ECONOMICS AS IF BANKS MATTERED: A CONTRIBUTