Sufficiency Behaviour in Germany". In: Santarius, Walnum and All.(Eds) *Rethinking Climate and Energy Policies*. Springer Publ. Switzerland.

Butters, C (ed,) (2010). *Signals – Success Stories in Sustainability*. The Ideas Bank, Oslo. www.thebalancingact.info/signals/SIGNALS.pdf.

Cross, G. (1993) *Time and money. The making of consumer culture*. London/ New York: Routledge.

Daily, G. C., Ehrlich A. H., Ehrlich, P. R. (1994) "Optimal population size". *Population and Environment*, 15(6), 469-475.

Daly, H. (ed.) (1973) *Towards a Steady State Economy*. San Fransisco, CA: Freeman

Daly, H. E. (1999) "Uneconomic growth in theory and in fact". FEASTA review 1 - The First Annual Feasta Lecture. Trinity College, Dublin 26th April, 1999. Downloaded September 2015 from http://www.feasta.org/documents/feastareview/daly1.pdf.

Daly, H. (2016) "Negative interest rates or 100% reserves: alchemy vs chemistry", *Real-World Economics Review*, # 76, September, pp. 2-4. http://www.paecon.net/PAEReview/issue76/Daly76.pdf.

Danish Energy Agency (2004) *Energy Saving Policies*. Copenhagen: Danish Energy Agency.

Danish Energy Agency (2015) *Energy Statistics 2013*. Copenhagen: Danish Energy Agency.

Demaria, F., Schneider, F., Sekulova, F., Martinez-Alier, J. (2013) "What is degrowth? From an activist slogan to a social movement". *Environmental Values* 22(2), 191-215.

Desveaux, M. (2008) Towards sustainable and optimal populations. http://populationmatters.org/resources/articles-reports-papers/population./Population Matters. 135-137 Station Road, London E4 6AG, UK

Dietz R., O'Neill, D. (2013) *Enough is Enough. Building a Sustainable Economy in a World of Finite Resources*. Oxon: Routledge.

Ehrlich, P., Holdren, J. P. (1971) "Impact of population growth". *Science* (New Series), 171 (3977), 1212-1217

Goldemberg, J., Johansson, Th. B., Reddy, A. K. N., Williams, R. (1985) "Basic needs and much more with one kilowatt per capita". *Ambio*, 14(4-5), 190-200.

Goldsmith, E., Allen, R. (1972) "A Blueprint for Survival". *The Ecologist*, 2(1).

Gordon, M. J., Rosenthal, J. S. (2003) "Capitalism's growth imperative". *Cambridge Journal of Economics*, 27(1), 25-48

Gorz, A. (1983) *Les Chemins du Paradis. L'agonie du Capital*. Paris: Edition Galilée.

Griethuysen, P. (2010) "Why are we growth-addicted? The hard way towards de-growth in the involutionary western development path". *Journal of Cleaner Production*, 18(6), 590-595.

Hajer, M. A. (1995*) The politics of environmental discourse: Ecological modernization and the policy process.* Oxford: Clarendon Press.

Harvey, D. (2010) *The Enigma of capital: And the crisis of capitalism*. London: Profile Books Ltd.

Hayden, A., (2000) *Sharing the work, sparing the planet. Work time, consumption and ecology*. New York: Zed Books.

Herring, H. and Sorrell, S. (eds.) (2009). *Energy Efficiency and Sustainable Consumption. The Rebound Effect*. UK: Palgrave Macmillan,

Hirsch, R. (1976) *Social Limits to Growth*. London: Routledge

Huber, J. (1985) *Die Regenbogengesellschaft: Ökologie und Sozialpolitik. Frankfurt am Main*, Germany: S. Fischer

Hunnicutt, B. K. (1988) *Work without end. Abandoning shorter hours for the right to work.* Philadelphia, USA: Temple University Press.

IPCC (2014) Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.)]. Geneva, Switzerland: IPCC

Jackson, T. (2009) *Prosperity without growth - economics for a finite planet*. London: Earth Scan

Kerschner, C. (2010) "Economic de-growth vs. steady-state economy". *Journal of Cleaner Production*, 18(6), 544-551

Kingholz, R., Töpfer, K. (2012) Das Trilemma des Wachstum. Discussion Papers. Berlin Institut für Bevölkerung und Entwicklung. www.berlin-institut.org.

Latouche, S. (2009) *Farewell to Growth*. Cambridge: Polity Press.

Lietaert, M. (2010) "Cohousing's relevance to de-growth theories". *Journal of Cleaner Production*, 18(6), 576-580.

Lovins, A. (1977) *Soft Energy Paths. Towards a Durable Peace*. London: Penguin Books.

Lütken, O. D. (1760) *Undersøgninger angaaende Statens almindelige Oekonomie (Investigations Concerning the State's general Economy*). Sorøe, Denmark.

Malthus, T. R. (2008) *An essay on the principles of population* (orig. publ. in 1798). UK: Oxford University Press.

Martinez-Alier, J. (2009) "Social metabolism, ecological distribution conflicts and languages of valuation". *Capitalism Nature Socialism*, 20(1), 58-87.

Martinez-Alier, J., Pascual, U., Vivien, F., Zaccai, E. (2010) "Sustainable de-growth: mapping the context, criticisms and future prospects of an emergent paradigm". *Ecological Economics* 69(9), 1741-1747.

Maxton, G. and Randers, J. (2016) *Reinventing Prosperity. Managing Economic Growth to Reduce Unemployment, Inequality, and Climate Change*. Vancouver/Berkeley, Greystone Books.

Meadows, D. H., Meadows, D. L., Randers, J., Behrens III, W.W. (1972) *The limits to growth*. New York: Universe Books.

Nicholson-Lord, D. (2008) "The silence of the greens". *Resurgence,* 251, November/December.

Næss, P., Vogel, N. (2012) "Sustainable urban development and the multi-level transition perspective". *Environmental Innovation and Societal Transition*, 4, 36-50.

Nørgård, J. S. (1979a) "Improved efficiency in domestic electricity use". *Energy Policy*, 7(1), 43-56.

Nørgård, J. S. (1979b) *Husholdninger og Energi. (Households and Energy)*. Lyngby, Denmark: Polyteknisk Forlag.

Nørgård, J. S. (2006) "Consumer efficiency in conflict with GDP growth". *Ecological Economics* 57(1), 15-29.

Nørgård, J. S. (2009) "Avoiding rebound through a steady state economy". Chapt.10 in: Herring, H., Sorrell, S**.** (Eds.), *Energy Efficiency and Sustainable Consumption. The Rebound Effect*. UK: Palgrave Macmillan, pp. 204-223 (Chapter 10).

Nørgård, J.S. (2013) "Happy Degrowth through more Amateur Economy". *Journal of Cleaner Production,* 38 pp.61-70.

OECD (2002) *Indicators to measure decoupling of environmental pressure from economic growth*. Paris: OECD.

OECD (2011) *Towards green growth. A summary for policy makers*. Paris: OECD.

Peters, A. and Dütschke, E. (2016) "Exploring Rebound Effects from a Psychological Perspective". Chapt. 6 in: Santarius, Walnum and All.(Eds.) *Rethinking Climate and Energy Policies*. Springer Publ. Switzerland.

Samuelson (2004) "The afflictions of affluence", *Newsweek,* March 22, 204.

Sanne, C. (2000) "Dealing with environmental savings in a dynamic economy". *Energy Policy*, 28(6-7), 487-495.

Santarius, T. (2016 a**).** Production-Side Effects and Feedback Loops between the Micro and Macro Level. Chapt. 5 in: Santarius, Walnum and All.(Eds) Rethinking Climate and Energy Policies, Springer Publ. Switzerland.

Santarius, T. (2016 b). "Energy Efficiency and Social Accelaration: Macro-level Rebound from a Social Perspective". Chapt. 9 in: Santarius, Walnum and All.(Eds) *Rethinking Climate and Energy Policies*. Springer Publ. Switzerland. .

Schmidt-Bleek, F. B. (2001) *The story of factor 10 and MIPS*. Carnoules, Provence: Factor 10 Institute

Schneider, F. (2008) "Macroscopic rebound effects as argument for economic degrowth". Proceedings of the First Degrowth Conference for Ecological Sustainability and Social Equity. Paris.

Schor, J. B. (1991) *The overworked American. The unexpected decline of leisure*. USA: Basic Books.

Schor, J. B. (2005) "Sustainable consumption and worktime reduction". *Journal of Industrial Ecology*, 9(1-2), 37-50.

Schumacher, E. F. (1973) *Small is beautiful. Economics, – as if people mattered*. London: Blond and Briggs

Slade, G. (2006) *Made to break. Technology and obsolescence in America*. Cambridge, Mass. USA: Harvard University Press.

Spaargaren, G., Mol, A. P. J. (1992) "Sociology, environment, and modernity: Ecological modernization as a theory of social change". *Society and Natural Resources*, 5(4), 323-344.

Steffen W, Richardson K, Rockström J et al. (2015) "Planetary boundaries: guiding human development on a changing planet". *Science,* 347(6223).

Suffolk, C. and Poortinga, W. (2016) "Behavioural Changes After Energy Efficiency Improvements in Residential Properties". In: Santarius, Walnum and All.(Eds) *Rethinking Climate and Energy Policies*. Springer Publ. Switzerland.

Tapio, P. (2005) "Towards a theory of decoupling: degrees of decoupling in the EU and the case of road traffic in Finland between 1970 and 2001". *Transport Policy*, 12(2), 137-151.

The Population Division of the UNs Secretariat (1998) "United nations world population projections to 2150". *Population and Development Review*, 24(1), 183-189. Population Council.
http://www.jstor.org/stable/2808146

Weizsäcker, E., Lovins, A. B., Lovins, H. (1998). *Factor Four: Doubling wealth, halving resource use. A Report to the Club of Rome*. London: Earthscan.

Weizsäcker, E., Hargroves, K.C., Desha, C., Smith, H.,and Stasinopulos,(2009). *Factor Five. Transforming the Global Economy through 80% Improvement in Resource Productivity*. London, Sterling VA.

Webster (1986) *Webster's New World Dictionary of American English*. Third College Edition. New York: Prentice Hall.

Wilkinson, R., Pickett, K. (2010) *The spirit level. Why equality is better for everyone*. London: Penquin Books.

WCED. World Commission on Environment and Development (1987) *Our common future*. Oxford: Oxford University Press.

WWF (2014) *Living Planet Report 2014*. Gland: WWF. www.factor10-institute.org

Xue, J. (2014) *Economic growth and sustainable housing: an uneasy relationship*. London: Routledge.

Author contact: jsn@byg.dtu.dk

SUGGESTED CITATION:
Jørgen Nørgård & Jin Xue, "From green growth towards a sustainable real economy", *real-world economics review*, issue no. 80, 26 June 2017, pp. 45-62, http://www.paecon.net/PAEReview/issue80/NorgaardXue80.pdf

You may post and read comments on this paper at https://rwer.wordpress.com/comments-on-rwer-issue-no-80/

real-world economics review, issue no. 80
subscribe for free

Split-circuit reserve banking – functioning, dysfunctions and future perspectives

Joseph Huber [Martin Luther University, Halle, Germany]

Abstract
This paper first provides a detailed outline of how the present money system works. This then serves as a backdrop to discuss a number of orthodox fallacies and heterodox flaws in money theory, followed by a summary of the dysfunctions of split-circuit reserve banking and a brief outlook on the perspective of a single-circuit sovereign money system.

JEL codes E42, E51, E52, E58, G21

Keywords monetary economics, money theory, credit creation, banking theory, fractional reserve banking, monetary policy, monetary reform

Introduction

This paper provides an up to date outline of the workings of the money and banking system – how money is created, how it circulates in the payment system, how it is temporarily de- and re-activated, and how it is finally deleted. This then helps clarify why a number of orthodox money and banking theories are obsolete, in particular the financial intermediation theory of banking in connection with the loanable funds model of deposits, the models of a credit multiplier, the reserve position doctrine, and other rather fictitious elements of present-day monetary policy.

However, some more advanced approaches also contribute to disorientation, for example, when describing the present system as a chartalist or sovereign currency system, or when defending the false identity of money and credit, or postulating an arbitrary notion of endogenous and exogenous money, or when denying the constraints on bankmoney creation.

Main elements of reserve banking today

The split-circuit structure of reserve banking

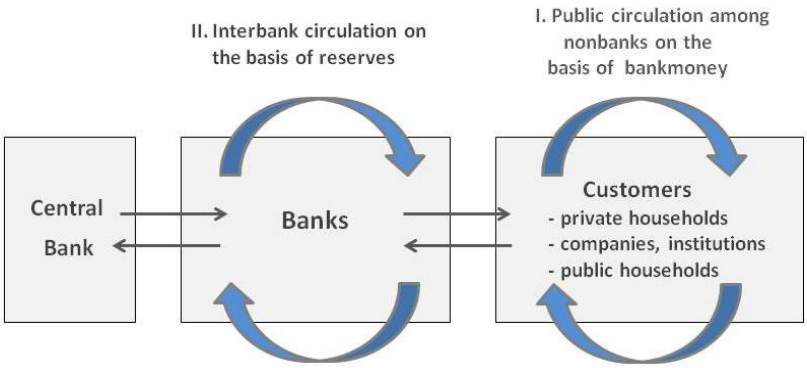
One of the first things to be read in most textbooks about money and banking is the two-tier structure of the system. One tier is the central bank of a currency area; the other is the banking sector. This is patently obvious were it not for some misleading views most often coming with two-tier explanations, for example, that in the first instance the money is created by the central bank, loaned to the banks, and loaned out from the latter to bank customers, or used as the basis for creating bankmoney as a multiple of the central bank money. As discussed below, nothing of this does apply.

What is more, the two-tier description of banking does not make explicit a most fundamental feature of the system, which is the split-circuit structure of modern reserve banking. The system consists of two different money circuits. One is the public circulation of bankmoney

among nonbanks. Bankmoney is another term for demand deposit or sight deposit or overnight deposit in a current account. The latter is also called a giro account, used in cashless payments. The term nonbank refers to the bankmoney-using public, including non-monetary financial institutions such as funds or trusts, non-financial firms, private households, and public households as far as the latter run bank accounts. Nonbanks run their accounts with banks. Except for some government bodies, nonbanks are not admitted to central bank accounts.

The other circuit is the interbank circulation of reserves among banks. Reserve is the technical term for non-cash central bank money on a bank's operational account with the central bank (see Figure 1). More precisely, the reserves referred to here are payment reserves, i.e. liquid excess reserves for making interbank payments, in contrast to basically illiquid minimum reserve requirements.

Fig. 1 The split circuit of reserve banking



The two circuits are separate and never mingle; however, the public circuit is technically tied to the interbank circuit, whereas the interbank circuit is basically independent, even though it helps mediate the cashless payments among nonbanks.

Reserves and bankmoney represent two distinct classes of money that cannot be exchanged for one another. Customers never obtain reserves in their current accounts, and bankmoney cannot be transferred into a bank's central bank account. Customer deposits (bankmoney) thus cannot be used by banks to make interbank payments, and cannot be lent by banks to whomsoever; only customers themselves can spend, or invest, or lend their deposits (bankmoney) to other nonbanks.

Modern money is non-cash

As far as traditional solid cash (banknotes and coins) is still in use, cash circulation represents a third circuit. In contrast to precious-metal coins, and like reserves, cash is token fiat money today. But rather than circulating between central bank accounts (reserves) or between bank accounts (bankmoney), traditional solid cash circulates from hand to hand in public circulation, without needing banks, or central banks respectively, as a trusted third party. Regarding the future of money, modern digital cash based on some form of blockchain technology might become a modern equivalent of traditional cash (notwithstanding the question of who will issue and control the stock of such digital currency). In any case, in a

basically cashless money system based on money-on-account, traditional cash is no longer of defining relevance.

Within the frame of reserve banking, cash and money-on-account must not be confused as is done by negligent speak, and even by official accounting standards.[1] At source, modern money is non-cash, a credit entry into a respective account. In the split-circuit structure, this applies to both bankmoney and central bank money. Traditional solid cash (coins, notes) has become a residual technical subset of the bankmoney in circulation, withdrawn from or exchanged back into a bank giro account.

Since about the 1920–60s, when bankmoney was definitely driving out solid cash in the course of the general dissemination of cashless payment practices, cash has no longer been constitutive of the money system. Cash now represents about 3–15% of the stock of money (M1), depending on the country, and a continued declining share in the long run. When referring to broad money aggregates (M2/3/4 which include, for example, deposit savings and money market fund shares) cash amounts to only 2–10%. Accordingly, cash can now largely be excluded from monetary system analysis (in spite of its present role as an effective hindrance to misguided negative interest rate policies of central banks). The means of payment that dominates everything today is bankmoney with its share of 90–98% in the entire money supply.

Credit extension and money creation in one act

Bankmoney and reserves are also called credit money or debt money because that money is created in one and the same act with crediting an account. Bankmoney is created when a bank enters previously non-existent currency units into a customer account. This creates a demand deposit. What makes the difference between a bank and a non-bank financial institution is a bank's ability to create primary credit that creates bankmoney, in contrast to secondary credit which is about lending or investing of already existing bankmoney among nonbanks. Central bank reserves are created in the same way through the central bank crediting a bank's account with the central bank. Central bank credit as well as bank credit are primary or originating, they are not about transferring already existing amounts of reserves or bankmoney.

“Credit creates deposits” has become a general teaching in post-Keynesianism and circuitism. The opposite of “deposits create credit” no longer applies to the bank-customer relationship in a predominantly cashless money system. This was already recognized in the bank credit theory of money from the 1890–1920s, but largely ignored by the mainstream, except for the Austrian School, the early Chicago School and German ordoliberalism.[2] Keynes’ writings are somewhat contradictory in this regard. He endorsed the bankmoney theory in his earlier writings, but fell back on the formula of “investment = savings” in his later *General Theory*. Under conditions of primary bankmoney creation the formula still applies to secondary credit among nonbanks, no longer, however, to primary bank credit.

[1] Cf. Financial Accounting Standards Board: FASB Accounting Standards Codification, Topic 305-2011, Cash and Cash Equivalents. The same in US GAAP (Generally Accepted Accounting Principles). For a critical assessment see Schemmann, 2012.
[2] Important contributions to the bank credit theory were made by Macleod, Withers, Hawtrey and Hahn, also by Schumpeter as well as von Mises.

Banks create credit and bankmoney whenever they make payments to nonbanks, for example when granting loans and overdrafts, or purchasing assets such as bonds, stocks, other securities or real estate, and also when paying salaries and bonuses to employees or nonbank service providers. However, the latter payments to employees or service providers are charged to the loss statement of a bank and thus its equity, whereas credit claims or securities are booked as assets.

Early private banknotes in Europe from the 1660s to the 19th century were promissory notes. The issuing bank promised the bearer to convert the note into silver coin anytime on demand. The banknote was a surrogate for the 'real thing' in the form of precious metal coin or gold bullion, until banknotes were made legal tender and the monopoly of a national central bank, thus the "real thing" by itself that consequently did not need gold coverage anymore and was finally taken off the gold standard.

In an analogous way, present-day bankmoney is a promissory ledger entry, in that the bank promises the customer either to cash out the bankmoney or to transfer the currency units to other bank accounts anytime on demand. Bankmoney is thus a claim of the customer on the bank, or the other way round, a liability of the bank to the customer. Bankmoney is a surrogate for solid cash and reserves.

It has now unfortunately become a drag on the further advancement of monetary theory that in various strands of post-Keynesianism today's credit and money creation in one act has been over-generalised into a doctrine of the alleged identity of money and credit, an identity that is seen as natural and functionally necessary.[3] This, in turn, has contributed to the strange phenomenon that many post-Keynesians are critical of financial capitalism, while at the same time standing up as fierce Banking School defenders of the present bankmoney regime, not recognising how that regime lies at the root of what they criticise.[4]

The fact that credit creation and money creation are done in one act today must not prevent us from recognising that money and credit are two different functions, albeit blended today. However, once bankmoney has been created, it circulates as an incoming and outgoing monetary asset only. Strictly speaking, credit creation creates but IOUs – which however we have adopted as the preferred means of payment. Credit does not create "money proper" as Keynes called it, such as precious metal money which, at source, did not involve credit and debt. That money was simply a monetary asset. It entered into circulation in that it was physically produced and then spent, not loaned. Rather than being an IOU in itself, it just facilitated payments, that is, the final settlement of a debt in financial and real transactions.

Moreover, one should be aware of the dual use of the word credit. It means (a) making a loan or financial investment, but then again it simply means (b) the accounting procedure of crediting/debiting some account, also figuratively speaking, for example, when students obtain credits for their exam achievements. The horizontal arrows in Figure 1 signify (a) credit creation adding to the stock of money. The thicker circular arrows signify (b) credit transfer, i.e. money circulation, not adding to the stock of money. Most crediting and debiting of accounts involves the circulation of already existing money, for example as earned income,

[3] For example, the so-called creditary economics of D. Bezemer and colleagues, which may have its merits in other respects: see Dyson, 2013, or R. Wray, 1998 and other MMTers, referring to Mitchell-Innes' wound-up credit theory of money from 1913/14 (Wray (ed.), 2004, pp. 14-78).
[4] For example, Dow, Johnsen & Montagnoli, 2015; Dyson, Hodgson & van Lerven, 2016 responding to Fontana & Sawyer, 2016.

sales proceeds, transfer payments, donations, etc. It is relatively rare that crediting an account coincides with extending bank credit (a bank granting a loan to a nonbank, or purchasing securities from nonbanks).

The money system is bank-led

Another term for reserves in the two-tier context is high-powered money. This is right and misleading at the same time. It is right because, in comparison with bankmoney, reserves are the money class of higher order and also the safer asset. However, this can be misleading if it obscures the fact that bankmoney is the dominant and decisive class of money today.

Within the present frame of split-circuit reserve banking, credit extension and money creation is bank-led. The initiative of money creation is with the banks, not with the central banks, as is most often assumed. It must be taken literally that central banks *re*-finance the banks, re-actively, upon or after the facts the banks have created beforehand. Central banks do not pre-finance the system by setting reserve positions first. The causation runs in the opposite direction. Central banks accommodate the banks' defining demand for central bank money (reserves and cash). This element was introduced into monetary economics by the accommodationist strand of post-Keynesianism.[5]

Through their pro-active lead in primary credit creation (bankmoney creation), banks determine the entire money supply, including the accommodating creation of reserves and cash by the central banks. Bankmoney is not the result of some sort of multiplication of central bank money. Quite to the contrary, the stock of central bank money is a follow-up quantity, a kind of sub-set of the stock of bankmoney.

Is bankmoney "endogenous", and are central banks "outside" the markets?

In post-Keynesianism bankmoney is considered endogenous, that is, created from within the economy according to demand, in contrast to exogenous money that is injected into the economy from the outside.[6] An analogous terminology thus distinguishes between inside and outside money.[7]

Although endogeneity of modern money can basically be endorsed, the distinction represents arbitrary labelling and contributes to mystification rather than clarification. If "exogenous" money has existed ever at all, it was the traditional metal money the supply of which depended on natural deposits of silver, gold and copper. Modern money, by contrast, consists of purely informational units, symbolic tokens, that are always created in response to economic needs and interests.

Furthermore, it is not just anybody "inside" the economy who can create their own money to use as a regular general means of payment. Only banks and central banks are relevant money creators, and to what extent the money they supply is endogenous or exogenous is

[5] Moore, 1988a, pp. 162-63; 1988b. The horizontal or accommodationist approach of post-Keynesianism became revised as the structuralist approach (Palley, 2013). The position contrasts with the verticalist view, which has it that central bank credit comes first. Also cf. Rochon, 1999a, pp. 155-201; 1999b; Keen, 2011, pp.309. Also Kydland & Prescott, 1990, have shown that the initiative is with the banks, not the central bank, and that the multiplier model thus is a myth.
[6] Cf. Moore, 1988a; 1988b; Rochon, 1999b; 1999a, pp. 15, 17, 155, 163,166. Rossi, 2007, p. 29, Keen 2011, p. 358. The notion of endogeneity of money goes back to Wicksell.
[7] Lagos, 2006; Roche, 2012.

actually open to debate. If endogeneity is understood as money creation upon market demand, automatically bringing about an optimal money supply, this reflects misleading Banking School doctrine. Modern fiat money can be created at discretion, "out of thin air", and market demand for money thus can be, and often is, excessive in self-reinforcing feedback dynamics of business cycles and financial cycles.

Both banks and central banks create credit and deposits in the same way, and both do it basically on market demand. The banking industry, however, does not just supply what is demanded. The banks supply bankmoney very selectively according to their own preferences. Ever more frequently they initiate business opportunities themselves, especially in investment banking. By contrast, the central banks today deliver the reserves as demanded by the banking sector; or, as needed in a banking and debt crisis to avoid pending bank insolvencies.

Consequently, if bankmoney is seen as endogenous in the economy, and banks as "inside the markets", the same must be said of central bank money. If central bank money is seen as exogenous to the economy and exerting control from the "outside", then this also applies to bankmoney.

Credit creation and balance sheet extension by cooperative bankmoney creation

The most widespread representation of bankmoney creation is by balance sheet extension of a single bank. According to this view, the respective bank makes a pairwise asset and liability entry on its balance sheet: on the asset side a credit claim on the customer securing their interest and redemption payments, and on the liability side an overnight liability to the same customer, obliging the bank to cash out or transfer the credits on demand of the customer.

This representation corresponds to an internationally widespread, but not uniform, accounting practice. Thus far, however, the representation does not really make sense. A customer does not take up money to keep it on account, but to make payments due – and as soon as the customer withdraws the bankmoney in cash or transfers the bankmoney to somewhere else, the balance sheet extension of the respective bank is reversed, in that the liability to the customer is closed out, and the cash account or reserves account of the bank is debited.

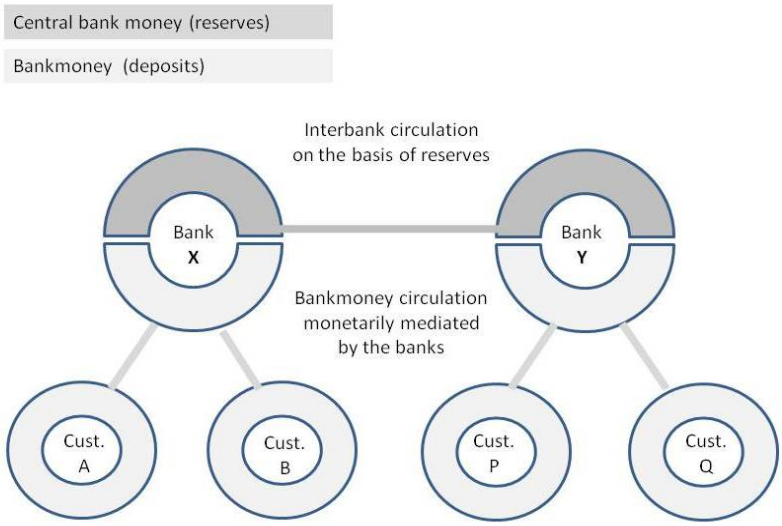
This reflects the fact that balance sheet extension by bankmoney creation is not an individual act by a single bank, but a cooperative process by many banks in the entire banking sector, in that a credit claim or other asset is added to the balance sheet of a credit-creating bank, while the related overnight liabilities (bankmoney) appear on the balance sheet of the recipient banks. All banks have to accept each other's liabilities transferred to them. Bankmoney creation could not otherwise work.

A balance sheet extension, both collectively and, in consequence, also individually, results from continued credit creation and mutual acceptance of bankmoney. The additional credit claims add to the balance sheet of banks as credit issuers to nonbanks, while the bankmoney liabilities add to the balance sheet of the banks as recipients of payments from the customers of other banks.

Bankmoney transfer via reserve transfer

Figure 2 below may help illustrate the transfer of money among banks and nonbanks. The figure again shows the separate circuits of interbank circulation (in darker shade above) and public circulation (in lighter shade below).

Fig. 2 Bankmoney transfer via reserve transfer



If, say, customer A at bank X wants to transfer bankmoney to customer Q at bank Y, this cannot be carried out directly by inter-customer transfer from A's account to Q's account. Instead, the transfer is carried out indirectly and involves the following steps:

- Bank X debits the current account of customer A, and
- transfers the respective amount in reserves to bank Y.
- Bank Y receives the reserves, and
- credits the amount of bankmoney into the account of customer Q.

The role of banks here is generally depicted as that of a trusted third party that carries out and documents payments among nonbanks. It can be seen this way, but can also be misunderstood as if a bank would transfer bankmoney, like cash, from A to Q. Such a transfer, however, takes place only as the transfer of reserves in the interbank circulation, where the central bank is the trusted third party that debits and credits the bank accounts on its balance sheet.

With regard to public circulation the process is somewhat different. Debiting the customer account at bank X actually means deleting the respective amount of bankmoney; while crediting the customer account at bank Y means re-crediting that amount. The banks are here in the role of active creators and extinguishers of bankmoney rather than just re-booking money on a single balance sheet. The process of bankmoney transfer may nevertheless be called a payment service or "intermediation" – on the understanding, though, that this refers to monetary, not financial intermediation, the latter being about the idea a bank would use its customers' bankmoney for making loans or purchases. In the split-circuit structure, however, a bank cannot use the customers' bankmoney for its own purposes, and a bank does not

need to do so, because a bank always creates credit by itself when it makes payments to nonbank customers, as it deletes credit when it receives payments from nonbank customers. In lieu of immediate payment in reserves, or shipping of cash historically, there has always been the practice of clearing of payments due to and from a bank, on an ongoing basis until further notice or some ceiling, or on a day-by-day basis, with final settlement of the resulting bottom line in reserves at the end of the day.

Today, most central banks and/or the banking industry run real-time gross-settlement (RTGS) payment systems. In the purest form, a payment order in such a system prompts an immediate debit from the reserve account of the remitting bank and a credit entry in the recipient reserve account. Other computerised payment systems involve immediate clearing of payments to and fro, so that the bottom line of each bank's payments is clear at any point in time, while the final settlement in reserves is carried out once a day.[8]

The situation is different again with payments among customers of a same bank. That bank is *vis-à-vis* its customers in the role of the trusted third party (rebooking deposits) as is the central bank *vis-à-vis* the banks (rebooking reserves). If, for example, in the figure above customer A of bank X wants to transfer bankmoney to customer B of the same bank, bank X simply debits the current account of customer A and credits the current account of customer B. Thus far, the bank neither needs cash nor reserves.

Were a bank to be huge and cover, say, half of all customers within a currency area, then about half of all cashless payments would be carried out by simple internal rebooking of overnight liabilities among the internal customers, without that bank needing central bank money and the central bank's cooperation. To a degree, this occurs in all banks, in large ones anyway, but even in small ones, and also in banks that participate in a banking union, pooling the participants' reserves. The latter practice is widespread among cooperative and municipal banks. Still, however, and despite the formation of banking oligopolies in many countries, the majority of domestic and international cashless payments include interbank transfers among different banks; and when transferring a customer deposit into an external account with another bank, the sending bank will need to have or obtain reserves which are transferred to the recipient bank. It applies nonetheless that the larger the bank, the more independent it is of central bank reserves. Clearly there is an ongoing concentration process in banking towards fewer and ever larger banks.

The banks' growing independence from central banks would be all the more pronounced by abolishing cash. Banks have to re-finance the bankmoney they create at only a small fraction, as explained below. The solid cash the banks still need, however, has to be financed in full since the banks have been stripped of their former privilege to issue private banknotes. The far more significant bankmoney privilege, by contrast, persists largely unimpaired and on an unprecedented vast scale.

[8] Examples of computerised payment systems include Fedwire = Federal Reserve Wire Network (USA, RTGS); CHIPS = Clearing House Interbank Payment System (USA, combines continual real-time clearing with daily final settlement in reserves); CHAPS = Clearing House Automated Payment System (UK, RTGS); TARGET2 = Trans-European Automated Real-Time Gross Settlement Express Transfer System (Euro/EZB); BoJ-Net = Bank of Japan Funds Transfer Network System; CLS = Continued Linked Settlement System, for international payments (combines, like CHIPS, clearing and final settlement).

Deletion of bankmoney and reserves

Deletion or extinction of credit money is simply the reverse process of its creation. As any payment from a bank to a nonbank creates bankmoney, any payment from a nonbank to a bank deletes bankmoney.

Consider the example of a bank's external nonbank borrower who pays interest and repays principal. In this case, the bankmoney is deleted, in that it is debited from the payer's account at the remitting bank, thus closed out, and not re-credited at the recipient bank. Instead, the recipient bank obtains the amount of interest and principal due from the payer's bank in reserves. The interest payment adds to the earnings account of the recipient bank, which in the end contributes to the bank's equity account. The repaid principal results in closing out, thus deleting, the respective credit claim on the customer. Closing out the liabilities there and the claims here represents a co-operative balance sheet reduction.

If the process is about an internal customer of a bank, this does not involve interbank transfer of reserves. Simply, the interest payment debits the customer account, deleting the bankmoney, and credits the bank's earnings account. The repayment of principal is reflected in the pairwise deletion of the bankmoney (liability) and the credit claim (asset), again representing a balance sheet reduction.

In the same way, any payment from the central bank to a bank creates reserves, as any repayment from a bank to the central bank deletes reserves.

De- and re-activation of bankmoney

Savings and time deposits are not involved in the processes discussed so far. What is that money used for? It isn't used at all, it is deactivated bankmoney, temporarily immobilised, so to speak, for a fixed period of time (maturity) or at notice. Deposit savings are not used, and actually cannot be used either by the banks or by the customers themselves as long as they are not re-activated by transferring them back to a current account.

Why then do customers run savings and time accounts? For customers, deposit savings represent a store of wealth, even if modest in most cases, a sort of near-term capital, also referred to as near-money, easy to reactivate if need be. Savings and time deposits can also serve as collateral. In normal times, savings yield the customers deposit interest, even if comparatively low.

But why do banks accept rather than deter interest-bearing deposit savings they cannot make use of? At first glance this seems to be an obsolete remnant from former times when the economy was largely cash-based and the banks needed their customers' cash deposits to fund the asset side of their business. Today, however, the banks need cash only for feeding the ATMs, which is no longer of relevance to a bank's lending and investment business.

The function of deposits has seamlessly changed. Today, deposits help maintain customer loyalty. Paying some deposit interest prevents the deposits from draining off to the competition. If this were to happen to a critical extent, the affected bank would face a severe liquidity problem, because more reserves would be going out without this being offset by incoming reserves. Borrowing the missing reserves on the interbank market would be costly, impairing a bank's business position. Offering savings and time accounts shields the banks

from such a liquidity risk, while allowing them to carry on with creating additional bankmoney at lending rates that are much higher than the deposit rates.

From a macroeconomic point of view, deactivated bankmoney, like all inactive money, does not contribute to effective demand, either to the money or capital supply on secondary credit markets. It thus has no effect on inflation and asset inflation. Even if deactivated bankmoney can hypothetically be reactivated within a short period of time, this does not happen in practice. Conversions of savings and time deposits into overnight deposits (liquid bankmoney) are normally more than compensated for by other customers converting overnight deposits into savings and time deposits. In times of crisis, too, most customers try to maintain rather than dissolve their deposit savings.

Fractional reserve banking and its operating conditions

The astounding thing about reserve banking now is that the quantity of reserves in interbank circulation is only a fraction of the amount of bankmoney in public circulation. In order to create and maintain an amount of deposits (bankmoney) in circulation, the banking sector needs central bank money of an amount that is only a small fraction of the credited bankmoney.

In the US, for example, that fraction is about 8.5% or less of the stock of bankmoney, comprising 1% cash for the ATMs, 0.1–0.5% liquid reserves (excess reserves) for the settlement of payments, and 10% obligatory minimum reserve minus the cash and further items. In the euro area, the fractional base of central bank money amounts to only about 2.5% of the stock of bankmoney, comprising 1.4% cash, 0.1–0.5% payment reserves, and 1% minimum reserve requirement.[9] Minimum reserves have been abolished in the countries of the British Commonwealth, Scandinavia (except euro-member Finland) and Mexico so that the remaining fractional base of excess reserves is tiny.[10]

How can it be that banks make do with such a very small base of central bank money, especially in view of pure RTGS systems where payments are not cleared before settlement, but settled real-time in full? There is no magic, just a number of operating conditions as follows.

Outflows equal inflows
Outgoing reserve payments of a bank are incoming reserves in other banks, so that the payments from and to the banks in the system are more or less offsetting each other. The reserves received can immediately be re-used in ongoing payments. In practice, the resulting payment balances represent some surplus or deficit, for example in international payments, and more likely to occur in smaller banks rather than large banking corporations where outgoing and incoming payments can offset one another even within minutes and seconds. Payment balances thus remain rather small and can easily be dealt with on the interbank money market as well as by the intraday-overdraft as it is provided in today's RTGS payment systems.

[9] Cf. Huber, 2017, pp. 71-72. Ryan-Collins, Greenham, Werner & Jackson, 2012, p. 75.
[10] Gray, 2011.

Cooperative bankmoney creation
This principle has already been touched upon. It means that the pace and rate of credit extension and bankmoney creation must take place roughly in step throughout the banking sector, so that the credit claims and liabilities of the banks largely grow in step with each other, and outgoing and incoming reserve payments do not result in significant imbalances. This, in turn, requires the banks to accept each other's transfers of deposits (bankmoney). In today's computerised payment systems mutual bankmoney acceptance can be taken for granted, in contrast to former cash-based economies when banks were often reluctant to accept the private banknotes of other banks.

Distributed transactions
Payments are spread over time and actors and do not include all of the actors' bankmoney at once. This means that only some part of the bankmoney is used at any point in time, with outgoing and incoming reserve payments largely offsetting one another.

Non-segregation of customer money
All outgoing and incoming payments of a bank are processed via one and the same operational central bank account of a bank, no matter whether the payments relate to customers or a bank's own dealings. Within the split-circuit system, the reserves related to a bank's own transfers and the reserves related to carrying out bankmoney transfers among customers cannot properly be distinguished. Given the fractionality of reserves, attempting to separate the reserves related to customer payments would not make much sense. Non-segregability of customer money is an additional advantage for the banks, amplifying the aforementioned conditions.

The combined effect of these conditions or mechanisms results in the frequency of reserve circulation in the interbank circuit being many times higher than is the case with bankmoney in public circulation. Put the other way round, the velocity of bankmoney circulation in the public circuit is many times slower than the high use frequency of reserves in the interbank circuit. This is the entire "trick" that enables the fractionality of reserve banking.

A bankmoney regime backed by the central banks and warranted by governments

Almost all schools of thought from both the neoclassical and Keynesian hemispheres of economics still depict the present money system as basically a sovereign currency system, with a mixed supply of sovereign money (central bank money) and bankmoney. Many bankers and academics even deny bankmoney creation, believing in deposits as a means of bank funding, and thinking that banks have to settle all expenses in full, not fully understanding the split-circuit nature of the reserve system and the different velocities of reserve and bankmoney circulation.

Moreover, as public circulation of bankmoney is tied to interbank circulation, and bankmoney thus still depends on a base, however small, of reserves and cash, this prompts most experts to misperceive the overriding importance of bankmoney, misunderstanding its status as a money surrogate as "subordinate" to cash and reserves. Solid cash, however, is now nearly irrelevant, and the reserves have become accommodatingly subservient to pro-active bankmoney creation. It is true that of the two classes of money, central bank money represents the safe and more reliable asset; and yet this is misleading, in that it conceals the fact that bankmoney is by far the dominant means of payment today, and that the pro-active

primary credit creation by the banks determines the entire money supply, including the central banks' re-active creation of reserves and cash.

In actual fact, the split-circuit bank-led reserve system is a bankmoney regime, based on fractional reserve banking backed by the central banks, the more so in times of crisis, and warranted by governments as if it were about sovereign money. Far from representing what would deserve to be called a sovereign money system, this is a state-backed rule of private bankmoney. Of the three sovereign monetary prerogatives – the currency (the national unit of account), the money (the means of payment) and the seigniorage (the gain from money creation) – the bankmoney regime has captured money creation and seigniorage-like benefits, especially in the form of refinancing costs avoided. Only the national monetary unit of account, the currency, is still defined by the state.

Old paradigms die hard

Against the background of the functioning of the money system discussed so far, it is easy to refute a number of old paradigms that have become obsolete or were misleading from the outset.

The piggy bank model

One of the oldest notions of banking, dating back to pre-modern cash economies, is the *piggy bank model.* This includes ideas such as "deposits are created by depositing cash" or "my money is in the bank". Today, the first variant is wrong and the second does not make sense. Even if using cash is still relatively widespread in some countries, cash is normally just about small transactions and is no longer the fundamental construction element of the monetary system it once was. Depositing cash today always means re-depositing it, in a subordinate follow-up transaction to primary bank credit and withdrawing some part of it in cash.

What, however, is in a bank account? Not "nothing" as die-hard believers in gold would have it, but not "the real thing" either, just a claim on it. The "real thing" is legal tender or central bank money, or sovereign money where the central bank has the legal status of a supreme monetary authority. In the split-circuit reserve system, however, the "real thing" – reserves on central bank account – remains a bank's possession that is never transferred into a customer account.

The loanable funds model of deposits

The loanable funds model conveys the idea that deposits are a means of bank funding and that "the banks are working with our money". This was certainly right throughout the metal age of money. It was an ongoing debate for centuries whether or not a bank's use of customer cash deposits – that is, banking on a fractional cash reserve – was "irregular" and fraudulent.[11]

However, with regard to *savings* accounts, their explicit purpose in former cash-based economies had always been to fund the banking business. Interest-bearing savings accounts were offered by Venetian bankers from the Middle Ages, regardless of the Church's ban on

[11] Huerta de Soto, 2009, chs. 1–3.

interest ("Venetian first, Christian second"). Other Italian banks by and by adopted the practice, and from around 1500 savings accounts spread across Europe, particularly when Pope Leo X officially lifted the canonical ban on interest. Leo, a Medici, was highly indebted, and the Vatican, too, needed funds for the continued construction of St Peter's Basilica. Fugger, at the time the preeminent banker of princes and cardinals, said the money could only be attracted by a 5% interest on deposit savings – and by selling papal indulgences, serving the redemption of Fugger's loans rather than of Christian souls.[12]

Today, deposits are no longer loanable funds, because in the split-circuit reserve system a bank cannot and need not make use of its customer deposits.[13] Rather than the banks using our money, we are using bankmoney. Cash remittances still fund cash withdrawals, but residually and on a diminishing scale. Among nonbanks, however, on secondary credit markets, bank deposits (bankmoney) are used as loanable funds among nonbanks further on, and on a large scale, for example by pension funds, investment trusts, private P2P lending platforms, or in crowdfunding.

The financial intermediation theory of banking

Since a bank in the split-circuit reserve system cannot pass on reserves to customers, and since a bank creates customer deposits, but cannot use the deposits thereafter for funding further banking business (for which they need a fractional base of reserves and residual cash), banks in the split-circuit reserve system can definitely not act as financial intermediaries. Assuming some such intermediation is presumably still the most widespread banking fallacy.

As deposits do not fund bank loans or other bank expenses, bank-financed lending and investment does not put in savings, but operates on self-created bankmoney. If there is a shortage of money or capital, this is not for monetary reasons. Modern money can be created anytime and, basically, at any amount if non-bank financial institutions, governments, firms and private households are willing to take up debt, and if the banks backed by the central banks are willing to extend credit. A modern bank is not a savings and loan association as, for example, cooperative mutual banks in the 19th century or building societies have formerly been. Present-day banks are monetary institutions, creators of primary credit and deposits, then de- and re-activators, and finally extinguishers of deposits (bankmoney).

By contrast, non-monetary or non-bank financial institutions are financial intermediaries indeed. They operate on bankmoney, taking it up from upstream savers and loaning or investing it downstream. Funds doing so may belong to a banking corporation, but this does not turn such funds into banks, as it does not turn the banks themselves into financial intermediaries. Referring to non-monetary financial institutions as shadow banks is most often confusing, except for money market funds whose shares are used since the 1980s especially on financial markets as a new deposit-like money surrogate, in fact a bankmoney surrogate.[14]

[12] Steinmetz, 2016 chs, 4, pp. 6-7.
[13] Also cf. McLeay, Radia & Thomas, 2014; Kumhof & Jacab, 2015; Werner, 2014 and 2015. Werner, though, fully right as he is with regard to the refutation of the loanable funds, multiplier and financial intermediation models, is misleadingly incomplete and partially wrong with regard to reserve circulation and his absolutized view of credit creation "out of nothing". A reservation regarding McLeay, Radia & Thomas relates to their still maintained belief in transmission of central bank policies to the banks and the economy by setting base rates. The same reservation applies to Bundesbank 2017.
[14] McMillan, 2014, pp. 54-79; Baba, McCauley & Ramaswamy, 2009, p. 68; Hilton, 2004, p. 180.

As a consequence for macroeconomics, the central identity of "investment equals savings" is in need of overhaul. It is still applicable to secondary credit markets among nonbanks, supposing that macroeconomics would finally be capable of introducing a distinction between GDP-contributing and non-GDP financial investment.[15] As far as banks are involved, however, "investment equals savings" does not apply because deposit savings are deactivated and thus irrelevant for investment or other expenditure, while the banks can create new additional bankmoney for funding any purpose at their own discretion.

The credit multiplier, the reserve position doctrine and other fictitious elements of monetary policy

Although rendered obsolete by the development of bank-led reserve banking, the multiplier model can still be found in most textbooks in numerous variants. Their common key feature is an amount of money (M) that is credited (Cr) to customers. The money is supposed to flow back to the banks, where it is thought to be re-used to extend credit to customers again, and so on. At each round, the banks retain a certain reserve rate of the money (Res). The amount of extendable credit thus is Cr = M (1-Res), and the total extendable amount of credit is Cr = M/Res.

This is a nice example of the proverbial "garbage in, garbage out" as it can occur in modelling.[16] The multiplier model does not include a distinction between cash, reserves and bankmoney. The model thus either starts from the misleading idea of a cash economy, or it wrongly considers customer deposits to be loanable funds and the banks to be financial intermediaries. The model does not take into account how the alleged creditary money recycling of the banks connects to the non-credit-creating circulation of the money.

Most importantly, there is neither addition to the money supply, nor deactivation and deletion of money. The model presupposes the amount of money as an exogenously pre-existing and basically invariable quantity, rather than building on an endogenous and variable money supply. The model puts the stock of money first, as this corresponds with the reserve position doctrine of monetary policy transmission, and the banks' credit creation second, rather than the reverse, which is actually what applies in the bankmoney regime. The fact that the split-circuit money system is bankmoney-led and re-actively accommodated by the central banks to only a small fraction invalidates most conventional wisdom on monetary policy and supposed transmission mechanisms.

For one thing, this applies to the reserve position doctrine, that is, monetary policy by setting reserve positions.[17] The doctrine has it that a pre-set quantity of reserves, in combination with the alleged multiplier mechanism, allows exerting control over the banks' credit extension and bankmoney creation. The doctrine may have had a point in the times of majority cash economies with a central bank monopoly on cash. Over the last century, however, the situation has completely changed. No wonder that conventional quantity policies failed, particularly the British gold standard in the period from 1914-1931, the US gold standard from

[15] The approach to subdividing equations of circulation into a real-economic and financial hemisphere, the latter consisting of GDP-contributing and non-contributing financial transactions, has been put forth by Huber (1998, p. 224) and Werner (2005, p. 185). In a similar attempt, Hudson 2006 has introduced the FIRE sector (Finance, Insurance, Real Estate) into his macroeconomic model.
[16] For a critique of the multiplier model also see Werner, 2005, p.191; Keen, 2011, pp. 306–312; Ryan-Collins, Greenham, Werner & Jackson, 2012, pp. 16-25.
[17] Bindseil, 2004.

1944-1971, and the subsequent failure of monetarist policies in the 1970-80s. Since then, central banks do no longer intend to exert control over the stock of money, including their own creation of reserves and cash.

The main lever of reserve-position policy is the minimum reserve requirement, as still imposed by most central banks. How is that transmission supposed to work if the banks have the pro-active lead and the central banks promptly provide the excess reserves the banks are demanding and the minimum reserves the banks are required to hold? From this angle, minimum reserves are utterly pointless, except for the related interest earnings of a central bank from lending the reserves to the banks when things are running normally.

Believers in reserve policies still think the price of the reserves will do the trick. Central banks thus completely changed over to quantity-disconnected interest rate policies, supposed to influence the real-economic inflation rate by way of some other nebulous transmission mechanism. The reserve position doctrine has been replaced with the short-term interest rate doctrine, that is, monetary policy by setting central-bank refinancing rates, in the eurozone misleadingly still referred to as “lead rates”. However, causation from central bank rates to the banks’ credit and bankmoney creation is not discernible, nor causation from central bank rates to the general inflation rate, only some limited and weak effect on the banks’ lending rates.

How should the interest on a 2.5% or 8.5% base of central bank money exert decisive control over the 100% of bankmoney, *all* of which can be issued at lending rates much higher than the base rate on only a small fraction of the bankmoney? This is all the more questionable as bankmoney creation is pro-active and the banks’ residual demand for central bank money is price-inelastic once the bankmoney facts have been created, at least in the short run. A supposed feedback effect in the longer run, though, is not discernible either.

The only interest rates a central bank can definitely set are its own rates on the small fraction of reserves. Central banks can also effectively influence interbank rates by expanding or reducing the fractional reserve base. The effect of this, however, is low for the same reasons as mentioned above. Thus, the transmission from interbank rates to public rates and bankmoney creation is not evident – more directly speaking, largely fictitious – at least in terms of market economics.

What central banks can effectively do in the split-circuit reserve system is create reserves by monetising debt. And this is what they are doing, all the more in times of banking and debt crises, presently by the insolvency-deferring policies of quantitative easing for banks and nonbank financial institutions, near-zero interest rates supporting heavily over-indebted governments, and negative interest rates. The latter, at the beginning, are debiting the banks, but if imposed on the banks’ customers, negative interest would burden the customers to the banks’ relief. Such policies not only support shaky banks and over-indebted public households, but also the non BIP-contributing stocks of financial capital, while burdening useful deposits and savings, overall adding still further to what is already too high (debt) or too low (real interest rates).

If there is no effective transmission from central bank to banks anymore, it follows that the conventional toolset of monetary policy does not really influence the economy and inflation rates in the direction of stated goals. However, by monetising vast amounts of financial debt, central banks contribute to asset inflation, particularly in stocks, commercial real estate,

private housing, commodities and other financially overstrained real assets, including derivatives on these underlyings.

If an effect of base rate policy on interest rates and bankmoney creation can clearly be identified, that effect is due to voluntary price administration by the banks. Quite a few banks tie certain lending rates – for example, for overdrafts or mortgages – to interbank rates such as the Fed Funds rate (dollar), LIBOR (pound) or EONIA (euro). Ironically, this is an exercise in corporatist planning rather than the market-borne price dynamics of supply and demand.

In view of the nearly vanished effectiveness of conventional monetary policy, forward guidance has become a preferred "tool" of monetary policy. Forward guidance is about information and communication rather than technical measures to be taken. A central bank notifies the public of what it expects in the near and more distant future in terms of interest rates, inflation, growth, employment etc., and how it might react. Central bank statistics are certainly among the best available. But why would central banks have more foresight than other agencies? Isn't this just a modern variety of augury? Even if many financial and economic actors believe in it, why would the augurs' divinations become self-fulfilling?

Lost control. Dysfunctions of the bankmoney regime

Quite a few economists consider the bankmoney regime a sophisticated and neutral system. It is assumed to work, and to do so without noteworthy effects on finance and the economy. But assuming the monetary system to be "neutral" is a rather strange doctrine of neoclassical economics – the more so in view of the effective exercise of power related to the creation, allocation and distribution of money, a power that shapes markets and lives and is in no way inferior to the legal authority to issue directives.

In actual fact, the bankmoney regime is not "neutral" and shows a number of recurrent and severe problems. In a sense one might say the bankmoney regime works "too well" in that it recurrently provides the fuel that reproduces that bipolar syndrome of boom and bust, manias and crashes, in recent decades more in financial cycles than business cycles. The dysfunctions of the bankmoney regime cannot be discussed in detail within the limited scope of this paper, but shall at least be touched upon so as to gain an understanding of the matter beyond the merely technical functioning of the present money system.

The key problem is the out-of-control creation of bankmoney. With the general spreading of cashless payment practices, central banks have lost control over money creation, and since about 1980 they have entirely given up some such ambition – as if the stock of money was not relevant anymore, and as if short-term central bank interest rates could significantly influence a complex result of the economic process as is the inflation rate.

In addition, money and capital markets permanently fail to reach what might be a self-limiting equilibrium. In any market, there is a demand curve and a supply curve working in opposite directions. This applies in money and capital markets too, but at the same time there is a positive feedback loop superimposing itself on the negative one: on balance, rising asset prices and asset volumes attract additional demand rather than deterring it.[18] This results in

[18] Also cf. Shiller, 2015, p. 225.

recurrent market failure in the form of overshooting money supply, over-investment and over-indebtedness, asset inflation, bubbles and crises.

To a degree, this applies to real business cycles, too. Even if overheating business cycles and inflation have now faded into the background since the take-off of the Great Financial Immoderation around 1980, inflation actually represents between a half and two thirds of the now modest growth of nominal GDP.

An often-quoted IMF study has identified 425 systemic financial crises from 1970 to 2007 in migratory hot spots around the world, intensifying in number and severity. Of these, 145 were sector-wide banking crises, 208 currency crises and 72 sovereign debt crises.[19] The bankmoney regime is not the only cause of such crises, but finance is conditioned by the monetary system, as the economy is conditioned by finance.

Many economists belittle the relevance of the monetary system, while at the same time putting hopes in present-day monetary policies for stabilising finance and stimulating the economy – apparently being unaware of the contradiction between the belittlement and the hopes. In contrast to preceding stages in the development of modern economies, the fundamental role of the monetary system is nowadays largely neglected. In view of the increased weight of banking and financialisation this is paradoxical.

Banks have a strong incentive to expand their balance sheets – that is, creating as much primary credit and bankmoney, also for proprietary purposes, as they dare to risk and can fractionally refinance. This is strongly facilitated by the banks lead in creating the money on which they operate. In consequence, and on balance of the ups and downs, there is over-expansion of the banking sector's balance sheet and overshooting primary bankmoney creation, also feeding overshooting of secondary financial intermediation.[20]

In any crisis of a bank, the more so in a systemic banking crisis, it becomes apparent that bankmoney is unsafe. Bankmoney is but a promissory credit entry on a bank's balance sheet, not the "real thing" that would be the safe possession of the customers. If banks fail, the positions in their balance sheet are largely nullified, making the bankmoney disappear and bringing money circulation and the entire economy to a corresponding standstill. Interbank deposit insurance (just a fig leaf) and government warranty (never tested) are simply proof of the non-safety of bankmoney.

Furthermore, a continued increase in financial assets in disproportion to GDP results in a biased income distribution.[21] The reason is that a growing share of current income and additional debt has to be devoted to servicing the claims of financial assets, disproportionately adding to financial income, which reduces the share of earned income, which in turn impairs aggregate demand and real output.

[19] Laeven & Valencia, 2008.
[20] For the criteria and empirical data of overshooting bankmoney supply see Huber, 2017, pp. 109-112. For related credit and debt bubbles see Shiller, 2015, pp. 70-97; Minsky, 1982; 1986, p.206, p. 218, p.223, p.294; Jordá, Schularick & Taylor, 2010; Schularick & Taylor, 2009.
[21] Inequality of income and wealth increasing again since about 1980 has been identified by many studies of late, among these Atkinson, Piketty & Saez, 2011; Atkinson, 2015; Piketty, 2013.

Finally, it has to be recalled that money is a creature of the legal system. In present-day economics this is much less reflected than was the case one to two centuries ago.[22] Control of the currency and money creation and benefitting from the seigniorage thereof are sovereign prerogatives – of constitutional importance and of the same rank as the prerogatives of legislation, public administration, jurisdiction, taxation and the use of force. Without these legal prerogatives, or monopolies respectively, a modern nation-state lacks sovereignty and functional capability, including, where applicable, the power to maintain the liberal rule of law.

Against this background, the present bankmoney privilege of the banking industry represents the illegitimate seizure of the sovereign prerogatives of money creation and seigniorage, rationalised by Banking School doctrine for 200 years, including untenable postulations such as the private-compact and market theory of money, and the false identity of money and bank credit.

The bankmoney regime pushes over-investment and over-indebtedness more pervasively than previous monetary regimes. But the banks and the financial industry cannot outsmart the gravitational force of productivity and economic output. Put differently, they cannot artificially extend the financial carrying capacity of the economy at a particular point in time. In ecology, carrying capacity means an ecosystem's capacity to provide and reproduce resources and sinks for a specified population. By analogy, an economy can carry only a limited volume of claims on income or economic output at any point in time. Financial carrying capacity relates to the sustainable levels of assets and debt as a ratio of GDP and additional non-Ponzi debt, notwithstanding the potential for extending such limits by sustainable productivity gains on the basis of innovation and structural change. If the carrying capacity of an ecosystem is overburdened, it breaks down or dies off; if the financial carrying capacity of an economy is overburdened, it crashes or declines.

The perspective: a single-circuit sovereign money system

If the bankmoney-led split-circuit reserve system is at the root of the problems discussed above, the solution is a single-circuit sovereign money system.[23] Sovereign money is legal tender, in most cases issued by the central bank of a nation-state or community of nation-states. Today, coins and central bank notes are sovereign money as well as central bank reserves, not, however, bankmoney.[24] In a sovereign money system, the customers' money-on-account will be sovereign central-bank money too, circulating among the public and the banks alike. This is a move beyond reserve banking, where reserves (central bank money-on-account) are reserved for the banks and withhold from the public. A single-circuit system thus is not another variety of 100% reserve banking in a split-circuit system.

[22] Money and the monetary system as a matter of public law and an institutional arrangement controlled by the government was a key feature in the British Currency School of the 1820-40s, the state theory of money around 1900 (Knapp, 1905[1924]) and Keynesianism since the 1940-50s (Lerner, 1943; 1947). The teaching dates back via Medieval Thomism to Aristotle: "Money exists not by nature but by law" (Aristotle*, Ethics*, 1133a30).

[23] There is a growing number of publications with the same or similar analyses and conclusions, for example, Huber & Robertson, 2000; Zarlenga, 2002; 2014; Jackson & Dyson, 2012; Benes & Kumhof, 2012; Yamaguchi, 2014; Positive Money, 2014; Sigurjonsson, 2015; KPMG Iceland, 2016; Huber, 2017.

[24] The question of what qualifies as sovereign money is less clear in countries where the central bank is not a state authority or an inter-governmental body as is the case with the ECB, but still a private joint-stock company or a private-public hybrid as is the case with the Federal Reserve of the US.

From a technical point of view, the transition from split-circuit bankmoney to single-circuit sovereign money can be achieved by converting demand deposits into central bank money and taking the respective accounts off the banks' balance sheet, enabling the direct money transfer between customers, without monetary intermediation by the banks. This then results in the separation of money and payment services from the banks' lending and investment business – which is a key feature of chartalist Currency School teaching.

Alternatively, a more gradual transition process could be triggered by introducing a new type of account for nonbank customers, that is, separate sovereign-money accounts, existing side by side with bank giro accounts. The new accounts can be managed in trust by banks and other payment service providers, similar to securities accounts *off* a provider's balance sheet. The entries in such accounts would represent central bank money and be the safe possession of the customers.

Something similar and potentially even more disruptive might soon be achieved by introducing central bank digital currency (CBDC) on the basis of distributed ledgers in a blockchain, set up and run by central banks as the monetary authority of a currency area.[25] Digital currency is a new class of money. CBDC would exist in addition and actually in competition not only to Bitcoin and Altcoins, but also, and more importantly, to bankmoney-on-account as well as digital currency possibly issued by the banking industry itself.

CBDC can be seen as a modern equivalent to traditional solid cash, while being cheaper and easier to handle than both solid cash and money-on-account. Making a payment in digital currency represents the direct transfer of an amount from digital "wallet to wallet", analogous to the traditional hand-to-hand circulation of solid cash. There is no need for a trusted third party and no counterparty risk, at least not in the sense of monetary intermediation as is the case with money-on-account.

Central bank money is still of a higher order and the safer and more trustworthy asset compared with private monies. Nonbank money users can thus be assumed to have a preference for central bank money-on-account or, alternatively, CBDC. A shift from bankmoney to central bank money, be it sovereign money-on-account or CBDC, would be tantamount to a monetary tide change, reversing the transsecular trend from central bank money (traditional cash) to bankmoney-on-account in favour of re-expanding the share of central bank money (sovereign money-on-account and CBDC). This would bring about a corresponding increase in the effectiveness of central bank monetary policies, be it re-enabled quantity policy or much enhanced interest rate policy.

A number of modellings – based on different approaches such as DSGE, system dynamics and stock-flow-analysis – came up with basically convergent findings. Sovereign money would not only be safe, but also bring about significantly more financial and economic stability, non-volatile normal interest rates, low inflation, financial assets and debt not growing in disproportion to GDP, and a higher level of output and employment.[26]

Sovereign money is about re-nationalising money creation and seigniorage – not, however, nationalising banking and the uses of money. Thus, money creation and control of the stock

[25] Recent contributions to central bank digital currency include Andolfatto, 2015; BIS, 2015; Broadbent, 2016; Barrdear & Kumhof, 2016.
[26] Yamaguchi, 2012; Benes & Kumhof, 2012; Lainà, 2015; van Egmont & de Vries, 2015.

of money would be made the sole responsibility of the central banks, provided these are national monetary authorities. Like the judiciary, central banks need to be independent and impartial, in fact representing a fourth branch of the state. A sound sovereign money system includes a thorough separation of monetary and fiscal state powers, and of both from banking and wider financial market functions.

The central banks in such a system would pursue discretionary and flexible monetary policies on the basis of a redefined legal mandate. The latter would include restated policy objectives and indicators such as comprehensive factor employment, interest rates, foreign exchange rates, inflation, asset inflation, asset and debt bubbles as well as financial assets-to-GDP ratios.

This does not mean overloading central banks and expecting too much from monetary policy. With sovereign money, too, the money system is a most important foundation of finance and the economy, not however a magic force that could attain desired goals by itself. But monetary policy can of course contribute much for the better or the worse. In a full-blown single-circuit sovereign money system central banks would have control of the stock of money. Their policies, unlike today, would thus be directly and fully effective.

References

Andolfatto, David. 2015. “Fedcoin – on the desirability of a government cryptocurrency”, *Macromania*, 3 Feb 2015, http://andolfatto.blogspot.de/2015/02/fedcoin-on-desirability-of-government.html.

Atkinson, Anthony B. / Piketty, Thomas / Saez, Emmanual. 2011.” Top Incomes in the Long Run of History”, *Journal of Economic Literature, 2011, 49:1, 3–71.*

Atkinson, Anthony B. 2015. *Inequality. What can be done?* Cambridge, MA: Harvard University Press.

Baba, Naohiko / McCauley, Robert N. / Ramaswamy, Srichander. 2009. “US dollar money market funds and non-US banks”, *BIS Quarterly Review*, March 2009, 65–81.

Barrdear, John / Kumhof, Michael. 2016. “The macroeconomics of central bank issued digital currencies, Bank of England”, *Staff Research Paper* No. 605, July 2016.

Benes, Jaromir / Kumhof, Michael. 2012. The Chicago Plan Revisited, *IMF Working Paper* WP/12/202. A revised draft dates from 12 February 2013.

Bindseil, Ulrich. 2004. “The Operational Target of Monetary Policy and the Rise and Fall of Reserve Position Doctrine”, *ECB Working Paper Series*, No. 372, June 2004.

BIS. 2015. *Digital currencies*, prep. by the BIS Committee on Payments and Market Infrastructures, Basel: Bank for International Settlements.

Broadbent, Ben. 2016. *Central banks and digital currencies*, http://www.bankofengland.co.uk/publications/Pages/speeches/2016/886.aspx.

Bundesbank. 2017. The role of banks, nonbanks and the central bank in the money creation process, *Deutsche Bundesbank Monthly Report*, Vol.69, No.4, April 2017, 13–30.

Dow, Sheila / Johnsen, Guðrún / Montagnoli, Alberto. 2015. “A critique of full reserve banking”, *Sheffield Economic Research Paper Series*, no. 2015008, March 2015.

Dyson, Ben / Hodgson, Graham / van Lerven, Frank. 2016. “A response to Critiques of ‘Full Reserve Banking’ or ‘Sovereign Money’ Proposals”, *Cambridge Journal of Economics*, Vol. 40 (5) Sep 2016, 1351–1361.

Dyson, Ben. 2013. Dirk Bezemer on Positive Money, http://positivemoney.org/2013/04/dirk-bezemer-on-positive-money-a-response/.

Fontana, Giuseppe / Sawyer, Malcolm. 2016. “More Cranks than Brave Heretics”**,** *Cambridge Journal of Economics*, Vol. 40 (5) Sep 2016, 1333–1350.

Gray, Simon. 2011. "Central Bank Balances and Reserve Requirements", *IMF Working Paper* WP/11/36, 55.

Hilton, Adrian. 2004. "Sterling money market funds, Bank of England", *Quarterly Bulletin*, Summer 2004, 176–182.

Huber, Joseph / Robertson, James. 2000. *Creating New Money*, London: New Economics Foundation.

Huber, Joseph 2017*. Sovereign Money. Beyond Reserve Banking*, London: Palgrave.

Hudson, Michael. 2006. "Saving, Asset-Price Inflation, and Debt-Induced Deflation", in Wray, L. Randall / Forstater, Matthew (eds.). 2006. *Money, Financial Instability and Stabilization Policy*, Cheltenham: Edward Elgar, 104–124.

Huerta de Soto, Jesús. 2009. *Money, Bank Credit, and Economic Cycles*, Auburn, Al.: Ludwig von Mises Institute, 2nd edition.

Jackson, Andrew / Dyson, Ben. 2012. *Modernising Money. Why our monetary system is broken and how it can be fixed*, London: Positive Money.

Jordà, Òscar / Schularick, Moritz / Taylor, Alan M. 2010. "Financial Crises, Credit Booms, and External Imbalances. 140 Years of Lessons", *NBER Working Papers*, No. 16567, Dec 2010.

Keen, Steve. 2011. *Debunking Economics*, London/New York: Zed Books.

Knapp, Georg Friedrich. 1905. *Staatliche Theorie des Geldes*, Leipzig: Duncker & Humblot. Engl. 1924. *The State Theory of Money*, London: Macmillan.

KPMG. 2016. *Money Issuance - Alternative Monetary Systems*, authored by Svanbjörn Thoroddsen and Sigurvin B. Sigurjonsson (KPMG), commissioned by the Icelandic Prime Minister, Reykjavik.

Kumhof, Michael / Jacab, Zoltan. 2015. "Banks are not intermediaries of loanable funds — and why this matters", *Bank of England Working Paper*, No. 529, May 2015.

Kydland, Finn E. / Prescott, Edward C. 1990. "Business Cycles: Real Facts and a Monetary Myth", *Federal Reserve of Minneapolis Quarterly Review*, Spring 1990, 3–18.

Laeven, Luc / Valencia, Fabian. 2008. "Systemic Banking Crises. A New Database", *IMF Working Paper* WP 08/224.

Lagos, Ricardo. 2006. "Inside and Outside Money, Federal Reserve Bank of Minneapolis", *Research Department Staff Report*, No. 374, May 2006.

Lainà, Patrizio. 2015. "Money Creation under Full-Reserve Banking: A Stock-Flow Consistent Model", *Levy Economics Institute Working Papers,* No. 851, Oct 2015.

Lerner, Abba P. 1943. "Functional Finance and the Federal Debt", in: Colander, David C. (ed.), *Selected Economic Writings of Abba P. Lerner*, New York University Press 1983, 297–310. First publ. in *Social Research*, Vol.10 (1943) 38–51.

Lerner, Abba P. 1947. "Money as a Creature of the State", *American Economic Review*, Vol.37, No.2, May 1947, 312–317.

McLeay, Michael / Radia, Amar / Thomas, Ryland. 2014. "Money creation in the modern economy", *Bank of England Quarterly Bulletin*, 2014 Q1, 14–27.

McMillan, Jonathan 2014. *The End of Banking. Money, Credit, and the Digital Revolution*, Zurich: Zero/One Economics.

Minsky, Hyman P. 1982. "The Financial Instability Hypothesis", in: Kindleberger, C.P./ Laffargue, J.-P. (Eds.). *Financial Crises. Theory, History, and Policy*, Cambridge University Press, 13–39.

Minsky, Hyman P. 1986. *Stabilizing an Unstable Economy*, New Haven: Yale University Press.

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

Moore, Basil J. 1988b. "The Endogenous Money Supply", *Journal of Post Keynesian Economics*, vol.10, no.3, 1988, 372–385.

Palley, Thomas I. 2013. "Horizontalists, Verticalists, and Structuralists: The Theory of Endogenous Money Reassessed", *Review of Keynesian Economics*, Vol. 1(4) 2013, 406–424.

Piketty, Thomas. 2013. *Le capital au XXI siècle*, Paris: Éditions du Seuil. Engl.

Piketty, Thomas. 2014: *Capital in the Twenty-First Century*, Cambridge, MA: President and Fellows of Harvard College.

Positive Money. 2014. *Creating a Sovereign Monetary System*, London: Positive Money.

Roche, Cullen. 2012. "Understanding Inside Money and Outside Money", *Pragmatic Capitalism,* www.pragcap.com/understanding-inside-money-and-outside-money.

Rochon, Louis-Philippe. 1999a. *Credit, Money and Production. An Alternative Post-Keynesian Approach*, Cheltenham: Edward Elgar.

Rochon, Louis-Philippe. 1999b. "The Creation and Circulation of Endogenous Money", *Journal of Economic Issues*, vol.33, no.1, 1999, 1–21.

Rossi, Sergio. 2007. *Money and Payments in Theory and Practice*, London/New York: Routledge.

Ryan-Collins, Josh / Greenham, Tony / Werner, Richard / Jackson, Andrew. 2012. *Where Does Money Come From? A guide to the UK monetary and banking system*, 2nd edition, London: New Economics Foundation.

Schemmann, Michael. 2012. *Accounting Perversion in Bank Financial Statements. The Root Cause of Financial Crises*, IICPA Publications.

Schularick, Moritz / Taylor, Alan M. 2009. "Credit Booms Gone Bust: Monetary Policy, Leverage Cycles, and Financial Crises 1780–2008", *American Economic Review*, 102 (2) 1029–1061.

Shiller, Robert J. 2015. *Irrational Exuberance*, revised and expanded 3rd edition, Princeton NJ: Princeton University Press.

Sigurjonsson, Frosti. 2015. *Monetary Reform. A better monetary system for Iceland,* Icelandic Parliament, submitted to the Committee on Economic Affairs and Trade.

Steinmetz, Greg. 2016. *The Richest Man Who Ever Lived. The Life and Times of Jacob Fugger*, New York: Simon & Schuster.

van Egmont, Klaas / de Vries, Bert. 2015. "Dynamics of a sustainable financial-economic system", *Sustainable Finance Lab Working Paper*, March 2015, Utrecht University.

Werner, Richard A. 2005. *New Paradigm in Macroeconomics*, London: Palgrave.

Werner, Richard A. 2014. "Can banks individually create money out of nothing? Theories and empirical evidence", *International Review of Financial Analysis* 36 (2014) 1–19.

Werner, Richard A. 2015. "A lost century in economics - three theories of banking and the conclusive evidence", *International Review of Financial Analysis*, Vol. 46, July 2016, 361–379.

Wray, L. Randall. 1998. *Understanding Modern Money*, Cheltenham, UK: Edward Elgar.

Yamaguchi, Kaoru. 2012. "On the Monetary and Financial Stability under A Public Money System", *Working Paper Doshisha University Kyoto*. http://bs.doshisha.ac.jp/attach/ page/BUSINESS-PAGE-JA-53/26829/file/DBS12-02.pdf.

Yamaguchi, Kaoru. 2014. *Money and Macroeconomic Dynamics. An Accounting System Dynamics Approach*, Awaji Island: Muratopia Institute/Japan Futures Research Center.

Zarlenga, Stephen A. 2002. *The Lost Science of Money. The Mythology of Money – the Story of Power*, Valatie, NY: American Monetary Institute.

Zarlenga, Stephen A. 2014. *The Need for Monetary Reform. Presenting the American Monetary NEED Act*, Valatie, N.Y.: American Monetary Institute.

Author contact: huber@soziologie.uni-halle.de

SUGGESTED CITATION:
Joseph Huber, "Split-circuit reserve banking – functioning, dysfunctions and future perspectives", *real-world economics review*, issue no. 80, 26 June 2017, pp. 63-84, http://www.paecon.net/PAEReview/issue80/Huber80.pdf

The coming revolution in political economy: money creation, Mankiw and misguided macroeconomics

Tim Di Muzio and Leonie Noble [University of Wollongong, NSW, Australia]

It is reputed that the German theoretical physicist and Nobel Prize winner Max Planck once said that: "a new scientific truth does not triumph by convincing its opponents and making them see the light, but rather because its opponents eventually die".[1] Whether this is a completely accurate assessment of scientific affairs can be debated, but what cannot is the fact that disciplines cling on to old beliefs as though they are steadfast and timeless truths. These "truths" are taken for granted and passed down from generation to generation unless they are sufficiently challenged with new and convincing evidence that demonstrates that things are indeed otherwise. Our aim in this article is to challenge one of the principal received truths in the field of Economics: the way that new money is created. We also aim to go further and argue that a proper understanding of how new money is created has such devastating consequences that it heralds no less than a coming revolution in how we understand political economy and future possibilities. Our main argument is that the received truth of the fractional reserve theory or "money multiplier" model taught in most economics textbooks cannot explain the expansion of the money supply by logic, simple math and by basic bank accounting practices. To demonstrate our argument we will rely on Mankiw's *Macroeconomics* and his chapter explaining to students how new money is created. We use Professor Mankiw's textbook because of its widespread use as an introductory text to macroeconomics. Our second argument is that the vast majority of new money is actually created by banks when they make loans. This has long been argued by heterodox economists who call this form of money creation "endogenous money". However, while other scholars have addressed the inaccuracy of the "money multiplier" model based on statements made by central bankers, there has been no demonstration of how the "money multiplier" model itself is logically and mathematically false.[2] To complicate matters, much of the heterodox literature that theorizes endogenous money has largely failed to develop a comprehensive critique of why commercial banks issuing new money as loans might be extremely detrimental to any political community and very favourable to a minority of dominant bank owners (but see Hudson, 2015; Huber and Robertson, 2000; Huber, 2014).[3] In

[1] Apple, Sam (2016) "An Old Idea Revived: Starve Cancer to Death" *New York Times*, May 12.
[2] Keen's *Debunking Economics* (2011: 306-314) provides an exposition of the limitations of neoclassical economics. While Keen addresses the myth of the "money multiplier" model he focuses on the available evidence which suggests that money is not created in this manner. This article will further explore the myth of the "money multiplier" model by systematically exploring the flaws within the model itself which make it an impractical, improbable and mathematically inaccurate method for expanding the money supply. Positive Money, a UK based not-for-profit research organization also disavows the money multiplier model as a myth. However, like Keen, they also base their critique on statements largely made by central bankers. This is of course admirable, but in this article we prefer to follow the math and accounting practices rather than rely solely on qualitative statements made by central bankers, however important these may be. http://positivemoney.org/how-money-works/advanced/the-money-%20multiplier-and-other-myths-about-banking/ (Accessed 14/10/2016).
[3] We have not had the chance to read Joseph Huber's (2016) *Sovereign Money: Beyond Reserve Banking*. (Basingstoke: Palgrave Macmillan), which was in press when we were writing this article. See also the debate on endogenous money in the *Review of Keynesian Economics*, Volume 2, Issue 3, 2014. Though there are precursors, the idea of endogenous money in the post-Keynesian tradition can be traced back to Basil Moore (1979) and is commonly held by heterodox post-Keynesians. But this

order to provide evidence for these arguments, we have organized this article in the following way. First, we provide a brief historical overview of how the money supply started to expand in England before moving on to discuss the two main theories of new money creation. Second, we focus closely on the arguments made by Mankiw as he explains the "money multiplier effect". We will demonstrate that it is incorrect on three levels: 1) by logic; 2) by simple math; and 3) by accounting standards. In the third section of this article, we will discuss how new money actually enters an economy and the leading consequences this act has for our political economies. But before we proceed, however, a few definitional issues must be addressed. First, while there are many possible definitions for the term "money" our understanding of money is that it is an abstract claim on society and natural resources measured in a unit of account (e.g. dollars, euros) and generally accepted for payment (see Ingham, 2004: 47-56). Second, while money can be represented by physical objects (e.g. notes, coins, cattle) and even takes the form of ownership claims over income-generating assets (e.g. stocks and bonds), we should not confuse the material representations of money with the concept of money itself. Last, while there are some national differences in how the money supply is counted, the money supply is typically subdivided based on a scale of liquidity – from most liquid (e.g. physical notes and coins) to less liquid (e.g. time deposits). With these definitional issues addressed, we can now turn to our political economy of money creation.

Historical overview and theories of money creation

Despite the fact that the goal of capitalists is to accumulate evermore money, the classical political economists largely took the analysis of money for granted.[4] To be sure, from Adam Smith to Karl Marx, we can certainly find passages on money but two things are of general note. First, the classical political economists as well as Karl Marx thought gold and silver were "real" money. In other words, money was understood as "commodity money" and therefore to expand the money supply meant finding new mines, plundering it from others, or selling goods or services on the world market to obtain it from others who possessed it. Indeed, a considerable portion of the history of slavery and colonial violence can be traced back to the elite concern for acquiring gold and silver (Di Muzio and Robbins, 2016; Graeber, 2012; Kwarteng, 2014; Vilar, 1986). Second, because gold and silver were thought to be money, the classics failed to understand the scale or level of credit creation that began with the institutionalization of the Bank of England in 1694. Many argue that the Bank of England was inspired by the Bank of Amsterdam and the success of Dutch finance. But this is not the case. While the Bank of Amsterdam did make loans from time to time, its primary function was to maintain the quality of the paper notes in circulation that represented coin. Moreover, the bank was owned by the city, not private social forces as came to be the case with the Bank of England (Wennerlind, 2011: 69; Vilar, 1986: 206; Zarlenga, 2002: 238ff). Whereas the notes issued by the Bank of Amsterdam mostly reflected the exact value of gold and silver in the city's vault, the Bank of England expanded the English money supply by extending paper notes as credit (Desan, 2014: 311ff).

---

literature is overwhelming concerned with system description in an effort to *better model* the economy whereas we, as critical political economists, are concerned with a radical critique of the way money is produced to *better society* (see also Di Muzio and Robbins, 2017).

[4] We do not deal with the origins of coined money which stretches back to King Croesus in Lydia minting the first standardized coins of electrum (a gold and silver alloy) in the 6th century BCE. It should also be mentioned here that there is no historical evidence for money emerging from barter or the double coincidence of wants as neoclassical theory holds (see Di Muzio and Robbins, 2017: 44-6 for a summary of the literature).

The Bank of England's largest customer was the Crown in Parliament who used the initial loan of £1,200,000 to finance war with France. Indeed, the main reason why the Royal Charter was granted to the Bank of England's 1509 investors was to provide the finance for organized violence against a dynastic rival (Davies, 2002: 261). The slave trade, colonization and continuous wars in the next two centuries lead to a mounting and unpayable "national" debt that solidified the Bank's role as the government's permanent debt manager. But the investors in the Bank of England did not only profit from war and debt, they also benefited from the interest received on loans to individuals and companies. As Wennerlind underscores, the Bank of England's notes became "Europe's first widely circulating credit currency" (2011: 109). Theoretically, however, the issued notes remained tethered to a metallic hoard of silver, and later only gold from 1861 (Davies, 2002: 315). No one knows for certain how much metal coin backed up the notes in circulation at any one time. In one study, Rubini argued that the Bank of England had a shifting reserve of silver for all notes in circulation of about 2.8 percent to 14.2 percent (1970: 696). Another study by Wennerlind argued that the founder of the Bank, William Paterson, proposed that 15 to 20 percent in silver for all notes outstanding would suffice to assure sufficient confidence in the Bank of England (2011: 128).[5] This ambiguity and the fact that the Bank of England was privileged by the government, likely helped the Bank gain confidence among the users of its notes. As long as citizens thought they could eventually cash in their notes for silver/gold coins, faith in this system of money creation could continue (Kim, 2011). This uncertainty need not delay us, for what is definite is that the notes in circulation were of a far higher value than the actual metallic hoard at the Bank. To sum up this brief history of the world's first widely circulating credit currency we can argue that new money was created as loans to customers – primarily to the British Crown in Parliament and primarily to finance an apparatus of international violence and Empire.

By the early 19th century, the British politician, Samson Ricardo, realized the absurdity of granting private social forces the power to create money:

> "It is evident therefore that if the Government itself were to be the sole issuer of paper money instead of borrowing it of the bank, the only difference would be with respect to interest: the Bank would no longer receive interest and the government would no longer pay it…It is said that Government could not with safety be entrusted with the power of issuing paper money – that it would most certainly abuse it... I propose to place this trust in the hands of three Commissioners" (Ricardo, 1838: 50).

Ricardo's proposal that the public take control of new money creation was ignored. In the 1844 Bank Charter Act, the Bank of England was given the exclusive right to issue banknotes in London. Country banks that were already issuing notes could continue to do so provided they were outside London (by a 65 mile radius) and backed their notes with some kind of credible security. Under this Act, the Bank of England was also divided into two distinct units, the Issue Department and the Banking Department. Davies highlights this important provision of the Act:

> "The Issue Department was to receive from the Banking Department some £14 million of government securities to back its fiduciary issue of notes, any

[5] To be historically accurate, the Bank of Sweden was the first bank of issue rather than a mere depository institution. It was created in 1661 (Zarlenga, 2002: 288).

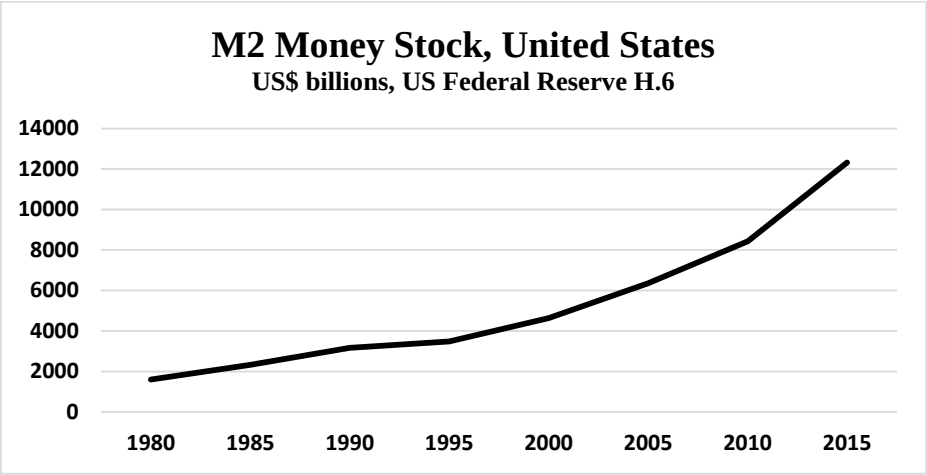
> issue above that [was] to be fully backed by gold and silver, the latter not to exceed one quarter of the gold" (2002: 315).

Thus, while the Bank of England had the exclusive right to issue banknotes in London, its ability to create new money *appeared to be* circumscribed by the new laws. Existing banks outside of London were also *seemingly* bounded in their ability to create money. However, while official note issuance was restricted, this did not stop the Bank of England and other provincial banks from merely recording new loans on their balance sheets and issuing cheques to borrowers (Davies, 2002: 317). In other words, the bankers found a convenient way around the legislation and continued to expand the money supply regardless of gold reserves which were never publically known anyway. This changed the nature of banking in Britain and as we shall discuss, its legacy largely remains with us today. With this in mind, we now move to examine two theories of money creation: the heavily taught fractional reserve theory known popularly as the money multiplier model and the underappreciated credit creation theory.

Money, the multiplier and Mankiw

In modern society, money is vital for survival and social reproduction. Yet, if money is so central to the daily lives of billions of people around the world, why don't we understand more about it? If we asked anyone where milk comes from, we would invariably get a correct answer from the vast majority of people: a cow, a goat or any other mammal. But when we ask people where money comes from or how new money is created, the answers are considerably varied and typically none of them are correct. Money is crucial to our contemporary civilization and the expansion of the money supply is constant and therefore we should have a convincing explanation for its growth. For example, consider Figure 1, which plots the growth of the money supply in the United States.[6]

Figure 1



[6] According to the Federal Reserve, "*M2* is defined as M1 plus savings deposits, small-denomination time deposits (those issued in amounts of less than $100,000), and retail money market mutual fund shares." M1 consists of " the sum of currency held by the public and transaction deposits at depository institutions (which are financial institutions that obtain their funds mainly through deposits from the public, such as commercial banks, savings and loan associations, savings banks, and credit unions)." https://www.federalreserve.gov/faqs/money_12845.htm (Accessed 10/1/2017)

Over the 35 year period illustrated in this figure, the money supply expanded by 674 percent. Similar charts can be constructed for the money supplies of most other states, suggesting that the money supply for most countries has increased over time. But how did this happen? Why is there no *accurate* general or popular knowledge about how new money is created or, put in another way, how the money supply grows? Why does there appear to be so much mystery and lack of knowledge surrounding the issue of money? One of the main reasons may simply be the division of labor both within and outside of the banking industry and the simple fact that people just take the current system as self-evident, natural and inevitable. Häring (2014) provides an additional answer in arguing that one of the reasons for the lack of knowledge is the fact that Economics textbooks teach a completely erroneous model of money creation. This error simply gets passed down from generation to generation without critical scrutiny.

Here we will examine the "money multiplier" theory which is outlined in N. Gregory Mankiw's popular textbook, *Macroeconomics* (2009) and taught in the vast majority of introductory courses to Economics. Importantly, Mankiw claims that banks are lending out existing deposits and that the banking system creates new money as money passes from bank to bank in the form of deposits and loans. We will demonstrate that this claim is mathematically and logically inaccurate.

Mankiw begins his narrative with a depositor entering a bank with money. The money is deposited into an account and we are told that the bank must keep a fraction of this money as a "reserve". This reserve remains in the account to allow for any withdrawals the depositor is likely to make. We are told that the bank has an incentive to loan out the portion of money not held in reserve to generate its own income stream through interest, fines and fees. To demonstrate this fractional reserve/money multiplier system, Mankiw sketches a balance sheet for what he calls Firstbank. Below we reproduce this balance sheet which assumes an initial deposit of $1000 and a 20 percent reserve ratio. The bank must retain $200 (20 percent of $1000) as a reserve, and will loan out the remaining $800:

Firstbank balance sheet

Assets		**Liabilities**	
Reserves	$200	Deposits	$1000
Loans	$800		

The books appear to balance and the narrative appears sound. Mankiw tells us that a fresh $800 has been created by this loan but, *very importantly*, that the original depositor still has a demand deposit for $1000. Supposedly the money supply is now $1800. Mankiw then introduces an additional bank to demonstrate how new money creation continues to build on the back of the original loan. He sets up a situation where the $800 loaned from Firstbank ends up in Secondbank via a deposit.[7] Below we reproduce Secondbank's balance sheet:

Secondbank balance sheet

Assets		**Liabilities**	
Reserves	$160	Deposits	$800
Loans	$640		

[7] For example, if the initial loan was used to make a purchase in a store, and the storeowner then deposited that money into their own bank account.

Once again, the books appear to balance and we are told that 20 percent of the deposit is held in reserve, and $640 in new money is loaned out and added to the money supply. So in this example, with the initial $1000 deposit, and the new money from Firstbank and Secondbank, there would now be a total money supply of $2440. To demonstrate this money multiplier process one more time, Mankiw adds an additional bank and assumes that the newly created loan of $640 gets deposited there. There, once again, 20 percent of the deposit is held as a reserve, and the remainder is loaned out.

Thirdbank balance sheet

Assets		**Liabilities**	
Reserves	$128	Deposits	$640
Loans	$512		

Mankiw tells us that this process continues until the reserve/loan ratio means that no further loans can be extended. In this example, the initial $1000 deposit will result in the addition of a fresh $4000 to the money supply, bringing the total money supply up to $5000.[8] While a bit convoluted and an odd way to create money, everything appears to make sense if we do not pay close attention. But if we do pay close attention, the model is fallacious on two counts.

The first issue is with the basic principle of double entry bookkeeping. Every accountant knows that *every* financial transaction has equal and opposite effects in at minimum two separate accounts. Based on this principle, Mankiw cannot record the loan (a financial transaction) as an asset without simultaneously recording the exact amount as a liability on the right side of the balance sheet. Not to do so violates the iron laws of double entry bookkeeping, making Mankiw's example erroneous (Gleeson-White, 2011: 93).

The second key issue with the money multiplier theory is that it is based on unsound math. If, as Mankiw claims, the bank lends out the portion of its deposits that it is not required to keep in reserve, then it is acting as a financial intermediary and is not in fact expanding the money supply at all: *it simply redistributes existing money*. In accordance with Mankiw's argument that the banks loan out fractions of its deposits (in this instance 80 percent, as 20 percent is kept as reserves), the accurate balance sheet operation should resemble this:

Firstbank

Initial Deposit $1000			
Assets		**Liabilities**	
Reserves	$200	Deposits	$200
Loans	$800	Credit	$800

What Mankiw fails to do in his example, which he must do if he is to remain consistent with the logic of the multiplier model and basic accounting standards, is subtract the loan from the depositor's account. It is not possible that the depositor can still have $1000 remaining in their account if $800 of that money has been loaned to someone else. In double entry bookkeeping, the money cannot be in two places at once; it cannot be in both the depositor's account and be loaned out simultaneously. Therefore, the $800 loan should be subtracted

[8] Mankiw's claim that this $1000 can increase the money supply to $5000 is calculated by adding together the initial deposit and the sum of all of the loans which are generated from this initial deposit.

from the depositor's account (as it has been removed and loaned to someone else) and the depositor's account should decrease to $200.

Now let us imagine the person who took the loan from Firstbank deposits the money in Secondbank. This means that Firstbank no longer has the liability for the $800 and its balance sheet would resemble the following:

Firstbank

Deposit $1000			
Assets		**Liabilities**	
Reserves	$200	Deposits	$200
Loans	$800	Credit	$0

After this transaction, the total money in the money supply is still $1000. $200 is sitting in Firstbank as a reserve, and $800 has just been deposited in Secondbank.

Of this $800, Secondbank must hold $160 as a reserve and can make a loan of $640. Once again, this means that $640 is subtracted from the Secondbank's liabilities as it is deposited in Thirdbank. After this second transaction, Firstbank's balance sheet remains the same as above, and Secondbank's balance sheet should like this:

Secondbank

Deposit $800			
Assets		**Liabilities**	
Reserves	$160	Deposits	$160
Loans	$640	Credit	$0

Again, we see that the money supply still totals $1000. Firstbank holds a reserve of $200, Secondbank holds a reserve of $160, and Thirdbank has just received a deposit of $640. It is clear to see, if we follow the logic of the money multiplier as well as proper accounting standards, that this system does not and cannot account for an expansion of the money supply. If we take the above examples as representing the entire economy, it is evident that the theory explains the redistribution of money, but there is always only the initial $1000 remaining in the economy. Put simply, rather than expand the money supply, the multiplier model merely demonstrates that portions of the initial $1000 gets transferred from person to person and bank to bank.

Let's look at another example starting with an initial deposit of $100 and a ten percent reserve ratio. The table below follows Mankiw's method of calculation, and it appears to demonstrate that after eight transactions the banking system will have actually created $469.53279 worth of additional or new money out of the initial deposit of $100.[9]

[9] This is calculated with the same method used in Mankiw's example of the $1000 deposit resulting in the creation of a new $4000 – all loans generated through the initial deposit are added together to ascertain "new money" in the supply, and the addition of the initial deposit determines the total money supply.

Initial Deposit $100	
Loans	**Reserves**
$90	$10
$81	$9
$72.9	$8.1
$65.61	$7.29
$59.049	$6.561
$53.1441	$5.9049
$47.82969	$5.31441
New Money created	**$469.53279**
Total Money Supply	**$569.53279**

But this is in fact not the case. Recall that Mankiw and other money multiplier adherents make the claim that *deposits are actually loaned out*. If this is the case, then the fraction of the deposit which is loaned out must be subtracted from the balance sheet as it is no longer held by the bank. In the table below, we highlight the money that must be subtracted on the balance sheet each time a loan is given by a bank and put money which remains in the money supply in bold.

	Deposits and Reserves	Money redistributed as Loans
Deposit 1	100	90
Reserve left of deposit 1	**10**	
Deposit 2	90	81
Reserve left of deposit 2	**9**	
Deposit 3	81	72.9
Reserve left of deposit 3	**8.1**	
Deposit 4	72.9	65.61
Reserve left of deposit 4	**7.29**	
Deposit 5	65.61	59.049
Reserve left of deposit 5	**6.561**	
Deposit 6	59.049	53.1441
Reserve left of deposit 6	**5.9049**	
Deposit 7	53.1441	47.8297
Reserve left of deposit 7	**5.31441**	
Deposit 8	**47.82969**	
Total money in supply (Reserves + final deposit)		$100

So by logic, basic accounting standards and simple math, the money multiplier model can in no way explain the *expansion* of the money supply. In the example above, there is only ever $100 in the economy until the $100 appears in reserves and there is no money left in the economy. Put simply, the way in which the money multiplier model is laid out actually explains the *diminution* of money and its eventual disappearance from the economy as it is held by banks as reserves. Here we have to remember that there is a key distinction to be made between claiming that the money supply expands versus the argument that the same money can be used more than once. The multiplier model shows the latter, not the former, and is therefore empirically false as an explanation for Figure 1 above.

The fact that the money multiplier is a myth was also confirmed by an empirical test performed by Professor Richard Werner (2014) at a German bank when he took out a loan of €200,000 and was allowed to observe the balance sheet operations. In no way were the accounting procedures similar to what we would expect if fractional reserve banking was indeed a real world practice. Rather, it provided evidence to support the credit creation theory of money expansion which we will discuss momentarily.

In addition to the factual inaccuracies and empirical evidence against the money multiplier model, the theory is also illogical. Firstly, if banks do indeed loan out other people's money, then how would they know which accounts to take money from in order to issue the loans they want to make? If a fractional reserve must be retained in each deposit account, how could banks keep track of account balances and ensure that their loans represented a specific portion of their total deposits? The mathematical difficulties associated with this model would be further complicated by the fluctuations caused by the constant transactions of the account holders themselves. Gone are the days when a person would have one bank account which they would deposit their savings into and would withdraw cash once or twice a week for daily living. These days, most people have several bank accounts which are constantly fluctuating as people access their accounts electronically several times each day. Electronic transfers, paying bills online, using digital debit to make purchases in stores or online – these practices all contribute to continuous fluctuations in our bank balances. It is near impossible to imagine how banks could perform all of the necessary calculations to ensure that there was always a fractional reserve available for each account of each account holder. Secondly, many people live paycheck to paycheck with very little savings and when savings do occur they are largely placed with pension and mutual funds to make higher returns than the meager interest banks pay on deposits. For example, in the United States, currently the world's largest economy by GDP, the savings rate as of 2016 was a mere 5.5% of total disposable income.[10] Thirdly, if banks were reliant on deposits to create loans, we would see far more competition between banks to attract deposits. Marketing and advertising campaigns to attract deposits would be ubiquitous and sustained with generous interest rate offers to encourage customers to save. Instead, daily experience teaches us the exact opposite: banks are far more interested in marketing and advertising their credit facilities – from mortgages to credit cards – then they are in encouraging new depositors or their existing customers to save. Finally, since the multiplier model assumes banks are borrowing other peoples' money, and therefore this money must be subtracted on the liability side of the balance sheet, bank balances would fluctuate wildly on a day to day basis. Anyone with a saving or checking account knows this does not happen. Moreover, if this truly was daily

[10] Personal Savings Rate https://fred.stlouisfed.org/series/PSAVERT (9/2/2016). Calculated as the ratio of savings to disposable personal income.

banking practice, there would be widespread panic amongst the population as people come to realize that a significant portion of the money they deposited and need on a daily basis is missing.

Finally, even if somehow adherents of the multiplier model could surmount these critiques, the process they describe would be an incredibly cumbersome method to create new money. If money was created in this way – by loaning out a percentage of deposits while not subtracting the amount from the bank's liabilities – the banking system would have no justifiable credibility. As previously discussed, in Britain of the 1800s, the moneyed population was satisfied with accepting notes in lieu of gold/silver because they believed they could exchange their notes for gold and silver if they wished. Similarly, people today are happy to leave their money in banks because they believe they can withdraw it if they wish. Imagine the panic and economic repercussions if the population discovered they could only access 20 percent[11] of their money because the bank had loaned out the rest. If the money multiplier theory was accurate, it would be creating liquidity based on the bizarre assumption that people will only ever want to access 20 percent of their funds. They would essentially be instigating a game of musical chairs where players keep joining, but there remains only 1 chair for when the music stops. With the money multiplier discredited, we have considerable reason to focus on the real world of banking practice. For this, we turn to the credit creation theory of money.

Commercial banks and credit creation

According to Werner's research, credit creation theory was prevalent in the late 19th century and the first two decades of the 20th century (2014: 2). Though a number of scholars and practitioners held this view, the work of the Scottish economist Henry Dunning Macleod stands out. Unlike most of his contemporaries Macleod is viewed as offering at least the outlines of a monetary theory of credit whereby loans make deposits rather than deposits enabling banks to make loans. An excerpt from his *The Theory of Credit* (1893) bears this out:

> "Thus we see that the very essence and nature of a Bank and a Banker is to create and issue Credit payable on demand: and this credit is intended to Circulate and perform all the functions of Money. A Bank is, therefore, not an office for 'borrowing' and 'lending' Money: but it is a Manufactory of Credit: as Mr. Cazenove well said, it is the Banking Credits which are the Loanable Capital: and as Bishop Berkeley said 'a Bank is a goldmine'" (Macleod, 1893: 594).

Despite this advance in realizing that loans make deposits, thereby expanding a bank's balance sheet and the money supply, Macleod's work was largely ignored by mainstream economists. According to Schumpeter the neglect of Macleod's work was because he failed to put his ideas in a "professionally acceptable form" (1986: 1081). Thus a great advance was lost and mainstream Economics and Finance textbooks came to teach two incorrect versions of money creation and the role of banks: the money multiplier model (fractional reserve banking) and financial intermediation (banks merely take in deposits and lend those deposits out to borrowers).

[11] Or whatever the reserve requirement may be.

After over a century of treating money as a "neutral veil" and largely unimportant in economic models, a small resurgence of interest in the credit creation theory is gaining headway (Bell, 2000; 2001; Graziani, 2014; Häring, 2013; Huber, 2014; Ingham, 1999; 2000; 2004; Kennedy, 1995; Lau and Smithin, 2002; McLeay et al., 2014; Mosler, 1997-8; Peacock, 2003-4; Rowbotham, 1998; Ryan-Collins et al., 2014; Werner, 2014a; 2014b; Wolf, 2014; Wray, 2004; Wray, 2015; Zarlenga, 2002).[12] The basic argument put forward by credit creation theory is that individual banks do indeed have the tremendous power to issue new money when they make loans to customers. Here it is important to note that most modern money is digital money that are numbers on computer screens rather than physical cash like notes and coins. For example, in the UK, notes and coins only make up 3 percent of the money supply, while *commercial bank money* registered on computers makes up the remaining 97 percent (McLeay et al., 2014: 2). A statement from the Bank of England leaves little doubt that new money is created by commercial banks:

> "Commercial banks create money, in the form of bank deposits, by making new loans. When a bank makes a loan, for example to someone taking out a mortgage to buy a house, it does not typically do so by giving them thousands of pounds worth of banknotes. Instead, it credits their bank account with a bank deposit of the size of the mortgage. At that moment, new money is created" (McLeay, 2014: 3).

The accounting identities on the balance sheet look as follows (Werner, 2014b):

Anybank

Assets		**Liabilities**	
Loans	$1000	Credit to Client A	$1000

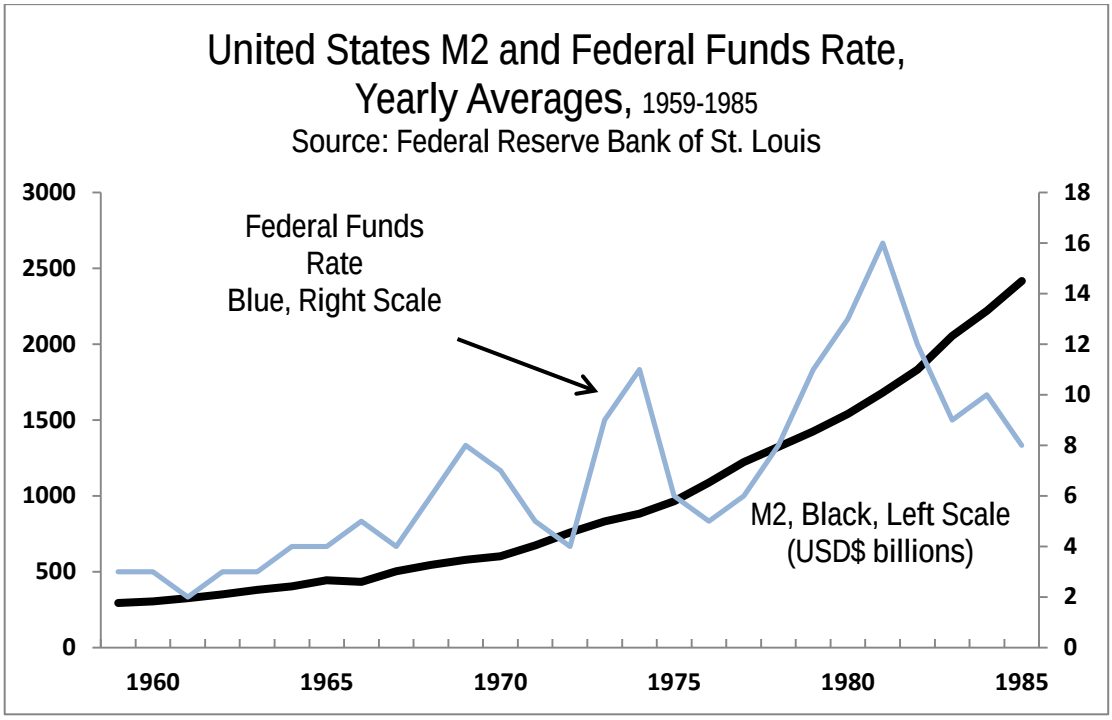
So as new loans are made, the balance sheet of a bank expands. It records the loan contract as an asset bearing interest and records a liability of $1000 to Client A in the example above. When the debt is repaid, the money ceases to circulate in the economy. What this means is that the vast majority of all new money entering the economy is issued as debt bearing interest. However, the banks do not create the interest when they create the loan, so there is always more debt in the system then there is the ability to repay (Di Muzio and Robbins, 2016; Robbins, 2013; Rowbotham, 1998). Thus, if everyone in the world stopped borrowing tomorrow, there would indeed be a *global* financial crisis. This is extremely problematic for a number of reasons and we explore the leading consequences of allowing banks to create money in the ensuing section. But first we must explore the limitations placed on new credit creation and the role of central banks.

The first limitation is that commercial banks must have willing borrowers that are likely to repay loans with interest added. Second, in uncertain economic conditions where the risk of repayment looks high, banks may stop lending. Third, borrowers may immediately use the loan to pay off old debts, effectively removing the new money from circulation. Last, it is

[12] It should also be noted that there is considerable disagreement among those who agree with credit creation theory or "endogenous money" about what this means for an economy and whether anything should be done about it. See, for instance volume 40, issue 5 of the *Cambridge Review of Economics*. See also the materials on the Positive Money website: http://positivemoney.org/

believed that new credit creation is constrained by the interest rate set by a central bank. This rate is the rate banks must pay should they need to borrow from other banks to balance their daily books. The interest rate for borrowers is set by commercial banks and is typically northward of the short-term interest rate set by the central bank. However, whilst it may seem intuitive that higher interest rates would slow the money supply, evidence from the United States suggests that this may not be a simple scientific truth. Figure 2 suggests that despite very high interest rates during the so-called "Volcker Shocks" in the late 1970s, businesses and households continued to borrow. M2 continues to increase despite record high interest rates.

Figure 2



Moreover, as Figures 2 and 3 suggest, the money supply actually expanded more over the period of record high interest rates. From 1960 to 1970, the percentage increase in the money supply was 105 percent while from 1970 to 1982, the period inclusive of the Volcker Shocks, the money supply increased by 179 percent. This is likely due to the fact that money is so central to the economy that businesses and households may continue to borrow even at record high interest rates.

This should not imply that very high interest rates do not slow the growth of the money supply by discouraging some potential borrowers, but from the empirical evidence, the money supply continues to increase despite historically high interest rates. However, as Figure 4 suggests, high interest rates do seem to slow down the pace of new money creation.

Figure 3

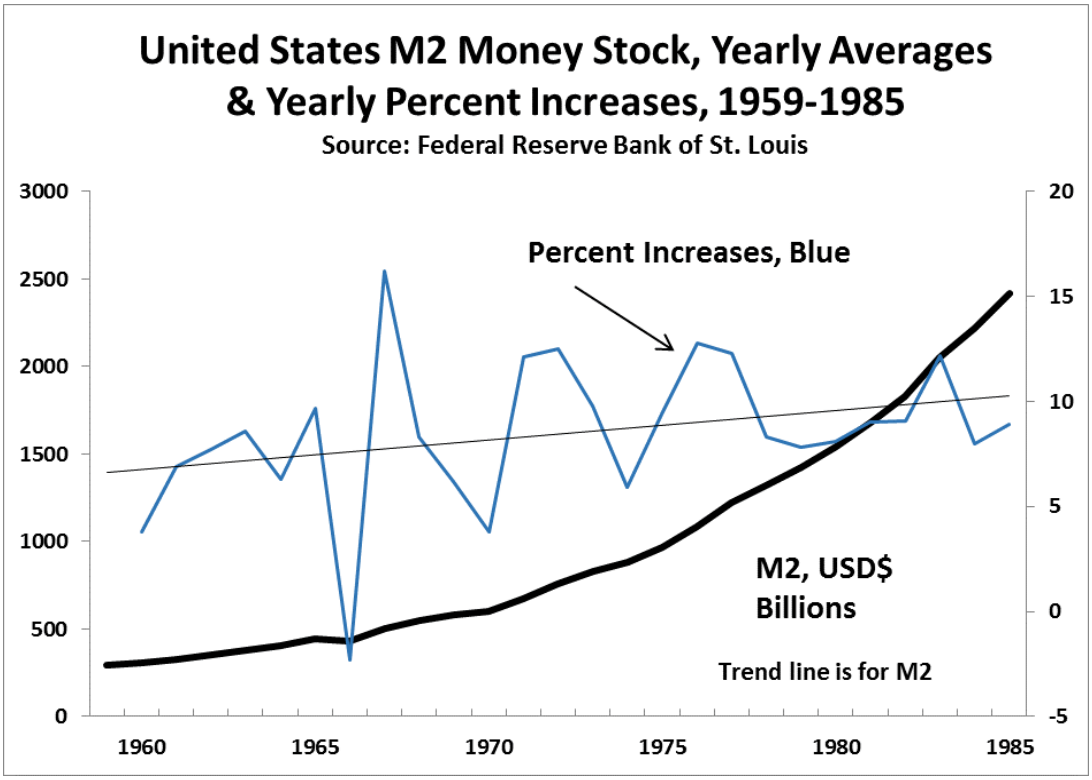
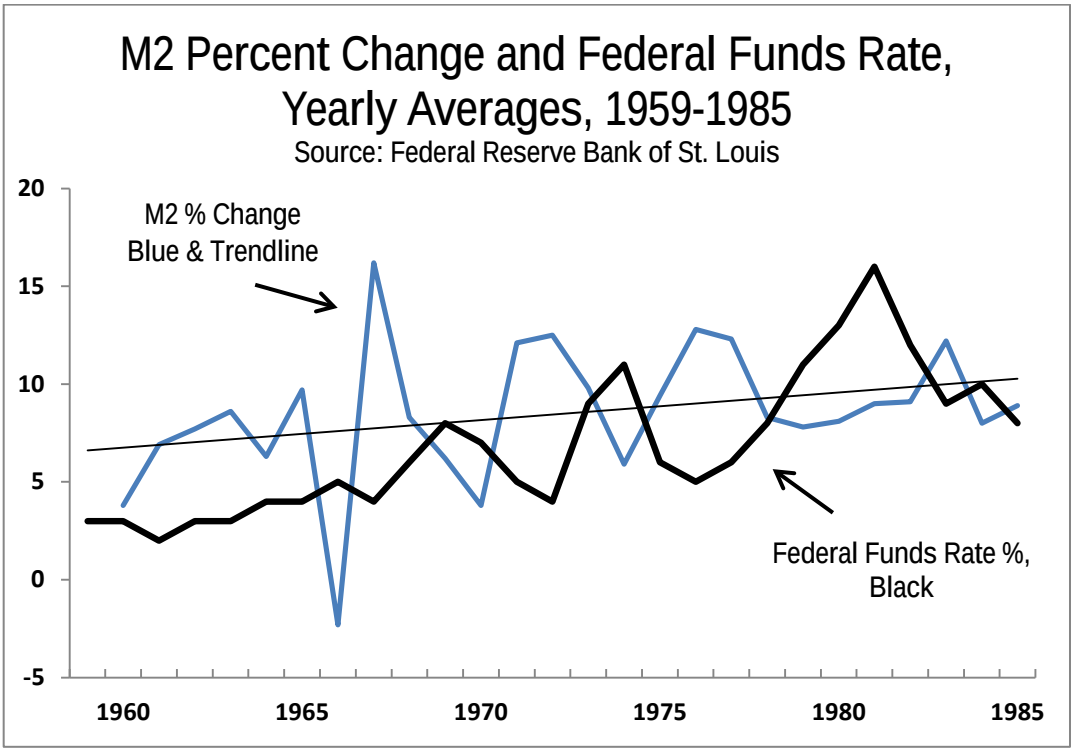


Figure 4



We can now turn to the role of central banks and in particular the role of reserves. Most large commercial banks hold a reserve account at the central bank. This account can be used to lend excess reserves to other banks or to purchase assets such as physical cash from the central bank. There are two types of reserves: 1) digital deposits in a commercial bank's account with the central bank and 2) actual cash. Because of the accounting identities used by the banking system as a whole, reserves can only change in one of three ways. First, the central bank either increases or decreases its assets. Second, the public wants to increase or decrease the cash they hold. Third, the government increases or reduces its deposits. Below we simplify the balance sheet for a central bank and a commercial bank following Sheard (2013).

Central bank

Assets	**Liabilities**
Government Bonds	Reserves
Foreign Exchange	Bank Notes
Gold	Government Deposits

Commercial bank

Assets	**Liabilities**
Reserves	Deposits
Loans	Equity
Bonds	

The double-entry bookkeeping used by the banking system means that the following balance sheet identities must hold:

$\Delta R = \Delta Assets - \Delta Bank\ Notes - \Delta Government\ Deposits$

To provide a quick example, suppose that reserves were at US$400:

Reserves $400 = Assets $1000 – Bank Notes $500 – Government Deposits $100 = $400

Should the customers of commercial banks want to increase their use of physical cash, we can see that reserves would go down since this is where bank notes/cash comes from on the balance sheet. In this scenario, it is likely that the commercial bank would have sold some of its assets to the central bank (typically bonds) in return for bank notes/cash.

Reserves $200 = Assets $1000 – Bank Notes $700 – Government Deposits $100 = $200

Similarly, if customers of commercial banks desire to hold less cash, reserves would increase.

Reserves $800 = Assets $1000 – Bank Notes $100 – Government Deposits $100 = $800

There is considerable confusion about reserves and we would suggest that this is because the money multiplier/fractional reserve model has been taught for so long combined with a failure to recognize that most money in advanced capitalist economies consists of digital deposits not bank notes and coins. It should also be emphasized that while some central

banks do issue a set reserve requirement, this has *little to no effect* on lending. Indeed, many countries such as Australia, New Zealand, the UK, Sweden and Canada have no reserve requirement. In other words, there is no reserve constraint on lending and banks *never* make loans to customers from their reserves. The only link between bank lending and reserves is that "reserves go down when banknotes increase" (Sheard, 2013: 7). Digital reserve deposits are only transferred from commercial bank account to commercial bank account at the central bank in the interbank money market. If there are excess reserves in one bank's account and not enough reserves in a second bank's account, the first bank may lend out its excess reserves to the second bank in order to earn interest. As a whole, these daily transactions help balance the books for the entire banking system.

The other important role played by the central bank is to set a target interest rate that ensures there are neither too few reserves (if so, the interest rate would go up) nor too many reserves (if so, the interest rate will go down) (Sheard, 2013: 8).[13] Thus, the short-term interest rate set by the central bank fulfills two main functions. First, it sets the price for borrowing reserves in the interbank money market. Second, it provides a base interest rate and acts as a guide to commercial banks when setting their own commercial interest rates its offers to clients. In no way does the interest rate affect a bank's ability to create new money as loans. It may be the case that interest rates get so high that households and businesses refuse to borrow, but this is not the same as claiming that a bank does not have the capacity to lend because of high interest rates. We are now in a position to examine the consequences of capitalized money creation by commercial banks.

Consequences of capitalizing credit creation

The empirical evidence all suggests that we live in a world of increasing individual, household, corporate and government debt.[14] Debt and the political cry to "balance the books" provides the major excuse for fiscal austerity, the privatization of public assets, layoffs, corporate outsourcing to cut labor costs and more (Di Muzio and Robbins, 2016; Hudson, 2015). The prevalence of debt across the global political economy is hardly surprising when we learn that banks create new money as loans that must be repaid with interest. But why do we need this credit in the first place? Why do citizens not have enough purchasing power to acquire the goods and services produced in an economy? Why are there periodic crises of "overproduction" or "underconsumption" and why do businesses often run under their capacity to produce?

As identified by C.H. Douglas in the early twentieth century, it turns out that the gap between available purchasing power and the total price quantity of goods and services on the market is structural or encoded into the very way in which goods and services are priced (1931: 16ff). It is quite strange that this has been forgotten but perhaps not surprising given the dominance of neoclassical economics and its fixation on supply and demand as the key to price formation. Let us be clear: to price their goods and services, businesses add up their costs and then add a markup or margin on top of the cost of providing a good or service. Generally, the more competitive the market, the lower the markup will be whereas in a less competitive market with a desirable product, the higher the markup will be. As Veblen underscored,

[13] Adjusting reserves, in normal times, allows the central bank to hit its "announced policy rate" (Sheard, 2013: 8).
[14] The discussion in this section draws on Tim Di Muzio and Richard Robbins (2017) *An Anthropology of Money: A Critical Introduction*. (New York: Routledge).

businesses try to charge "what the traffic will bear" (1923: 85). A simple example should suffice to demonstrate Douglas' point about a structural gap between purchasing power and the price of goods and services available on the market.

Suppose we started an orange juice company and we will do one run of production. Our cost may look something like the following:

Oranges	$100
Machines	$1000
Utilities	$100
Packaging	$100
Loans/interest payments	$300
Tax	$200
Labor	**$200 Purchasing Power**
Total cost of 1 run of production	$2000

During this first run of production, suppose we created 2000 jugs of orange juice. How should we price our goods? Well, if we sell each jug for under $1, then we will lose money, certainly not the end goal of business. If we sell each jug for $1, then we simple break even, again, not the goal of capitalist enterprise. We would certainly want to sell each jug for over $1. The market for fruit juices is quite competitive but we estimate that we can get a ten percent markup on our costs or $200. This means we would sell each jug for $1.10 to realize a total profit of $200. So the total dollar amount of our goods in the marketplace is $1.10 x 2000 = $2200.

We can start to see the problem immediately. The only purchasing power created during the first run of production is $200. C.H. Douglas, an engineer by trade discovered that there is a *perpetual structural gap* between purchasing power and the price of goods outstanding on the market. Since all businesses price their goods and services in a similar manner (figure out costs then add a markup), the gap can never be overcome by the producers of goods and services themselves should they want to remain successful in business by earning a profit. Increasing wages cannot get over this problem simply because labor is always a cost of production. This is why C.H. Douglas argued for a social credit to be dispersed to citizens by the government. Instead of listening to C.H. Douglas, what has happened in practice is that the gap has been partially filled by commercial bank credit and as Douglas recognized, the banks have a monopoly over it. In other words, to avoid severe recessions or depressions, credit and debt are structurally necessary in a capitalist economy that employs a cost-plus manner of pricing.[15] This is a giant gift to the owners of commercial banks since they ultimately profit not only from interest but also from the fees they charge on *needed* credit. This helps us to explain why the commercial bank sector of the global economy is the most

[15] One reviewer of this article asked why the pre-modern world economy was not in permanent recession given that capitalism and the division of labor are older than commercial banking. First, it is difficult to talk about "recession" in the past since it can only be discussed in terms of gross domestic product, a modern calculation introduced by Simon Kuznets and his associates in the early 20th century. Second, historians of energy and economic growth have argued that there was never *sustained* economic growth until Europe, and England in particular, started to use coal in significant historical quantities. Third, at least in England, but likely in Europe as a whole, there was a constant recognition of the dearth of money to circulate goods, suggesting low economic growth until commercial banking and coal energy became more prominent. The connection between energy and the expansion of the money supply is crucial for explaining greater economic growth and the globalization of capitalism (for a summary of this literature see (Di Muzio, 2015:27-46).

heavily capitalized at US$4.8 trillion.[16] It also helps to explain how recessions and depressions can occur if banks stop lending or customers refuse to borrow. If all borrowing ceased tomorrow, the global economy would be in freefall.

In addition to the lack of purchasing power created in the economy by capitalist production, there is a further problem: there is always more debt in the system than there is the ability to repay the debt (Rowbotham 1998). When banks create money as loans, they create the principal, never the interest. The following example should suffice to illustrate the problem.

Loan A	$1000	%10 interest	$100
Loan B	$1000	%10 interest	$100
Loan C	$1000	%10 interest	$100
Total Money	$3000	Total Interest	$300

Thus the bank that made the loans expects to be paid $3300 dollars, however there is only $3000 circulating in the economy. Unless new loans are given with a longer maturity date, at least one of the loans must fail. What this suggests is that bankruptcy and the failure to pay debt is a structural feature of modern capitalism and the way new money is created by commercial banks. The money has to be obtained from elsewhere, which is also a key trigger for the need for economic growth and the greater commodification and monetization of nature (Di Muzio and Robbins, 2016; Robbins, 2013). The sabotage of the possibility of public or sovereign money and the private ownership of the capacity to create new money leads to an inevitable need for credit/debt when incomes do not meet spending expectations or a desired lifestyle. For example, most people are forced into debt if they want to buy a home or car. But as Soederberg (2014) points out in her wonderful book, many low income groups have been turning to consumer credit just to make ends meet. Thus, there are two structural features of capitalism that ultimately benefit a small amount of bank owners: 1) the lack of purchasing power in the economy because of cost-plus production, leading to the need for credit; and 2) the fact that there is always more debt in the economy than there is the ability to repay it. But while we argue that these are the main structural features that benefit a tiny minority of dominant bank owners and disadvantage everyone else, there are additional causes for serious concern.

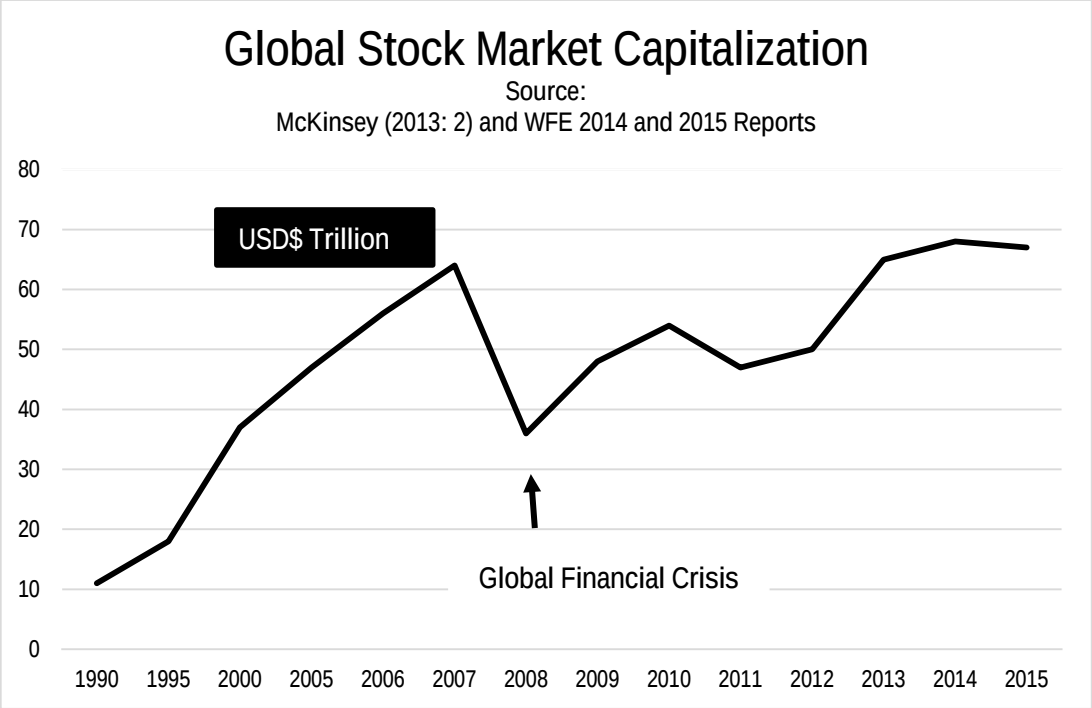
First, democratic governments are not in control of most of the money supply and like individuals and businesses, they are structurally forced into debt to a minority of private social forces who profit immensely from this relationship. The fact that the state has the power to tax the population allows for dominant bank owners to capitalize on this power process and direct a stream of income to themselves through government securities. As Creutz (2010: 139-158) pointed out long ago, it is a mathematical certainty that due to the unequal ownership of government securities (the minority) and the payment of taxes (the majority) more money will be received by the minority of the bondholders from the majority of taxpayers. Though we await more empirical studies from additional states, Sandy Brian Hager's research on the United States' public debt empirically demonstrates that ownership of government securities is highly concentrated with the top one percent owning a little over fifty five percent of the public debt (2016: 41). Currently, there is an ever increasing amount of public debts totaling

[16] https://www.ft.com/content/1fda5794-169f-11e5-b07f-00144feabdc0 Data are from the *Financial Times* Global 500, a list of the world's largest corporations by market capitalization.

slightly below US$60 trillion.[17] Even at 1 percent simple interest, the peoples of the world would fork over US$600 billion, mostly to a minority of creditors.

Second, while governments do set spending, distribution and allocation priorities based on a budget, previous spending, and a public mandate, it is largely commercial banks that set allocation/distribution priorities for society given that they are the primary institutions of new money creation. Banks need not create money for productive purposes and can create money to speculate on securities and real estate. The quantity of money that was created to reflate global stock markets after the crash began in 2007 is difficult to ascertain. But within the span of two years, global market capitalization in publicly listed firms recovered to slightly over what it was at the peak of the crash.

Figure 5[18]



Third, the issuance of new money as loans is based on creditworthiness and tied to assets and income. What this means is that the already rich can borrow more money, leading to greater inequality. Those with higher incomes and good credit scores can also borrow more money to buy assets including stocks and real estate. They can typically do so at better interest rates than people with lower incomes who tend to rely on high interest rate loans out of necessity (Bolton and Rosenthal, 2005). These unequal interest rates are often highly gendered and racialized, as evidenced by the subprime mortgage crisis in the United States (Gill and Dymski et al., 2013; Gill and Roberts, 2011: 154-171). Thus, access to credit and on what terms helps to explain the social reproduction of class inequality (Ingham, 2000). Another example is hedge fund managers. Though most money managers leverage their

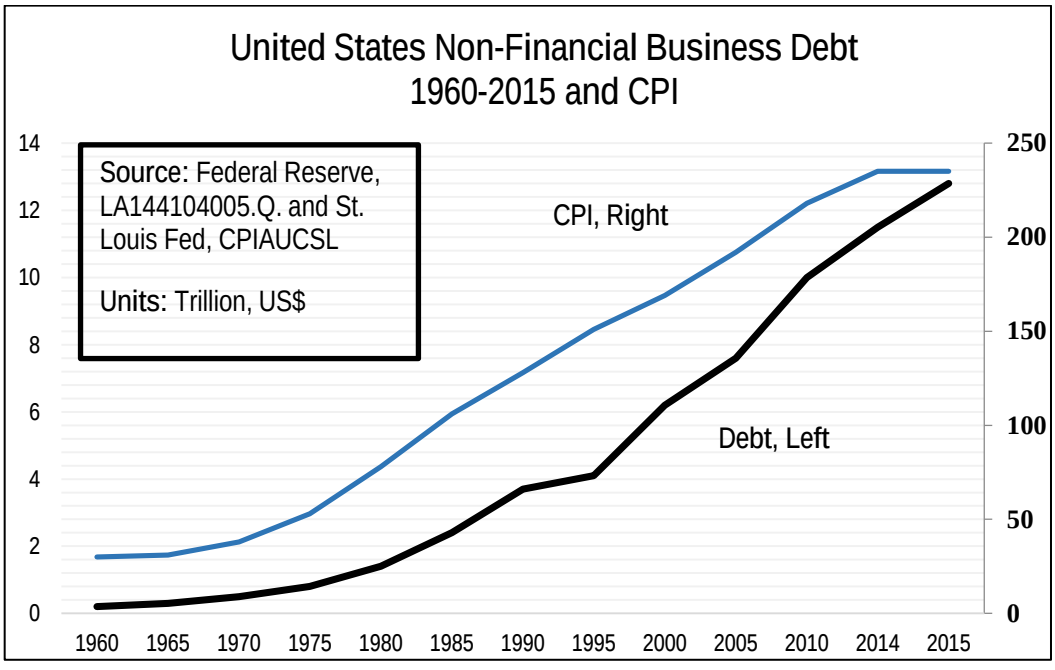
[17] http://www.economist.com/content/global_debt_clock Data from the *Economist*'s public debt clock. (Accessed 7/9/2016).
[18] World Federation of Exchanges (2014 and 2015) Market Highlights. file:///C:/Users/Tim/Downloads/2014%20WFE%20Market%20Highlights.pdf and file:///C:/Users/Tim/Downloads/2015%20Market%20Highlights.pdf (Accessed 14/9/2016)

capital by two times, some can leverage their assets up to ten times. What this means is that when managers control capital of US$1 billion, they can borrow another US$10 billion from commercial banks to speculate on income-generating assets (Mallaby, 2010: 12). We have to recall that a five percent return on US$10 billion is far greater than a five percent return on US$ 1 billion. This helps to explain the proliferation of hedge fund billionaires since the turn of the second millennium.

Another significant fact to note about the current way we produce new money is that the owners of banks essentially profit from a fraud. Fraud is typically understood to be a deliberate deception in order to secure an unfair gain or advantage. Since the banks create new money and do not act as intermediaries between savers and borrowers, they are indeed deceiving the public and certainly are securing unfair financial gains in an economy where purchasing power can never equal the total cost of all goods and services outstanding.

What is more, interest on money/debt is a key driver of differential inflation. Interest is a cost to business and gets pushed on to consumers. So consumers not only pay for the base costs of a good or service, but also a portion of the interest the business owes to the banks as well as whatever mark-up on costs the business feels it can get away with. This is interest inflation and markup or profit inflation. Though it may have been the case in the early to mid-twentieth century, it is interesting to consider that most businesses to do not finance their expansions or operations out of their retained earnings as some might expect.[19] Consider Figure 6 which plots the level of nonfinancial corporate debt in the United States with the consumer price index.

Figure 6

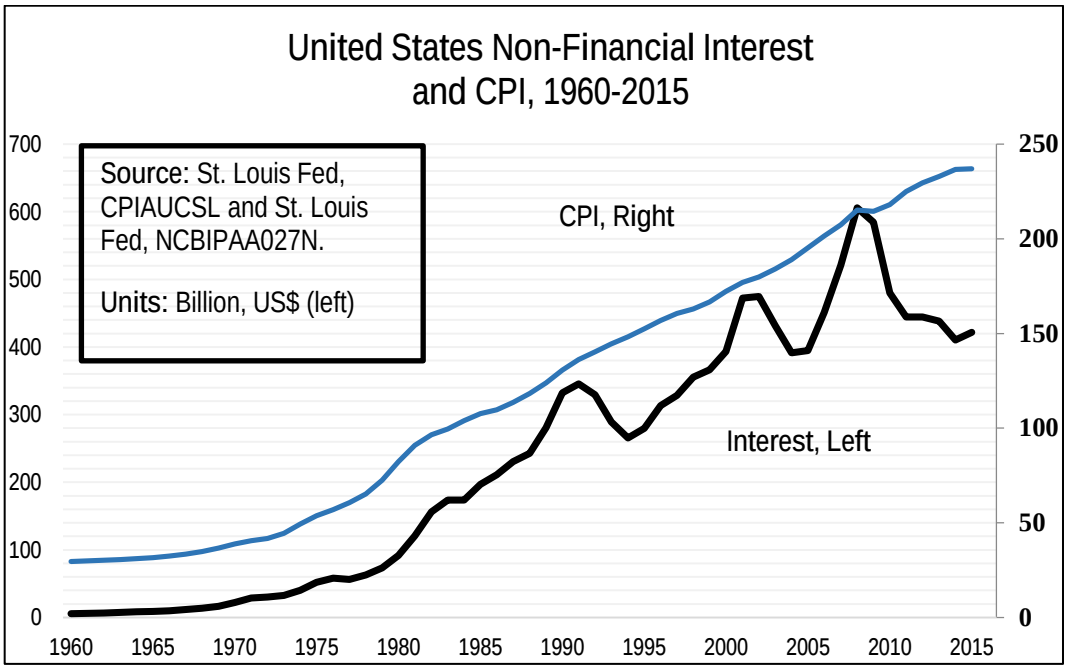


As Figure 6 reveals, non-financial corporations in the United States currently hold just over US$13 trillion in debt. Not surprisingly, inflation throughout the period under scrutiny largely mirrors the growth of nonfinancial business debt which grew from US$56 billion in 1945 to

[19] In the United States, nonfinancial business debt was first recorded in 1945 and stood at US$56 billion compared to today's whopping US$13.4 trillion in debt.

US$13.4 trillion in 2015. Over the period, this is an incredible increase in corporate debt loads of 23,811 percent. In Figure 7 we chart the consumer price index with the amount of interest paid by nonfinancial corporations in the United States. Though there are peaks and troughs in the amount of interest paid by nonfinancial corporations (due largely to shifting interest rates rather than overall yearly debt loads), the uptrend is clear with an overall increase for the period at 8,320 per cent. What this means in dollar terms is that nonfinancial corporations paid just over US$13 trillion in interest to commercial banks from 1960 to 2015.

Figure 7



Taken together, what Figure 6 and 7 suggest is that differential inflation is another structural feature of the way new money is produced as commercial banks loans. This is because the cost of interest bearing loans taken out by nonfinancial corporations gets passed on in consumer prices – in the US case, just over US$13 trillion in 55 years. We can call this "differential inflation" because different companies will have different debt loads, interest rates and costs of production. Thus, each corporation with debt will contribute to price inflation differentially. However, we should be clear here that *we are not* making the argument that interest on commercial bank loans is the *only* driver of inflation, but it does appear to be a major contributor and should be the subject of future research.

What this analysis also suggests is that government fiscal policy is incredibly important and has more to do with monetary policy than the monetary policy of central banks – which basically regulates the inter-bank market. This is the case because, should an economy stagnate with low or negative growth and high unemployment then it is only governments that can help create effective demand by borrowing and spending into the economy. The only problem with this solution is that, at present, thanks largely to Keynes' denial of sovereign or publically issued money, governments are *structurally forced* into debt at interest when the cost of their priorities exceed the revenue they receive from taxes, fines, fees and additional measures like the sale of public assets. Interest free currency such as the Greenback in the United States and the Bradbury notes created during World War I in the UK are evidence of how public, interest free money, can be created.

There is another consequence for entrepreneurs who may have a great idea but not enough money to invest in their business to make it viable. Since banks typically do not lend to new small businesses without collateral or some other guarantee, this means that entrepreneurs have to turn to venture capitalists and the like for an investment and therefore give up equity in their companies. This can act to further concentrate wealth and ownership. Though more research would need to be undertaken, the research that does exist shows that the degree of corporate control held by US and UK financial institutions is extremely concentrated, suggesting that these financial institutions have created money over their existence to purchase ownership claims over the entire universe of multinational companies (Vitali et al., 2011). As Ingham has argued, Hilferding and the Marxist tradition have vastly undertheorized the power of finance/banking (2004: 63).

There is a final implication when we realize the multiplier model explains the diminution of money whereas credit creation theory explains new money creation correctly. That is, we need to abandon the notion that savings can lead to investment for the economy as a whole. This is largely false. No saving **as deposits in banks** has to take place before new money can be issued by a commercial or public bank. Furthermore, more saving means less money in an economy, not more.

Conclusion

The shroud of secrecy around money and banking is as enormous as the general confusion experienced by most of the population when it comes to explaining what money is and how the supply of it expands. In this article, we have done our best to help demystify modern money creation. To do so, we began by offering a brief historical account of how the money supply started to expand over and above a supply of gold/silver in England. We did so to suggest that banks were creating new money as credit, albeit backed by a publically *unknown* hoard of silver and later gold. We then considered one of the most popular accounts of money creation offered by Harvard academic N. Gregory Mankiw in his first year textbook, *Macroeconomics*. We have demonstrated by simple math, logic and basic accounting identities used in modern banking why the fractional reserve or money multiplier model cannot explain the expansion of the money supply. Instead, the model actually demonstrates how money gets transferred from borrower to borrower until the money supply disappears as the cumulative reserves held by the banks in the system. With the multiplier model/fractional reserve theory discredited, we moved to discuss how commercial banks create new money when they make loans to customers. In the credit creation theory of monetary expansion, loans create deposits rather than deposits creating loans. We have argued that the capitalization of the money supply by the owners of banks is not only a tremendous power to possess but also one that has severe consequences for society, not least radical inequality in incomes, wealth and life chances.

We are of course not the first to suggest that the public should control the issuance of money, but we strongly believe that this should be the case and that it is completely feasible to create non-interest bearing money without creating inflation. In fact, the current way commercial banks create money guarantees a considerable level of inflation since interest on loans is a cost to business and gets pushed on to consumers as shown in Figures 6 and 7.

We must always keep in mind that the current monetary order was not established by divine intelligence or to better society but to enrich a few state creditors in a particular era where

"real" money was thought to be gold or silver and their paper representatives. Once we realize that current fiscal-monetary mechanics is not "natural" or "inevitable" but a historical creation, this makes change towards a sovereign monetary and new fiscal system possible. How this process should be institutionalized in a democracy can be debated, but the fact that honest scholars need to press this debate should be clear given the evidence we present in this article. Mathematically, there is a structural fault line in capitalist economies identified by C.H. Douglas: there is never enough purchasing power generated to clear the goods and services produced in any given economy. This fact, combined with the way in which the fiscal system has been set up so that governments must borrow instead of finance their own expenditure out of money creation and taxation, leads to a perpetual need for credit. This helps to explain the extension of credit to governments by private social forces, but also the invention and explosion of consumer credit along with other debt instruments such as corporate and financial bonds. Under the present system, debts will continue to mount, contributing to the need for austerity policies and an uncertain and politically dangerous future for most political economies. It should be clear that business failure, structural inequality and financial crises are structurally embedded in the very math of the present monetary and fiscal order.

In conclusion, if there is any novelty in our own argument, it has been to demonstrate, in our understanding for the first time, why the multiplier model cannot explain new money creation, combined with laying out in plain English, the consequences and structural imperatives of the present way new money is produced. Without doubt, if we are to overcome the unfair and historically created structural deficits in the current monetary and fiscal order, there is much work to be done, academically, politically and practically.

References

Bell, Stephanie (2000) "Do Taxes and Bonds Finance Government Spending? *Journal of Economic Issues* Vol. 34, No. 3: 603-620.

Bell, Stephanie (2001) "The Role of the State in the Hierarchy of Money" *Cambridge Journal of Economics* Vol. 25: 149-63.

Bolton, Patrick and Howard Rosenthal (ed.) (2005) *Credit Markets for the Poor*. (New York: Russell Sage Foundation).

Creutz, Helmut (2010) *The Money Syndrome: Towards a Market Economy Free from Crises*. (London: Robert Middelstaed).

Desan, Christine (2014) *Making Money: Coin, Currency and the Coming of Capitalism*. (Oxford UK, Oxford University Press).

Di Muzio, Tim (2015) *Carbon Capitalism: Energy, Social Reproduction and World Order*. (London: Rowman Littlefield International).

Di Muzio, Tim and Richard Robbins (2016) *Debt as Power.* (Manchester: Manchester University Press).

Di Muzio, Tim and Richard Robbins (2017) *An Anthropology of Money: A Critical Introduction*. (New York: Routledge).

Douglas, C.H. (1931) *The Monopoly of Credit*. (London: Chapman and Hall).

Dymski, Gary, Jesus Hernandez, and Lisa Mohanty (2013) Race, Gender, Power, and The Us Subprime Mortgage And Foreclosure Crisis: Ameso Analysis." *Feminist Economics* 19.3: 124-151.

Gleeson-White, Jane (2011) *Double Entry: How the Merchants of Venice Created Modern Finance*. (New York: W.W. Norton & Company).

Graziani, A. (2009). *The Monetary Theory of Production* (Cambridge, U.K., Cambridge University Press).

Graeber, David (2012) *Debt: The First 5,000 Years*. (Brooklyn: Melville House Publishing).

Hager, Sandy Brian (2016) *Public Debt, Inequality and Power: The Making of a Modern Debt State*. (Oakland: University of California Press).

Häring, Norbert (2013) "The Veil of Deception over Money: How Central Bankers and Textbooks Distort the Nature of Banking and Central Banking" *Real-World Economic Review*, I. 63: 1-18

Huber, Joseph and James Robertson (2000) *Creating New Money: A Monetary Reform for the Information Age*. (London: New Economics Foundation).

Huber, Joseph (2014) "Modern Money Theory and New Currency Theory: A Comparative Discussion, Including an Assessment of Their Relevance to Monetary Reform" *Real-World Economics Review*, No. 66: 38-57.

Hudson, Michael (2015) *Killing the Host*. (Petrolia: Counterpunch Books).

Ingham, Geoffrey (1999) "Capitalism, Money and Banking: A Critique of Recent Historical Sociology", *British Journal of Sociology*, Vol. 50, no. 1: 76-98.

Ingham, Geoffrey (2000) "Class Inequality and the Social Production of Money" *The Sociological Review*, Vol. 48, I.1: 66-86.

Ingham, Geoffrey (2004) *The Nature of Money*. (Cambridge UK: Polity Press)

Kim, Jongchul (2011) "How Modern Banking Originated: The London Goldsmith-Bankers' Institutionalization of Trust", *Business History*, 53:6, 939-959.

Kennedy, Margrit I., and Declan Kennedy (1995). *Interest and Inflation Free Money: Creating an Exchange Medium that Works for Everybody and Protects the Earth*. (Philadelphia, PA: New Society Publishers).

Kwarteng, Kwasi (2014) *War and Gold: A Five Hundred Year History of Empires, Adventures and Debt.* (London: Bloomsbury).

Lau, Jeffrey Y.F and John Smithin (2002) "The Role of Money in Capitalism" *International Journal of Political Economy*, Vol. 32, No. 3: 5-22.

Macleod, Henry Dunning (1893) *The Theory of Credit*. (London: Longmans, Green).

Mankiw, N. Gregory (2009) *Macroeconomics, 7th Edition*. (New York: Worth Publishers).

McLeay, Michael and Amar Radia and Ryland Thomas (2014) "Money Creation in the Modern Economy", *Quarterly Bulletin Q1* Bank of England.

Moore, Basil J. (1979) "The Endogenous Money Stock", *Journal of Post Keynesian Economics*, Vol. 10, No. 3: 372-385.

Mosler, W. (1997–8) "Full Employment and Price Stability", *Journal of Post Keynesian Economics* 20, 2: 167–82.

Peacock, Mark S. (2003–4) "State, Money, Catallaxy: Underlaboring for a Chartalist Theory of Money." *Journal of Post Keynesian Economics* 26, No. 2: 205–225.

Ricardo, Samson (1838) *A National Bank: The Remedy for the Evils Attendant on our Present System of Paper Currency*. (London: Pelham Richardson).

Robbins, Richard (2013) *Global Problems and the Culture of Capitalism* (New York: Pearson).

Roberts, A., Gill, S., "Macroeconomic Governance, Gendered Inequality and Global Crises" in Young, B. (Ed.), Bakker, I. (Ed.), & Elson, D. (Ed.) (2011) *Questioning Financial Governance from a Feminist Perspective.* (London and New York: Routledge), pp. 154-171.

Rowbotham, Michael (1998) *The Grip of Death: A Study of Modern Money, Debt Slavery and Destructive Economics*. (Charlbury: Jon Carpenter Publishing)

Ryan-Collins, Josh, Tony Greenham, Richard Werner and Andrew Jackson (2014) *Where Does Money Come From? A Guide to the UK Monetary and Banking System*. (New Economics Foundation).

Soederberg, Susanne (2014) *Debtfare States and the Poverty Industry: Money, Discipline and the Surplus Population*. (London: Routledge).

Veblen, Thorstein (1923) *Absentee Ownership: Business Enterprise in Recent Times*. (New Brunswick: Transaction).

Vitali, Stefania, James B. Glattfelder and Stefano Battiston (2011) "The Network of Global Corporate Control", *PLoS ONE* Vol. 6 No. 10: 1-36.

Werner, Richard A. (2014a) "Can Banks Individually Create Money out of Nothing? – The Theories and the Empirical Evidence" *International Review of Financial Analysis*, Vol 36: 1-19.

Werner, Richard A. (2014b) "How do Banks Create Money, and Why Can Other Firms Not Do the Same? An Explanation for the Coexistence of Lending and Deposit-Taking" *International Review of Financial Analysis*, Vol 36: 71-77.

Wolf, Martin "Strip Private Banks of Their Power to Create Money" *Financial Times*, April 24, 2014.

Wray, L. Randall (ed) (2004) *Credit and State Theories of Money*. (Cheltenham: Edward Elgar)

Wray, L. Randall (2015) *Modern Money Theory*, 2nd Edition. (Basingstoke: Palgrave Macmillan)

Zarlenga, Stephen A. (2002) *Lost Science of Money: The Mythology of Money: The Story of Power*. (Valatie: American Monetary Institute)

Author contacts: tdimuzio@uow.edu.au and lnoble@uow.edu.au

SUGGESTED CITATION:
Tim Di Muzio and Leonie Noble, "The coming revolution in political economy: money creation, Mankiw and misguided macroeconomics", *real-world economics review*, issue no. 80, 26 June 2017, pp. 85-108,
http://www.paecon.net/PAEReview/issue80/DiMuzioNoble80.pdf

real-world economics review, issue no. 80
subscribe for free

A diagrammatic derivation of involuntary unemployment from Keynesian micro-foundations

Philip George [India]

Abstract
This paper lays the micro-foundations for Keynesian macroeconomics. It shows that there is only one geometrical demand curve which does not assume that aggregate demand is constant. This curve may therefore be regarded as the Keynesian demand curve. Further, it shows that this curve aggregates without distortion. This means that the macro-economy can be constructed from heterogeneous agents without invoking a representative agent. Finally, it shows how the aggregate labour demand curve interacts with the labour supply curve to yield involuntary unemployment.

JEL Classifications E12, E13

Keywords Keynes, macroeconomics, micro-foundations, involuntary unemployment, representative agent

Introduction

One of the principal charges against the economics of Keynes is that it lacks micro-foundations. This charge is usually hurled by the New Classical school, whose own micro-foundations consist of a single representative agent, which, many would say, amounts to no micro-foundations at all.

Keynesians themselves have, however, by and large, not attempted to build micro-foundations for the Keynesian macroeconomic superstructure. The argument generally advanced is that it is unnecessary or impossible or both.

In the conclusion to a previous paper (George, 2016) I suggested that General Equilibrium theory and Marshallian analysis are mathematically equivalent and that this had implications for Keynes's position on involuntary unemployment. In this paper I not only construct micro-foundations for Keynesian macroeconomics but show that they negate one of the principal assertions of New Classical economics, viz, that there is no such thing as involuntary unemployment.

Keynes' claim in *The General Theory* was that classical economics dealt with the special case in which aggregate demand was constant, and that his own book dealt with the more general case in which no such constraint was placed on aggregate demand. I show that linear and other demand curves implicitly assume that aggregate demand is constant, and that there is only one geometrical curve which does not make this assumption. That curve is the rectangular hyperbola. Since it meets Keynes' criterion of generality, it may be considered as the Keynesian Demand Curve.

Linear demand curves have the additional feature that they are distorted under aggregation. The Keynesian demand curve, on the other hand, scales up without distortion, maintaining its shape from the level of the individual or firm right up to the aggregate economy. With the

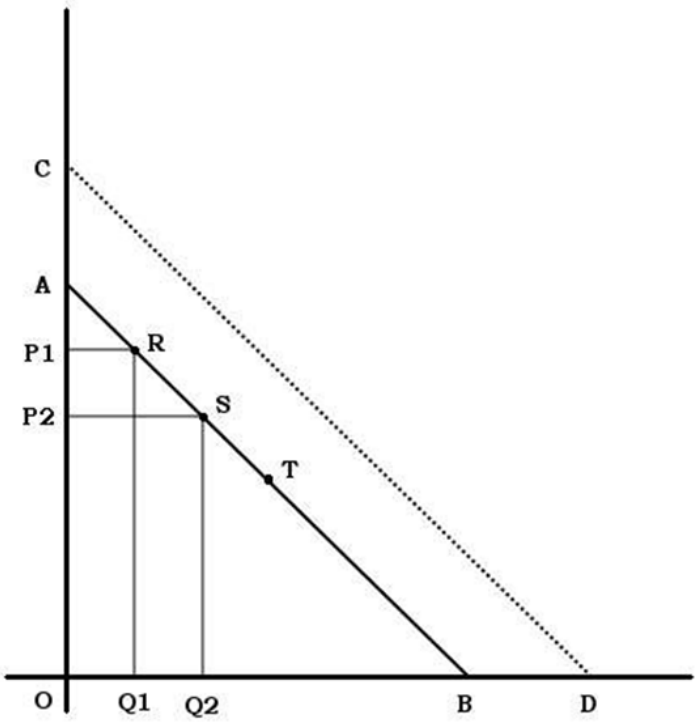
Keynesian demand curve it is therefore possible to construct the macro-economy by aggregating heterogeneous individuals and firms in the economy. My analysis thus calls into question the concept of the representative agent, which constitutes the micro-foundations of New Classical Economics.

Finally, I show how the aggregate labour demand curve, constituted by adding up the labour demand curves of all the firms in the economy, interacts with the labour supply curve to yield involuntary unemployment, a result quite at variance with other schools of macroeconomics.

A. The demand curve

The demand curve is one of the building blocks of economics. In textbooks the demand curve is sometimes depicted as a line, sometimes as a curve, sometimes both. The assumption is that the actual shape of the curve does not matter so long as the slope is negative throughout. The figure below shows the market's demand curve for fish as a line.

Figure 1 Linear demand curves for fish



The original demand curve is AB. The demand curve is drawn assuming that people's incomes are constant as are their tastes. If people's incomes increase then at every price they can buy more fish and the demand curve moves to CD. Similarly, if people develop an increased taste for fish the demand curve moves to CD.

T is the midpoint of the demand curve. The segment AT is the elastic zone. The segment TB is the inelastic zone.

Assume that the initial equilibrium is at point R (in the elastic zone) where the price is P1 and the quantity sold is Q1. Assume also that at this point individuals spend all their income and

do not save anything. Next suppose that because of a movement of the supply curve the price falls to P2, also in the elastic zone, as a result of which the equilibrium moves to S. The quantity of fish bought increases to Q2. We can also see from the graph that the money spent on fish rises; the initial amount spent is the area of O-P1-R-Q1 and the final amount spent is the area of O-P2-S-Q2.

But here we run into a problem. We had assumed that individuals spent all their income at the first equilibrium point R and now we find them spending a larger amount on fish at S. This can happen only if they spend a smaller amount on some other good or goods so as to maintain their spending constant.

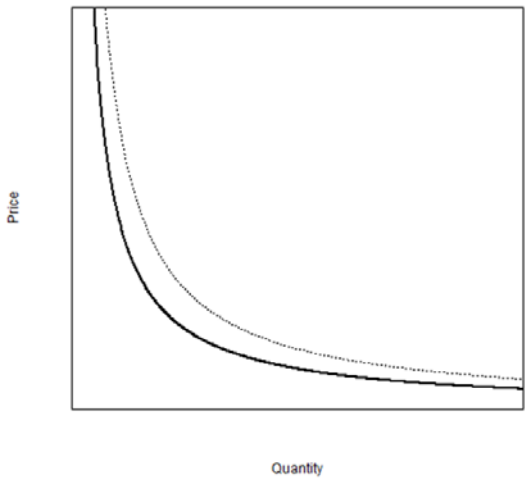
To summarize, along the elastic portion of a linear demand curve, when the price of fish falls not only does the quantity of fish bought increase but the money spent on fish also increases. So the money spent in other markets has to fall so as to maintain our initial assumption of a constant income. In general, the money spent at any point on each half of the linear demand curve is different from that spent at any other point. To compensate for this difference the money spent in other markets, and therefore the price and demand in those markets, has to change.

This property of linear demand curves is also shared by demand curves of other shapes, with a solitary exception: the rectangular hyperbola PQ = constant shown in Figure 2.

When the price of fish falls between two points on a rectangular hyperbola, the quantity of fish bought increases. But the money spent on fish at the two points remains the same since PQ is a constant. Though the quantity of fish bought rises the increase does not come at the expense of spending on any other good. Between two points on a linear demand curve spending is in general different. In a family of rectangular hyperbolas such changes of spending are effected by movements *between* demand curves, not *along* a demand curve.

The rectangular hyperbola is thus the only demand curve which does not assume that aggregate spending is constant. Keynes' claim for the *General Theory* was that unlike classical economics, it dealt with the general case in which aggregate spending was free to change. The rectangular hyperbola meets this criterion. It may thus be considered as the Keynesian Demand Curve.

Figure 2 Demand curves as rectangular hyperbolas



B. The addition of demand curves

Figure 3 shows two linear demand curves. To add two demand curves we note that prices cannot be added but quantities can. So for each price on the y-axis, we note the respective quantity for each demand curve and add them to get the quantity at that price on the aggregated demand curve. Figure 4 shows the demand curve formed as a result of the addition. The first point to note is that unlike the constituent demand curves the aggregate demand curve is discontinuous. If many linear demand curves of a variety of magnitudes are added the aggregated demand curve will be an assembly of disconnected line segments that get closer and closer to a curve. If every linear demand curve is identical to CD then the combination of n demand curves will be a line segment hinged at C. The intersection with the x-axis will be at a point whose distance from the origin is equal to n times OD

Figure 3 Two linear demand curves

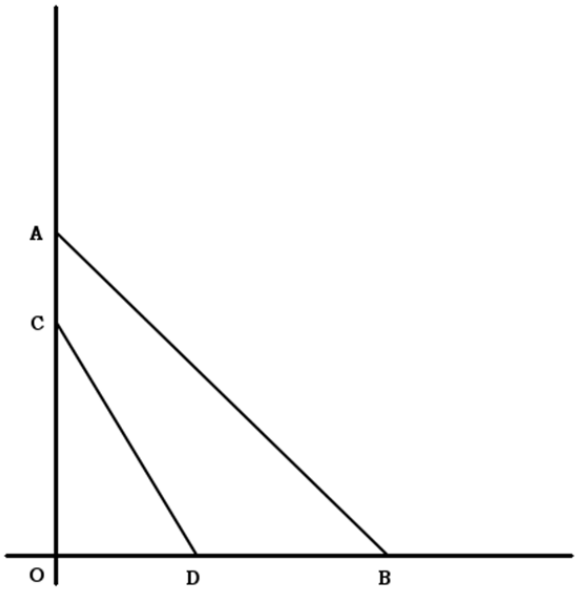
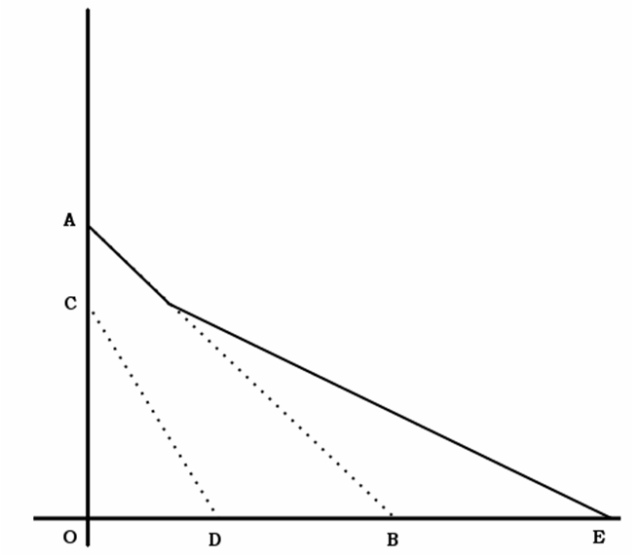


Figure 4 Addition of two linear demand curves



If the representative agent is taken as the smallest (or lowest) demand curve (which is a line) then the behaviour of the aggregate demand curve (which is a curve) will be completely different from the behaviour of the representative agent. Even if the representative agent is taken as the average demand curve the behaviour of the aggregate demand curve will be quite different. Only if the representative agent is a scaled down version of the aggregate demand curve can its behaviour be regarded as mirroring the aggregate demand curve. But in that case, it is not the aggregate demand curve which is being constructed from the representative agent but the representative agent that is being constructed from the aggregate demand curve. Or, put another way, we would be constructing the micro-foundations from the aggregate economy, not the other way round. If every linear demand curve in the economy were identical, then the aggregate demand curve would approach closer and closer to the horizontal; its behaviour would thus be completely different from that of the demand curves which constitute it.

The rectangular hyperbola, which has been previously introduced as the Keynesian demand curve, suffers from no such limitation. Rectangular hyperbolas when added yield rectangular hyperbolas. It is easy to prove this. Consider three rectangular hyperbolas PQ = a, PQ = b, and PQ = a + b. At price P1, the first rectangular hyperbola gives us Q1 = a/P1. At the same price, the second rectangular hyperbola gives us Q2 = b/P1. Adding the two gives us Q1 + Q2 = (a + b)/P1, which is of course Q3, the quantity at that price for the third rectangular hyperbola.

What this means is that no representative agent is required in order to construct the Keynesian macro-economy. Every individual's demand for fish is a rectangular hyperbola. The small consumer's demand curve will be a rectangular hyperbola close to the origin; the large consumer's demand curve will be a rectangular hyperbola farther from the origin. Adding every such demand curve will give us the market demand curve for fish, which will be a rectangular hyperbola. And adding the demand curves of every good and service, all of which are rectangular hyperbolas, will yield the aggregate demand curve, which will again have the shape of a rectangular hyperbola. The aggregate demand curve will be the farthest away from the origin.

From the arithmetic of addition used above, it should be clear that the rectangular hyperbola is not the only curve that aggregates without distortion. For example, QP^2 = constant or Q.logP = constant also aggregate to yield new curves of the same type. They do not, however, meet Keynes's criterion of generality because the product of P and Q varies from point to point along a demand curve.

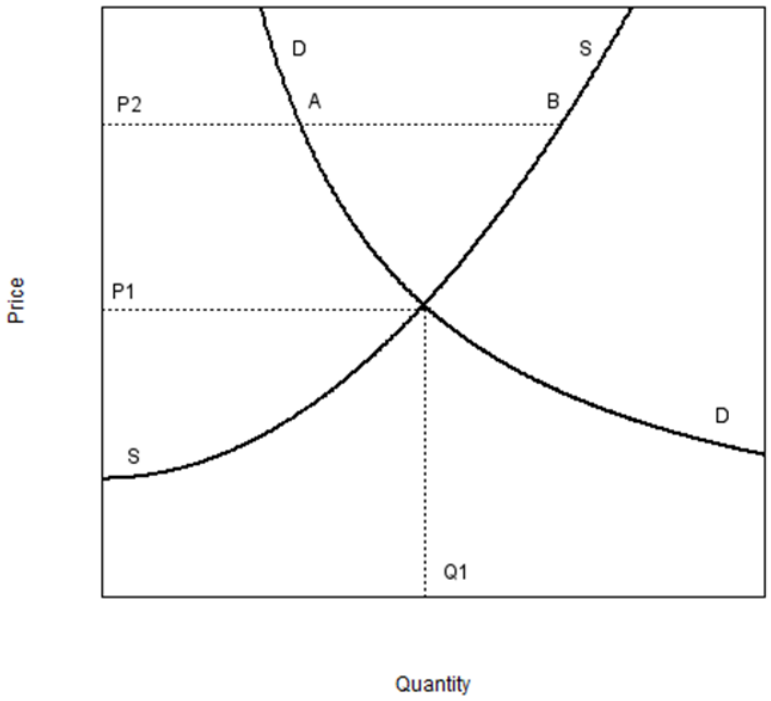
With the basic groundwork in place we now examine the classical view of unemployment.

C. The classical analysis of unemployment

Figure 5 shows demand and supply curves for labour, DD and SS respectively. The two intersect at a point where the wage is P1 and the quantity of labour demanded is Q1. This is an equilibrium point. If the wage rises to P2, the demand for labour falls and the supply of labour rises. More workers are willing to work at that wage than firms are willing to employ. The length of the segment AB is a measure of the unemployment generated at that wage. So workers reduce their wage demands until equilibrium is again reached at the original wage P1.

This is the classical analysis of the labour market. Unless a regulatory authority sets the minimum wage at P2 or trade unions set their wage demand at P2, which is higher than the equilibrium price, involuntary unemployment is impossible.

Figure 5 The wage floor

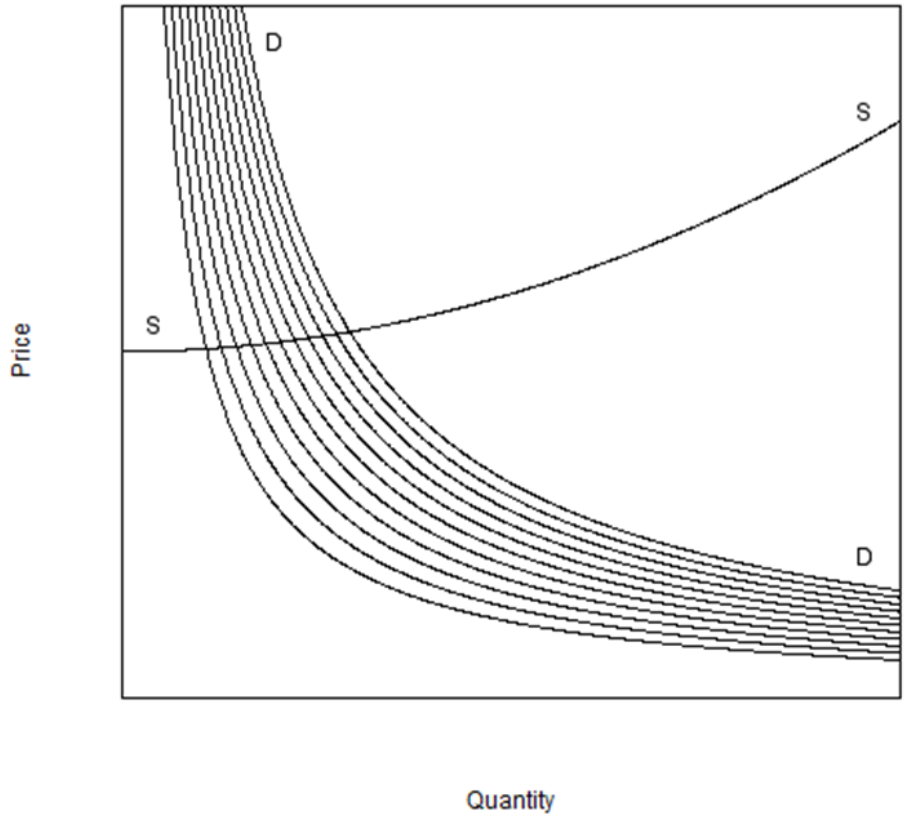


Most modern schools of economics generally accept this basic analysis. For instance, the New Keynesian school believes that sticky wages, not a wage floor per se, are the cause of unemployment. In the Stiglitz-Shapiro model (Shapiro & Stiglitz, 1984) too it is a wage above the equilibrium wage as in Figure 5 that is the cause of unemployment. In order to prevent workers from shirking firms pay them a wage that is higher than the market clearing wage. It is this differential that is the cause of unemployment during recessions. The New Classical school simply assumes that there is no such thing as involuntary unemployment; workers voluntarily choose to trade leisure for employment.

D. Involuntary unemployment resulting from the Keynesian demand curve

In section B we showed that every market in the economy from the smallest to the largest (aggregate demand) can be plotted on a single graph as rectangular hyperbolas. Small markets are located on demand curves close to the origin; large markets are located far from the origin. The labour market has one feature that sets it apart from other markets; it is by far the largest market in the economy. An idea of its size can be had from the fact that compensation paid to employees in the US amounts to about 44% of GDP at present (St Louis Federal Reserve, 2017). Labour is a necessary input in every good and service. This means that even if other markets shrink by small amounts, those small amounts are cumulative so that the demand curve for the labour market (which is a rectangular hyperbola) falls by the largest magnitude among all markets. Only the aggregate demand curve falls by a greater amount. Figure 6 depicts the fall and the interaction with a supply curve.

Figure 6 A labour supply curve cutting through a family of demand curves



It will be seen that at every consecutive intersection of the supply curve with a falling demand curve the part of the rectangular hyperbola at the intersection becomes closer and closer to the vertical.

But when the demand curve is vertical a change in wages (nominal or real) will not result in any change in employment. There is nothing that workers can do, individually or as a class, that will increase aggregate employment. The problem lies with the demand for labour and there is nothing that suppliers of labour can do to remedy the situation. By using a demand curve which does not assume that aggregate demand is constant (or, in other words, by using a Keynesian Demand Curve) we have thus arrived at a theoretical situation consistent with real-world involuntary unemployment.

Figure 6 is somewhat at variance with what Keynes himself felt about involuntary employment in *The General Theory*. In Chapter 2 he wrote:

> "Men are involuntarily unemployed if, in the event of a small rise in the price of wage-goods relatively to the money-wage, both the aggregate supply of labour willing to work for the current money-wage and the aggregate demand for it at that wage would be greater than the existing volume of employment."

There are two parts to this. In the second part Keynes argued that a fall in real wages will result in an increase in the aggregate demand for labour whereas Figure 6 suggests that when the demand curve is vertical a fall in wages will have no impact on employment demand. Three years later Keynes (Keynes, 1939) changed his mind, though with some

reservations, on the subject after studies showed that money wages and real wages tended to move in the same direction, not in opposite directions as Keynes believed. “That I was an easy victim of the traditional conclusion because it fitted my theory is the opposite of the truth,” Keynes wrote. “For my own theory this conclusion was inconvenient, since it had a tendency to offset the influence of the main forces which I was discussing and made it necessary for me to introduce qualifications, which I need not have troubled with if I could have adopted the contrary generalisation favoured by Foxwell, Mr Dunlop and Mr Tarshis.” He added in a footnote that Chapter 2 “is the portion of my book which most needs to be revised”.

If Keynes was reluctant to abandon the relationship between falling real wages and increasing employment demand it was because he believed in the correctness of what he called the first fundamental postulate of economics: “The wage is equal to the marginal product of labour.”
At first sight it indeed appears to be incontrovertible. Assume that the marginal product of labour is greater than the wage. In that case the capitalist can maximize his profit by hiring additional workers and raising output until the point that the marginal product of labour equals the wage.

Beyond that profit will fall. So the first postulate appears to be correct.

What Keynes does not seem to have realized is that his own theory disproved the first postulate. Remember that the aggregate demand curve is the only rectangular hyperbola that lies further out from the origin than the labour demand curve. During recessions it falls by an even larger magnitude than the labour demand curve. Therefore, the aggregate demand curve too is vertical at the point where it intersects the aggregate supply curve during recessions. Hence our argument that the capitalist can hire additional workers and increase output is incorrect; a vertical aggregate demand curve means that any additional output would remain unsold. The first fundamental postulate is therefore best regarded as an inequality: The wage is equal to or less than the marginal product. Note also that to disprove the first postulate all we need is the fact that aggregate demand has hit a wall. It does not matter whether the marginal product of labour is increasing or decreasing, a question which exercised Keynes a lot. We also note in passing that our argument casts doubt on the principle of profit maximization by firms.

Studies that show output and employment falling when wages are stickier than prices do not prove causation. When a capitalist finds his widgets remaining unsold (because aggregate demand is falling though he is not aware of this) he does not react by cutting wages; doing so will not help him get rid of his unsold inventory. He believes his widgets are not selling because he has overpriced them. So he first cuts prices and, perhaps, output. It is only when lower prices do not result in higher sales that he cuts jobs and wages in the hope of lowering prices further. The arrow of causation runs from falling aggregate demand to sticky wages and not the other way round.

The second part of Keynes’ definition of involuntary unemployment is interesting. It says that when there is involuntary unemployment the aggregate supply of labour will increase even when real wages are falling. The purpose of stating this appears to be to counter the argument that if unemployment is greater during a recession it has nothing to do with demand but is because workers choose not to supply their labour at the lower price or, in other words, voluntarily substitute leisure for work. This is contradicted by the fact that, as the aftermath of every recession shows, employment expands by large quantities even when real wages are

falling. See graph of median real earnings for workers, 16 years and over (St Louis Federal Reserve, 2017 (b)).

That employment rises for a long time after recessions without significant increases in the real wage suggests that the labour supply curve is horizontal during that period.

It is instructive to compare Figure 6 with Figure 5, which constitutes the classical proof that involuntary unemployment is impossible. Two differences are apparent. Figure 5 only considers movements along a demand curve whereas Figure 6 considers movements across demand curves. A classical demand curve for any good or service implicitly assumes constant aggregate demand, so it is only natural that it cannot result in involuntary unemployment, which is a feature of economies in recession. But even if Figure 5 considered a series of falling demand curves it would not result in "involuntary" unemployment if the demand curves were parallel to each other, which is how most textbook demand curves are drawn. For involuntary unemployment to occur it is necessary to have a family of curves that neither intersect nor are parallel so that the intersections of the supply and demand curves take place at points where the demand curves are closer and closer to the vertical.

We stated earlier that Keynesian demand curves are not the only curves that aggregate to give other curves of the same family. Figure 7 shows a family of curves of the kind Q.logP = constant.

Figure 7 Curves of the kind Q.logP = constant

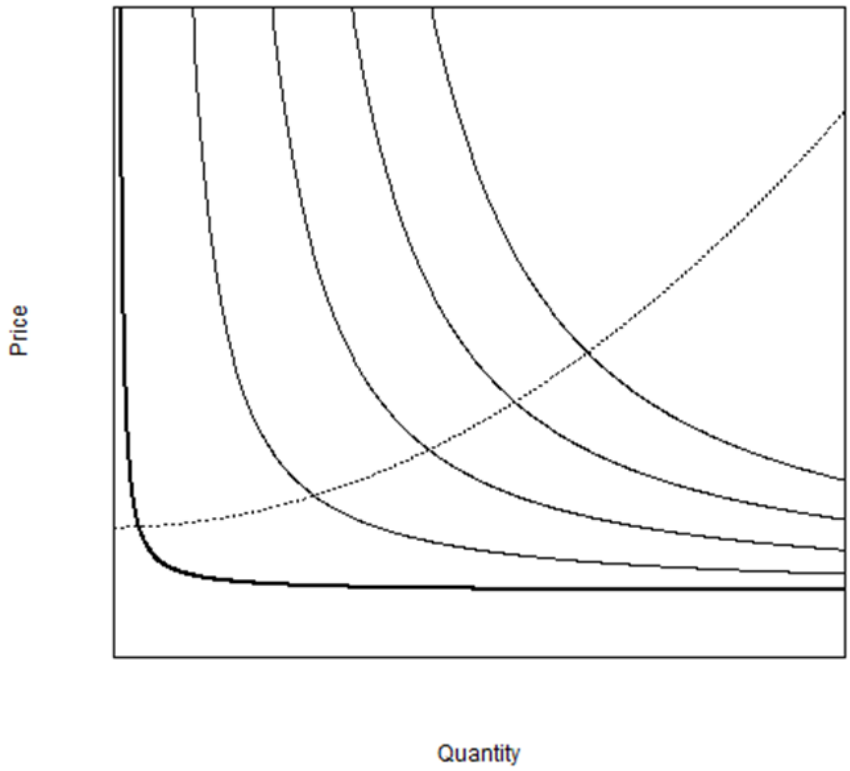
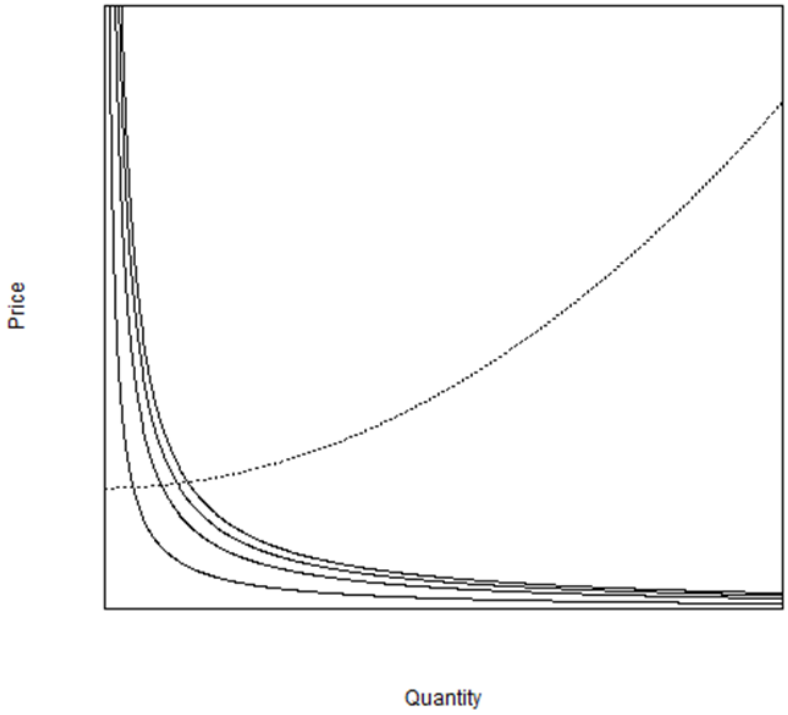


Figure 8 shows curves of the kind $Q.e^{P}$ = constant



Both families of curves on interaction with the labour supply curve show involuntary unemployment. Neither curve of course meets Keynes' criterion of generality. We depict them only to show that many curves that aggregate without distortion (and thus do not need the invention of a representative agent) result in involuntary unemployment. Demand curves that result in involuntary unemployment are quite common, not the impossibility that the New Classicals claim.

E. The minimum wage and elasticity of labour demand

Economic theory is clear that when the minimum wage is increased low-wage employment is negatively affected. Empirical evidence about the relationship is not, however, so unambiguous.

One review of a large number of studies (Neumark & Wascher, 2006) found that measured elasticities varied from well below -1 to well above zero although the authors concluded that the overall evidence supports the idea that increasing the minimum wage causes disemployment effects.

The problem with most of these studies lies in the metric used to measure the effect of a change in minimum wage on employment: the elasticity of demand for labour. It is defined as $\frac{\Delta Q/Q}{\Delta P/P}$ where P refers to the wage and Q to the quantum of employment. Since wages and employment move in opposite directions this is theoretically always negative. On the face of it using the elasticity of demand for labour to measure the effect of minimum wage changes on employment is plain common sense. The assumption is that if two studies in two separate locations show the same elasticity of demand for labour this means that employment is equally responsive to a change in the wage rate at both locations. Similarly, a higher elasticity

of demand, it is assumed, means a greater responsiveness of employment to a change in the minimum wage than a lower elasticity of demand.

But although this idea is universally accepted it is incorrect as can be seen from a look at the rectangular hyperbola. The elasticity of the rectangular hyperbola is equal to 1 everywhere. This suggests that at every point on the rectangular hyperbola, employment is equally responsive to a change in the wage rate. But we know that the rectangular hyperbola at points close to the origin along the x-axis is nearly vertical and that at points far from the origin along the x-axis it is nearly horizontal. In the nearly vertical portion a large change in the wage produces only a small change in employment. In the nearly horizontal portion of the rectangular hyperbola a small change in the wage produces a very large change in employment. So clearly elasticity is the wrong metric to measure the responsiveness of employment to a change in the wage rate. The correct measure to use is $\frac{\Delta Q}{\Delta P}$, which is the reciprocal of the slope.

In non-recessionary periods the labour supply curve cuts the demand curve at points where it is relatively horizontal. During recessions the labour supply curve cuts the demand curve at points where it is relatively vertical. The slope of the rectangular hyperbola (and thus its reciprocal) at the intersection with the supply curve traverses a range of values. This implies that the responsiveness of employment to a change in the minimum wage can vary widely, depending on when the study is performed, which is what the empirical data suggest. If studies show a wide range of effects of changes in the minimum wage it is not because some of those studies are badly designed but because reality is that way.

Interestingly, the graph says that in recessionary periods the minimum wage can be increased by a small amount without having a negative effect on employment and that in non-recessionary periods increasing the minimum wage can have deleterious effects on employment. This may appear to violate common sense. The explanation is that in non-recessionary periods, firms with very small profit margins can survive in the market. Increasing the minimum wage harms their margins and causes them to go out of business, thus contracting employment. During recessions the firms that survive are likely to have larger margins and are thus in a better position to weather an increase in the minimum wage.

Note, also, that changes in demand (measured in dollars) reveal themselves in movements of the rectangular hyperbola, not in movements along the curve. So large increases of the minimum wage may cause some firms to go out of business. This may result in the aggregate labour demand curve moving downward. But equally, small increases in the minimum wage can result in the demand curve moving up because of positive changes in income and spending and thus in aggregate demand.

Conclusion

In this paper we showed that the rectangular hyperbola is the only geometrical demand curve that meets the Keynesian condition of generality, i.e. it does not assume aggregate demand is constant. Moreover, it has the interesting property that it aggregates without distortion. Thus, it is possible to construct the Keynesian macro-economy by aggregating the demand curves of every agent in the economy, whether consumer or firm. A representative agent is therefore

rendered unnecessary. Finally, when this family of Keynesian demand curves for labour, which happens to be the largest market in the economy, interacts with the supply curve for labour involuntary unemployment follows as a matter of course. Most other goods and services constitute smaller markets, and in their case, the Keynesian analysis approximates to the Marshallian demand analysis.

Humans tend to view reality through the lens of theory. Before Copernicus it was evident to every child that the sun revolved round the earth. After the Copernican theory gained acceptance it has been apparent to every child that the earth revolves around the sun.

Before Keynes every economist was convinced that there was no such thing as involuntary unemployment. After Keynes, for some decades at least, there really seemed to be such a thing as involuntary unemployment. After the rise of New Classical Economics it was again obvious to most economists that involuntary unemployment did not exist. This paper, which demonstrates involuntary unemployment in theory, will, it is hoped, again change the way that economists view reality.

References

Keynes, J M (1936) "The General Theory of Employment, Interest and Money", London, Macmillan

George, Philip (2016), "The Mathematical Equivalence of Marshallian Analysis and 'General Equilibrium' Theory", *real-world economics review*, December, pp 73-84, http://www.paecon.net/PAEReview/issue77/George77.pdf

Shapiro, Carl and Stiglitz, Joseph (1984) "Equilibrium Unemployment as a Worker Discipline Device", *American Economic Review*, June

St Louis Federal Reserve Bank (2017), Graph of ratio of labour compensation to GDP in the US, https://fred.stlouisfed.org/graph/?g=75e6

Keynes, J M (1939) "Relative Movements of Real Wages and Output", *The Economic Journal*, Vol 49, No 193, March, pp 34-51

St Louis Federal Reserve Bank (2017, b), Graph of median usual weekly real earnings: Wage and salary workers, 16 years and over, https://fred.stlouisfed.org/series/LEU0252881600Q

Neumark, David and Wascher, William (2006, revised 2007) "Minimum Wages and Employment: A Review of Evidence from the New Minimum Wage Research", National Bureau of Economic Research, Working Paper 12663, http://www.nber.org/papers/w12663.pdf

Author contact: philip.george@gmail.com

SUGGESTED CITATION:
Philip George, "A diagrammatic derivation of involuntary unemployment from Keynesian micro-foundations", *real-world economics review*, issue no. 80, 26 June 2017, pp. 109-120,
http://www.paecon.net/PAEReview/issue80/George80.pdf

real-world economics review, issue no. 80
subscribe for free

Asymmetric price adjustment: the missing link in Keynesian macroeconomics[1]*

Victor A. Beker [University of Belgrano and University of Buenos Aires, Argentina]

Abstract

In the real world, prices do not behave symmetrically. Usually, nominal wages and prices are sticky downward but a lot more flexible upward; the latter is illustrated by inflationary and hyperinflationary processes.

However, most mainstream economics is built upon the assumption that nominal prices are equally flexible *in both directions*. This leads to quite unrealistic and erroneous predictions as far as downturn in economic activity is concerned.

Given price asymmetry, it is necessary to do separate analyses: on one hand, full-employment macroeconomics (price equilibrium macroeconomics) and, on the other, the macroeconomics of recession and depression (Keynesian macroeconomics). Prices play a role in the first case, but not in the second one.

The present paper aims at pointing out the need for reconstructing macroeconomics from a realistic point of view. It argues that price downward stickiness must be a fundamental assumption in any economic model which tries to explain and predict real-world market behavior as well as recommend economic policies. It also claims that Keynesian macroeconomics has to be the point of departure of a realistic reconstruction of macroeconomic theory. Finally, it maintains that price downward rigidity fits perfectly well the Keynesian model, while this does not happen with either the New Keynesian or the Post-Keynesian models.

JEL B22 B41 E12

Keywords Keynes; price asymmetry; realisticness; unemployment

1. Introduction

> "Most macroeconomics of the past 30 years was spectacularly useless at best, and positively harmful at worst" (Paul Krugman, *The Economist*, July 18-24, 2009: 58).

In the real world, prices do not behave symmetrically. Usually, nominal wages and prices are sticky downward but a lot more flexible upward; the latter is illustrated by inflationary and hyperinflationary processes.

> "As documented by many authors for many countries (e.g. Cover, 1992), positive demand shocks give rise to inflation without affecting output significantly, while negative ones reduce output without affecting inflation" (Dobrynskaya, 2008).

[1] With this article I try to pay homage to the late Argentine economist Julio Olivera who passed away in 2016. From him I learned the critical role that price asymmetry plays in today's real world economy.
* This paper benefited from helpful comments by an anonymous reviewer.

In fact, examining quarterly U.S. post-war data, Cover (1992) concluded that positive shocks in the money supply have had no effect on output, whereas negative shocks reduced output. De Long and Summers (1988) had reached similar conclusions in their investigation of annual pre- and post-Second World War U.S. data. The same results were obtained by Rhee and Rich (1995) and Karras and Stokes (1999) for European countries, by implementing the method regarding asymmetry first introduced by Cover (1992).

After studying over 240 markets for consumer and producer goods, Peltzman (2000) concluded prices rise faster than they fall, and price asymmetries are persuasive, substantial, and durable and exist in periods of low inflation and high inflation. The author asserted that

> "This asymmetry is fairly labeled a 'stylized fact'. This fact poses a challenge to theory. The theory of markets is surely a bedrock of economics. But the evidence in this paper suggests that the theory is wrong, at least insofar as an asymmetric response to costs is not its general implication" (Ibid.: 493).

So, it is a very well-established empirical regularity that positive shocks have a tiny effect on output and they basically pass to prices, while negative shocks are, to a larger extent, passed to output. Nominal prices are sticky-down and the main effect of a negative shock is absorbed by output.

However, most mainstream economics is built upon the assumption that nominal prices are equally flexible in both directions. This leads to quite unrealistic and erroneous predictions as far as downturn in economic activity is concerned. In fact, what we usually see – and empirical evidence corroborates – is that, in the presence of a negative shock, quantities – not prices – fall. However, if nominal prices are downward rigid, they cannot clear markets in the presence of an excess supply, as is usually assumed in mainstream economics.

Given price asymmetry, it is necessary to do separate analyses: on one hand, full-employment macroeconomics (price equilibrium macroeconomics) and, on the other, the macroeconomics of recession and depression (Keynesian macroeconomics). Prices play a role in the first case, but not in the second one; the mechanisms at work in the first case are quite different from the ones in the second case.

Since the 1970s, macroeconomic theory has been developed under the implicit assumption that the main economic problem is inflation. However, unemployment is now back and it should be given at least the same importance as inflation received in the research agenda during the last 50 years.

The present paper aims at pointing out the need for reconstructing macroeconomics from a realistic point of view. Why do we need realism in economics? There has been a long-lived discussion on the subject of realism in economics. Its main milestones have been Milton Friedman's contribution in his 1953 essay "The Methodology of Positive Economics" and Paul A. Samuelson's response 10 years later in "Problems of Methodology: Discussion".

In their famous manifesto from June 2000, the Parisian graduate students who led the struggle against "autistic science" complained, "This disregard for concrete realities poses an enormous problem for those who would like to render themselves useful to economic and social actors."

The main fact that must be taken into consideration is that economics, and especially macroeconomics, are supposed to be a guide for economic policy. This is the context in which the issue of realism must be discussed. Let us give a simple example of what happens when we use unrealistic assumptions.

If we assume that lions are herbivorous, we will predict that any human being will be safe in the presence of a lion. Unfortunately, for the human beings in the real world, lions are carnivorous. That assumption may be useful to depict what an ideal world of peaceful coexistence between both species would be, but it is a very dangerous guide for human being action.

In the same way, many assumptions in mainstream economics are adopted only because they facilitate the analytical treatment of the problem; in some others, it is just because of elegance.[2] As with the case of the lion, those assumptions unfortunately lead to predictions far away from what happens in the real world and with equally tragic consequences, as the 2008 crisis pointed out. Mainstream economic theory did not even consider the possibility of the type of collapse that the subprime mortgage meltdown unleashed, let alone the appropriate ways to deal with it.

It is true that the premises of any economic model are always "inexact" because they abstract of numerous causal factors that are present besides those effectively taken into consideration. This is the method used in all sciences when there is a complexity of causal factors (Hausman, 1992: 148).

As Mäki (2005: 304) rightly states,

> "…models serve as 'substitute systems' of the target system they represent. They are substitute systems in the sense that one does not directly examine the target systems, rather one focuses on the properties and behavior of the representatives as substitutes of the targets."

The target system is too complex to be understood in its entirety, so a simpler model is constructed to explore it (Hodge, 2007: 26). However, not just any substitute system will do; the representative model must adequately resemble the target system where "adequately" depends on the intention or purpose of the model (ibid). As Mäki asserts, "thought experiments" replace the "material experiments" of the natural sciences.

Therefore, by definition, every model implies a certain degree of unrealism in its assumptions – it is a simplification of the real world. But it is one thing to simplify reality and quite another to overtly tergiversate it. The representative model must resemble the target system.

I have referred elsewhere (Beker, 2015) to several of the unrealistic neoclassical assumptions (price symmetric flexibility, unbounded rational expectations, no coordination problems, the representative agent, *etcetera*).

[2] Of course, there is also an ideological component. As De Vroey (2011: 7) points out, "The split is between those who want to give competition its full rein, the defenders of the self-regulating characteristics of markets (or 'free marketers') and those, the Keynesians, who think that the market economy, although the best economic system, can buttress failures, in particular an insufficiency in aggregate demand, which it is the state's role to remedy."

This time, I deal with the price flexibility assumption. This paper argues that price downward stickiness must be a fundamental assumption in any economic model which tries to explain and predict real-world market behavior as well as recommend economic policies.

I go on arguing that realisticness is precisely one of the features that distinguishes Keynes' unemployment analysis. For this reason, Keynesian macroeconomics has to be the point of departure of a realistic reconstruction of macroeconomic theory. Keynes' economic theory was designed to explain the causes of the 1930s economic crisis and fight against its consequences. In principle, this theory seems better equipped to deal with subjects like unemployment, recession, and depression than the neoclassical one, which does not even consider a deep crisis as a possibility.

I add that price downward rigidity fits perfectly well the Keynesian model, while this does not happen with either the New Keynesian or the Post-Keynesian models.

The rest of the paper is organized in the following way. Section 2 is devoted to arguing that in the real-world price behavior is not at all symmetric. In Section 3 the superiority of Keynesian economic theory is defended because of the realisticness of its assumptions. Section 4 is devoted to presenting the essence of the Keynesian model. In Section 5 a distinction is introduced between price and non-price equilibrium economics. Section 6 and Section 7, respectively, present the New Keynesian and Post-Keynesians contributions and the reasons why they should not be considered truly Keynesians. In Section 8 the microeconomics behind the Keynesian macroeconomic model is introduced. Section 9 discusses why prices do not fall when there is excess supply. In Section 10 I introduce what I call the Fundamental Microeconomic Assumption. Section 11 is devoted to the analysis of price asymmetry on the wealth effect. Section 12 reviews the discussion on the Phillips curve, while Section 13 is devoted to the analysis of the impact that the appearance of the phenomenon of stagflation in the 1970s had on economic theory. Section 14 makes a short review of Keynes' monetary and fiscal policies. Section 15 briefly discusses price stickiness in the European crisis. Section 16 concludes.

2. Price symmetry is not a realistic assumption

In the real world, price behavior is not at all symmetric. It is a curiosity that even Friedman (1953: 165) recognized that asymmetry: "At least in the modern world, internal prices are highly inflexible. They are more flexible upward than downward." He did it in the context of an argument in favor of flexible exchange rates but he ignored it when dealing with goods market clearing.[3]

Asymmetry in price behavior requires one approach to analyze an economy in full-employment and another one to study an economy facing unemployment.[4] We need an economic theory that can explain involuntary unemployment but which, at the same time, allows for the existence of inflation.

[3] Had he taken into consideration downward stickiness in his monetary analysis he would have arrived to the non-monetary theory of inflation, developed by Latin American economists using precisely that assumption to arrive at a theory of inflation opposed to Friedman's one.
[4] This is something I have been insisting since long ago. See Beker (1985).

In traditional mainstream economic theory there is no room for involuntary unemployment. As wages are assumed to be downward flexible any labor market surplus will be removed by a fall in nominal wages.

3. Realisticness and Keynesian macroeconomics

Realisticness is precisely one of the features that distinguished Keynes' analysis. Keynes was a practical-minded economist. In contrast to many past and present economic theorists, he had great practical experience in economic policy. He did use simplifications of economic reality–the propensity to consume is one of them–but they allowed him to reach significant practical results. Recovering the original Keynes' legacy and pointing out its relevance to deal with current economic problems is the starting point for a realistic economic theory. Let me emphasize: I'm just talking of the point of departure. Of course, many things have changed since Keynes published his General Theory in 1936, and there are also some gaps in his reasoning that need to be filled.[5] In this paper I just want to argue that Keynes' approach to macroeconomics is still a relevant model of how to deal with economic issues. Taking it as a departure point, a research program has to be developed in order to update Keynesianism to the 21st-century realities.[6] But as in any building, foundations play a decisive role. I argue here about the foundations of macroeconomics and, particularly, on the role of the asymmetric price behavior in them.

While Keynesian macroeconomics considers full employment as a particular and unusual case, mainstream neoclassical macroeconomics deals only with full-employment macroeconomics. Many of its predictions and recommendations are flawed because they are based on the assumption that prices are downward flexible. They resemble the recommendations made for a world where lions are herbivorous.

4. Keynesian macroeconomics

Let us make a quick review of Keynes' ideas. This is necessary in order to distinguish what Keynes really contributed to the economic analysis from what the anti-Keynesians and the different kinds of "Keynesians" interpreted. The starting point of Keynesian macroeconomics is that the labor market does not necessarily clear. There is no self-adjusting mechanism in the labor market that ensures full employment. This has been the key contribution from Keynes. The most likely situation in the labor market is one of involuntary unemployment, where labor supply exceeds labor demand.

Involuntary unemployment has nothing to do with real wage rigidity; for Keynes there is no real wage rigidity. On the contrary, as Keynes argued, although workers will usually resist a nominal wage reduction, they will not resist moderate reductions in real wages because of an increase in prices (Keynes, 2016: 13). While nominal wages are downward rigid, real wages are flexible.[7]

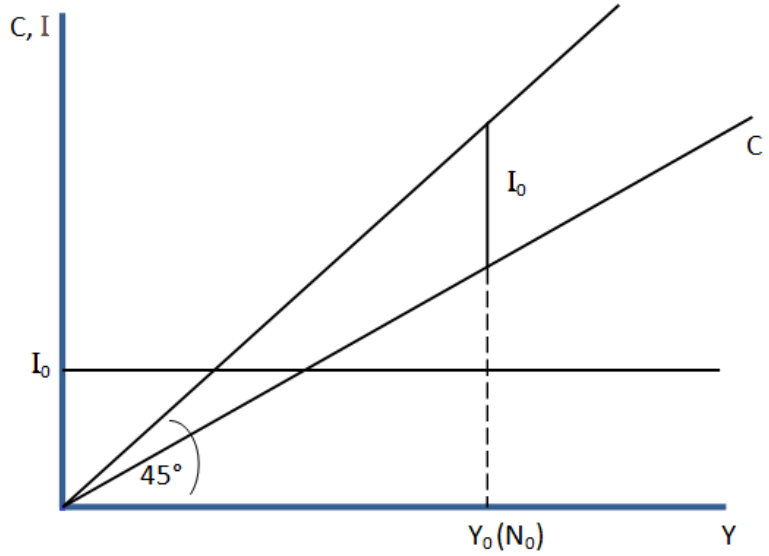
[5] Some of them are mentioned in Section 14.
[6] The globalization and financialization processes as well as the role of supranational institutions are some of the new realities an updated macroeconomics has to take into consideration.
[7] For a discussion of this issue, see Beker (2015).

The huge fluctuations in employment studied by Keynesian macroeconomics are related to fluctuations in the level of output, not with the level of real wages.

In the same way, Keynes also disregarded the role of prices in eliminating any discrepancy between aggregate supply and demand. The equilibrium[8] in the goods market is attained when demand (consumption plus investment) equals aggregate supply. If there is a general glut, firms would reduce their supply (and employment), not prices, until equilibrium is reached.

Investment (I_0 in Figure 1) plays a key role in determining the level of employment. Employment is determined in the goods market at the intersection point between the aggregate supply and aggregate demand for goods (Y_0).

Figure 1 Equilibrium in the goods market



Fluctuations in investment are responsible for fluctuations in aggregate output (Y) and thereby in employment.

In short, the aggregate demand function is

$$\mathbf{D(N) = C(N) + I} \qquad (1)$$

where **N** is the level of employment, **C(N)** is consumption, and **I** is investment. The equilibrium in the goods market requires excess aggregate demand to be zero at some level of employment:

$$\mathbf{D(N) - Y\,(N, K_o) = 0} \qquad (2)$$

[8] I use equilibrium with the usual meaning: a state where there is no economic force which can change the value of economic variables.

where $\mathbf{Y\ (N,\ K_o)}$ is the aggregate supply function for a given level of capital stock $\mathbf{K_o}$. So, employment is determined as the inverse of the excess demand function for given values of investment, namely the exogenous variable:

$$\mathbf{N} = g\ (\mathbf{I},\ \mathbf{K_o}) \qquad (3)$$

Given the organization, equipment, and technique of production, employment is a function of the level of investment. Unfortunately, no mechanism guarantees the level of investment will be the one that leads to full employment.

If $\mathbf{Q} = h\ (\mathbf{N},\ \mathbf{K_o})$ is the aggregate production function and $\mathbf{Y} = \mathbf{Q} \times \mathbf{p}$ where $\mathbf{p}$ is the general level of prices, the real wage is given by:

$$\mathbf{w/p} = \mathbf{Q_N}(\mathbf{N},\ \mathbf{K_o}) = h_N\ (\mathbf{I},\ \mathbf{K_o}) \qquad (4)$$

where $\mathbf{w}$ is the nominal wage and $\mathbf{Q_N}(\mathbf{N,\ K_o})$ is the marginal productivity of labor for a given level of capital $\mathbf{K_o}$.

The real wage rate is a function of the level of employment or, ultimately, of the level of investment.

5. Non-price equilibrium economics vs. price equilibrium economics

In the previous section, no role was ascribed to prices in reaching equilibrium. In orthodox economics they are the magic instrument that clears markets.

In Keynesian macroeconomics, quantities (income, consumption, investment, savings, et cetera) are related to other quantities, while the role of prices is de-emphasized.

For example, in the Keynesian model, a decline in the investment goods demand would have a direct impact on the level of the aggregate output *via* the multiplier. The reduced level of investment will equal the level of saving at a lower rate of interest. The fall in the rate of interest will only have a second-order effect on the level of consumption, if any.

On the contrary, in the neoclassical model the fall in investment would be followed in the first place by a decline in the rate of interest, which will stimulate consumption. Therefore, the level of aggregate demand will remain unchanged: only its composition would change.

Phelps (1970) and Lucas (1976) introduced the need for proper microfoundations in macroeconomics. By that they understood the use of the Walrasian microeconomics from the *Arrow* and *Debreu model,* where prices are the tool that clears all markets including the labor market. The Walrasian general equilibrium model is the cornerstone of mainstream economics which, for that reason, should be called *price* equilibrium economics – not just equilibrium economics, as Lucas baptized it.

Once the price clearing markets assumption is introduced in macroeconomics the possibility of involuntary unemployment disappears. Excess labor supply will push wages down until unemployment vanishes.

The main difference between the Keynesian and the Walrasian approach is that adjustments are made *via* quantities in the former, while they are made through prices in the latter.

If you accept the Walrasian approach, prices clear all markets. If so, wage/price stickiness seems to be the only line of defense available to justify Keynesian unemployment in a Walrasian context; the New Keynesians resorted to it.

6. The New Keynesian contribution

The New Keynesian program was mainly interested in proving the non-neutrality of money in the short run. If money is not neutral, its expansion or contraction will have an impact on output, at least in the short run; then, there is a role for monetary policy. This is the point New Keynesians were interested in making. For this purpose, the New Keynesians needed a microeconomic model in which prices would not respond to excess supply or demand. Their starting point had to be imperfect competition, which implies that firms set prices and the demand chooses quantities. Thus, changes in demand always cause changes in output in the same direction.

They had empirical support in Blinder et al (1998). Blinder and his colleagues interviewed, between 1990 and 1992, two hundred randomly selected firms about their pricing behavior. They collected “what may be the first evidence on price stickiness ever derived from a random sample of the whole economy” (Blinder, 1994: 120). Moreover, this stickiness did not appear to be asymmetric.

From the theoretical point of view, there were several theories arguing why prices might be sticky. For instance, Ball and Mankiw (1994) use a menu cost model to explore a possible explanation for price adjustment asymmetry. In order to sell their products, firms have to write prices on menus, catalogues, and tags; changing prices may be rather costly. If we assume positive trend inflation, a firm that wants to lower its relative price may save the menu costs just waiting for inflation to do the work. However, as Mankiw (1985) himself had already admitted, these menu costs are small and, therefore, they provide a very weak foundation for fixed-price models.

It is true that menu costs are usually small but it is also true that in an inflationary context, inflation may itself adjust relative prices. The real argument for price asymmetry is not menu costs but the fact that if you need to reduce a relative price, you can rely on inflation to do it. You do not need to change the price: time will do the job. With trend-expected inflation, prices are sticky when a firm's optimal price falls. Note that it is the presence of trend-expected inflation that generates the asymmetric price response. This is an important reason for price adjustment asymmetry in the real world.

Another New Keynesian argument is coordination failure (Stiglitz, 1999; Cooper and John, 1988; Ball and Romer, 1991). Stiglitz argues that the risks associated with wage and price adjustments may well be larger than those associated with output adjustments, at least for goods that could be stored. Stiglitz develops this and other arguments, but all of them have to do with symmetric rigidities. The same can be said about the article by Cooper and John as well as about Ball and Romer’s paper that unify the coordination failure and the menu costs approaches.

With Ball and Mankiw's article as the only exception, most of the New Keynesian contributions tried to identify reasons for the lack of response of prices both up- and downward. In their models prices are sticky and quantities are flexible both up- and downward. The New Keynesian point of view is summed up by Ball et al. (1988: 12): after recognizing that "traditional Keynesian models often imply asymmetric effects of demand shifts" they argue that "asymmetric effects of shocks could arise from asymmetric price rigidity – prices that are sticky downward but not upward – but this is another appealing notion that is difficult to formalize."

However, we need a model where prices do not respond to excess supply, although they may well respond to excess demand; this was the original Keynesian assumption and it provides a better approximation as to what happens in the real world. We need a model that is apt to explain involuntary unemployment but allows for the existence of inflation and hyperinflation. It is true that Blinder et al. did not find asymmetric stickiness.[9] This might be true for environments of very low inflation, but it seems rather difficult to make sluggishness in upward price adjustment compatible with medium and high inflation. At best, to assume symmetric stickiness may be valid for only the particular case of very low inflation, but in this case the difference between flexibility and stickiness blurs.[10]

New-Keynesian economics was supposed to be the answer to the Lucas critique. But "New-Keynesian economics is the art of finding Keynesian results in a New-Classical framework" (Melmies, 2008: 4). The difficulty is that once you take New-Classical assumptions you get New-Classical conclusions.

This is what happened to the New-Keynesian program. It devoted a lot of effort to finding microeconomic foundations for wage and price rigidity. However, strictly speaking, unemployment due to rigid wages is the (classical) voluntary kind of unemployment. A reduction in real wages will reduce/eliminate the kind of unemployment found in New Keynesian models; it has nothing to do with Keynes' involuntary unemployment. Let us remember Keynes' definition of involuntary unemployment:

> "Men are involuntarily unemployed if, in the event of a small rise in the price of wage-goods relatively to the money-wage, both the aggregate supply of labor willing to work for the current money-wage and the aggregate demand for it at that wage would be greater than the existing volume of employment" (Keynes, 2016: p. 14).

Therefore, in the Keynesian framework, involuntary unemployment persists even if real wages are reduced, which is contrary to the New-Keynesian conclusion. Unemployment in New Keynesian models is not at all Keynesian. Involuntary unemployment means, by definition, a non-optimizing behavior; that is why it is called "involuntary". New Keynesians tried to show that unemployment is compatible with optimizing behavior, while Keynes was talking about a non-optimizing behavior.

[9] The interview method used by Blinder and his colleagues has two difficulties. First, answers can be influenced by the precise wording of the questions, and secondly, people may have no incentive to respond truthfully or thoughtfully.
[10] Using data from Australia and US, De Abreu Lourenco and Gruen (1995:16) find that "the inflationary impact of relative price shocks depends strongly on expected inflation. When expected inflation is high, a rise in the economy-wide dispersion of shocks is inflationary in the short-run. By contrast, when expected inflation is low, a rise in the dispersion of shocks has minimal impact on inflation."

The New Classical answer to the New Keynesian arguments was straightforward: remove rigidities and you will have no unemployment. In one way or the other, unemployment is voluntary. In Lucas's words, "there is also a voluntary element in all unemployment, in the sense that however miserable one's current work options, one can always choose to accept them" (Lucas, 1978: 354).

Unemployment is, in this context, voluntary by definition: it exists only because there is some constraint which avoids real wages to find their equilibrium levels. Of course, this means that, for example, you have to call unemployment rates of 25%, as Spain and Greece witnessed, "voluntary".

Then, if you do not feel comfortable calling a 25% rate of unemployment "voluntary", you had better get rid of the wage market clearing assumption. If you want to get involuntary unemployment, you cannot assume that the labor market clears. This is what Keynes did. As we have seen, in the Keynesian model the level of employment depends on the goods market, where the volatile, and thus determinant factor, is investment. Whenever there is a deficiency in investment – and hitherto in aggregate demand – there will be unemployment. There is no force in the economy that pushes the aggregate demand to its full-employment level. In Keynes' words

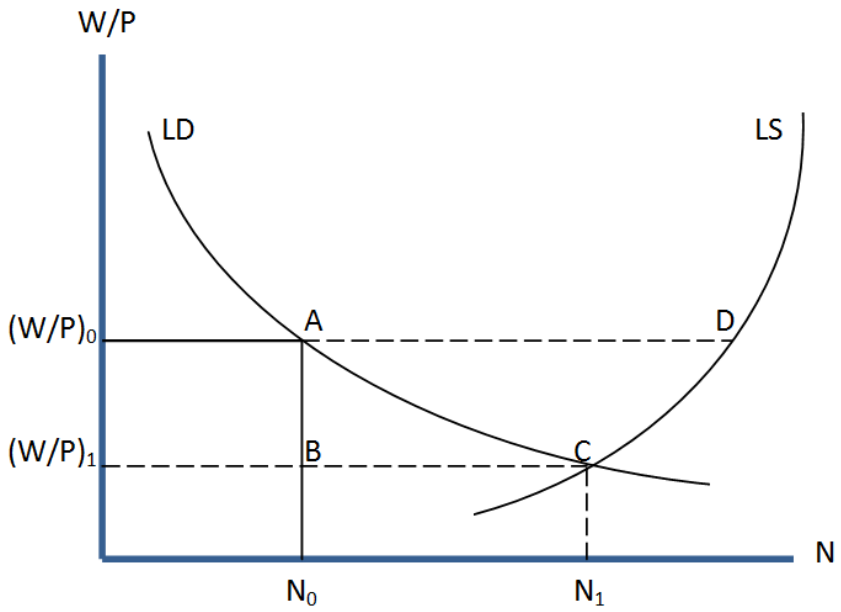
> "The propensity to consume and the rate of new investment determine between them the volume of employment, and the volume of employment is uniquely related to a given level of real wages – *not the other way around"* (ibid: 27, emphasis mine).

Thus, there is no self-adjusting mechanism in the labor market that ensures full employment. In the Keynesian model, it is not true that real wages and the level of employment are determined by the intersection of the labor demand with the labor supply. The level of employment (N_o in Figure 2) and the marginal productivity of labor $\mathbf{Q_N}$ define an equilibrium point *on* the labor demand schedule (*A*). Involuntary unemployment is measured by the distance *AD*. Workers earn a real wage $(w/p)_o$ which equals the marginal productivity of labor, as shown in equation (4) above, but it does not necessarily equal the marginal disutility of labor.

The mainstream economics argument is that in such a situation wages would fall under the pressure of excess supply. In fact, in the labor market there will be excess supply – there is involuntary unemployment measured by *AD* – but even if real wages fall – for example, from $(w/p)_o$ to $(w/p)_1$ – the volume of employment will not increase; point *B* depicts this situation but it is not an equilibrium point because firms are prepared to pay higher real wages $(w/p)_o$ for that level of employment.[11] *A* is an equilibrium point, but neither *B* nor *C* are. If there has been no change in the goods market, there is no reason why firms should hire N_1 instead of N_o, whatever the real wage is. There are no economic forces at work to drive the labor market to *C*.

[11] Unless the decline in the real wage is a consequence of a downward shift in the demand for labor curve, in which case B would be an equilibrium point and *BC* the involuntary unemployment at the new real wage level.

Figure 2 The level of employment in the Keynesian model



7. The Post-Keynesian approach to price behavior

Post-Keynesians economists offer different microeconomic foundations to Keynesian macroeconomics. They argue that prices are sticky because firms prefer stable prices and that is why they do not react to changes in demand. Firms act in a world of fundamental uncertainty, and they want stable prices to cope with that sort of world. "In 'Post-Keynesian markets,' price rigidity comes from price stability which is desired by firms on decentralized markets" (Melmies, 2010: 14). Firms prefer price stability to maximizing profits. "Firms set prices they keep unchanged for a certain period. If costs increase during this period, profit margins will decrease" (ibid.). It is output rather than price which fluctuates over the cycle.

Prices reflect both production costs at the normal level of output and the demand for retained profits to finance the planned level of investment expenditure. The latter determines the size of the mark-up or margin on costs.

Although there are different versions of Post-Keynesian price theory the essential idea is that "firms fix prices based on some measure of costs, rather than as a reaction to demand fluctuations" (Lavoie, 2001: 21)

For our purposes – which have to do with how to model price behavior in macroeconomics – this argument shares the difficulty pointed out above: prices are sticky both up and down. But sticky prices never prevented inflation from happening in the real world. And Keynes never ignored inflation as a possibility. Therefore, the Post-Keynesian approach is, in this respect, subject to the same critique as the New Keynesian one, although it has been conceived as an alternative to it.

8. How do markets reach a non-price equilibrium?

Can markets reach their equilibrium without the intervention of a price mechanism, as we have suggested in Section 5?

Let us start with the aggregate demand and supply functions. The propensity to consume and the amount of investment determine between them the aggregate demand.

Let us start with investment. The amount that firms decide to invest in a given year determines the amount of capital goods that they demand and suppliers sell in that year. If the planned supply of capital goods exceeds that amount, this oversupply will remain unsold in the short run.

The consumption component of aggregate demand, as we have seen above, is a function of income. Given the amount of consumption goods demanded, any excess supply will remain unsold, as in the capital goods case.

In the long run, some firms will downsize and others will shut down; this process will go on until planned supply equals demand. In the long run the excess supply disappears. Of course, those firms that remain in the market are those that are profitable.[12]

So, in the short run there will be two sorts of markets: those where demand equals supply, as taught in the textbooks, and those where supply exceeds demand; in the latter case, price downward rigidity means there is no force that can remove the excess supply. It will disappear only in the long run.

Formally, calling **z** the vector of excess demands, in the short run $\mathbf{z} \leq 0$, where $z_i = 0$ for $i = 1 \ldots m$ and $z_i < 0$ for $i = m+1 \ldots n$. This is because $dp_i/dt \geq 0$ for $i = 1 \ldots n$. In the long run, $\mathbf{z} = 0$ because quantities adjust until the excess supply is eliminated.

Some scholars may miss the aesthetics, beauty, and elegance of the Walrasian general equilibrium approach, where the symmetric adjustments of prices do all the magic. But are these aesthetics, beauty, and elegance enough to adopt an approach that has no empirical content?

Other economists may feel uncomfortable because they may interpret that, in Barro's words, we are leaving opportunities for mutually desirable trades or, as Lucas often repeated, we are leaving a $500 note on the sidewalk. Well, this is the meaning of involuntary: the economy lacks the necessary investment to attain full-employment equilibrium. The more than 6 million unemployed in Spain in 2013 were leaving many $500 bills on the sidewalks but a fence prevented those people from collecting them. The name of that fence was deficient demand.

9. Why prices do not eliminate the excess supply?

If planned output exceeds effective demand, why is it that prices do not go down until markets clear?

[12] This is the general rule in the capitalist world. If costs systematically exceed income firms cannot survive in the long run, unless they are state-owned or state-backed.

First of all, let us recall the distinction introduced by Okun between auction and customer product markets. In auction product markets, prices clear competitive markets as the Walrasian general equilibrium model assumes. In costumer product markets, firms set prices that do not necessarily equate demand and supply.

With a few exceptions (commodity and asset markets), in the XXI century real-world customer markets are absolutely predominant. Therefore, there is no reason to base the edifice of microeconomics on the assumption that markets behave as competitive auction markets when they are the exception rather than the rule. It was reasonable in the XIX century, but not now.

What microeconomic theory did Keynes have in mind when he wrote the *General Theory*? The *General Theory* is held to be compatible with both perfect and imperfect or monopolistic competition (Davidson, 1962; 2002). Indeed, in the *General Theory*, there is no explicit reference to the underlying assumption about the degree of competition. However, in a later paper, Keynes (1939: 46-50) admitted that perfect competition was not a realistic hypothesis and accepted imperfect competition as the benchmark to analyze what he called "the modern quasi competitive system" (ibid: 46).

Getting back to the question asked at the head of this Section, there are many arguments that explain the aforementioned behavior at the firm's level once one leaves the golden realm of perfect competition and enters the intricate labyrinths of the real world. Let us list some of them.

1. If goods are heterogeneous and each supplier faces an inelastic demand function, there is little or no incentive to reduce prices.[13]
2. Firms do not change prices because they do not want to start price wars. "Do not do what can be easily mimicked by your competitor" is a practical rule in business. Facing a decrease in demand, an individual firm will not reduce its price if it expects other firms will follow suit.[14] Cutting prices is such an unusual practice that the entrance in the taxi market of a new competitor charging lower rates became front-page news in many countries.
3. As it was pointed out above, if we assume positive trend inflation, one can rely on it to reduce a relative price without any need of changing its nominal price.
4. Price reductions are a source of conflict with customers. "Why did I pay more for the same product yesterday?" is a question that firms find difficult to answer if the decrease is not associated with some special event like, for example, a clearance sale.
5. Firms fear that customers may associate a fall in price with a fall in quality.
6. Lowering prices may be interpreted as a signal of difficulties and a lack of confidence in the product. It may have a deleterious effect on the brand.

We have here six solid practical reasons why firms do not cut prices in the real world when they face a contraction in demand. They prefer to resort to other tools like improving advertising and marketing, reinforcing the sales task force, stretching the payment period, etc.

[13] Moreover, if after a shift down the demand function becomes more inelastic it is profitable to increase the price.
[14] In his seminal article on oligopoly, Rothchilds (1947: 310) underlined that firms refrain from lowering prices to avoid retaliation from the competitors.

These are all arguments that explain why prices do not response to excess supply. They are specific to explaining downward price rigidity: they cannot be invoked to justify upward price stickiness, if it exists at all.

The result is, if there is a fall in aggregate demand, prices will not fall; there will be an excess supply in the short run.[15] In the long run, the aggregate supply will adjust to the level of aggregate demand. Prices play no role in this process.

Mainstream economists argue that assuming prices downward stickiness is ad hockery because it lacks microfoundations. If a firm does not adjust its price when relevant state variables change it will not be maximizing profits, they argue. This may be true in the frictionless world of theory.

However, none of the six reasons which justify price downward rigidity contradicts the profit maximization assumption. On the contrary, they explain firms' behavior in order to maximize profits in the long-run in the *real*-world under *real*-world constraints.

On the contrary, I don't find convincing reasons to assume perfect price flexibility except the elegance of symmetry. But elegance should not be confused for truth. If economic models are ever going to provide realistic guides to policy the real issue is what assumption is a better approach to what happens in the real world.

In empirical sciences, when there is a conflict between theory and empirical evidence, it is theory which is in trouble. For mainstream economics it is the opposite as if economics were a branch of applied mathematics and not an empirical science. That is why internal consistency, rather than external consistency – in the sense of conformability with empirical evidence – becomes the criteria for model admissibility (Wren-Lewis, 2009).

If there is someone who still remains unconvinced by the above arguments, remember Solow's AEA presidential address reflection:

> "I remember reading once that it is still not understood how the giraffe manages to pump an adequate blood supply all the way up to its head; but it is hard to imagine that anyone would therefore conclude that giraffes do not have long necks. At least not anyone who had ever been to a zoo" (Solow 1980: 7).

However, a mainstream economist would argue that, for the sake of internal consistency, zoologists should better assume that giraffes have short necks.

Although it is, of course, always desirable to go on improving the arguments that explain price change asymmetry, it seems absolutely much more reasonable to assume asymmetric rather than symmetric behavior anyway, at least for anyone who studies the real-world economy.[16]

[15] Even so, markets will be in equilibrium because, in the short run, excess supply does not generate any movement to eliminate it.
[16] In physics, Newton's theory of gravity tells us how bodies attract each other but it fails to identify the mechanism responsible of the motion of bodies.

Moreover, as recalled above, even Friedman agreed that price asymmetric behavior is a well-established fact in modern economies. However, it is likely he would add that assumptions in economics do not necessarily have to be based on real-world behavior.

10. The fundamental microeconomic assumption

For all the reasons given above, I propose the adoption of the following Fundamental Microeconomic Assumption (FMA): nominal prices display downward rigidity. Formally, $dp_i / dt \geq 0$ for $i = 1 \ldots n$.

This should be the starting point for the microfoundations of Keynesian macroeconomic analysis.

From FMA two corollaries immediately follow:

Corollary I. Any change in relative prices entails an upward change in the general price level.[17] In fact, FMA means that any change in relative prices implies an upward change in at least one price of the economy while, by assumption, none falls.

Corollary I implies that, for every economy, there would be a *natural rate of inflation*. This is the rate of inflation necessary to allow for the changes in the relative prices of that given economy.

Corollary II. In the absence of an increase in the money supply equivalent to the natural rate of inflation, there will be a fall in output.

In fact, from MV = PT, it follows that if there is an autonomous increase in P and MV is constant, T should necessarily fall.

This means that if the money supply rises less than the natural rate of inflation – assuming that V is relatively constant in the short run – the economy will be condemned to stagflation. That is why the natural rate of inflation can also be called the NAURMI (non-accelerating unemployment rate of money increase) because it measures the minimum rate of growth in the money supply necessary to keep the level of output (and employment) constant.

Once the FMA is accepted, it follows that if there is an excess supply, the equilibrium will be reached by an adjustment in quantities, and not in prices, just as Keynes assumed.

To sum up, let us suppose there are *n* good markets; let us assume that in *m* markets there is excess demand. Then prices will increase until supply equals demand, as in the classical model. In the *m-n* markets where there is excess supply, as monetary prices are downward inflexible, the excess supply will remain unchanged in the short run. Prices remain in these markets at the historically attained levels. Strictly speaking, all markets will be in equilibrium as there is no economic force that can change the value of the economic variables in the short run. As stated above, in the long run some firms will lay off some part of the labor force and others will shut down until planned supply equals demand.

[17] This is the essence of the non-monetary theory of inflation as explained, among others, by the recently deceased professor Julio Olivera. See Olivera (1960; 1964).

11. The wealth effect and price asymmetry

Keynes never thought the decline in prices could be a way out of involuntary unemployment. He did not consider the possibility of a real balance effect on the goods market, just as nobody did before Pigou (1943). The experience of inflation after World War II, which was attributed to the excess liquidity built up *during* the war, paved the way for the inclusion of the wealth effect as an argument in the consumption function.

For Keynes, the real balance effect was limited to the money market, the so-called Keynes effect. That is to say, an increase in real balances would have a reduction in the interest rate as its main effect.

Keynes was a practical-minded economist. He was very skeptical about downward nominal wage and price flexibility in the real world. That is why he insisted that real wages, in practice, can be lowered only by the increase in wage-good prices, not by the contraction of nominal wages.

Mainstream orthodox economics has used the wealth effect as the key instrument that leads the economy to full employment: excess supply in the goods market lowers prices and the consequent wealth effect reestablishes the level of aggregate demand at its full employment level.

However, if prices are downward rigid there is no wealth effect at all and there is no magic key to the full employment kingdom. Although the wealth effect may be useful in analyzing inflationary processes, it is of no practical relevance when dealing with recession and unemployment, the subject matter of Keynesian macroeconomics.

12. The Phillips curve

The *General Theory*'s main concern was unemployment. Its aim was to show why an economy can be stuck in unemployment and how to get out of it. The appearance of chronic inflation as an economic problem in the 1970s triggered the anti-Keynesian revolution. It was argued that demand stimulus to raise employment would always be associated with higher inflation.

Popular folklore has it that Keynes was largely unconcerned with inflation. As a matter of fact, Keynes (2016: 271) admitted that wages and prices would rise gradually as employment increases: "(…) we have in fact a condition of prices rising gradually as employment increases" and "an increasing effective demand tends to raise money-wages though not fully in proportion to the rise in the price of wage-goods" (ibid.: 275).

This was the origin of the idea behind the Phillips curve: there is always a trade-off between alternative levels of unemployment and inflation: the lower the level of unemployment, the higher the level of inflation is. The Phillips curve provided a link between the level of employment and the general wage level.

The debate on the Phillips curve became a turning point in the development of macroeconomics.

Actually, Phillips investigated the relationship between unemployment and the rate of change of money wages along one century, from 1861 to 1957, in the United Kingdom. Phillips (1958) found an inverse relationship between the rate of changes in the money wage rate and the rate of unemployment. He argued that

> "when the demand for labor is high and there are very few unemployed we should expect employers to bid wages rates up quite rapidly […] On the other hand it appears that workers are reluctant to offer their services at less than the prevailing rates when the demand for labor is low and unemployment is high so that wage rates fall only very slowly."

Therefore, the relationship between the two variables is not only inverse but also highly non-linear. These findings fit perfectly well within the original Keynesian model. It depicts the consequences of shifts in the demand for labor curve together with downward wage stickiness.

These results were hijacked by Solow and Samuelson, who substituted the rate of price inflation for the change of money wages. This substitution led to the policy conclusion that there exists an exploitable trade-off between inflation and unemployment. They presented this relationship as a policy menu to determine the costs of full employment. The Modified Phillips Curve version became highly popular during the 1960s. Decision-makers used it to estimate the costs of lowering unemployment in terms of the increase in the inflation rate.

However, in the 1970s, the modified Phillips curve was challenged from both the theoretical and the empirical points of view.

From the theoretical point of view, *F*riedman (1968) and Phelps (1967; 1968) pointed out that it was real, not money, wages which varied to clear the labor market. This thesis was sharpened with the help of the rational expectation hypothesis by Lucas. They proposed an expectations-augmented Phillips Curve and argued that, since all expectations are fully realized in the long run, a "natural rate of unemployment" will prevail. The long-run Phillips curve would be vertical at the natural rate of unemployment and no trade-off would exist. However, if an expansionist monetary policy is unanticipated, the general price increase that follows will be interpreted by each agent as an increase of relative prices. Consequently, monetary policy will have a real effect, by increasing output and employment. "Because prices are sticky, faster or slower monetary growth initially affects output and employment. But these effects wear off."[18] This explains the existence of a short-run Phillips curve. The faster workers' expectations of price inflation adapt to changes in the actual rate of inflation, the faster unemployment will return to the natural rate, while inflation will remain at the new higher level. The natural rate of unemployment is compatible with any rate of inflation: the long-run Philips curve is vertical.

As stated above, the trick consists of introducing wages as the clearing mechanism in the labor market. Any unemployment above the "natural" unemployment will be eliminated by a fall in the real wage rate. Any remaining unemployment is natural unemployment.

[18] Milton Friedman "Defining Monetarism" *Newsweek*, 12 July 1982, p. 64

The concept of the natural rate of unemployment was later replaced by the NAIRU (non-accelerating inflation rate of unemployment), which is understood as the level of unemployment at which inflation stabilizes. Strikingly, Phillips's article discusses an anticipation of the NAIRU concept, which would be coined 20 years later. He states that "if aggregate demand were kept at a value which would maintain a stable level of product prices the associated level of unemployment would be a little under 2½ per cent" (Phillips, 1958: 299).

NAIRU was considered the equilibrium rate of unemployment to which the system would return after any disturbance. "Only if the real wage desired by wage-setters is the same as that desired by price-setters will inflation be stable. *And the variable which brings about this consistency is the level of unemployment*" (Layard et al., 1991: 12, emphasis in original).

From the empirical point of view, the 1970s witnessed a simultaneous increase in both unemployment and inflation, which apparently contradicted the Modified Phillips Curve. I say "apparently" because the presence of a simultaneous rise in unemployment and inflation after the OPEC oil price hikes could have been interpreted as just an outward shift in the existing Phillips curve. For example, Lipsey (1960) has already argued that the Phillips curve shifted in the period between 1923–1939 and 1948–1957, in comparison to the pre-World War I period. With many more reasons the same could be argued with reference to the 1970s after OPEC substantially changed its oil price policy.

The main interest in the 1970s had shifted from unemployment to inflation. Hahn (1980: 285) noted that "even ten years ago one would have taken it for granted that a government should and could have a policy designed to reduce the average level of unemployment. Now this is no longer so."

The New Classical counterrevolution was, in Blinder's (1988: 278) words "a triumph of *a priori* theorizing over empiricism, of intellectual aesthetics over observation and, in some measure, of conservative ideology over liberalism. It was not, in a word, a Kuhnian scientific revolution."

After burying the Phillips Curve, Friedman (1956) revitalized the Quantity Theory, restating it in terms of a demand for money function, which now included an expected inflation term, which affects the expected nominal returns on the various classes of assets.

People demand a certain real quantity of money. If the quantity of money unexpectedly increases, people will seek to dispose of their excess money balances. Prices will increase until the real quantity of money held by people coincides with that which they want to hold.

This meant that in the long run all analyses could be conducted in real terms because the price level is proportionate with the stock of money.

13. Stagflation

Economic theory and economic policy faced a quite new phenomenon in the 1970s: stagflation.

As stated above, the 1970s witnessed a simultaneous increase in both unemployment and inflation after the first OPEC oil shock in 1974. The simultaneous presence of those two

phenomena marked the end of the Phillips curve's popularity and, with it, of the Keynesian era.

The New Classical economists demolished the Keynesian model with the argument that it could not explain the phenomenon of stagflation. The paradox is they gave no explanation of it at all, unless one considers as such the following argument. There is always a natural rate of unemployment; it is compatible with any rate of inflation: a vertical Phillips curve in the long run implies a natural rate of unemployment consistent with any given rate of price-increase. Then, the problem with stagflation is the inflation component and not the stagnation one. Lucas rejected the short-term–long-term split of Friedman in the analysis of the Phillips curve. Agents are endowed with rational expectations and they efficiently use all the available information. The economic agent acts as if s(he) were an econometrician and estimates the model parameters. Authorities cannot "fool" economic agents, not even in the short run. Expectations play a critical role in New Classical economics. If people expect an expansionary monetary policy, they will adjust their behavior instantaneously and prices will go up even if the Central Bank does not expand the monetary supply.

The actual unemployment is the natural one. And it is absolutely voluntary: it is the result of the households' choice because they find the actual real wage rate too low to motivate them to supply their labor services. Policy makers should be concerned with inflation, not with unemployment. Unemployment stays at its natural rate and any increase in monetary creation, if it is anticipated, is inflationary.[19] Policy makers who expand the monetary supply to fight unemployment increase inflation without any long-run effect on unemployment. Expansionary policies just push inflation higher and unemployment rises because inflation decreases the real wage; workers prefer to work less because leisure is cheaper. That is why stagflation exists. This is a very ingenious explanation; the problem is that during a recession, *unemployment is the result of an increase in layoffs, not of a decision by workers to stay at home*. The New Classical school gave no valid explanation for the stagflation phenomenon.

The New Classical economists were interested in, and were successful at, changing the focus of economic analysis from unemployment to inflation.

14. Keynes on monetary and fiscal policies

In order to complete the picture of Keynesian economics, let us briefly look at Keynes' point of view on monetary and fiscal issues.

Keynes' monetary theory has as a starting point his theory of liquidity preference. This preference has an opportunity cost: the rate of interest. Thus, the quantity of liquidity *demanded* is inversely related to the *interest rate.*

When the quantity of money is increased, its first impact is on the rate of interest, which tends to fall.

[19] If it is unanticipated, it may increase output and employment but soon agents will realize there has been no change in the relative prices and unemployment will return to its natural level while inflation will remain at the new higher level attained.

However, Keynes warned that "whilst an increase in the quantity of money may be expected, *ceteris paribus*, to reduce the rate of interest, this will not happen if the liquidity preferences of the public are increasing more than the quantity of money" (Keynes, 2016 : 270).

Given the marginal efficiency of capital, a fall in the rate of interest will increase the volume of investment. The increased investment will raise effective demand through the multiplier effect, thereby increasing income, output, and employment. As we move from unemployment to full employment, prices gradually rise as employment increases. (Keynes, 2016: 271).

So, there are a number of "positions of semi-inflation" (Keynes, 2016: 275), "a succession of earlier semi-critical points at which an increasing effective demand tends to raise money-wages though not fully in proportion to the rise in the price of wage-goods" (ibid.).

The early transmission of money increasing into prices and the possibility of a "liquidity trap" (ibid.: 187) were reasons that explained Keynes' skepticism on the monetary policy's ability to deal with unemployment (ibid.: 242). He was much more confident on the effectiveness of fiscal policy to cope with it. He argued in favor of public construction, building houses, or even digging holes in the ground if narrow-minded statesmen couldn't set in motion the former two alternatives. "I expect to see the State […] taking an ever greater responsibility for directly organizing investment" (Keynes, 2016: 147).

Public expenditure was conceived as the best tool to fill the gap created by deficient demand.

After the 2008 financial crisis, Keynes' prevention was confirmed.

The Great Recession revealed the limitations of monetary stimulus alone to overcome a severe recession. The Fed doubled the monetary base between September and December of 2008 but that money didn't reach the people: it only increased bank reserves. The federal funds rate was cut from about 5% in mid-2007 to nearly 0% in late 2008, yet the economy continued to suffer from inadequate aggregate demand for goods and services. As Samuelson very graphically said:

> "'You can lead a horse to water, but you can't make him drink.' You can force money on the system in exchange for government bonds [...] but you can't make the money circulate against new goods and new jobs [...] You can tempt businessmen with cheap rates of borrowing, but you can't make them borrow and spend on new investment goods" (Samuelson, 1948: 354).

Calvo (2016: 33) calls this situation a "Supply-Side Liquidity Trap": a point maybe reached where printing money increases real monetary balances but has little effect on real liquidity.

As Koo (2016: 24) rightly points out, when "private-sector borrowers sustain huge losses and are forced to rebuild savings or pay down debt to restore their financial health," they have no choice but to pay down debt or increase savings regardless of the level of interest rates in order to restore their financial health. We are here in the presence of an economy in which everyone wants to save but no one wants to borrow, even at near-zero interest rates. Under these circumstances, "there is very little that monetary policy, the favorite of traditional economists, can do to prop up the real economy" (Koo, 2016: 34). It is the time for fiscal policy.

As Romer (2011: 3) recognizes, "we need instruments of discretionary fiscal stimulus as part of the macroeconomic toolkit" because monetary policy is not enough to stabilize an economy facing a large shock.

Meltzer has criticized Keynes' recommendations in favor of using fiscal policy against unemployment. Meltzer (1988: 309) argues that short-term fiscal policy has not proven to be an effective tool of stabilization. He mentions that "attempts to lower unemployment by short-term policy adjustment have been followed by rising prices and capital outflow or currency depreciation."

The recent experience, as Romer pointed out, vindicates the importance of fiscal policy. Moreover, he mentions a very conclusive example: the fact that the major increases in government purchases in the two world wars and the Korean War were associated with booms in economic activity.

It is true that Keynes modeled a closed economy. The challenge, then, is to extend Keynes' model to an open economy to take into consideration a phenomenon like capital outflows mentioned by Meltzer. This implies including foreign trade as another component of aggregate demand as well as adding the possibility of acquiring foreign assets/liabilities.

For the open economy, the initial Keynesian workhorse model has been the Mundell–Fleming one. More recently, New Keynesian open-economy models with nominal price rigidities and intertemporally maximizing agents have been designed to understand the transmission of shocks across countries, exchange rate pass-through, and the effects of different pricing rules. New Keynesian models, particularly dynamic stochastic general equilibrium models (DSGE), became popular among central banks that use them in their job of setting an appropriate interest rate. Nominal price rigidity is introduced to ensure that shocks and central-bank interventions go beyond mere price effects. However, as they assume that wages and/or prices are upward and downward equally sticky, these models are only useful provided inflation is negligible. On the other hand, the recent financial crisis damaged the reputation of the New Keynesian DSGE models whose contribution has been of minimal value in addressing the greatest macroeconomic crisis in three-quarters of a century (Romer, 2011: 1/2)

Post Keynesians have extended their models to economies that are open to international trade and financial flows. Adding open economy features alters the potential outcomes of post-Keynesian models in several important ways. The analysis has focused on the effects of changes in the rate of exchange on output and trade balance. Blecker (2010) surveys the empirical results for several Post Keynesians models. Unfortunately, the results show that different studies using different methodologies have found different results for the same countries. This outcome is qualified as "disconcerting" by the survey's author; however, this is a very common situation in economics. Beker (2005) has drawn attention to this fact: "given a certain econometric result, in many cases it is enough to just include another variable, or to slightly modify the model assumptions or the estimation method to get different, and even opposite, results." There is nothing like a crucial experiment in economics. That is why models accumulate and remain available inside a big toolbox to be used according to the case under analysis and the analyst's expertise.

New and Post Keynesians have developed some valuable instruments that may be useful for making further progress in the understanding of open economy macroeconomics from a true Keynesian perspective.

Openness imposes restrictions on monetary and fiscal policies. If prices increase, the substitution of foreign assets for domestic money is an alternative and currency depreciation may be a likely result, aggravating inflationary pressures.

This means that inflation should be carefully watched and should not be neglected as it was in some so-called Keynesian policy experiments Meltzer refers to. Unlike several so-called Keynesians, Keynes did not favor inflation: "The money wage level as a whole should be maintained as stable as possible […] This policy will result in a fair degree of stability in the price level" (2016: 245). This puts a narrow limit on the use of monetary expansion as has been underscored by Keynes himself. The policymaker must watch unemployment with one eye and inflation with the other.

However, the recent European experience shows it is not so easy for inflation to gain momentum in a weak economy even when this is the central bank's target. And it also shows that, as has been pointed out above, it is equally difficult to refloat an economy only with monetary policy, even when resorting to negative interest rates.

Anyway, the monetary and fiscal areas in Keynesian macroeconomics demand an updating and deepening in order to take into account the role of the financial system in the present economy, explain current economic phenomena, and allow economists to formulate sensible economic policy recommendations.

In this respect, the 2007–2008 financial crisis is a superb natural experiment that provides very rich empirical material for the analysis.

15. Price stickiness: lessons from the European crisis

We have proposed to adopt as FMA the statement that nominal prices display downward rigidity. However, the recent European experience shows that not only inflation remained far below the rate targeted by the ECB, but that it was even negative in some months. Does this not make our FMA as unrealistic as the flexibility neoclassical assumption?

The first observation is that the rate of core inflation has always remained positive, showing that deflation has had to do with the price of commodities – energy and unprocessed food – that is to say with goods whose prices are quoted in world auction-like markets. For the rest of the prices downward inflexibility has been the rule. Even in a country suffering from a very deep recession, like Greece, core inflation remained positive. Although Greece lost 25% of its GDP between 2007 and 2015, it accumulated an inflation of 11.26% over this same period. Compare this figure with a 2013 Goldman Sachs's study mentioned in Sinn (2013: 5)) in which they estimated that Greece's prices would have to come down by 25-35% to achieve external debt sustainability. The Greek case is a very clear example that, due to price and wage downward inflexibility, internal deflation is not a way out of the crisis. In spite of the huge loss in GDP, prices did not decline and even kept rising. Price asymmetric behavior means that the only way for Greece, or Italy, to get cut its relative price level is by means of an increase in the EU price level average.

This is just one example of how FMA is far closer to reality than the usual mainstream economics assumption of symmetric price behavior. It is time to adjust economic theory to reality instead of waiting for reality to adjust itself to economic theory.

16. Conclusions

After the crisis in economic theory triggered by the 2008 financial crisis, which exposed many of the fallacies of orthodox economic thinking, it is time to rebuild the theoretical edifice of economics.

Realisticness is a necessary condition economic theory must fulfill if we intend to make meaningful predictions for the real world and sensible political recommendations. In this respect, it is argued that price downward stickiness has to be a fundamental assumption of any model that tries to reach these goals.

It is argued that realisticness is precisely one of the features that distinguishes Keynes' analysis. For this reason, Keynesian macroeconomics must be the point of departure of a realistic reconstruction of economic theory. And price downward rigidity fits perfectly well the Keynesian model, while this does not happen with either the New Keynesian or the Post-Keynesian models.

Given price asymmetric behavior, it is necessary to conduct separate analyses; on one hand, full-employment macroeconomics (price equilibrium macroeconomics) and, on the other, the macroeconomics of recession and depression (Keynesian macroeconomics). In the first case, prices play a role, while in the second case they do not. We need a theory apt to explain involuntary unemployment as well as inflation. However, for this purpose, we need two different approaches to the economy's behavior because it is quite different when the economy faces a positive shock than when it faces a negative one.

I propose the adoption of the assertion that nominal prices display downward rigidity as the Fundamental Microeconomic Assumption (FMA). This should be the starting point for the microfoundations of Keynesian macroeconomic analysis. The *natural rate of inflation* is defined as the rate of inflation necessary to allow for the changes in the relative prices of a given economy. It can also be called NAURMI, as it measures the minimum rate of growth in the money supply necessary to keep the level of output (and employment) constant.

If prices are downward rigid, there is no positive wealth effect at all and there is no magic key to the full employment kingdom. Many of the predictions and recommendations by traditional mainstream economics are flawed because they are based on the assumption that monetary prices are downward flexible.

The New Classical economists demolished the Keynesian model with the argument that it could not explain the phenomenon of stagflation. The paradox is that while, from a Keynesian point of view, it could be explained as a shift outwards in the existing Phillips curve, the New Classical theory gave no explanation of that phenomenon at all.

On the other hand, stagflation appears every time the money supply grows less than the natural rate of inflation. Therefore, it is crucial for the monetary authorities to estimate that rate in order to avoid inducing stagflation by an excessively rigid monetary policy.

The recent European experience under the effects of the financial crisis show that price downward inflexibility has been the rule, even in countries like Greece which lost 25% of its GDP between 2007 and 2015, while inflation was 11.26% over the same period.

The discussion in this paper has argued that Keynesian macroeconomics (the study of unemployment, recession, and depression) should again be at the top of the agenda of economic research. What are the next steps? I single out here some leading examples of where I believe progress can be made.

As I mentioned above, Keynes' model has to be updated in order to take into consideration today's real world. This is a first area for future research.

Keynes modeled a closed economy. The first challenge is to extend Keynes' model to an open economy. Moreover, the globalization and financialization processes as well as the role that supranational institutions play in the contemporary economy have to be taken into consideration in order to update Keynes' ideas to the present-day reality.

New and Post Keynesians have developed valuable instruments that may be useful for making further progress in the understanding of open economy macroeconomics from a Keynesian perspective.

The financial sector now includes not only banks but also other intermediaries such as life insurance companies, investment funds, leasing companies and other finance companies like those which constitute the so called shadow banking system. All of them play a decisive role in the contemporary economy which cannot be ignored as the American and European crises prove. Among other things, this new institutional environment has a crucial role in conditioning saving and investment behavior which are key factors in macroeconomics.

Institutions like the IMF, ECB or the EU did not exist at the time Keynes wrote his General Theory. These institutional actors have to be taken into consideration at the time of analyzing economic policy alternatives and recommendations.

A second area for research concerns the empirical estimation of the national natural rate of inflation and what factors affect it.

A third area for research consists in analyzing what role monetary policy has played in different stagflation processes and what lessons have emerges from that analysis.

A fourth area for research concerns learning how the 2007–2008 Great Recession was prevented from turning into something like the 1930 Great Depression and, at the same time, the reasons why the recovery has been so slow and feeble.

References

Ball, L. and Mankiw, N. G. (1994). Asymmetric Price Adjustment and Economic Fluctuations. *The Economic Journal*, 104 (March), 247-261.

Ball; L. and Romer, D. (1991). *The American Economic Review*, Vol. 81, No. 3 (June), 539-552

Ball, L., Mankiw, N. G. and Romer, D. (1988). The New Keynesian Economics and the Output-Inflation Trade-off. *Brookings Papers on Economic Activity*, Vol. 1, 1-65.

Beker, V.A.(1986). Asimetría y teoría económica. *El Trimestre Económico*, Mexico, Vol. 53 No. 2, April-June,. 215-223.

Beker, V.A. (2005). Is Economics a Science? A Discussion of Some Methodological Issues. SSRN. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=839307
or https://papers.ssrn.com/sol3/papers.cfm?abstract_id=839307

Beker, V. A. (2016a). From the Economic Crisis to the Crisis of Economics. In Moro, B. and Beker, V.A. (eds.) Modern Financial Crises, Chapter 9. Springer International Publishing, Switzerland.

Beker, V. A. (2016b) Rethinking Macroeconomics in Light of the Great Crisis. In Moro, B. and Beker, V.A. (eds.) Modern Financial Crises, Chapter 10. Springer International Publishing, Switzerland.

Blecker, R. A. (2010). Open economy models of distribution and growth. In Eckhard H. and Stockhammer, E. (eds.), *A Modern Guide to Keynesian Macroeconomics and Economic Policies*, Cheltenham, UK: Edward Elgar.

Blinder, A. (1994).*On Sticky Prices: Academic Theories Meet the Real World*. In N. Gregory Mankiw, ed., Monetary Policy. The University of Chicago Press, 117 - 154.

Blinder, A.S., Canetti, E.R.D., Lebow, D.E., and Rudd, J.B.(1998). Asking About Prices: A New Approach to Understanding Price Stickiness. Russell Sage Foundation, New York.

Calvo, G., (2016). "From Chronic Inflation to Chronic Deflation: Focusing on Expectations and Liquidity Disarray Since WWII," NBER Working Paper 22535.
http://pubdocs.worldbank.org/en/766231464121854489/Calvo-Macro-PAPER.pdf

Cooper, R. and John, A. (1988). Coordinating Coordination Failures in Keynesian Models. *The Quarterly Journal of Economics.* Vol. 103, No. 3 (Aug.), 441-463.

Cover, J.P.(1992). Asymmetric Effects of Positive and Negative Money-Supply Shocks. *Quarterly Journal of Economics*.Vol. 107, No. 4, 1261-1282.

Davidson, P. (1962). “More on the aggregate supply function.” Economic Journal, 72 (286), 452-457.

Davidson, P. (2002). *Financial markets, money and the real world*. Cheltenham: Edward Elgar.

De Abreu Lourenco, R. and Gruen, D. (1995). Price Stickiness and Inflation. Economic Research Department, Reserve Bank of Australia. http://www-dr.rba.gov.au/publications/rdp/1995/pdf/rdp9502.pdf

De Long, J.B. and Summers, L.H. (1988). How Does Macroeconomic Policy Affect Output? *Brookings Papers on Economic Activity*. Vol. 19, No. 2, 433-494.

De Vroey, M. (2011). Lucas on the Relationship between Theory and Ideology. *Economics: The Open-Access, Open-Assessment E-Journal*, 5 (2011-4).
http://dx.doi.org/10.5018/economics-ejournal.ja.2011-4

Dobrynskaya V.V. (2008) Asymmetric price rigidity and the optimal interest rate defense of the exchange rate: some evidence for the US. *Journal of Policy Modeling*. October, 30(5),713–724.

Friedman, M. (1953). *Essays in Positive Economics*. The University of Chicago Press, Chicago.

Friedman, M.(1956).”The Quantity Theory of Money—A Restatement.” In Friedman, M. (ed.) *Studies in the Quantity Theory of Money*. The University of Chicago Press, Chicago.

Friedman, M., (1968). “The Role of Monetary Policy”. *American Economic Review*. Vol. 58, No. 1, March, 1-17.

Hausman, D.M. (1992). *The inexact and separate science of economics*. Cambridge University Press, Cambridge.

Hahn, F. (1980). “Unemployment from a Theoretical Viewpoint.” *Economica*. Vol. 47, No. 187, August. Special Issue on Unemployment, 285-298.

Hodge, D. (2007). "Economics, realism and reality: a comparison of Mäki and Lawson". Working Paper Number 63, UNISA. https://core.ac.uk/download/files/153/6428994.pdf

Karras, G. and Stokes, H (1999). "On the asymmetric effects of money-supply shocks: international evidence from a panel of OECD countries". *Applied Economics*. Vol. 31, No. 2, 227-235.

Keynes, J.M. (1939). "Relative Movements of Real Wages and Output". *The Economic Journal*. Vol. 49, No. 193, March, 34-51. http://piketty.pse.ens.fr/files/Keynes1939.pdf

Keynes, J.M. 2016 [1936]. *The General Theory of Employment, Interest and Money*. Atlantic, New Delhi.

Koo, R.C. (2016). "The Other Half of Macroeconomics and Three Stages of Economic Development". In Lima. G. P. and Madi, M.A. (eds.) *Capital and Justice*. World Economics Association Book Series Vol. 4.

Lavoie, M. (2001). "Pricing". In R. Holt & S. Pressman (Eds.), *A new guide to Post Keynesian economics*, 21-31. Routledge

Layard, P.R.G., Nickell, S.J., and Jackman, R. (1991). *Unemployment: Macroeconomic Performance and the Labour Market.* Oxford University Press, Oxford.

Lee, F. (1998). *Post Keynesian price theory*. Cambridge

Lipsey, R.G. (1960). "The Relation between Unemployment and the Rate of Change of Money Wage Rates in the United Kingdom, 1862-1957: A Further Analysis. *Economica*, Vol. 27, No. 105, February, 1-31.

Lucas Jr., R.E. (1976). "Econometric Policy Evaluation: A Critique", in Brunner, K. and Meltzer, A. (eds.), *The Phillips Curve and Labor Markets* , Carnegie-Rochester. Conference Series on Public Policy, Vol. 1, 19-46.

Lucas Jr., R.E. (1978). "Unemployment Policy". *The American Economic Review.* Vol. 68, No. 2, Papers and Proceedings of the Ninetieth Annual Meeting of the American Economic Association, 1978, pp. 353-357.

Mäki, U. (2005). "Models are experiments, experiments are models". *Journal of Economic Methodology* 12 (2):303-315.

Melmies, J. (2008). New-Keynesians Versus Post-Keynesians On The Theory Of Prices. http://www.boeckler.de/pdf/v_2008_10_31_melmies.pdf

Meltzer, A. H. (1988). *Keynes's Monetary Theory.A different interpretation*. Cambridge University Press, Cambridge.

Olivera, J. H. G. (1960). "La teoría no monetaria de la inflación". *El Trimestre Económico*, *Octubre-Diciembre, 626- 628. México D.F.*

Olivera, J. H. G. (1964). "On Structural Inflation and Latin American 'Structuralism'". *Oxford Economic Papers*, 16 (3), November, 321-332.

Peltzman, S. (2000). "Prices rise faster than they fall". *The Journal of Political Economy*, Vol. 108, No. 3 (Jun., 2000), 466-502.

Phelps, E. S. (1967) "Phillips curves, expectations of inflation and optimal unemployment over time". *Economica* 34(August), 254-281.

Phelps, E. S. (1968). "Money wage dynamics and labor market equilibrium". *Journal of Political Economy*, 76(July-August), 678-711.

Phelps, E. S., ed. (1970). *Microeconomic foundations of employment and inflation theory*. New York, Norton.

Phillips, A. W. H. (1958) "The relation between unemployment and the rate of change of money wage rates in the United Kingdom, 1861-1957". *Economica* 25 (November), 283-299.

Pigou, A. C. (1943). "The Classical Stationary State". *Economic Journal*, LIII, 343-51

Rhee, W. and Rich, R.W. (1995), "Inflation and the Asymmetric Effects of Money on Output Fluctuations". *Journal of Macroeconomics,* 17: 683-702.

Romer, D. (2011). "What Have We Learned about Fiscal Policy from the Crisis?" http://eml.berkeley.edu/~dromer/papers/What%20Have%20We%20Learned%20about%20Fiscal%20Policy%20from%20the%20Crisis%20May%2020.pdf

Samuelson, P. A: (1948). *Economics: An Introductory Analysis*. McGraw-Hill.

Samuelson, P. A.(1963). "Problems of Methodology: Discussion". *American Economic Review*, 53(2): 231- 236.

Sinn, H. W. (2013) "Austerity, Growth and Inflation. Remarks on the Eurozone's Unresolved Competitiveness Problem". CESifo Working Paper No. 4086.

Solow, R. M. (1980). "On Theories of Unemployment". *American Economic Review*, 70 (1), March, 1-11.

Stiglitz, J. E. (1999). "Toward a General Theory of Wage and Price Rigidities and Economic Fluctuations". *American Economic Review*, 89 (2), May, 75-80.

Wren-Lewis, S. (2009). Internal Consistency, Nominal Inertia and the Microfoundation of Macroeconomics. University of Oxford.
http://www.economics.ox.ac.uk/materials/working_papers/paper450.pdf

Author contact: victor.beker@ub.edu.ar

SUGGESTED CITATION:
Victor A. Beker, "Asymmetric price adjustment: the missing link in Keynesian macroeconomics", *real-world economics review*, issue no. 80, 26 June 2017, pp. 121-147, http://www.paecon.net/PAEReview/issue80/Beker80.pdf

real-world economics review, issue no. 80
subscribe for free

Chinese economics as a form of ethics

Kazimierz Poznanski [University of Washington, USA]

Abstract
This article offers the first full account of Chinese economics. That China has its own economics is not realized. It exists, but still only informally, as part of the Chinese mindset. Being as different culturally as China is, her approach to economics is very different as well. It certainly departs from Liberal economics that dominates Western thinking. It is a form of ethics, while Liberal economics is basically value free. Its eight principles can be derived from Confucian thought. Called by me "Confucian (or – Chinese) economics", it assumes that people's goal is not an "instant gratification" but a pursuit of continuity, or legacy. The Chinese also disregard the notion of scarcity, and claim that – if allowed to regenerate – resources are abundant, even exuberant. It is not the struggle for resources with each other, but earnest work that is the superior strategy. Individuals work not for themselves, but for others, mainly family. Motivation doesn't depend on the enforcement of property rights, but on morality, understood as responsibility for others. The facts are, during two-thirds of their lives, individuals depend on transfers from other – working – family members. Since morals come from family, not the market but the family is the principal institution. But Confucianism is not hostile towards markets. It appreciates them but not in the Liberal "free market" form. It favors "familial markets", animated by households. Confucians don't advocate a "authoritarian state" but a limited one, called "moral state". Its role is to propagate virtue by setting an example. The highest priority is to ensure equality. This is to avoid social conflict, as the worst enemy of growth. During the post-1978 reforms, China phased out of its Soviet-style system in order to revive her "Confucian system", based on "familial market" and "moral state". Produced by the reforms, the longest "economic miracle" in history is only the latest evidence of the advantages of China's approach to the economy.

Introduction

As two major civilizations, China and her Western counterpart follow largely separate views of the world, including their respective approaches to economics as a study of wealth. This intuitively obvious fact is not accepted by economists, whether Western or Chinese. They deny there is such a thing like a distinctly Chinese school of economics. Mine is the first attempt to outline China's complete approach to economics (compare: Chen, 1911). Unlike the Liberal (or Classical) economics that dominates Western thought, the Chinese approach is still at an informal stage. It is a part of Chinese mindset contained in Confucian writings. This body of knowledge can be called "Confucian (or Chinese) economics". Confucianism, of course, can't be equated with Chinese civilization, which has many sources. Nor is Confucian outlook a monolith. It is subject to internal rifts. However, a casual observation suggests Chinese intellectual landscape is more monolithic than the Western one. Confucianism is constructed to ensure continuity rather than exploration. But this is where it finds its incredible resilience. Admitting these nuances, my article offers only a general assessment of the core of Chinese economics. It also doesn't deal in necessary depth with the question of the application of this theory and thus its relative worth. Only at the end of the article do I raise some points on how Chinese thinking is reflected in real life, i.e. how the economic agents act and how the system – that frames their actions – operates.

Searching for the core of Chinese economics, I discern a lineup of its eight principles that together constitute a logical sequence. They are all found to differ from the Liberal approach; actually they are often almost a "mirror image" of Western perspective on economy. This is true about the opening question that economists have to ask while constructing their lines of theory. Namely, what is the individual's main objective function? To Confucians, the primary goal of the individual is the continuity of family line through procreation – or reproduction – and not an "instant gratification", meaning consumption. Individuals are said to work not for themselves, but on behalf of their family. The motivation for individuals to pursue their goals is thus the moral responsibility for other family members. In contrast, in the Liberal approach, the individual is driven by the security of its property rights. Since it is the family – as a "work unit" – where the key moral choices are made (e.g., giving birth to a child), it is also the family where morality mainly originates from. And because growth of wealth is driven mainly by a sense of morality, it is not the market, but rather the family that is the basic institution economies rely upon. It follows that the key difference here is that Liberal economists assume that economy and morality are detached from each other, while Chinese see them inseparable.

True, some non-liberal Western approaches do recognize the importance of morality and/or family. But none treat economics as ethics, the way Confucianism does. Not even the "Evolutionary economics" of the Austrian-born Schumpeter (1942). In his analysis of capitalism, he argues that capitalism couldn't emerge without injection of values by three feudal classes, i.e. peasants, clergy and aristocracy, jointly called "protective strata". They brought a sense of property, acceptance of hierarchy and honor. Of course, capitalism is not possible without capitalists, but to him a capitalist is not a value-free – rational – individual. It is a head of family working for his spouse and children. And, with its "bourgeois ethics" he/she is in pursuit of continuity, i.e., building "industrial dynasty". This is their function in a capitalist economy, and also the precondition of prosperity. As Schumpeter writes, in a way:

> "…the capitalist order entrusts the [investment – KP] interests of society to the … bourgeoisie. They are really entrusted to the family motive …" (Ibid. p. 160).

However, as successful as it is, capitalism can destroy itself by eroding the "protective strata". But also by – as he calls it – "rationalization" of everything in life, making capitalists increasingly unready to sacrifice immediate personal consumption for investment in the future – of family.

Ethical criticism of Liberalism originates also from outside of the broad Liberal circle, most notably from among Marxists. Among them is another Austrian scholar Polanyi (1994), who similar to Confucianism, but with no direct references, assumes that until the advent of capitalism, societies were squarely based on two principles. The one that relates to "family" is the principle of "house holding", i.e., production for family's use. The other, related to "morality," is the principle of "redistribution" of wealth. By allowing the capitalist "free market" to uproot both principles, introduction of the capitalist system released – to use his phrase – the most "devastating destruction" in human history. It is not so much economic deprivation as a psychological one, relating to as he calls it "commoditization", involving the loss of their sense of humanity (or – dignity). To Polanyi, the most detrimental aspect of introducing capitalism was the bringing in of the fictitious concept of the "economic man" – or "rational actor". This – central for Liberalism – concept defined an individual as one who by birth – instinctively – needs "free market." Polanyi summarized his critical opinion this way: ". . . no

misreading of the past provided was more prophetic of the future" (Ibid., p. 44). And, it is only because of the periodic reversals in the implementation of the Liberal agenda that the Western world has been able to last.

But can individuals "afford" being moral? To Confucians, with their view of nature, resources that people need to survive – whether land, water or energy – are abundant, even excessive. They are excessive as long as people respect the rhythm – or logic – of nature, and allow nature to regenerate. With the so defined "economic condition" faced by individuals, the challenge is to work diligently. It makes the economic survival of individuals principally a function of the part of general ethics that is called "work ethics". Under conditions of abundance, individuals have enough wealth to share their riches and thus can "afford" being moral. This theoretical postulate is a rather drastic, probably the sharpest departure from Liberalism, which is assuming that universally – in any place and at any time – in their efforts, people are faced with the economic problem that Liberals call "scarcity". The implications of this "scarcity" for human behavior, are enormous. Under these conditions, there are not enough resources for all individuals to survive. Survival becomes not a matter of work, i.e., amount of effort expended, but rather a matter of forceful – if not violent – acquisition of resources from others. It follows, under scarcity, that being moral is, simply speaking, not "affordable". How could it be otherwise, if giving away wealth you own makes others a greater threat to one's individual chances of physical survival.

To clarify, treating family as a key agent does not imply that Chinese economics rejects markets. Like Liberals, Confucians are very amicable to markets. But they stand against the Classical concept of the "free market" as a basis of capitalism, where "prices" of goods and not moral "principles" (or "values") rule the choice. Because of its fundamentally ethical nature, Chinese economics favors a type of arrangement that I will call the "familial market" based on moral principles and coming from the family. Contrary to the Western stereotypes of Chinese political history, Confucians do not call for a "maximal (or authoritarian) state". The facts are, similar to Liberals, Confucians advocate a "minimal state" that respects independence of individuals to the fullest. But Confucians have a very different concept of state agency, one where family freedom rather than individual freedom is of main concern. In further contrast, Confucians view the state not as a nexus of physical or coercive power but as a moral nexus, with the state providing moral leadership, mostly by setting an example. A more fitting name for such an agency is a "meritorious state", charged mainly with preventing practices that corrupt society, such as establishing a monopoly on goods and collecting undeserved rent from the buyers. As defined, this Chinese alternative to Liberal capitalism can be labelled a "Confucian system" (Poznanski, 2012).

While Classical economics serves as an ideological base for the capitalist system, even its most orthodox applications greatly deviate from its theories. As an alternative to Liberal economics, Confucian (or Chinese) economics can't serve as a frame of reference for setting up some kind of capitalist system and it doesn't. This means that Chinese economics must favor some type of system that would represent its opposite. The question is where in the Western typology of economic systems such a design with a "familial market" and a "meritorious state" would logically fit in. In widely accepted Western typology, there is the "capitalist system" and, in contrast, the "socialist system", both coming in a range of variants. Therefore, if not capitalist, the Chinese system would have to be seen as a form of the "socialist system." This means that there might be some affinity between Confucian economics and Western approaches that advocate the "socialist system", such as the German "Social-Market economics" with the so-called "Rhine capitalism". The other would

have to be of course the much more widely-adopted Marxist economics, with the Soviet (or Stalinist) model as its extreme application. While calling the Chinese system "socialist" seems acceptable, I suggest we go beyond this dichotomous typology and – to avoid confusion – continue calling it "Confucian system".

After examining this – by its very nature complex – system, it is not certain that for practical (and efficiency) reasons, Liberalism can be an attractive option for China, with her focus on the ethics of responsibility ("ren") and preference for family held together by the uniquely Chinese concept of "filial piety". It appears, time and time again, that Confucian approach offers some potentially substantial advantages compared to its Liberal alternative. For instance, by focusing on continuity of family line as an objective function, "Chinese economics" not only induces population growth and its renewal. It also induces businesses to take a considerably longer time horizon and favor deferred consumption, with higher savings. By stressing the centrality of family as an organizing form, it also ensures stronger motivation to work, including the higher level of mutual trust as an antidote for risk aversion. Due to the greater intimacy, family also lowers the so-called "transaction costs", and thus the total costs as well. The call for modesty, might serve as an additional burrier to the dangerously mounting pressures on natural resources. Confucian contempt for the destabilizing "monopoly practices" – as a source of wealth inequality (and thus – poverty) is another likely advantage that seems to come with the application of Confucian alternative vision of the economic life.

I. Individual's objective function: pursuit of "legacy"

The opening issue in constructing any complete approach to economics is the assumption about the individual's sense of purpose, or his/her objective, and this is the issue where my reconstruction begins (see Table I). Addressing this critical issue, Confucians make it clear that to resolve it one needs to remember that people are an integral part of nature. Therefore, to understand how people operate, one needs to examine nature first. And, to define the individual's primary objective in economic life, it is necessary to ask first what nature's purpose might be. In Confucian eyes, nature is equated with the life process, or life itself. Each living creature is said by them to be animated by a sort of code of life, which consists of two indispensable elements, called *yin* and *yang* (see: Mote, 1971). These terms refer not to female and male genders but female and male features, or characteristics. Combined, they are programmed to ensure continuity of life, and since the continuity of life is the aim of nature at large, this must be the goal of individuals as well.

In fact, the best way to conceptualize Chinese economics is to recognize that as with the assumption about the principal objective of economic life – but also of life in general – all its "building blocks" address one and the same central issue of the extension of life through procreation. Confucius himself makes it absolutely clear that the giving of birth to another – consecutive – generation of people is the ultimate mission of the living. This declaration makes Confucian economics an "economics of life", with a special role for sons in pursuit of it. This is especially a mission for sons, since it is through them that the legacy under the traditional Chinese form of paternalism is transmitted. On this issue, Confucius makes the following cardinal point:

> "Since the parents have given birth to a son, it is the perpetuation of the human race, and there is nothing greater than this" (from Chen, 1911, p. 328).

The above prerogative for individuals also applies to the state, whose major role is to ensure as rapid population growth as sustainable, including by the pursuit of economic policies that support large families. It is within that context that despite its leftist, Marxist leaning, the Chinese leader Mao glorified large-size families and phased in pro-natal policies.

This Confucian focus on – let's call it – "distant gratification" that procreation represents, stands in direct contrast with Classical economics. It assumes that the primary objective function is individual consumption, or "instant gratification". Before they meet any other desires, people have to satisfy their basic needs. For this, they must first get food to nourish them and shelter to protect them from physical harm. Confucians recognize the importance of basic needs but argue that one can get more from economic analysis by assuming that procreation is the primary goal. It is not so much about what comes first, but rather about what gives more meaning to life. Their thinking is that having children (or – leaving "legacy") is the meaning of life. Without children the individual's life is terminal, with a defined end. But not if individuals procreate, since then they are survived by children. This must be the ultimate objective, since by being survived their life – or "blood" – flows into an endless future.

With the distinctions of these two human objectives comes the Confucian separation of goods into two categories, material and spiritual. "Instant gratification" involves mainly the former and the provision for posterity relates basically to the later. On the "material" goods Confucians call for showing a necessary restraint. Mencius wrote that:

> "For cultivating heart, nothing is better than having few [material – KP] desires. If someone has few desires, although there will be times when he does not persevere, they will be few. If someone has many desires, although there will be times when he perseveres, they will be few" (from Norden, 2009, Book 7B, 35.1).

Regarding "spiritual" needs there is no threat of scarcity, since they mostly involve human relations (e.g. the comfort of the household, familial love). "Spiritual" goods are thus superior to "physical" goods, a point forcefully made by Mencius:

> "The best life is characterized by simple, everyday pleasures and rich relationships with family, friends, and members of one's community."

Table I Basic assumptions: Confucian and Classical Economics

	Confucian Economics	*Classical Economics*
1. ***Principal Objective***	Continuity (Procreation)	Consumption (Existence)
2. ***Resource Condition***	Abundance (Excess)	"Scarcity" (Deficit)
3. ***Survival Strategy***	Earnest Work	Resource Acquisition
4. ***Motivating Force***	Moral Attitude	Property Rights
5. ***Basic Institution***	Family Unit	Free Market
6. ***Income Distribution***	Equality (No Poverty)	Inequality (With Poverty)
7. **State Prerogatives**	Meritorious State	Minimal State
8. ***Preferred Method***	Common Sense	Theoretical Models

The practical effect of the Confucian focus on continuity of kinship is that it blurs the difference between perceptions of the present and the future, something that is neatly woven into the whole Chinese culture. This blending is consistent with the Chinese idea that time is a continuum, with no beginning and no end. Everything alive is seen as being caught in cyclical returns. This sense is also encoded by the Mandarin language. It lacks tenses, so the present and future are equated (as argued in Chen, 2013). This way of encoding time differs from the British practice, where explicit time markers are required. The respective concepts of time also differ, since the British people have a linear notion of time, where each moment brings one closer to an inevitable end. The effect is that Chinese feel that the future is closer than it is for the British. Given this, they more readily accept choices with "distant gratification" in sight, that is, to save for future consumption. Besides, the primary focus on seeking future rewards makes people work harder (Harrell, 1985).

II. "Economic condition": abundance of resources

The next issue for any economics to settle is under what conditions people pursue their goals. In the Classical economics this issue is called "economic problem (or question)". However described, such a problem gives rise to economic activity, meaning the economy. To settle

this defining issue, the Chinese turn again to nature as a "designer" of human life, both the taught process and the action taken. The Confucian argument is that if people are coded with the need to pursue continuity, nature must have created the material conditions that allow for it to take place. Accordingly, Chinese economics rests on an assumption that universally – in time and space – nature provides people with sufficient resources, and that it supplies an abundance of resources, such as land and water. This is not to say that people never face a lack of resources and related economic hardships. But, as Mencius argues, while this lack can become a reality to suffer from, even with devastating consequences, nature can't be blamed for it. Only people can cause scarcity or famine by acting not in accordance with but against a nature that is generous. He explains:

> "If one does not disrupt the farming seasons with building projects, but only waits until after the crops have been harvested, the grain will be inexhaustible. If overly fine nets are not used in the ponds, so that sufficient fish and turtle are left to reproduce, they will be inexhaustible. If people bring their axes into the mountain forest only in the proper season, the wood will be inexhaustible" (Norden, 2009 Book 1A, 3.3 – 3.4).

Of course, it is not all the fault of human "economic errors" that people may occasionally be confronted with an insufficient amount of goods to continuously support their lives. Nature can be even very capricious and bring flood or draught, so that there might be – say – less food available than needed for a tough year. But as Confucians argue, even if such nature-driven fluctuations of food supplies may lead to starvation, this can't be viewed as an evidential proof that nature is unforgiving. Mencius writes on this issue of such cases of inadequate supplies of food that: "…not the harvest is responsible but people" (from Norden, 2009, Book 1A, 3.3 – 3.4). People are responsible since starvation is avoidable by storing and releasing food in time for the most disadvantaged to be taken care of. This Confucian view on the causes of starvation is echoed by a key figure in the field of the modern economics of development, the Indian scholar Sen (1999). In his view, poverty of any kind is not a curse but a product of failing economic institutions, e.g., market- or state-type speculation with supplies and prices of foods.

The Confucian belief in the abundance of resources is antithetic to the thinking of Classical economists. This is probably the most divisive issue separating the two approaches. The liberal view is that as a "law of nature" people are universally made to live under the condition of the so-called "scarcity of resources". The introductory material in most Western textbooks on economics usually offers a definition of this inescapable condition as one that stems from the combination of two forces, namely the unlimited needs of individuals and the insufficient supplies of resources available to meet them. How accurate this interpretation of the Liberal definition of scarcity is might be debatable but what is certain is that liberal economists don't share the Confucian outlook on what humans' fate is. As described by Confucians, humans have emerged in the world and continued because and only because the world is easy to survive in and thus should be viewed as a "happy place" Liberals view the world as one of the limited resources and hardship, and is thus a "dismal place". And delivering this scary message Liberals themselves characterize their discipline – as a "dismal science".

On the surface, the Liberal assumption of unlimited needs might seem almost self-evident, and so the related consequences. With an unconstrained appetite for goods, people can't ever satisfy all their needs. Consequently, they have to go through painful substitution of one product on their "consumption function" by another rather than enjoy benefits of having

access to all of them at once. Confucians admit that individuals always make choices within finite budgets, but they don't accept the implied notion that this itself makes people fall into such a dismal situation. As Confucian economics claims, if the ethical dimension is added to the equation, there is an easy solution to this apparent spending dilemma. People can simply show a necessary moral restraint and prioritize their spending without undergoing any kind of "pain" at all. Mencius explained this particular point plainly:

> "I like fish, but I also like bear's paw, but if I can't have both at the same time, I will forgo the fish and eat the bear's paw. I love life, but I also love righteousness, and if I can't have both at the same time, I will sacrifice life to have righteousness" (from Norden, 2009, p.260)

Turning to the related question of availability of physical resources, already the early British luminaries of Liberalism such as Hobbes (1651) and Malthus (1798) made a strong claim that resources are always inadequate, so that economies are confronted with a problem of so-called "surplus of population"**.** This idea has also attracted historians, particularly those working on migrations. Their argument is that when scarcity emerges, a surplus of the population surfaces up and in response, people move to other locations. The example of choice is Britain at the time of the Industrial Revolution when her population multiplied. The Brits actually moved abroad in large numbers, with the whole episode called in Western literature the "Great Migration". But the outflow is no proof of "scarcity", since as much as 4/5 of the demographic increase stayed on the supposedly overpopulated island. And how can "scarcity" be a universal phenomenon, if the migrants moved to North America, which was a relatively sparsely-populated, huge continent with an abundance of land and water. The reality is, the British didn't run away from scarcity as understood in classical economics, but decided to find "greener pastures" and/or to bring a thrill of adventure to their lives.

Back to China, there is here no special character for "scarcity" in her language, as if "scarcity" didn't exist. The reason for this linguistic omission is that indeed – literally speaking – it did not exist. Through her extended history, China has not engaged in large-scale territorial expansion. Accordingly, China's land mass hasn't changed much over centuries, even though China has continuously enjoyed military superiority over her neighbors. An exception has been the period of Mongol invasions that extended to Tibet and Vietnam and involved failed attempts to conquer Japan. When Chinese expanded their reach it was by migrating, as in the case of Manchuria, where in few years, millions of Chinese brought their families to this sparsely populated region. While largely confined to the same territory, her total population magnified tremendously. The last time it did was after 1949 when "forced" industrialization started. Within the same official borders, China's population increased during these decades by no less than three times to become the most populous society on earth. Knowing of these cases, how can one claim that universally there is "shortage" of resources. China actually represents one of the most important cases that points exactly to the opposite.

III. Prevailing economic strategy: earnest work

The next step in building any line of economics is to determine what the prevailing strategy for meeting goals is. For Confucians to argue that resources are inexhaustible doesn't mean the world is a sort of dreamland or paradise, where all is handed over on a plate for good people to enjoy. To meet their needs people have to convert resources into goods. In other words, they need to work, and work – labor – is thus the only source of value, or wealth. Meeting

needs through work is not a big problem, since at least the "basic needs" are modest. Besides, resources are a prerequisite for meeting "material needs" but not so much for the "spiritual" ones, which happens to be the superior of the two. Importantly, Confucians goes as far as to view work itself as a significant source of joy, and thus on its own work becomes a "spiritual good" itself. For this to materialize one needs to choose the profession that he/she loves most, and most individuals tend to do so. This must be why in reality, people hate laziness and tend to work more than they need.

Differentiation within society is recognized in Chinese economics, but this is only with respect to the types of occupations people hold. The lowest in the hierarchy are merchants and money lenders who work for profit, or rent. That they are put at the bottom of the social taxonomy is not because Confucians have contempt for making money, which is often stressed by liberal critiques of Confucianism as a detriment to prosperity, or progress. The Confucian intention is only to prevent the levers of power from being handed over to those members of society whose inclination is to meet pecuniary interest first. The rise of these groups in the social hierarchy is discouraged, since their financial success might be confused by people as living evidence that they are better people who should be listened to in matters of state. Not to forget, there is one more group and it has the lowest status – they are soldiers, together with those who supply them with deadly hardware. For clarity, Confucians talk about the professional soldier and not the ordinary people – historically peasants – who join army not for money. Or any other reason except for the defense of their households, i.e., families that they belong to (or communities that families are an intricate part of).

In the world of scarcity assumed by Classical economists, work is not so much a strategy of choice. It cannot really be a preferred response to scarcity, because work alone can't ensure an adequate supply of finite resources. People who are faced with scarcity as it is defined thus have no alternative but to secure their survival by taking resources from each other. This is often described by Liberals as a ruthless "struggle for survival", where only winners stay in the game. This is certainly how Hobbes and Malthus understood the essence of the free markets. Here, due to competition, higher cost producers declare bankruptcy. Through the bankruptcy procedure, resources are captured by the lower cost producer. The ultimate weapon is war, where not individuals – meaning producers – but nations fight for physical resources, i.e., territory. By assuming the above, Liberalism provides a kind of "theory of war", where repetitive bloodshed is something imposed on people by nature. Being an opposite approach, Confucians seem to offer a "theory of peace", where war is not a curse but rather a matter of choice for people to make. Or more precisely, it is a matter of what ideas people share about potential benefits and damages from waging wars (point made by Wight, 1992)

The concept of the "struggle for survival" had been given a theoretical form already by the first Classical thinkers, who came up with the concept of an oversupply of people that – for the sake of global balance – must be periodically removed. The already mentioned prominent predecessor to Smith, Hobbes (1588-1679) wrote in his famous book "Leviathan": "When the world is overcharged with inhabitants, then the last remedy of all is war …" Partly overlapping with Smith, the Reverend Malthus (1766-1834) produced a model to explain this apparent phenomenon and that harsh remedies are unavoidable. Called the "Malthusian Trap", this model claimed that due to their inborn instinct people strive to multiply and have as many children as women are capable of delivering. When for whatever reason material wealth increases, the gains are used disproportionally to have – meaning feed – more children. A related rise in population will outpace the change in production level. Historically, as the theory goes, the resultant surplus of people has been dealt with either by starvation or by war.

The only way to avoid such brutalities is a moral way, namely the voluntary reduction of the birth rate, but history provides little evidence of that.

The Malthusian theory was almost immediately challenged by the classical economists themselves, and it is still widely treated in the profession as a failed theory. It has been debunked for misrepresenting both sides of the equation, both human fertility and the fertility of the soil. But in spite of this categorical dismissal by the Liberal economists, Malthus is still doing quite well. He still enjoys great name recognition, and this first economics professor ever is regularly mentioned in the contemporary literature. While his demography is dismissed as counterfactual, the Liberal concept of the "struggle for survival", with its nasty contents is kept alive by the voices of many scholars, like the widely received and widely revered academic Dawkins (1978). In the spirit of his economic predecessors, this British evolutionary biologist produced a most widely received theory that in the world of scarcity, genes that animate the robot-like people are obsessed with maximizing their number of replicas by making individuals engage in what he repeatedly calls a "ruthless" struggle for survival.

IV. Principal motivating force: moral sentiment

Once the Confucian argument is made that economic wealth is a function of the work effort rather than of taking control of others' physical resources, the question arises what motivates individuals to adopt this particular strategy. In Confucian economics, work effort is assumed to be driven by peoples' concern for others. Confucius categorically stated: "You live for others, not for yourself." This is also the correct meaning of Mencius's related general point on the ethics of life: "To use the world to care for someone is the ultimate in care" (Norden, 2009, Book 4.3) He meant by this that individuals are motivated by their concern not for themselves but for their family. More precisely, it is about obligations to an extended family, or clan, that traditionally has been a dominant form of organization in China. Such an assumption turns Chinese economics into ethics, understood as an art of living. This should come as no surprise, since nature is seen by Confucians as a moral order, so every aspect of human life, including the process of wealth creation, has its moral root as well.

This is another instance of when Liberals will categorically disagree with the Confucians. To Liberals, the primary economic motive is not to care for others but to secure the most resources for yourself. Accordingly, what is said to make people act are property rights (North, 1981; Levy, 1988). Property rights are understood as legal and/or casual rules that ensure appropriation of rewards by individuals for putting resources to use. This assumption might seem to make sense, but there is a simple proof that a moral outlook and not economic security is the key motivator. It is self-evident that people go through a "life cycle", with three stages consecutively as children, adults and elders. Let's say that each stage lasts one-third of the lifespan. People can work only during this one-third when they are adults. During the remaining two-thirds of their life, as children and elders, they can't work. To survive, they need to rely on wealth transfers coming from working family members. Their survival is possible due to working adults' sense of responsibility, which is the exact meaning of morality. It follows that at least two-thirds of the management of wealth is driven by ethical forces. For Liberalism to ignore this two-thirds of economic activity that relates to transfers of wealth is a measure of its internal limitations as a discipline.

Morality is ignored by Liberal economists as an economic variable, since bringing subjective concerns, like care for others, into their model of behavior only detracts from the assumed

rational individual's ability to make optimal choices. But how it is that morality – as a rule – undermines rationality, since morality is developed spontaneously to assist people in survival? Let's assume that within the simple Confucian model of life introduced above, the working adult refuses to help an elderly parent. While the adult's instant gratification will double, this refusal won't go unnoticed by any children that the adult supports. When this utterly heartless adult becomes elderly, the child won't feel any obligation to help its parent to survive. How rational could the adult's refusal could be, since without a financial transfer from the child its lifespan as an elderly will be abruptly cut by as much as one-third. Not surprisingly, the central tenant of the Confucian concept of a virtuous life is the uniquely Chinese concept of so-called "filial piety", i.e. a feeling of complete devotion to one's parents at whatever cost to the child, even its life. This wisdom is well preserved in the following passage from Confucius:

> "Filial piety is about pleasing parents – by understanding them – to achieve 'peace of mind'" (from Lin, 1938).

One can look differently at the transfers and notice that at each point in their work as an adult, a given person must consider transferring a sufficient portion of their generated wealth to other family members. If the working member is single and there is one child and one elderly parent – to support them materially, he/she would have to sacrifice 2/3 of their earned material wealth. And, in a larger family, with two working adults and their two children plus two elderly, the ratio will be 4/6, or 2/3. And, this burden is not in a single year but during the whole one-third of his/her life of an adult. This, by all means, is an extreme financial sacrifice, so it would require some extreme moral attitude. As Confucians argue, this sort of extremity requires an unconditional love, one with no gain in sight other than a "gain" from displaying love itself (compare with Fromm, 1956). This is the reason, why Confucius himself said the words: "Love is man" and Mencius repeated them (Classics, vol. I, p. 405; and vol. II, p. 485). But man to him is life, so it follows that "Love is life". "Life" in turn is the energy "qi", which comes in two forms, female "qi-yin" and male "qi-yang". Thus, the most accurate way of looking at the empowering force of "qi" is to say that "Qi is love".

The logic of the Confucian argument is that while as a factor of production labor – work – is the source of value, or wealth, all ingredients or threads in the human life rest on ethics. Therefore, it is ultimately ethics – its condition or strength – that is responsible for peoples' motivation to seek wealth by doing work. It can also be said that it is not so much work, but it is morals that generate wealth. Therefore, best understood as a factor of production, it is morals that are wealth. In the words of Confucius: "… the material prosperity of a nation does not consist in its material prosperity, but in righteousness" (from Lin, 1938, pp.136-7) He expressed this idea also in a shorter line: "… doing the good is our only treasure" (after Lin, 1938, p. 135). This idea is echoed by Mencius who interpreting Confucius passage from in the "Great Learning" – chosen by Chen, (1911, p.293) – wrote that according to his master:

> "Having virtue there will be a man. Having the man, there will be the land. Having the land, there will be wealth. Having the wealth, there will be its use" punctuated with the memorable sentence: "Virtue is the root, and wealth is only its outcome" (Classics, vol. I, p. 375).

V. Basic economic institution: the "family unit"

Each approach to economics requires that it specifies the type of institution, or system that is most important for wealth creation. This is a critical question, since in their wealth creation individuals set up rules – or routines – that make their economic pursuits most effective. The system that secures the best conditions for individuals to maximize their contributions to wealth creation must be the most effective. The principal concern of liberals, with their focus on "solitary individual" who strictly follows the dictates of his/her reason, is that the system maximizes the individual's autonomy. The individual should have complete access to information and no shackles to limit his/her choices from the menu of alternatives available. The reason shouldn't be, for sure, disturbed by subjective considerations, such as moral concerns, or sensitivities. The institution that to them guarantees complete autonomy is a "free market", and where competition determines both prices and the fate of players, whether they stay in the game or go bankrupt.

In "Confucian economics", as it should be rather clear by now, the principal institution must be the family as an alternative to the free market, where individuals act on their own. Given the highest importance of morality in the pursuit of wealth for the institution of the family to play such a vital role in society as a "work unit", the family has to be most critical for the strength of morality. This is exactly the point made by Confucians about the role of the family in any economy, Chinese or not. They argue that family is so essential, because the family is where codes of morality are coming from. To be precise, people don't "invent" morality, but rather learn its intricacies from nature. As a "moral order" nature provides people not only with resources but also with moral rules of economic conduct, waiting to be realized. Family is where this takes place, since this is where the various most important moral choices are made, for example, to bring a child into this world (i.e., birth). Besides, the strongest incentive for acting morally comes from within the family itself, since no other type of associations are as close – or intimate – as the "kin (blood) ties".

What is largely missed in Western literature on China is that while Confucians assume family as superior to markets they still find markets indispensable. It follows from the fact that production is diverse in terms of what people make and in what size, so that there should be an exchange of goods for the benefit of each other. Mencius gives this account of markets:

> "If you do not exchange goods for raw materials so as to make up for what people lack with what people have in surplus of, then the farmer will have an excess of grain, and the woman will have an excess of cloth. If you do exchange them, then the wheelwright and carpenter will all get food from you" (Norden, 2009, p.37; Book 3 B, 4.3).

People will willingly engage in market exchange on their own, unless the state is mismanaged and allows a collapse of social order. To remedy economic failings the only effective way is for the state agency to:

> "… simply return to root … to bestow benevolence in governing. This would cause … all those who plow to want to plow … all traveling merchants and shopkeepers to want to place their goods in … markets" (Norden, 2009, p. 8; Book I A, 7.18).

The "Confucian system" is different from the "capitalist system" not because only one of them, the capitalist system, is based on the market as a principal institution. Throughout the whole of human history, markets have come in different versions to meet different needs that drive peoples' economic actions. The real difference between the two systems discussed here is that they operate different types of markets. "Free markets" – advocated by Liberals as an apparently superior institution – are based on competing individuals, while the markets advocated by Confucians are based on competing families. By assumption, the former type is unrestrained by ethical rules and follows the logic of competition itself, while in the latter type competition is "embedded" in ethics with its own logic. To capture these distinct institutional features, it is best to call such a Confucian systemic design a "familial market" Keeping this distinction in mind, markets are thus not so much a Western invention, but are much – or more – of Chinese origin (see: Pomeranz, 2013).

Speaking of the centrality of family, it is not all about the critical impact of "work ethics" on the growth of production, but also about the fact that the "familial market" can produce goods that the "free market" can't. These happen to be products that fall in the already identified category of "superior goods". Superior goods are those that, when their production cost goes up, rather than reduce their demand people demand more of them. Confucians argue that such goods are not provided by "free markets" but by "familial markets", even more from outside of the market context, that is, by the family unit itself. Among them is sense of security, parental affection etc. The evidence of their superiority is that with growing real wages in the "market sector" people don't elect to work more hours. They cut their work week in the "market sector" (e.g., down to four days in France). This is to spend more time in the "family sector" with the loved ones they care most for. Not surprisingly, though still having a status of a developing economy, the Chinese, being family centered, in both the formal and the informal sectors, enjoy a very low retirement age, both for males and females.

As a rule, there is no perfect information, but in the case of "instant gratification", transactions are frequent and prediction can be relatively easy. The risk of missing an optimum solution is thus not excessive. In these situations, risk-averse individuals can act at least approximately like liberal "rational actors". But in the case of seeking continuity of life, information required for "rational choice" is simply not available. In this pursuit of, as I will call it "distant gratification", couples would have to have access to ample information on costs and benefits for the remainder of their life. Of course this is not possible, so if couples were acting like "rational actors" and avoiding risk at any price, they would be discouraged and wouldn't ever choose to have children (Morse, 2001). In real life, as we all well know even from our casual observations, people have children, so they can't be "rational" in the strict – classical – sense. That this is a different kind of rationality, doesn't mean that continuity as an objective is not an economic issue and thus not a subject for economics to seriously analyze.

For "rational choice" to lead people to an "equilibrium", a fine-tuned "free market" must generate perfect information in the form of the so-called "free prices". While this precondition can be met, though imperfectly, in the market for goods satisfying instant gratification, there is no such a thing as a market for procuring children as a way to achieve one's posterity. Rare attempts have been made in Western economics to model such a market by assuming children are tradable like any other goods (e.g. Becker, 1989). Of course, there can be no such markets, though in places like traditional China there were some – limited to relatives – "exchanges" of children. Children are not tradable since there is no sizable supply of them, nor is there a significant enough demand for them to make a market operate. If legacy is an objective, disincentives for selling ones' offspring must be enormous. Buying is also out of the

question, because people want children of their "kin" which they can't ensure only by breeding their offspring.

VI. Egalitarian distribution: seeking social harmony

The next issue to examine in order to develop a full picture of the economy is to resolve the paradigmatic issue of income equality. On this issue, the Confucian assumption is that for economic success a very strict cross-society equality is required. Speaking of the merits of equal distribution, it is critical to look first at political consequences. Confucians believe it is imperative for wealth creation that there is "social peace" and because, as moral beings, people develop a sense of "social justice", there is no peace without justice. Since the differences among individuals in work ability are relatively small, what is found just by people is that distribution of income is equitable. Only then a "social peace" is ensured, or in Confucius's words: "When wealth is equally distributed, there is no poverty … [and] there is no dissatisfaction, [and] the country is secure", meaning stable. Conversely, the noticeable inequalities are perceived as a proof that abusive practices were allowed. This leads to moral resentment or even "social disorder", as the worst enemy of wealth creation.

Confucians argue that the major way for people to decouple their wealth from their work effort is by establishing a monopoly. With a dominant position as suppliers, monopolists can manipulate amount of goods offered. By artificially creating shortages, they can raise prices above the "free market" level. The difference in price levels is called "superior profit", or "rent". Those who have to pay rents resist doing so. As a consequence, "social peace" can be disturbed. For Confucians, to avoid such an adverse outcome the agency of the state – be it emperor and his bureaucracy or some other institution – has the moral duty to prevent rents from occurring and taxing them out existence if they have appeared. Such form of income equalization is postulated in order to correct "immoral behavior". This happens to be the only reason for Confucians to advocate income equalization. For fuller clarity, in this case income equalization is not about implementing some calculated – socially – optimal pattern of wealth distribution but rather about denying certain sources of incomes.

While the general rule is that the state should collect low taxes, no higher than a tenth of farmer income earned by all classes, including farmers. An exception to this rule is made in Chinese economics for the cases when some monopoly abuse is involved, regardless of whether this is private or public monopoly of supplies. Confucius articulates this taxation principle this way: "Riches … are what men want. But if they are obtained in an improper way, they should not be held" (Classics, vol. I, p.166). Mencius captured the same principle for designing taxes in the following quote:

> "When the ancients had markets, they were for exchanging what they had for what they lacked. The officials merely kept order. But there were some base fellows there who would seek for a "vantage point" and climb up on it. They would gaze left and right monopolizing the profit from the market. Everyone thought they were base, so they followed up by fining them. Taxing merchants had its origin in dealing with these base fellows" (Norden, 2009, Book 2B 10.7).

The Confucian call for a strict equality of income sharply departs from the Classical view, where there is no such arguing in favor of equality of wealth on similar grounds. Nor does it

call for modesty to limit apparently unlimited appetites for goods. Liberals are basically against the concept of "social justice", as a moral construct. To them, "social justice" implies some kind of theoretical optimum, where transfers of wealth among people raise the total welfare precisely to the highest level possible. However, as the Liberal argument goes no such optimum can ever be calculated by anyone with necessary precision. The most precise mechanism for allocating wealth among people is the "free market", where no one is a judge of what is socially just and what is not. There is thus no need for a separate theory of distribution and having a production theory is enough. This is claimed to be the case, since under the Liberal rules of "free market" all "contributions" to wealth are rewarded proportionally. Therefore, there are guarantees that the pattern of wealth distribution doesn't upset the search for most efficient use of physical resources, which are assumed to be scarce.

That calculating the socially optimal pattern of distribution can't be perfect, wouldn't mean to Confucians that there is no justification for redistribution across society. This is inescapable, since, however foggy in their minds, people always have some "sense of justice". Besides, to sensibly argue that allocation of wealth should be left to a free market, presumes that as defined this agency can guarantee that proportionality will be perfect. For distribution to be perfect by Liberal standards, the free market as envisioned would have to be perfect. But how can the distribution be left to "free market", since "free market" has a built in tendency to become imperfect? Here producers compete on the basis of their efficiency and the reward for the more efficient players is that – through bankruptcy procedure – they take over production from the less efficient, so that monopolies emerge. Freed from "free market" discipline, they can choose to be "alert" and further thrive on efficiency or to be "lazy" and earn "rents", which is anathema to the strictly "egalitarian" Confucians.

VII. "Meritorious state": providing moral guidance

Another "building block" for any school of economics is the concept of state, in relation to two other agents, i.e. the individual and the family. The Liberal view – with Smith (1776) taking the lead – is that as an instrument of power the state should be a "minimal state" with very little intervention. This is a logical implication of treating "fee market" as the principal institution since it provides most "autonomy" to individuals. Western historians believe that, as a manifestation of progress, there is a tendency over time for this individual "autonomy" to expand by replacing state control with market control of economic life. They also argue that the role of the state in China has always been excessive since operating a "maximal state" that grabs all the power available is here not a matter of choice but a primal instinct. The records of the past reveal that this claim is clearly a gross mischaracterization. The Chinese concept of state, like the Western one, also fits into the category of a "minimal state". But a "minimal state" doesn't have to be a powerless – or near-absent – and the case of China proves this particular point (compare with Vries, 2015).

While in the Liberal economics a "minimal state" is preferred for the sake of freeing individuals to pursue their own destiny, in "Chinese economics", the reason for a minimal state is to ensure "autonomy" of the family. States are expected to serve families and to serve them best. States must therefore take on themselves only the least amount of duties – tasks – to aid families. The state should undertake only what families can't do more effectively for themselves. With the wealth of knowledge about specifics, families should be able to run their own life better than the state can. This must be particularly true for a country as huge

territorially and as populous as China has always been. As an evidence of this preference for a "minimal state", there are records of imperial instructions to provincial authorities to try to handle local issues or face penalties, even jail time if they pass an undue burden onto the state. In Confucian thinking, the states' "hands-off" policy itself enhances state power to rule the country, since by concentrating on fewer issues, the state can obviously be more diligent in taking on economic issues that matter the most.

What then is the main role of such a minimal state, which requires low taxes and is in some sense "withdrawn"? The Confucian thinking is that what is most important for the economy, must also be most important for the state. Since within this doctrine morality is the main factor behind wealth creation, the principal role of the state must be also to ensure that ethics are upheld. For the state to propagate morality among the ordinary people, the state itself must follow the ethical prescriptions, i.e., be incorruptible. Accordingly, similar to the Confucian concept of market, the Confucians concept of state is a moral one. It is the contents of the underlying ethics that makes such a state – as much as the "familial market" – uniquely Chinese. The goal of providing a moral example for the society at large might look like a rather limited kind of duty for the state to carry on, except that by stressing moral aspect this concept helps the state to increase further its power to control practical aspects of economy. Since the state's power is a function of the trust that it earns with the people based on the moral example it sets, once the trust is widely earned running the country should be an easy exercise.

Going further, the Confucian doctrine of state agency makes it clear that if there would to be no attempts to violate the protection of individuals as owners or otherwise, legal protection by the state bureaucracy wouldn't be necessary. To accomplish compliance by fear through purely legal means would be cost prohibitive and thus detrimental to the wealth creation that it is supposed to enhance. In this context North (1981) mentions morality – but also the ideology, and religion – as ways for reducing transaction costs. This is also the Confucian position, except that to Confucians the purpose of societies should be to reach the stage when interactions are not regulated by legal means but rest fully on moral resolve. If there is a moral harmony people will voluntarily comply and society will avoid or even eliminate any costs of enforcement. Confucius said emphatically that:

> "The thing is we should make it our aim that there may be no lawsuits at all, so that people who have done wrong will be too ashamed of themselves to indulge in words of self-defense" (Lin, 1938, p.129)

Thus, as the highest value in Confucian doctrine, moral harmony can stop destructive litigation, which will bring in the ideal outcome, the "end of law".

The other distinctive feature that allows the Confucian version of the minimal state to be an effective actor is the utmost stress on the professionalism of state officials. The Confucian ideal is that of a "meritorious state", or meritocracy (Bell, 2008). For Confucians, authority to rule over others should be given only to individuals selected from among the largest pool of talent. Selection of the most capable must be left to a strict examination system that anyone, regardless of position or wealth can join at any point. The selection process involved must guarantee complete impartiality, e.g. by conducting multiple blind tests. Massive reading material for the tests required candidates to devote years to absorb and memorize. From the earliest days, China has operated such a scheme, often called in literature a "mandarin system". This system was formally abandoned in 1911, but even after the revolutionary shift

in 1949, China's state has retained its general rules, though without explicitly appealing to Confucianism. What was there to erase from collective memory, given that the Confucian meritocratic state was a primary invention, and which, hundreds of years later influenced Europe, in part by way of the German sociologist Max Weber and his ideal of a "modern bureaucracy".

The primacy of the family relative to the state agency is repeatedly stressed by Confucius and his followers. Whatever the importance of the state for economic life, what matters most is that – in his view – it is family where primarily or ultimately the strength of the state comes from. This assumption about the family is expressed in the following quotation by Mencius:

> "The root of the world lies in the state; the root of the state lies in the family; the root of the family lies in oneself."

In this passage, "oneself" doesn't refer to the autonomous individual of Liberalism, but an individual as a family member, one who squarely operates within the confines of the family (after Norden, 2009, Book 4A, 5.1). What is meant by this specific hierarchy of power where family overrides the state, is that because to succeed the state needs to be virtuous – or uncorrupt – the families that the state rests upon must be virtuous. This critical point was made by Confucius:

> "… when the family life is regulated, then national life is orderly, and when national life is orderly, then there is peace in the world" (after Lin, 1938, p.106).

VIII. Preferred methodology: use of "common sense"

The final "building block" for each school of economics is a choice of the preferred methodology, including the definition of the proper research agenda. To Confucians this agenda is a moral aspect of life. More specifically, they call for the examination of threats to the core of ethics, which comprises "*moral laws*". Nature provides all individuals with what is needed to ensure continuity so that in addition to abundant supplies of resources, nature also provides people with "moral laws". Mencius states it plainly: "There is no human who does not tend towards goodness … (from Norden, 2009, p. 69; Book 6 A, 2.2-2.3) Coming as a birthmark, this predisposition for what can be called "*natural goodness*" can be lost due to two kinds of threats. One is the failure to cultivate virtuous behavior, a task that can never be finished. The by far more important threat is the submission by individuals to wrongful ideas. Mencius stated: "Moral capacity can be threatened by "pernicious doctrines" (e.g., the theory of "egoism" of Yang Zu (440-360 BC). He also used a more aggressive wording, repeatedly calling such threats "*evil doctrines*", with evil consequences. In Mencius view: "Evil doctrines and cruel practices …" cause political chaos, which, as said earlier, Confucians view as the main cause of economic misery (after Norden, 2009, Book 4A, 4.9).

With the research agenda set, the related methodological question for economics is that of validation of assumptions made and in the case of Confucian thinking – for the reasons just mentioned – this validation is about methods for identifying "pernicious doctrines". According to Confucians, the method of choice is for individuals to use "common sense", i.e. to examine personal experience without turning for help to specialized – or theoretical – knowledge (compare with Popper, 1972). Mencius stresses this idea:

> "The best teachings are those that discuss what is near but with significance that is far-reaching" (from Norden, 2009, Book 2A, 37.12).

Motte (1971, p. 26) captured in Chart 3 this principle of Confucianism: "The Chinese world view kept man's attention on life here and now". Discerning "moral laws" as singularly correct is to them an empirical issue of whether one's actions serve others or harm them. There is no place for theory in addressing such moral issues by creating categories and building models, which is the Western method of choice for addressing social issues. For Confucians, while nature speaks with one voice and has one set of rules for all, theories can be multiple and may also contradict each other. While theories can meet standards of logic, they may question the merits of "moral laws". It is because theory is powerless in finding which one is "true" that Confucians detest theory, as it is understood in Western thought.

To most Western scholars, this – seen as ostensibly anti-scientific standing – is a sign of China being in a "pre-scientific" stage of knowledge that is ruled by superstition. This opinion was expressed by Weber (1951) mentioned earlier, who was one of the first major Western academic major figures to venture into examination of Chinese tradition. He attempted to explain why China failed to instantly join the "Industrial Revolution". In his view, Britain gained advantage by switching from superstition to "science". China failed to engineer this shift, because:

> "In spite of the logical qualities of the language, Chinese thought has remained rather stuck in the pictorial and the descriptive" (Weber, 1951, p.125).

He continues his critical assessment with:

> "The power of logos, of defining and reasoning, has not been acceptable to the Chinese … Chinese philosophy did not give birth to scholasticism because it was not professionally engaged in logic … based on Hellenistic thought. The very concept of logic remained absolutely alien to Chinese philosophy … oriented to purely practical problems" (Ibid. p.127).

The escape from the entrapments of the theory is one method for identifying threats, and Confucian disinterest in history is another. This disinterest in history shouldn't be confused with disinterest in tradition. Tradition is to be greatly cherished, but not necessarily when the past is filtered by biased historians. In the Confucian view the practice of history as a discipline rests on shaky grounds, since material from days gone is sparse and of course can't be related to direct experience. Consequently, history can be interpreted in many ways that can't be tested for full accuracy. The worst that can happen is when theory is brought to historic analysis to imitate "modern physics" and like this brand of physics looks for "iron laws" of history, with fixed patterns of change and precise predictions of the future (see: Popper, 1957). Again, there is an issue of how to separate a "true" from the "untrue", because the application of logic to these laws can't really help. It is better to be skeptical about uses of history, particularly for judgment on issues that are of a moral nature. And, as the Chinese traditionally do it, they use history only sporadically to illustrate "moral lessons" for the living and do so descriptively, mainly for the purpose of archiving what is left of the past (e.g., record books – and genealogies – kept by the heads of family).

There is one more preferred method for people to safeguard against threats, namely to "filter" from whom ideas come. For Confucians, the most trustworthy sources are "sages". Educated, they are not mainly scholars but rather moral teachers fully versed in Confucianism. Ascribed the highest rank in society, sages are considered more important than emperors. Expecting the most from emperors, Confucius claims that only an emperor who is also a sage can rule (compare with Vries, 2011). The emperor is not law, nor is he above the law of the land. The same rules apply to all, for there is one "moral order" that nature provides. Emperors' authority, or the "Mandate of Heaven", is bestowed not by "heaven" but by the people. Their verdict is based on how he the ruler treats people. In Confucius's words: "Loving what people love, and hating what people hate: this is he who is called the parent [emperor – KP] of the people" (from Lin, 1937, p. 77). Otherwise, when this requirement is violated, people should rebel and dethrone an unjust ruler (rebellion rather than – state – war being the main source of political "havoc" in China's millenniums of history (Lin, 1938).

Conclusions

Demonstrating that Chinese have their own complete brand of economics is the first step; the next one is to discuss its real-world application and establish what the relative worth of this approach to economy is. Left with limited space, I will use my conclusions to point to the evidence of influence that Chinese economics exerts on today's China. Ideally, these remarks could stimulate future research on Chinese approach to economy. Beginning with the opening issue, one wonders whether the Chinese still show overwhelmingly preference for continuity over "instant gratification". One might doubt this being the case, since in 1979 China adopted "family planning", i.e. "one-child" policy. Yes she did, but this is no evidence that a major change in mentality has taken place. The idea to restrict procreation didn't emerge from within China, but was picked up by the state from Western academics prematurely alarming that "population explosion" is nearing and will soon destroy the globe. Control of birth was installed by the state, but the population has never accepted it. Even some officials have voiced concerns, mainly that it will lead to "population implosion" From the outset, the policy has been opposed by the masses, particularly by the still dominant peasants. Though people have been getting richer, the opposition didn't subside, so that the state finally decided in 2015 that the restriction will be lifted, starting with the educated class.

Turning to the second issue of the "economic problem", the question is how "Confucian" China is in terms of using its resources. To follow Confucian principles, China would have to show interest in giving resources a "breathing space" to be replenished. The evidence loudly speaks to the opposite. As a latecomer, China is clearly replicating – or even exceeding – the damage experienced by European countries at the time of their own – early – industrialization. The drop in water level in many, or most rivers and lakes is quite alarming, and air quality in most of the major cities – agglomerations – is alarming as well. The pictures from the capital city Beijing make one wonder how much harm these conditions cause to the health of inhabitants, and – thus – what must be the cost of failing health to the economy. This sorry status is very well documented, particularly in Western research (recent example: Smith, 2015) But this deterioration of environment – land and water – is no evidence of scarcity that Liberals attribute to "population expansion". It isn't nature that is the cause, but people, who violate the Confucian principle to "follow the seasons" and allow nature to regenerate (e.g., through poor urban planning, underspending on water purification). The only area where China has made some advances is energy. Low on conventional sources – coal

and oil – in almost no time, China has become the world leader in alternative sources, solar and wind energy (as well as in the related manufacturing branches).

Is there any evidence that today's China follows Confucian – "puritan" in spirit – attitude to work? The answer is, there is ample evidence that – earnest – work is viewed by people as a way of life, often even a meaning of life. Any visitor can see it in today's China, where people typically start their work at 8-9 am and stay there even for as late as 10pm. They work also on Saturdays, even though – not that long ago – free Saturdays were introduced cross-China. Free Sundays were instituted in 1911, when dynastic system was put to an end. Mostly state officials took advantage of this new law, but not the peasants. Until now, free Sundays, or two-day weekends are relatively rare among them. It is a common practice that officials are expected to attend meetings during weekends. In the private sector, dominated by small shops, people use fronts of their hoses for business – work – and the back is for family life. Thus, in practical terms, they are constantly open for business. And there is still no tradition of having holidays, except for two per year official holidays, called festivals. Incidentally, these one-week long festivals, are family gatherings, devoted to rekindling the Confucian spirit (to digress, about 300 million children travel to meet with parents, in these largest modern-time migrations).

There is also evidence of strong respect for Confucian ethics, with "filial piety" as its cornerstone. In practical sense, it is a requirement to transfer wealth from children to elderly. As documented by studies on income distribution in China, it is common nowadays for children to transfer as much as one third of their income to retired parents, who used to earn much less. Following tradition, parents typically live under one roof with their children. To be specific, as much as 60 percent of elderly parents – older than 65 – reside together with their children (Zhang, 2014). This fact explains why in China most of the new apartments for young couples are designed with extra rooms for parents /and why Chinese living in America prefer houses with facilities for in-laws/. But the transfers are not one-sided, since parents sacrifice for the benefit of their children as well. The major help for children is provided by parents through sharing the burden of raising their grandchildren, which further explains why three generations tend to live under one roof. By taking care of their grandchildren, parents allow their own children devote more time to work and contribute to the overall – family wealth. Importantly, it is this wide-spread family practice that allows hundreds of millions of migrant workers – mostly from among married couples – to assume urban jobs (and return home when jobs are gone).

China also seems to have followed Confucian "recipe" for prosperity by using the post-1978 reforms to hand over the economy mainly to families, or – to put it differently – by delegating its control to the "familial market". Like in many other countries, "family" is not a category used for national economic statistics, so to give a definite answer on how much China's production comes from family businesses isn't easy. However, there are many – I would add sufficient – indications that the choice by the reformist leaders was to shift management of resources to families rather than individuals, of which the later would be a Liberal choice. Returning the state-owned farms – or communes – to families (and their traditional clans) was one step toward a family based system (Naughton, 1996). Initially, peasant families were allowed to use state land for a payment and later they were also given right to acquire land and thus enjoy limited ownership rights. Accordingly, an internal market for land was rather quickly reestablished and speculation became rampant. These, and other dynamic changes in the farm sector eventually reached the straggling industry. The subsequent decision by the party-state to assign the task of industrial revival to the so-called "township" – in fact family-run –

firms, is another valid proof that the recent China has followed the Confucian "recipe" for success (see: Rawski, 2011).

There is then the issue of Confucian concern for equality. Before the post-1978 reforms, difference in wealth (but not status) were insignificant, and there was a wide-spread poverty. Judging by the formal coefficients, China has reached extreme inequality. In 2016 she reached the US ratio of 0.5, compared with 0.3 reported back in 1980. This would suggest a sharp departure from Confucian principles, but one should be cautious. Chinese economics has given the highest priority to eliminating poverty. During the recent decades, as a rare exception among developing economies, China eliminated most of the endemic poverty, numbering over 300 million people. Average Chinese experienced constant growth in real wages at rates close to the productivity growth. For much of this time, real wages were going up at 6-7 percent annually. By contrast, in other countries that around the same time decided to reform a similar Soviet model – under liberal guidance i.e. the so-called "Washington Consensus" (Kolodko, 2012) – real gains from reforms for the wage-earners were much lower. In extreme cases, even with some production growth, by the year 2015 real wages were lower than in 1990 (e.g. Bulgaria, Serbia, but also Ukraine), or near stagnant (e.g., Russia) (Poznanska & Poznanski, 2015)

There are also indications that – generally speaking – China has preserved the quintessentially Confucian concept of "meritocratic state", where the state is mainly a "bystander" acting as a moral guide. While the imperial examination system was abandoned in 1911 by the Republican China, the political regime formed after 1949 has continued this tradition (not counting the upheaval of Cultural Revolution, when college entry exams and public service tests were temporarily suspended). As before, the brightest recruits are promoted to high party/state positions. To reach upper echelons, members of the party have to pass through a near dozen-step venting procedures (many times more than, for example, currently in the United States, Lee 2011)). Out of concern that officials "serve people" rather than themselves, they are systematically screened for ethics. Nevertheless, the corruption is rampant and represents a serious problem. The major stimulants for corruption are the "connections". Called in Chinese "guanxi", they are built mainly around family ties. This is a price for assigning – as Confucianism requires – the key role to family, rather than individuals or state. And, it is not that corruption goes unpunished since in contrast with most of the developing countries China operates a vast and harsh apparatus to police abuses of public office. Continuing the traditional dislike for mixing politics with business interests, Chinese party/state restricts the ability of business people to join the ranks.

Bibliography

"*Analects*" (translation Arthur Waley), New York: Alfred Knopf, 2000 (first edition London: Allen & Unwin, 1938).

Bell, D. ed., (2008) *Confucian Political Ethics*, Princeton: Princeton University Press.

Bell, D. (2014) *China Model: Political Meritocracy and the Limits of Democracy*, Princeton: Princeton University Press.

Brandt, L, D. Ma and T. Rawski, (2014), "From Divergence to Convergence: Reevaluating the History behind China's Economic Boom", *Journal of Economic Literature*, vol. 522, no. 1.

Chen, H. (1974) *The Economic Principles of Confucius and His School*, vol. 1 and vol. 2, Honolulu: Honolulu University Press (first print 1911).

Chen, K. (2013) "The Effect of Language on Economic Behavior", *American Economic Review*, 103 (2).

Friedman, M. (1970) "The Social Responsibility of Business is to Increase its Profits", *NY Times*, 32 (13).

Fromm, E. (1956) *The Art of Loving,* New York: Harper & Row.

Fukuyama, F. (1995) *The Trust: Virtues and the creation of prosperity*, New York: Free Press.

Fukuyama, F. (2011) *Lunch with FT: Francis Fukuyama,* by Wolf. M, (May 27).

Harrell, S. (1985) "Why Chinese work so hard? : Reflections on Entrepreneurial Ethics", *Modern China*, April.

Hobbes, T. (1651) *Leviathan*, New York: Dover Publications, 2014 edition.

Hsu, L. (1932) *The Political Philosophy of Confucianism*, London: Curzon Press.

Kolodko, G. (2012) *Truth, Errors and Lies: Politics and Economics in the Volatile World*, New York: Cambridge University Press.

Lardy, N. (2014) *Markets over Mao: The Rise of Private Business in China*, Washington D.C.: Paterson Institute for International Economics.

Lee A., 2011, *What the US Can Learn From China*, New York: Berrett-Koehler Publisher.

Levy, M. (1988) *Off Rule and Revenue*, Berkeley: University of California Press.

Lin, J. (2011), *China's Economic Development and Cultural Renaissance in the Multipolar Growth of the 21st Century*, Peking: National School of Development.

Lin, J. (2012) *Demystifying the Chinese Economy*, New York: Cambridge University Press.

Lin, J. (2012) *Benti & Changwu: Dialogues on Methodology in Economics,* Beijing: Cengage Learning Asia.

Lin, Y. (1937*) The Importance of Living*, New York: Reynolds & Hitchcock.

Lin, Y. (1938) *The Wisdom of Confucius*, New York: Random House. The Modern Library.

Mahbubani, K. (1998) *Can Asians Think*? Singapore: Times Books International.

Mahbubani, K. (2015), "When China becomes Number One", *Horizons*, vol. 4.

Malthus, T. (2003) *An Essay on the Principle of Population*, Cambridge: Cambridge University Press, (first print in 1798).

Martin, J. (2009) *When China Rules the World: The End of the Western World*, London: The Penguin Press.

Milbank, J. (1990), *Theology and Social Theory*, Cambridge: Wiley and Blackwell.

Milbank, J at all (eds.) (2009), *The Future of Love: Essays in Political Theology*, Wipf & Stock Publishers.

Morse, J. (2001), *Love & Economics. Why the Laissez-Faire Family Doesn't Work*, Dallas: Spence Publishing Company.

Mote, F. (1971) *Intellectual Foundations of China*, New York: Alfred A. Knopf.

Naughton, B. (1996) *Growing out of the Plan. Chinese Economic Reform 1978-1993*, Cambridge University Press.

Norden, B. (2009) *The Essential Mengzi. Selected Passages with Traditional Commentary*, Indianapolis: Hacket Publishing.

North, D. (1981) *Structure and Change in Economic History*, New York: Norton & Norton.

Pomeranz, K. (2013) *The Great Divergence: China, Europe and the Making of the Modern World Economy*, Princeton: Princeton University Press.

Popper, K. (1957) *Poverty of Historicism*, London: Routledge.

Popper, K. (1972) "*Objective Knowledge: An Evolutionary Approach*, Oxford: Oxford University Press.

Poznanska, J. and K. Poznanski, 2015 "Comparison of Patterns of Convergence among "Emerging Markets" of Central Europe, Eastern Europe and Central Asia", *Comparative Economic Research*, (1).

Poznanski, K. (1996) *Protracted Transition: Institutional Change and Economic Growth in 1970-1994*, Cambridge: Cambridge University Press.

Poznanski, K. (2005), *Negative Globalization: Privatization of National Capital in Eastern Europe*, Beijing: Chinese Academy of Social Sciences.

Poznanski, K. (2011)"*Market and Ethics*, Vienna: Vienna Institute of International Economy.

Poznanski, K (2012), "State Condition, Foreign Influence and Alternative Models of Market Reforms in China, Russia and Eastern Europe", Beijing: Chinese Academy of Social Sciences, *International Critical Thought,* vol.1, No. 3.

Poznanski, K (2015), "Confucian Economics: The World at Work", Beijing: Chinese Academy of Social Sciences, *World Review of Political Economy*, vol. 6, No. 2.

Poznanski, K (2017), "Confucian Economics: How Different is Chinese Thinking", Beijing: Beijing University, *China Economic Journal*, No. 3 (Summer).

Rawski, T. (2011) "Human Resources and China's Long Economic Boom", *Asia Policy*, vol. 12.

Schumpeter, J. (1935) *The Theory of Development. An Inquiry into Profit, Capital and Business Cycle* .

Smith, A (2003) *Wealth of Nations*, New York: Bentham House (first print in 1776).

Smith, R. (2015) "China's Communist-Capitalist Ecological Apocalypse", *Real-World Economics Review*, No. 71 (May).

Weber, M. (1951) *The Religion of China*, Glencoe: Free Press.

Wight, M. (1992) *International Theory: The Three Tradition*, London: Holmes and Meier Publishers.

Vries, P. (2015) *State, Economy and the Great Divergence: Great Britain and China*, London: Bloomsbury Academic.

Xi, J. (2014) *The Governance of China*, Peking: Foreign Language Press.

Zhang, Z. et. al. (2014) "Co-residence with Elderly Parents in Contemporary China", University of Michigan, PSC Research Reports, No. 14.

Zurndorfer, H. (2004) "Capitalism: Culture as Impediment and/or Stimulus to Chinese Economic Development", Third Global Economic History Network Workshop: Konstanz, June.

Author contact: kazpoz@uw.edu

SUGGESTED CITATION:
Kazimierz Poznanski, "Chinese economics as a form of ethics", *real-world economics review*, issue no. 80, 26 June 2017, pp. 148-170, http://www.paecon.net/PAEReview/issue80/Poznanski80.pdf

real-world economics review, issue no. 80
subscribe for free

A brief history of Pakistan's economic development

Muhammad Iqbal Anjum and Pasquale Michael Sgro [International Islamic University Islamabad, Pakistan; Deakin University Australia, Australia]

Introduction

This paper briefly describes the development of Pakistan's economy during seven distinct decades, from the late 1940s to the end of the 2000s. Pakistan has progressed from a low-income to a lower middle-income developing economy ─ GDP per capita increased from US$294 in 1980 to US$1,017 in 2010 according to the World Bank (2012).

The late 1940s: era of emergence of a new national economy

At its inception in 1947, Pakistan had a predominantly agrarian economy – agriculture contributed 53% of GDP in 1947, and 53.2% of GDP in 1949-50.[1] Then, Pakistan had a population of 30 million with 6 million people living in urban areas, 65% of the labour force working in the agricultural sector, and agricultural output contributing 99.2% of exports and about 90% of Pakistan's foreign exchange earnings.[2] Pakistan's resources, in East Pakistan and West Pakistan, were an immense reservoir of land and mineral resources – natural gas, crude oil, coal, limestone, and marble.[3] Pakistan had a per capita income of almost $360 (1985 international dollars) in 1950[4] and a literacy rate of 10%,[5] amidst economic crises – absence of economic infrastructure, financial resources, and industrial base.[6] Then, poverty incidence ranged from at least 55% to 60% in the West Pakistan.[7] Keeping in view the scarcity of capital in Pakistan's small private sector, the Government relied on the public sector for building the economic-cum-industrial base.[8] In 1949-50, Pakistan registered a national savings rate of 2%, foreign savings rate of 2%, and an investment rate of 4%.[9] Then, manufacturing contributed 7.8% of GDP and the services/trade/other sectors contributed 39% of GDP – a result of policy of imports-substituting industrialization.[10] Pakistan's trade balance of payments was in deficit by 66 million Rupees (Rs) during 1949/50-1950/51.[11]

[1] State Bank of Pakistan (2010), Husain (1999) and Zaidi (2005).
[2] Husain (1999), Khan (2002), Fasih-Uddin and Swati (2009), and Zaidi (2005).
[3] Fasih-Uddin and Swati (2009).
[4] Khan (2002).
[5] Fasih-Uddin and Swati (2009), and Zaidi (2005).
[6] There were small industries and few services, and there were almost no large-scale industrial units at all in 1947 [Husain (1999), Fasih-Uddin and Swati (2009), Zaidi (2005) and The World Bank (Undated)].
[7] Hasan (1997) and Hasan (2004, p. 63).
[8] Fasih-Uddin and Swati (2009), and Zaidi (2005).
[9] Hasan (1997).
[10] Hussain (2003), Husain (1999), and Hasan (1997).
[11] Zaidi (2005).

The 1950s: era of a traditional economy in transition

The 1950s was the first decade of planning.[12] After launching the Colombo Plan in 1951, Pakistan instituted a series of Five-Year Plans during the period 1955-1998 and a Ten-Year Perspective Plan alongside a rolling Three-Year Development Plan.[13] Pakistan continued its policy of imports-substituting industrialization during the 1950s.[14] During the Korean War (1950-1953), Pakistan's public and nascent private sector thrived on spectacular merchant profits, which were transformed into industrial capital that accelerated industrialization.[15] Pakistan banned the imports of cotton textiles and luxury goods in 1952 and regulated virtually all imports in 1953.[16] Consequently, Pakistan joined the group of the most rapidly growing countries in the 1950s.[17] But, anti-agriculture policy biases and anti-agriculture terms of trade between industry and agriculture caused the annual growth rate of agriculture to decline from 2.6% in 1949/50-1950/51 to 1.9% in 1957/58-1958/59.[18] After achieving self-sufficiency in cotton textiles in the late 1950s, export development assumed vital significance,[19] amidst inflow of the US military and economic aid of US$500 million during 1955-58.[20] Consequently, Pakistan entered a phase of foreign aid-dependent growth in the 1950s.[21] In 1959 − after military *coup d'état* in 1958, the martial law regime introduced export bonus vouchers, which were treated as import licenses, and free list of the goods, which could be imported without having any import license.[22] Balance of trade deteriorated from -831 million Rupees in 1950/51 to -1043 million Rupees in 1959/60 due to sharp decrease in exports from 1,038 million Rupees in 1950/51 to 763 million Rupees in 1959/60.[23] Agriculture grew at a rate of 1.6% per annum and manufacturing grew at a rate of 7.7% per annum in the 1950s.[24] In 1959-60, Per Capita GNP was Rs.355 in West Pakistan and Rs.269 in East Pakistan.[25]

The 1960s: era of economic growth

Amidst massive inflow of American aid, political stability enabled Pakistan to sustain high rates of growth in the 1960s.[26] Poverty incidence (poverty headcount ratio expressed as a percentage of population) ranged from almost 50% in the early 1960s to 54% in 1963-64.[27] In the 1960s, Pakistan achieved an agricultural growth rate of 5% per annum by achieving

[12] Zaidi (2005), Khan (2002), and Hasan (1997).
[13] Fasih-Uddin and Swati (2009).
[14] Hasan (1997), Husain (1999), and Zaidi (2005).
[15] Merchant profits referred to the profits realized from the Korean War-induced dramatic growth in exports of raw materials to the war-panicked countries, which were then piling up raw materials during the war [Zaidi (2005), Papanek (1996), and Hussain (2003)].
[16] Hasan (1997), Khan (2002), and Zaidi (2005).
[17] Zaidi (2005) and Husain (1999).
[18] Even negative growth rates in the agricultural sector were observed – that is, -9.1% in 1950/51-1951/52 and -0.8% in 1953/54-1954/55. With the then 75% of the population of Pakistan living in the rural area, the prolonged stagnation of the agricultural sector in the 1950s restricted further growth in the manufacturing sector (Zaidi, 2005).
[19] Hasan (1997).
[20] This US aid − a result of Pakistan-United States Mutual Defense Pact signed in 1954 − reduced the heavy burden of public expenditure on the budget of the public sector (Hasan, 1997).
[21] Nafziger (2012) and Hussain (2003).
[22] [Hasan (1997) and Zaidi (2005).
[23] Hasan (1997), Zaidi (2005) and Fasih-Uddin and Swati (2009).
[24] Fasih-Uddin and Swati (2009).
[25] Thus, the West-East Disparity Ratio was 1.32 (Zaidi, 2005).
[26] Chenery and Strout (1966), Papanek (1996). Zaidi (2005) and Khan (2002).
[27] Hasan (1997) and Hasan (2004, p. 63).

significant investments in water resources, increased incentives for farmers, mechanization of agricultural production processes, increased usage of fertilizers and pesticides, and the increased cultivation of high yielding varieties of rice and wheat.[28] The large-scale manufacturing grew at a rate of 16% per annum during 1960/61-1964/65 due to protection of domestic industry from imports and subsidies for exporters.[29] In the wake of the Pakistan-India War of 1965, the reduced foreign economic assistance caused the large-scale manufacturing to grow at a lower rate of 10% per annum during 1965-70.[30] Pakistan achieved an average annual growth rate of 6.7% in GDP during 1960-1970.[31] In 1969-70, poverty incidence declined to 46%[32] and Per Capita GNP was Rs.504 in West Pakistan and Rs.314 in East Pakistan – indicating a widening of the regional economic disparity noted earlier.

The 1970s: era of socialism and its aftermath

Because of growing interregional economic disparity, East Pakistan revolted against West Pakistan and became independent (Bangladesh) in 1971. Then, the martial law authorities empowered the socialist Pakistan People's Party amidst very difficult macroeconomic circumstances[33] − poverty incidence rose to 55% in 1971-72,[34] there was an increase in Pakistan's import bill due to the October 1973 world oil price shock, a serious post-1973 global recession during 1974-77, failures of cotton crops in 1974-75, pest attacks on crops, and massive floods in 1973, 1974, and 1976-77.[35] Pakistan experienced the worst inflation during 1972-77, when prices increased by 15% per annum.[36] During 1973-77, annual average fiscal deficit/GDP ratio was 8.1%.[37] Trade balance deficits were US$337 million in 1970-71 and US$1,184 million in 1976-77.[38] A military coup d'état occurred on 5 July 1977, and the martial law regime accomplished denationalization, deregulation, and privatization.[39] Agriculture grew at a rate of 2.4% per annum and the large-scale manufacturing grew at a rate of 5.5% per annum in the 1970s.[40] While the large and medium-scale private manufacturing contributed 75% of the value-added and 70-80% of the total investment in manufacturing in the 1970s, the remainder of the 25% of the value-added was contributed by the small-scale manufacturing.[41]

The 1980s: era of revival of economic growth

[28] Husain (1999) and Hasan (1997).
[29] Hasan (1997).
[30] Hasan (1997).
[31] Hasan (1997), Zaidi (2005), and Fasih-Uddin and Swati (2009).
[32] Hasan (1997).
[33] Socialism was instituted by nationalizing private industries/banks/schools/colleges and enforcing socialist reforms in the land tenure system [Husain (1999), Zaidi (2005) and Hasan (1997)].
[34] Hasan (1997).
[35] Hasan (1997) and Zaidi (2005).
[36] Hasan (1997).
[37] High fiscal deficits were financed primarily by means of inflationary money creation in the 1970s.
[38] Zaidi (2005).
[39] Due to a significant decline in the popularity of the ruling socialist party caused by both the high inflation and the nationalization policy, the ruling socialist party won the general elections of 1977 with simple majority, amidst the opposition's accusations that the general elections were rigged. Then, the opposition launched an anti-socialist regime political movement, which culminated in Pakistan's second military coup d'état. The martial law regime launched a program of accelerating economic growth via a massive inflow of the foreign remittances and the foreign aid from the Western countries in the wake of Soviet invasion of Afghanistan in 1979 [(Zaidi, 2005), (Lansford, 2012)].
[40] Fasih-Uddin and Swati (2009).
[41] World Bank (Undated).

Hallmarks of the 1980s were the reversal of the nationalization regime of the 1970s[42] and the revival of private sector's industrial investment, which led to high rates of growth.[43] Poverty incidence declined to 29.1% in 1986-87.[44] Unemployment rate declined from 3.7% in 1980 to 2.6% in 1990.[45] During 1985-88, the Government tried to implement the Islamic interest-free banking system, which introduced Islamic business partnerships between entrepreneur and the owner of capital based on the principle of sharing profits and losses.[46]

Pakistan achieved a national savings/GDP ratio of 16% in 1986-87 amidst massive inflows of worker remittances from the Middle East.[47] However, Pakistan experienced the problems of negative public savings and declining public investment/GDP ratio throughout the 1980s and used a large portion of the additional national savings to finance the enlarged fiscal deficits – a result of both the steep growth in the public sector's non-development expenditures and the tendency of the tax revenue/GDP ratio to decline – since the 1980s.[48] The increasingly enlarged budget deficits in the early 1980s were financed mainly via non-bank domestic borrowing.[49] Domestic debt grew from Rs.58 billion in mid-1981 to Rs.521 billion in 1988.[50] Consequently, the public debt/GDP ratio was 77.1% in 1988, 81.9% in 1989, and 82.6% in 1990.[51] This explosion of the domestic debt resulted in large interest payments, public expenditure, and fiscal deficits.[52] Democracy was restored in 1985.[53] During 1980-1990, Pakistan's average annual growth rate of GDP was 6.3%.[54] A manufacturing exports' boom occurred in the 1980s, with an annual large scale manufacturing growth rate of 8.8% and an annual agricultural growth rate of 5.4%.[55]

The 1990s: era of debt crisis

In the 1990s, Pakistan confronted the problems of declining worker remittances and rising external deficits.[56] In the wake of declining growth rates of GDP, there occurred the second worst inflation in the 1990s.[57] Unemployment rate sharply increased to 5.9% in 1991 and 7.2% in 2000.[58] Pakistan financed the enlarged current account deficits via the sustained increases in her residents' Foreign Currency Deposits.[59]

---

[42] Husain (1999) and Fasih-Uddin (2008).
[43] Fasih-Uddin (2008).
[44] Fasih-Uddin and Swati (2009).
[45] World Bank (2012), *World development indicators 2012.*
[46] Nafziger (2012).
[47] Hasan (1997), Husain (1999) and Zaidi (2005).
[48] Hasan (1997).
[49] Hasan (1997).
[50] Hasan (1997).
[51] Fasihuddin (2008).
[52] Hasan (1997).
[53] After both the dismissal of this democratic government by the President of Pakistan in 1988 and general elections of 1988, the newly formed government was dismissed in 1990. General elections of 1990 resulted in a new government. This scenario depicts higher political uncertainties, fast changing economic policies, and higher economic risks for investors (Lansford, 2012).
[54] Pakistan's average GDP growth rates remained higher than the average GDP growth rates of India, Bangladesh, and Sri Lanka for the periods 1960-1970, 1970-1980, and 1980-1990 [(Hasan,1997) and (Lansford, 2012)].
[55] Hasan (1997) and Fasih-Uddin and Swati (2009).
[56] Hasan (1997).
[57] Prices grew at a rate of 12% per annum from mid-1993 against a backdrop of explosive growth in money creation for supplying credit to the public sector, especially during the period 1990-96 (Hasan, 1997).
[58] World Bank (2012), *World development indicators 2012.*
[59] Hasan (1997).

In 1995, external debt amounted to US$30 billion – external debt tripled during 1980-1995.[60] During 1980-1995, the external debt/GDP ratio increased from 42% to 50%, the external debt/exports ratio increased from 209% to 258%, and the debt service ratio sharply increased from 18% to 27%.[61] Due to a seriously deteriorating profile of Pakistan's external liabilities – the prime cause of her foreign exchange difficulties after the first half of 1996 – domestic debt rose to Rs.909 billion, and the domestic debt/GDP ratio rose to 42%.[62]

A serious debt crisis occurred in the late 1990s, when the public debt/GDP ratio rose from 57.5% in 1975-77 to 102% in 1998-99.[63] Similarly, the public debt/revenues ratio rose to 624% in 1998-99, interest payments/revenues ratio rose to 42.6%, and the public debt became unsustainable.[64] Likelihood of external debt default emerged in 1996, and 1998, due to the Western economic sanctions imposed in reaction to Pakistan's nuclear tests on 28 May 1998.[65] Sanctions triggered massive capital flight.[66] This debt crisis occurred despite an agricultural growth rate of 4.4% per annum and a large-scale manufacturing growth rate of 4.8% per annum in the 1990s.[67] Poverty incidence sharply increased to 30.6% in 1998-99.[68]

The 2000s: era of economic crisis

In 2001, the official Debt Reduction and Management Committee judged the high public debt as a major cause of decline in the growth rate to less than 4 per cent per annum.[69] Debt crisis of the 1990s was followed in the 2000s by an era of macroeconomic crises[70] − in spite of improvement in the growth rate until 2004-05 when the growth rate was 8.6%, the subsequent years were characterized by growth slowdown and low growth along with high inflation, energy crisis, and deterioration in fiscal and balance of payments positions.[71] Poverty incidence increased to 34.5% in 2000-01 and then decreased to 22.3% in 2005-06.[72]

[60] Amidst this external debt fiasco, Pakistan met almost one-third of her foreign exchange gap via the use of volatile short-term liabilities in the form of the resident and non-resident foreign currency accounts (Hasan, 1997).
[61] That is, principal and interest payments/foreign exchange earnings ratio (Hasan, 1997).
[62] Hasan (1997).
[63] Debt overhang caused a decrease in the rate of investment to 15% of GDP in 1998-99/1999-2000. In the wake of this debt crisis, the third *coup d'état* empowered the third military regime on 12 October 1999 for the period 1999-2008. In spite of realization of some debt relief by Pakistan via her agreement with IMF regarding the rescheduling of her debt payments obligations, the possibility of her external debt default was not ruled out [(Hasan, 1999) and (Lansford, 2012)].
[64] Hasan (1999).
[65] Lansford (2012), Fasih-Uddin (2008) and Hasan (1999).
[66] Irfan-ul-Haque (2010).
[67] Fasih-Uddin and Swati (2009).
[68] Fasih-Uddin and Swati (2009).
[69] Pakistan's debt entrapped her in a vicious circle of high debt servicing, which caused stagnation in investment and growth as well as limited her capacity of debt servicing (Debt Reduction and Management Committee, 2001).
[70] Namely, a rise in the fiscal deficit/GDP ratio to 4.3% in the fiscal year 2000, a dire deterioration in the balance-of payments position in 2001, a reduced average annual economic growth rate of 2%, a negative trade balance alongside the signing of IMF's standby agreement with an exceptionally tough conditionality amidst the looming prospect of default on external debt during 1998-2001, the failure of public and private sectors to effectively use the inflows of US$62.2 billion realized during the period 2002-2007 as a result of the increased strategic significance of Pakistan in the post-9/11 scenario for establishing a robust basis for sustainable growth, the sharp deterioration in Pakistan's international investment position after the fiscal year 2005, the increasing conspicuous consumption, and a steep hike in the import bill in the fiscal year 2008, which caused unsustainable trade deficits [(Irfan-ul-Haque, 2010) and (Zakaria, 2012)].
[71] Fasih-Uddin and Swati (2009).
[72] Fasih-Uddin and Swati (2009).

Unemployment rate increased to 7.8% in 2002 and then declined to 5% in 2008.[73] Adult literacy was 55% in 2007-08. Pakistan experienced economic crisis in 2008[74] and the prime effect of global financial crisis in 2009-10.[75] In 2009-2010, the inflation-adjusted growth rate was a respectable 4.1%, the agricultural growth rate was 2%, industrial output growth rate was 4.9%, large-scale manufacturing growth rate was 4.4%, and the services growth rate was 4.6%.[76] In March 2010, the total public debt amounted to Rs.8,160 billion with a total public debt/GDP ratio of 56%, while the foreign-currency denominated debt/GDP ratio was 25%.[77]

Pakistan experienced remarkable development-oriented structural transition ─ GDP share of agriculture declined from 53% in 1947 to 21.2% in 2010, GDP share of industry rose from 9.6% in 1949-50 to 25.4% in 2010, and GDP share of the services rose from 37.2% in 1950 to 53.4 % in 2010.[78]

Conclusions

The history of Pakistan's economic development highlighted the key role played by the manufacturing sector. Pakistan progressed from its status as a low-income to a lower middle-income country and achieved her objective of poverty reduction.[79]

For sustainable growth, Pakistan needs to significantly increase national saving and investment rates, achieve budget surpluses for minimizing her domestic and external debt burden, and have political stability to promote a healthy investment climate for domestic and foreign investors, high levels of investment in human capital, and greater openness to international trade and private foreign investment.[80]

References

Ali, SS & Tahir, S 1999, "Dynamics of growth, poverty, and inequality in Pakistan", *The Pakistan Development Review*, vol. 38, no. 4 (Part II), pp. 535–50.

Arif, GM 2007, "Pakistan's poverty reduction strategy: successes, barriers and challenges", Proceedings of the conference on "Pakistan economy: an assessment with special reference to quality of life", Hanns Seidel Foundation and Institute of Policy Studies, Islamabad (Pakistan), pp. 43-57.

Chenery, HB & Strout, AM 1966, "Foreign assistance and economic development", *American Economic Review*, vol. 56, no. 4 (Part I), pp. 679-733.

Debt Reduction and Management Committee of the Government of Pakistan 2001, *A debt burden reduction and management strategy*, Finance Division (Government of Pakistan), Islamabad (Pakistan).

Fasih-Uddin 2008, *Pakistan under IMF shadow*, Institute of Policy Studies, Islamabad (Pakistan).

Fasih-Uddin and Swati, A 2009, *Pakistan's economic journey*, Institute of Policy Studies, Islamabad (Pakistan).

[73] World Bank (2012), *World development indicators 2012.*
[74] 2008 was the year of transition from the military regime into a democratically elected Government of Pakistan.
[75] Irfan-ul-Haque (2010).
[76] Ministry of Finance, Government of Pakistan (2010).
[77] Ministry of Finance, Government of Pakistan (2010).
[78] State Bank of Pakistan (2010) and Fasih-Uddin and Swati (2009).
[79] World Bank (2011).
[80] Papanek (1996).

Hanns Seidel Foundation & Institute of Policy Studies 2007, "Pakistan economy: an assessment with special reference to quality of life", in *World Economics*, Islamabad (Pakistan).

Hasan, P 1997, "Learning from the past: a fifty-year perspective on Pakistan's development", *The Pakistan Development Review*, vol. 36, no. 4 (Part 1), pp. 355-402.

Hasan, P 1999, "Pakistan's debt problem: its changing nature and growing gravity", *The Pakistan Development Review*, vol. 38, no. 4 (Part I), pp. 435–70.

Hasan, P 2004, "Poverty and social justice: some challenges for Islamization in Pakistan", in RM Hataway & W Lee (eds) 2004, *Islamization and the Pakistani economy*, Woodrow Wilson International Centre for Scholars, Washington D.C. (USA), pp. 57-63.

Husain, I 1999, *Pakistan: the economy of an elitist state*, Oxford University Press, New York (USA).

Hussain, A 2003, *Pakistan national human development report 2003: poverty, growth, and governance*, Oxford University Press and United Nations Development Programme Pakistan, Karachi (Pakistan).

International Monetary Fund 2012, *Balance of payments statistics 1960-2010*, International Monetary Fund, Washington (USA).

Irfan-ul-Haque 2010, *Pakistan: causes and management of the 2008 economic crisis*, Third World Network, Penang (Malaysia).

Kemal, AR 2001, "Debt accumulation and its implications for growth and poverty", *The Pakistan Development Review*, vol. 40, no. 4 (Part I), pp. 263-81.

Kemal, AR, Musleh-ud-Din & Qadir, U 2002, *Global research project: Pakistan country project*, Pakistan Institute of Development Economics, Islamabad (Pakistan).

Khan, MH 2002, "When is economic growth pro-poor? Experiences in Malaysia and Pakistan", IMF Working Papers Series, no. WP/02/85, International Monetary Fund Institute, Washington D.C. (USA).

Khan, MZ 1996, "Pakistan: Prospects for private capital flows and financial sector development", *The Pakistan Development Review*, vol. 35, no. 4 (Part II), pp. 853-83.

Lansford, T (ed.) 2012, *Political handbook of the World 2012*, SAGE Publications Limited, Washington D.C. (USA).

Ministry of Finance 2010, *Pakistan economic survey 2009-10*, Ministry of Finance (Government of Pakistan), Islamabad (Pakistan).

Nafziger, EW 2012, *Economic development*, 5th edn, Cambridge University Press, New York (USA).

Naqvi, SNH 2007, "Chair's remarks", in Hanns Seidel Foundation and Institute of Policy Studies (ed.), *Pakistan economy: an assessment with special reference to quality of life*, Hanns Seidel Foundation and Institute of Policy Studies, Islamabad (Pakistan), pp. 89-92.

Pakistan Bureau of Statistics 2011, *Pakistan statistical year book 2011*, Pakistan Bureau of Statistics, Islamabad (Pakistan).

Papanek, GF 1996, "Pakistan's development and Asian experience", *The Pakistan Development Review*, vol. 35, no. 4(Part I), pp. 343-82.

State Bank of Pakistan 2010, "Handbook of statistics on Pakistan economy 2010", database, State Bank of Pakistan, Karachi (Pakistan).

Tahir, P 1998, "The debt of a nation", *The Pakistan Development Review*, vol. 37, no. 4, pp. 331-53.

World Bank 2011, "World Bank list of economies (November)", database, The World Bank, Washington (USA).

World Bank 2012, *World development indicators 2012*, retrieved January 2012, database, The World Bank, Washington (USA).

World Bank Undated, *Pakistan and the World Bank: partners in progress*, The World Bank, Washington

(USA).

Zaidi, SA 2005, *Issues in Pakistan's economy*, 2nd edn, Oxford University Press, Karachi (Pakistan).

Zakaria, M 2012, "Inter-linkages between openness and foreign debt in Pakistan", *Dogus Universitesi Deris*, vol. 13, no. 1, pp. 161-70.

Author contact: miqbalanjum@iiu.edu.pk and pasquale.sgro@deakin.edu.au

SUGGESTED CITATION:

Muhammad Iqbal Anjum and Pasquale Michael Sgro, "A brief history of Pakistan's economic development", *real-world economics review*, issue no. 80, 26 June 2017, pp. 171-178,
http://www.paecon.net/PAEReview/issue80/AnjumSgro80.pdf

real-world economics review, issue no. 80
subscribe for free

Negating 1984: Michael Hudson's antidote to doublespeak vocabulary in economics

Michael Hudson, *J is for Junk Economics: A Guide to Reality in an Age of Deception* Dresden: ISLET-Verleg, 2017, 403pp. ISBN: 978 3 9814842 5 0

Jamie Morgan [Leeds Beckett University, UK]

Introduction

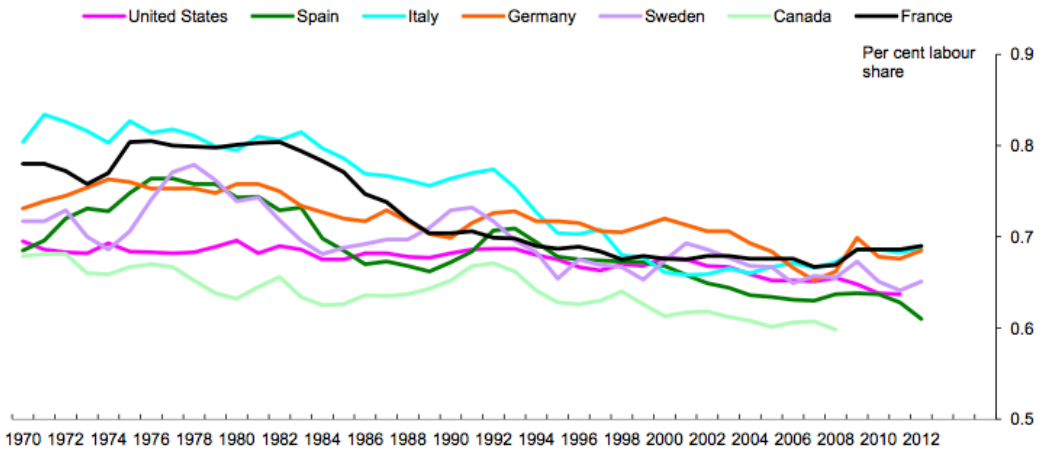
Michael Hudson is a professor of economics at University of Missouri, Kansas City, and Levy Institute associate. He is perhaps best known for his recent book *Killing the Host* (2015) and for his Harper's articles (2005; 2006) that identified key aspects of the financial crisis that was to manifest in 2007/8. *J is for Junk Economics* is ostensibly a dictionary, but a highly unusual one. The typical entries do not deal with terms commonly used in mainstream economics as they are used, but rather seek to clarify these in terms of the fallacies and mismeanings they encapsulate, whilst also providing entries that state the foundations from which the clarifications arise. That is, classical economics, post-Keynesianism, MMT and elements of new theories of money creation and financial system activity, as well as some insights from Marx. The range is eclectic and personal and yet also systematic, in so far as consistently thematic with reference to the structure, dynamic, and logic of actual contemporary financialised economies. The "dictionary" is in many ways an excellent work. It contains numerous pithy statements that concisely express important insights. It does so according to themes that follow directly from Hudson's more general concerns and that are set up in the introduction and developed throughout the text entries (and also in five appended previously published essays). The preface provides a brief account of the format of the book and how it came into being:

> "I drafted this dictionary and its accompanying essays more than a decade ago, for a book to have been entitled *The Fictitious Economy*. It did not find a publisher. My warnings about how debt leveraging would lead to a crisis hardly qualified as a timely how-to-get-rich manual of the sort that publishers consider to be popular 'economics books'. Most readers were making easy money in the stock market and real estate… Nobody wanted to hear that the gains couldn't be permanent" (2017: p. 7).

It is, of course, both tragic and ironic that the work would have been genuinely timely had it been published over a decade ago and that it was delayed precisely because of its relevance and point. This, as Hudson notes, extended to the incomprehension expressed by an anonymous reviewer who read the whole in terms of the kinds of misappropriation and distortion of classical economics that has helped to inform modern mainstream economics. For Hudson, the driving intention of classical political economy was to *free* markets from exploitative rent extraction, whilst today free market ideology, theory and policy consequences have become ways to facilitate that extraction. Hudson in general essentially justifies and contextualises *J is for Junk* as applied sociology of knowledge, and does so by initial reference to Orwell's *1984*. Economics now seems to have become a Ministry of Truth producing doublespeak and newspeak that reverse, invert, occlude or obfuscate regarding

actual states of affairs. According to Hudson, this is not purposeless, it serves the interests of a minority and enables the perpetuation of *caused* inequality. There is a real and continuing problem here that maintains the relevance of *J is for junk* even though the project was initially mooted more than 10 years ago. This is easily demonstrated (see also 2017: p. 179). For example, one can track the decline in labour share of national income over the neoliberal period:

Labour share selected countries, 1970-



Source: Bank of England/OECD in Haldane 2015: p. 28.

Economics that speaks falsity *for* power rather than truth *to* power

For Hudson, the causes of inequality include an ideational framework that involves a "learned ignorance" (see 2017: p. 141) or "trained incapacity" articulated by mainstream economics. This capture creates widespread passivity, confusion, and a sense of powerlessness (in so far as the way things are, is naturalised and becomes ultimately either a claim that it is to everyone's benefit or without reasonable alternative, and so suppresses alternatives). The capture has its own language:

> "The academic curriculum has been hijacked to replace classical political economy with a seemingly de-politicized but actually pro-*rentier* ideology. Mathematical symbolism is given the sanctifying role once afforded by Latin. Aping the natural sciences, economists take refuge in abstruse modes of expression. The more complex the math, the more simplistic and banal the postulated relationships and conclusions tend to be. Most of the math refers to choices between different 'menus' of goods and services, without much analysis of how these come to be produced, or the long-term economy-wide consequences of buying on credit instead of cash. Economic theories that focus on the exchange of goods and services without discussing the means of acquiring control over wealth divert attention from examining what is most important in shaping the economy" (2017: p. 18).

And so part of the point of *J is for Junk* is to clarify what has been confused, and this in particular revolves around:

- The distinction between productive investment and asset price expansion/inflation (where money creation can occur towards either of these ends);
- The important positive and constructive role of the state in sponsoring productive investment (especially in infrastructure) where the state can also provide an effective institutional check and balance on elites and corporations, rather than simply delegate power to them through deregulation (creating an unelected concentration of power and decision-making that whilst oligarchic and oligopolistic is articulated as free market individualism),
- The key distinction between earned and unearned income.[156]

Without this final distinction rent extraction becomes conflated with economic growth and wealth creation. This is a point Hudson notably pursues with reference to, heavily influenced by a language more familiar to Americans than others, the FIRE acronym (Finance, Insurance and Real Estate sector, 2017: p. 103). The FIRE sector receives transfer payments and yet has become a core aspect of modern economies whose real role and consequences (see below) are not widely understood, as Hudson states in his introduction:

> "The following A-to-Z guide aims at providing the vocabulary and concepts for a more effective diagnosis of today's economic (and by extension, psychological) depression, by thinking in terms of compound interest, debt peonage, *rentier* economies, unearned income, zero-sum activities and economic parasitism. Without such concepts in the forefront of one's mind, today's neoliberalized economies are prone to succumb to the virus of Orwellian Doublespeak. Junk Economics and its euphemistic vocabulary aim to limit the tools of thought by distracting attention from the causes – and hence, the needed remedy – by trickle-down economics weaving a cloak of semantic invisibility around the phenomena of *rentier* parasitism" (p. 20).

An illustration of the main themes of *J is for Junk*

J is for Junk is an extended "thought crime" in Orwell's sense of dangerously transgressive. Its pithiness is also indignant and sometimes polemical but not in the pejorative sense of unreasoned. One gets a clear sense of this by chaining together a selection of entries. For example, from "As if argument" and "Asset-Price inflation":

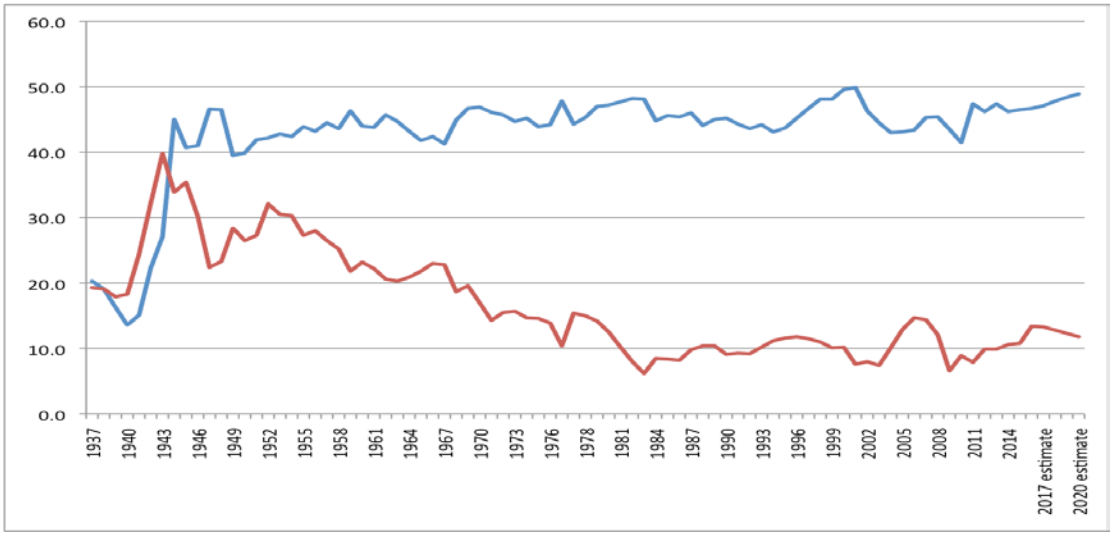
> "A parallel universe is presented as a set of assumptions. As in novels, the key is to get observers to suspend disbelief. Mainstream economics, for instance, reasons as if all wealthy individuals earn their income by playing a productive role and put their savings in banks or bond markets – which are assumed to increase prosperity by lending these savings to entrepreneurs to build factories and employ labor. *Rentier* income, junk mortgage lending and corporate takeover loans play no role in this 'as if' picture… The reality is that banks don't lend for new direct capital investment, and only a small

[156] *Inter alia* a whole range of mainstream theory (based on comparative advantage) exclusions are stated. For example, "Free trade theory leaves out of account structural problems leading to chronic trade deficits, food and trade dependency, non-cost related rake-offs such as economic rent, emigration as a result of poverty, war and the effects of financing trade deficits by running interest-bearing debt and losing domestic political autonomy to international financial institutions" (2017: p. 52). For an extended exploration of related issues see Wade (2017).

> proportion is lent for consumer goods. Banks lend mainly against assets in place… This credit for buyers of real estate, stocks and bonds inflates debt-leveraged windfall gains" (2017: p. 33).

For Hudson, a FIRE focus creates an emphasis on capital gains (2017: p. 50), which in turn creates motives to minimise taxation on such gains and to create preferential tax regulation (noting debt has various tax benefits including through real estate and fictive depreciation). This, as many have acknowledged (Piketty, Keen, Wray, Palley, Galbraith, Kelton, Tcherneva etc.), benefits those who own the vast majority of financial assets and thus creates a feedback loop for inequality. Inequality, meanwhile, is exacerbated by debt leveraging, creating a Minsky process of unstable financial expansion (2017: p. 154). The majority population are co-opted into a system of debt peonage (2017: p. 71) through a system that claims house "ownership" and small scale equity participation in stock markets is an entryway to real wealth, rather than a means to yoke the majority to long term debt; something that is further exacerbated as the state devolves pensions to private investment based on equity markets, and the wealthy lobby to reduce corporation tax, and top rate income tax whilst favouring the transfer of taxation from land, assets, and high incomes to the rest of labour and VAT. Consider how corporation tax and income tax have diverged in the US over the neoliberal period:

US corporation tax (red) and income tax (blue) as % of total tax receipts, 1937-



Source: Table 1: http://www.whitehouse.gov/omb/budget/Historicals

As Hudson notes, new obfuscations then become possible; Laffer curve claims (2017: p. 138), trickle down claims (2017: p. 231), and:

> "It is argued that hurting corporate profits would leave pension funds with lower gains, making it harder to pay retirees. Investing pension funds in the stock and bond market instead of financing direct investment leaves pensioners (along with middle-class savers) hostage to the financial sector. Its lobbyists claim that reforms to help consumers by regulating monopoly pricing and product safety, improving working conditions or paying better

> wages would hurt pension funds by eroding corporate profits and hence stock-price gains" (2017: p. 177)

As more of the system is diverted into financial activity less productive investment occurs, and so there is less in the way of actual returns on real productive investment to pay down debt; the system becomes more dependent on internal asset price inflation, debt accumulates, and diversifies – extending to student debt as a "necessary" investment in human capital to become employable, car loans, credit cards etc. An increasing proportion of current income of the majority is given over to servicing debt (so real income is not as it seems in much the way that national income accounts are also misleading, 2017, p. 165). Eventually, a point is reached where financial crisis manifests. These are *real* processes that mainstream economics does not appropriately describe or explain. This too becomes socio-politically useful.

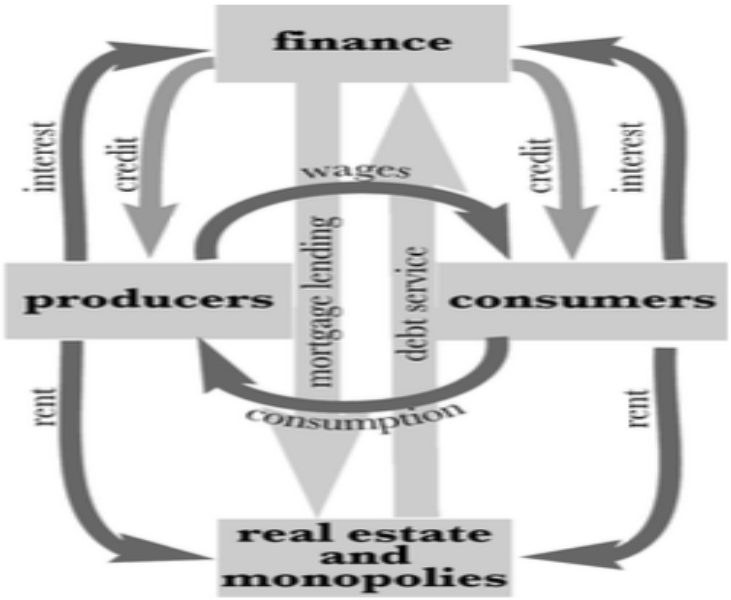
Based on the power of finance and the interests that arise precisely because of the instability of financial processes, banks and owners of large volumes of financial assets are preferentially positioned. They are bailed out and compensated, and are essentially treated as victims rather than victimisers (2017: p. 37). The majority, meanwhile, find that their debt carries over from one financial crisis to the next. This is personally problematic and also systemically problematic, since debt overhang affects the scope for economic activity in the next period, and leads also to further financialisation-related solutions (via QE and other means, which simply create renewed asset price inflation). As Hudson notes, in his entry on business cycle theory it:

> "Fails to explain the exponential build-up of debt from one recovery to the next, and hence fails to see the ultimate crisis. Anti-labour, anti-government neoliberals have hijacked 'business cycle' theory by depicting downturns as being caused by rising wages and raw-material prices as full-employment and full-capacity operations are reached, cutting into profits so that growth tapers off. But the key factor spanning business cycles is the growth of debt and rising interest charges that stifles profits. Debt service absorbs the income hitherto spent on new direct investment and consumption, so employment and production fall off" (2017: p. 46).

It follows within the core theme that pervades the entries in *J is for Junk* that a financial crisis can be one in a sequence of such crises, and also indicative of or constituted within a more basic structural problem with contemporary financialised economies as *political* economies. In his entry, the Two Economies, Hudson states:

> "Domestic private sectors are composed of two distinct systems. These are conflated to mean 'The Economy', but their dynamics are quite different. 1) The 'real economy' of current production and consumption, wages and industrial profits… 2) The FIRE sector consists of land, monopoly rights and financial claims that yield *rentier* returns in the form of interest, financial fees, economic rent (unearned income) and monopoly gains, plus asset-price gains ('capital gains')… Most of the FIRE sector's financialized wealth – the asset side of its balance sheet – is held by the *rentier* class… Its debt counterpart on the liabilities side of the balance sheet consists mainly of mortgage debt… Since World War II, the 'real economy' has spent more and

more income on real estate, insurance and payments to banks, pension funds and other financial transactions" (2017: p. 232).



One could go on but it should be clear that much of *J is For Junk* sets out a consistent theme. Thereafter, perhaps the other important point to highlight is the purpose to which Hudson directs his dictionary. In so far as debtors are the genuine victims of financial crises, and in so far as debt overhang from one crisis to another is a cumulative problem, then in conjunction with a more constructive role for the state, Hudson argues for debt forgiveness as a key constituent in the transformation of current problems of financialised economies (2017: pp. 59, 72 and 131). This, of course, contests the current positioning of victimisers as victims.

Some final comments

Whilst *J is for Junk* is an excellent work – eminently readable, nailing in a few brief sentences here and there some important insight that many have thought but few have articulated quite so well – it is also something of an oddity as a dictionary. I by no means intend to discourage any potential readers. I sat and read the whole work in two or three sittings, and not just because I was reading it to review it. Conversely, I wondered whether I would read it in any other way, by which I mean would I consult it as a dictionary is *ordinarily* consulted, rather than read it as a single extended work? Given that the point of a dictionary is typically to provide a clear statement of the *common usage* of terms, and often, if used in an academic context, to provide a quotable source of authority (in writing an essay etc.), a dictionary whose actual point is common misuse of language creates something of an odd status, at least as a point of reference in the instrumental sense.

Concomitantly, if I required a technical account of the concepts and terms Hudson criticises, the information contained in *J is for Junk* would likely be insufficient. If I required a significant substantive critique of those concepts and terms then the entries would still likely be insufficient. The entry on Efficient Market Hypothesis illustrates both these points (2017: p. 87). This sounds more critical than it is intended to be and should rather be taken as a reminder of what *J is For Junk* is intended to be. If you are looking for a general reminder of the underlying logic, absurdity, irrelevance or harm created by some important terms and concepts in relation to others, then reading *J is for Junk* is rewarding. That said, *J is for Junk* is, however, in so far as it works to a theme, also repetitive in places, somewhat under-developed in some of its historical points (why Margaret Thatcher supported policy that favoured financialisation in the long term is sociologically complex in a way that a brief dictionary entry cannot capture, if one is interested in the background to the UK's "Big Bang" of the mid-1980s, 2017: p. 223), and begs the odd question regarding who and what is omitted, given the selection of what is included (for example, if Larry Summers why not Frederic Mishkin?).

In the appropriate context, however, little of this matters. If you want to read something genuinely popularising rather than simplistically popularist then I would recommend *J is For Junk*. It is by no means intended to be derogatory to state this book is ideal bathroom literature. A great deal of economics may be fecal, but this is not. It is a reminder of what really matters for the many rather than the few.

References

Haldane, A. (2015) "Labour's share", transcript of speech to Trades Union Congress, London November 12th, Bank of England

Hudson, M. (2005) "The $4.7 trillion Pyramid: Why Social Security Won't Be Enough to Save Wall Street", *Harper's*, April, pp. 35-40.

Hudson, M. (2006) "The New Road to Serfdom: An illustrated guide to the coming real estate collapse", *Harper's*, May, pp. 39-46.

Hudson, M. (2015) *Killing the Host* Dresden: ISLET-Verleg

Wade, R. (2017) "Is Trump wrong on trade? A partial defence based on production and employment", in E. Fullbrook and J. Morgan, editors. *Trumponomics: Causes and Consequences* London: WEA/College Books

Author contact: jamiea.morgan@hotmail.co.uk

SUGGESTED CITATION:
Jamie Morgan, "Negating 1984: Michael Hudson's antidote to doublespeak vocabulary in economics", *real-world economics review*, issue no. 80, 26 June 2017, pp. 179-185,
http://www.paecon.net/PAEReview/issue80/Morgan80.pdf

Deserving economics

The Nobel Factor, by Avner Offer and Gabriel Söderberg, 2016
Economics in the twenty-first century, by Robert Chernomas and Ian Hudson

Peter Radford [USA, The Radford Free Press]

Do people deserve what they get? Or, do they get what they deserve? This is a dilemma that divides economics into two distinct sub-groups. The answer in what we can call the mainstream or majority sub-group is unequivocal: people get what they deserve. To put it in the words of Milton Friedman: "To each according to what he and the instruments he owns produce". This certainty of outcome, which bears no relation to the levels of luck, opportunity, or assets that an individual enters the economic arena with, is regarded as both ethical and natural by economists who subscribe to this majority view. In more vernacular language we are reduced to expressing this as "stuff happens" or simply "tough". There is no question that, in this view, each one of us is on our own. We are individuals who plough through life and whose welfare is entirely in our own hands. There is no greater nor lesser level for analysis. The individual is all there is, and whatever happens to that individual, in their economic existence, is attributable solely to their own efforts. There is no recourse, no assistance, no succor, and no relief. You are what you are. And you are alone.

Majority economics has taken this isolation and made it a cornerstone of its thinking. Everything inside the many theories and ideas of the majority is based upon a reflexive belief that aggregate behavior is simply that: an accumulation of lots of individual behaviors. There are no properties or behaviors of the aggregate that cannot be explained at the lower isolated individual level. Indeed, for the past few decades it has been accepted by the majority that all discussion and analysis of aggregate phenomena must be firmly rooted in what Robert Lucas called strong "micro-foundations". While to mainstream economists this stricture appears sound – they tend to think of their understanding of individual behavior as theoretically secure – it immediately introduces a bias difficult to ignore. It precludes, or at least it appears to preclude, theories that focus on aggregate phenomena: it downplays or even denigrates social or collective action that cannot be explained as the summing up of individual actions. This makes it extremely difficult to create space for a role for government in everyday economic activity.

Thus, majority economics provides the intellectual backdrop for the famous dictum of Margaret Thatcher that there is no such thing as "society" but only a collection of individuals. The coincidence of contemporary neoliberal policy making and the relentless emphasis on micro-foundations in economics with its elevation of market based thinking to dominant status is no accident. The two are obviously interlinked, and majority economics continues to impress its market bias into its students throughout the world whilst maintaining that it is "positive" or scientifically based.

Yet there is a manifest alternative, the existence of which belies this claim of the majority. Avner Offer and Gabriel Söderberg call this alternative Social Democracy in their recent book.

Unlike the majority view, Social Democracy recognizes the legitimacy of collective action and takes the position that people deserve what they get.

One practical way of understanding the essential difference between the majority view and its Social Democratic alternative is by looking at the way in which it handles the vicissitudes of life.

According to the majority view individuals have enough information, calculating capacity, and understanding to shift their income and wealth about within their lives. They can, thus, predict lean times when their earnings do not cover their expenses and compensate for the consequent deficiency by taking out insurance or by saving when times are good. So even though an average person is clearly going to experience great variations in their ability to provide for themselves throughout their lives, they have a sure method of coping. The key component of this view is that security is a commodity and that solutions to insecurity can be bought and sold in markets in the form of personal contracts. There is no uncertainty, everything is known, so the only problem is of suitable calculation and response.

In contrast to this, the Social Democratic view recognizes the central role of uncertainty. It acknowledges, too, that individuals experience periods of personal resource insufficiency, but because of uncertainty it assumes that both the period and extent of that insufficiency is unknown and unknowable. Contracting in markets therefore becomes problematic and the resulting logical response is to allow for social insurance in which all citizens receive assistance in times of insufficiency simply by being a citizen. Everyone deserves what they get.

History tells us that the Social Democratic view, articulated variously in the ideas of thinkers such as Maynard Keynes and Gunnar Myrdal, dominated in the decades immediately following World War II, a period in which many industrial countries built large social programs to cover the insecurities of health, education, and retirement. The legacy of that period lingers on in the continued existence and popularity of those programs and in the economic structure of the countries involved, but that continued existence is hotly contested and often subject to attack because of the overthrow of the Social Democratic view in economics in the late 1970s and the subsequent rise of the modern majority alternative.

It is the supplanting of the older system with the later individualistic market-based system that Offer and Söderberg seek to explain. Their primary question is: how did the market-based view come to acquire so much credibility that it has effectively eclipsed all other forms of economic thinking? One answer they provide is the legitimacy given it by the introduction in 1969 of the “Prize in Economic Science in Memory of Alfred Nobel”. This prize, colloquially known as the Nobel Prize in economics, becomes the centerpiece of their discussion of the rise in influence of market biased economic thought.

The majority view within economics needs support in asserting its legitimacy because it often appears more a body of fantasy than a serious attempt to understand actual economies. Whilst it is rich with arcane and hard won insights that are of immense value to economists it is also difficult to believe its credibility. Its many components require a mastery of what sometimes feels more like make-believe than science. It has the air of a religion rather than a field of genuine scientific study. One acute observer within the world of economics was reduced to describing economics, not as the study of economies, but as “what economist do”. While that comment was made in the context of the methodologies economists apply to the

subject we can take it in a broader setting: majority economics is very much concerned with itself and not with real world economies.

Take, for example, the great swathe of majority economics that involves the search for "general equilibrium". This search has its roots in a few almost carelessly uttered words of Adam Smith who wondered out loud how it was that the apparent order we see around us could possibly co-exist with the equally apparent selfish pursuit of personal goals. How, in other words, did all that selfishness contrive to end up as a socially beneficial whole? It was as if, Smith opined, an invisible hand organized our economic affairs and turned greed into good. Lead into gold. Alchemy. Whereas modern chemistry has moved on from alchemy, majority economics has not. It is still deeply immersed in the attempt to turn lead into gold.

Some of the greatest names in the history of economics, from Walras on, dedicated themselves to the question of general equilibrium, so when, in the early 1950s, Ken Arrow and Gerard Debreu produced a proof of its existence the entire profession gasped in pride. This was, surely, a massive step forward. That the proof was based on truly absurd assumptions that could never be realized in practice was set aside, and general equilibrium whatever its ridiculous foundation was moved into the majority view's intellectual arsenal. The reality was, however, that the contortions necessary to arrive at the Arrow-Debreu conclusion stood as an indictment not a validation. Even the authors of its proof have, along the way, said as much. Yet, nothing daunted, and immune to criticisms of its other-worldliness, most economists accept general equilibrium as central to their theorizing. This ability to deny reality at any cost to hold onto an intellectually fascinating notion bedevils economics. Apparent magic and supernatural powers abound in its textbooks.

Another example is more practical: the sheer computational power needed by an average consumer to perform in the manner that most economists argue they do. Lurking inside the theories economists use to describe individual behavior is a series of assumptions that defy any observation of real people. Consumers, according to such theories, not only have access to all information, from everywhere, about everything, and for all time, but they have the amazing capacity to use this information in instantaneous, and we must admit, spectacular feats of calculation. That no computer exists to match these feats is not a barrier to most economists carrying on the argument and extending the discussion of economic transacting to include such prowess.

Reality, however, does intrude, but rather than retreat and come up with ideas that match the cognitive skills of actual people the majority view in economics persists and, instead, adds new wrinkles as if two acts of folly produce one insight of value. Milton Friedman, an early father of the majority view, argued that consumers "act as if" they were doing these prodigious calculations. After all, he said, billiard players do not articulate the mathematics of hitting ball into the side pocket, so why would we expect consumers to be aware of their ability to undertake the kind of calculation economics ascribes them? Economics is riddled with this kind of special pleading. Subsequent, but related, efforts to reduce the amount of calculation needed to make economic models tractable produced an even more odd assumption. This was the heroic reduction of the world into a combination of "representative agents". One such a household doing the consuming, and another doing the producing. So, rather than engage reality, economists trying to mimic economic behavior created a caricature so simple that it eliminated reality entirely. All this to preserve the pre-eminence of clever and appealing intellectual wizardry.

If an economics of this sort was to compete with and then replace the existing Social Democratic view it needed bolstering. A Nobel Prize was just the trick.

Getting a Nobel Prize was, however, a complicated problem, and here Offer and Söderberg provide us with invaluable information. The establishment of the Nobel Prize in economics was a result of political conflicts in Sweden.

Economic policy in Sweden had, since World War II, been informed almost exclusively by the Social Democratic tradition. Social programs had been developed to cover the consequences of the uncertainties of life. Taxes were high enough to provide for these programs, with such taxation being recognized as a necessary aspect of social cohesion. Per Albin Hansson, a member of the Swedish Parliament, called Sweden "The People's Home" arguing that "The basis of the home is commonality and mutuality … In the good home you find equality, compassion, cooperation, helpfulness." Contrast this sentiment to that of Milton Friedman when he was asked about the establishment of the Mont Pèlerin society that he and Hayek created as an early conservative think tank. He said: "The threat to a free society that we envisaged at the founding meeting of the Mont Pèlerin Society is very different from the threat to a free society that has developed over the intervening period. Our initial fear was of central planning and extensive nationalization. The developing threat has been via the welfare state and redistribution." The clash between the two views is stark, with Hansson arguing for a collective perspective and the apparatus needed to support it, with Friedman condemning that apparatus as an intrusion on individual liberty and an anathema to the efficiency based miracle-working machinery of what has become the majority view.

Swedish business interests and the Swedish Central Bank objected, for less theoretical reasons, to the levels of taxation and the activity of government implied by Social Democracy and hankered for a way to give more market-centric ideas sufficient credibility that Hansson's home could be shrunk, if not eliminated. Serendipity presented them an opportunity: the three-hundred-year anniversary of the Central Bank could be celebrated in a typically Nordic way with the establishment of a prize in economics.

The struggle within Swedish domestic politics to roll back Social Democracy exactly mirrored the efforts to achieve the same objective within the economics profession. The coincidence in timing was perfect. Both sets of insurgents would benefit. Swedish business interests and the Central Bank would receive the backing of newly celebrated and suitably honored ideas, and economists seeking to further the cause of those market based ideas would add the embellishment of worldwide recognition to their intellectual revolution. The rest, as they say, is history. Economics is saddled with a prize that adds luster largely to its majority market based view, thus aggrandizing and embedding that viewpoint within the profession, whilst creating an additional obstacle for those seeking to attain credibility for alternative ideas.

This is not to say that the prize has always gone to mainstream thinkers. Indeed, the prize-giving committee has had to balance its ideological bias with a substantial number of awards to heretics. Offer and Söderberg go to great lengths to analyze the predispositions of the award recipients and one of their book's great contributions is to highlight just how balanced the award giving has been. Combatants on either side of the divide will recognize the names of many foes from the other in the award winners list. This balance has allowed the prize committee to retain credibility within the profession hence enhancing the award's prestige. There are oddities of course, such as when the urge to balance ended up with perverse outcomes. More than once have awards been given to two economists with contradictory

ideas in the same year. The most recent example of this was in 2013 with the awards to Eugen Fama and Robert Shiller: the former an advocate of the purity of markets, the latter being a well-known critic of that purity.

So, if the origins of the prize were in the fight between Social Democracy and the market based majority view, and yet the prize has been awarded frequently across that divide, is there a different, subtler, message that having a Nobel Prize conveys to the public at each annual award time? Yes, there is. The prize confers scientific stature on the entire enterprise of economics, and here Offer and Söderberg end their book with a more profound question: is economics more like physics or more like literature?

Unlike the rest of their book the last chapter is a meditation on the state of economics, its manifest problems, and its ethical stance with respect to society. Up until this point the book is more a history and background to the establishment of the prize, accompanied by a standard recapitulation of the defiance the majority view has for the constraints of reality. By concluding in this way Offer and Söderberg give us all a great deal to ruminate on.

The establishment of a Nobel Prize in economics, albeit a prize that the Nobel family objected to and which is not conferred through the same channels as those for the physical sciences, was not simply a matter of creating an honor for the enterprise of economics in its entirety. It was, at least in part, an effort to give the then insurgent market-based faction the imprimatur of intellectual and scientific superiority. It was designed to give extra license to the increasing formalization of economics and to downplay the Social Democratic thinking that the insurgents wanted to overthrow. This was necessary because much of the work and opinion of the insurgents took a remarkably anti-democratic turn. To impress these ideas onto society it was essential that these anti-democratic principles be sanitized as being immutable “laws” akin to the laws of physics. Society had to be led to believe that resistance to these laws would be futile. And so the clockwork like predictability of those laws is said to grind on relentlessly banishing choice and dehumanizing key social relationships, and any attempt to mitigate the result of the operation of the market place is doomed to be a failure.

This unpalatable core message of the majority view runs directly counter to the more hopeful and optimistic notions of Social Democracy where society is quite able to deal with the uncertainties of life by acting as a collective and offsetting life’s insecurities through cooperative rather than competitive means.

This dismal core message, laden as it is with a dismissive attitude towards democracy, is not some inconsequential sidebar to the whole of economics. On the contrary: it is the centerpiece. Take for example what is known as public choice theory which is antithetical to democracy precisely because of the fear that majority rule might encourage the invasion of the rights of the minority. It is built from the ground up as a repudiation of democracy. As Offer and Söderberg note, the ideological passion with which notable anti-democratic economists such as Friedman, Stigler, and Hayek is breathtaking, with the latter having been quite clear that he would trade away democracy to preserve what he called freedom – by which he meant the sanctity of private property.

Once this ideological driving force is exposed all the absurdities of economics and its ability to turn a blind eye to reality become easily explained. Its willingness to crush individuals into a rational straight jacket whilst pretending to give them choice; its deliberate ignorance of human cognitive limitations; and its stubborn loyalty to its core ideas despite those ideas

having been demonstrably disproved, all make sense. It is a narrative not a science. It more akin to a religion than anything else. It is what mainstream economists would prefer the world to be like rather than an attempt to describe the world as it is. It is "what economists do".

So, is there any hope for the future?

This is the subject that Robert Chernomas and Ian Hudson attempt to tackle. We need briefly to touch upon their discussion because it fits neatly with that of Offer and Söderberg. Published last year, their book is a study of the work of younger economists all of whom have been awarded one of the other big prizes in economics: The John Bates Clark Medal. This prize is given only to young economists and is thus a guide to the future of the discipline. After all, the current crop of younger stars is likely to be a fertile source for future Nobel Prize winners.

The first thing we notice in looking at the list of Medal winners is their depressingly similar concentration in the same corners of the academy as dominant in the professional literature. Whether this concentration reflects the quality of the universities the young guns teach and research at, or whether it is some vestigial pecking order reflecting bygone power and prestige rather than performance we can leave for another day.

The second is that the work that they do is slightly away from the core base of economics and more often in topics of great contemporary interest. Things like income inequality, health care systems, and development seem to have captured their attention and they steer assiduously away from tackling those central elements of the majority view that so vex Offer and Söderberg. On the contrary, with varying degrees of explicitness these young stars export and utilize most of the equipment required to build the majority view. Rarely do they stray away from an individual perspective, rational behavior lurks in the background, and nowhere are the major influences of power relationships in society clearly built into their analysis.

So, at least in this group, there is little evidence to suggest a willingness to look inward and question openly the anti-democratic structure of mainstream economics. This is true even when they advocate social welfare policies, aggressively progressive taxes, or take a view on the role of institutions.

These two books, taken together, leave us contemplating mainstream economics as an often contradictory and other-worldly web of theories, more ideological than scientific, internally focused rather than facing the real world, and increasingly unwilling to face up to its anti-democratic core. Its younger stars appear uninterested in correcting that taint and yet simultaneously express a desire to think about issues of great social consequence.

We leave the tour these books give us more skeptical than ever that the mainstream in economics has either earned the right, or can be trusted to be objective, in the provision of advice to public policy makers. The conclusion in economics that outcomes are deserved and that we are powerless to invert them, or that we deserve more, is a matter of faith not science, let alone social science. Once we recognize this we are open to include other narratives in our thinking about the economy. Mainstream economics thus forfeits its exclusive grip on public policy making and becomes one of many competing narratives all equally important and valuable. Many of those other narratives benefit from not having the yoke of prizes weighing them down and preventing the importation of both novelty and open enquiry.

Perhaps economics would benefit from being free of its prizes while it re-acquaints itself with reality.

This is not a message of denigration of the efforts of economists, far from it. As Offer and Söderberg argue: let economists do what economists do. Let them theorize and ponder the arcane inner workings of economics. Rather, it is a message of humility and the recognition that economics is still in its infancy and that it cannot answer fully many of the questions its sets itself. Not least because it asks those questions only after having answered them. Ideology dressed as science is merely narrative. The full story of economics is yet to be told.

Author contact: pacr@mac.com

SUGGESTED CITATION:
Peter Radford, "Deserving economics", *real-world economics review*, issue no. 80, 26 June 2017, pp. 186-192,
http://www.paecon.net/PAEReview/issue80/Radford80.pdf

real-world economics review, issue no. 80
subscribe for free

Trade imbalances are undesirable: a note

Leon Podkaminer [The Vienna Institute for International Economic Studies, Austria]

Abstract
This Note provides a very simple argument suggesting that the buildup of trade imbalances may be slowing down global output growth. The observed slowdown of global economic growth (since the mid-1970s) may have been a consequence of globalization-driven emergence of large and persistent trade imbalances.

JEL classifications F15, E12

Key words economic growth, globalization, imbalanced trade

1. Introduction

The last 50 years not only have produced a series of revolutionary technological changes which should have accelerated global growth. These decades have also witnessed a truly revolutionary systemic change (gradual at first, accelerating later on) on the global level. The change started with stepwise internal liberalizations and deregulations in major industrialized countries. The developed countries' socio-economic models, which had sought to balance the interests of labor and business while relying on fiscal and incomes policies, were gradually replaced by the neoliberal and monetarists ones. The internal systemic changes have been synchronized with the consecutive waves of liberalization of international economic relations. Trade liberalizations (cuts in tariff levels, progressive removal of many non-tariff barriers to trade) were followed by the wholesale liberalization of capital flows, to a large degree completing the process of globalization. The phenomenal rise in international trade has been the most obvious effect of globalization. But, globalization – and the globalization-driven expansion of international trade – appears to have been associated with a *slowdown* of real growth at the global level (Podkaminer, 2014, 2016.)

According to the classical, neoclassical and contemporary theories of international trade, "more trade" (and especially more *free* trade) should bring output gains. Why are such positive effects not showing up in the available data? There may be two major reasons.[157]

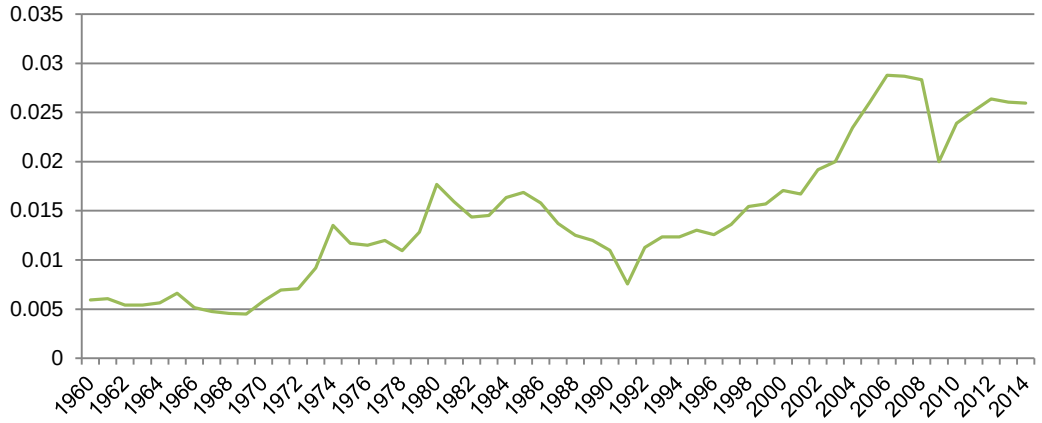
Firstly, the expanding internationalization of production (which has been made possible by the liberalization of trade and capital flows coupled with advances in transportation technologies) seems to be generating, or at least supporting, the tendency for the global wage shares to decline – and thus for the global profit shares to rise (see e.g. Palley, 2009.) This development may be closely related to the development of inequality on the global level. While the impact of globalization on global inequality remains a controversial issue, there is also a possibility of a reverse impact: from higher inequality to slower growth. The global shift in income distribution from wages to profits can account for the weakening of global growth because such a shift raises the

[157] The supply-side, or structural, developments cannot be made responsible for the secular global growth slowdown. The natural resources have become more abundant secularly (as it is evidenced by their prices trending downwards in relative – and often absolute – terms.) The weakening pace of labor productivity growth is also an unlikely cause of the slowdown of global growth. Rather, the weakening pace of productivity growth an *effect* of output growth slowdown rather than its cause (Podkaminer, 2017.)

overall saving propensity – without raising the propensity to invest. The tendency for the slowdown of growth of global output could then be an end effect of both developments: rising global profit share/profitability and falling propensity to invest.

Secondly, it may be argued that rising world trade *could* have been productive on the global scale if output growth in individual countries had been at least approximately balanced most of the time – and not only sporadically, in response to the severe debt/payments or exchange rate crises.[158] The negative output effects of rising trade may have emerged under the huge and persistent trade *imbalances* that have developed under progressing globalization (Figure 1.)

Figure 1 Ratio of global trade surplus to global output, 1960-2014[a]



Source: Own calculations based on WDI (August 2016 edition.)
[a] The global trade surplus is defined as the sum of national trade *surpluses* (positive trade balances) across the world.

Such imbalances may have acted as brakes on sustained output growth in *both* the persistent deficit and the persistent surplus countries. Under a different international economic order, somehow enforcing more balanced trade among nations – with major nations not allowed to compensate deficient domestic demand with huge trade surpluses that destabilize their partners – global trade may assume the positive role assigned to it by the conventional trade theories. The classical Bretton Woods system (terminated in 1973) was an example of such international arrangements limiting persistent and large trade imbalances. It is worth remembering that the global output kept rising, in per capita terms, on average by over 3.4% per year from 1961 through 1973. The average yearly p.c. growth rate for the period 1974-2015 is 1.5%. This Note will provide a formal (but very simple) "proof" that the buildup of trade imbalances is likely to slow down global output growth.

[158] In the neoclassical (and derivative) trade theories countries engage in *barter* trade – very much like the individual "agents" populating the microeconomics textbooks. The barter trade is assumed to culminate in a Pareto-optimal, *balanced*, equilibrium. But in the real world nations do not engage in barter trade, but in trade involving *money*, or debt. Germany, or China, does not earn export revenues in order to spend them, immediately and completely, for paying for the imports urgently desired.

2. A simple "formal proof" of undesirability of trade imbalances

Let us start with the basic GDP identity:

$GDP = C + I + B$

where C is consumption, I is investment and B is trade balance. (All items are assumed to be "private". The public sector is ignored, for the sake of simplicity.)

Add the ("micro-funded") behavioral relationship linking consumption (C) to income (GDP.) The simplest such relationship is just C = c*Y, with c being the propensity to consume out of income. Combining the GDP identity with the consumption function yields the immortal "multiplier" formula:

$GDP = (I + B)/(1-c)$ (1)

where (1-c) is the saving propensity denoted as s (0< s < 1.) 1/(1-c) = 1/s is the simplest version of the "multiplier" and (1) is equivalent to:

$GDP = (I + B)/s$ (2)

Now, let us consider the "global economy". For the sake of simplicity "our" global economy will consist of two countries, indexed 1 & 2.

The formula (2) for country 1 is: $GDP_1 = (I_1 + B_1)/s_1$

and for country 2: $GDP_2 = (I_2 + B_2)/s_2$

In a two-country global economy B_1 must equal *minus* B_2 (because $B_1 + B_2 = 0$.) Assume country 1 runs surplus B. $B = B_1 > 0$. Then $B_2 = - B$.

The question: what is the *global* income? The answer is elementary:

$(GDP_1 + GDP_2) = (I_1 + B)/s_1 + (I_2 - B)/s_2$ or, equivalently:

$(GDP_1 + GDP_2) = (I_1/s_1 + I_2/s_2) + B(1/s_1 - 1/s_2)$ (3)

Under balanced trade (B=0), global GDP equals the first bracketed on the right-hand side of (3.) The same obtains when $s_1 = s_2$. However, if B > 0 the second bracketed term may be *negative,* depending on the values of s_1 and s_2. Specifically, B (the country's 1 surplus and country's 2 deficit) *raises* the global income (from $I_1/s_1 + I_2/s_2$) when $1/s_1 > 1/s_2$ that is when $s_1 < s_2$, i.e. when $(1-c_1) < (1-c_2)$, i.e. when $c_1 > c_2$. Otherwise B *lowers* the global income.

Is "everything possible" then? Not quite. The condition $c_1 > c_2$ (which is necessary for the imbalance to be "productive globally") means that the consumption propensity in the *surplus* country is higher than in the *deficit* country. But such a constellation seems highly unlikely in practice (though perhaps imaginable "in theory".)

Concluding remarks

It remains true that output of some countries may heavily rely on the expansion of *their* exports. Moreover, productivity growth (and growth of potential output) in many cases may critically depend on rising imports of capital goods and intermediate inputs. It is equally true that rising *net* exports may contribute substantially to overall GDP growth in some nations. But rising net exports may well be achieved at the cost of overall domestic GDP growth stagnation. This is the case in Germany where high trade surpluses (achieved through the sustained repression of wages and domestic demand) have been associated with secularly anemic GDP growth. Moreover, it must be remembered that for each country relying for GDP growth on the improvement of net exports there must be some other countries whose net exports necessarily contract – thus depressing their GDP growth. The existence of a club of countries following such "export-led" growth paths implies the existence of a club of "import-fed" countries whose GDP growth must sooner or later be held back by falling net exports. The *global* economy – being an autarchic system – cannot follow the export-led growth path based on trade surpluses.

The final "policy conclusion" could be that the basic paradigms of the international economic order need to be changed. The reformed international order should be capable of enforcing more balanced trade among nations. The major trading nations must not be allowed to compensate deficient domestic demand (and stagnant wages) with huge trade surpluses that destabilize their partners. Under the reformed world economic order the expansion of global trade could then be expected to support global growth. Of course, the basic paradigms of domestic macroeconomic policy-making in major countries would have to be overhauled too if these countries were to follow the externally balanced growth paths (Laski and Podkaminer, 2012 and Podkaminer, 2015.)

References

Laski, K. and Podkaminer, L. 2012. "The basic paradigms of EU economic policy-making need to be changed", *Cambridge Journal of Economics*, vol. 36, no 1, 253-71.

Palley, T.I. 2009. "A Labor Perspective on Globalization", IPG/FES, no.2, 22-39: http://library.fes.de/pdf-files/ipg/ipg-2009-2/2-09_a_palley_us.pdf

Podkaminer, L. 2014. "Does trade drive global output growth?" *Bank & Credit*, vol. 45, no 4, 311-330.

Podkaminer, L. 2016. "Has Trade Been Driving Global Economic Growth?" Working Paper no 131 of the Vienna Institute for International Economic Studies:
https://www.google.at/?gws_rd=ssl#q=wiiw+working+paper+131

Podkaminer, L. 2015. "The euro area's secular stagnation and what can be done about it. A post-Keynesian perspective", *Real-World Economics Review*, Issue 70, 1-16.

Podkaminer, L. 2017. "The slowdown in labor productivity growth is an effect of economic stagnation rather than its cause", *Acta Oeconomica*, vol. 67, forthcoming.

Author contact: podkaminer@wiiw.ac.at

SUGGESTED CITATION:
Leon Podkaminer, "Trade imbalances are undesirable: a note", *real-world economics review*, issue no. 80, 26 June 2017, pp. 193-196, http://www.paecon.net/PAEReview/issue80/Podkaminer80.pdf

---

ISSN 1755-9472

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

---

Subscriptions to this email journal are free to all members of the World Economics Association (WEA). WEA memberships are free at http://worldeconomicsassociation.org/

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

To subscribe to this journal go **here** (http://www.feedblitz.com/f/f.fbz?Sub=332386)

To unsubscribe to this journal go to
http://www.feedblitz.com/f/f.fbz?EmailRemove=_NDUzNDkzMHwzMzIzODZ8cGFlX25ld3NAYnRpbnRlcm5ldC5jb218MTA5OTU4_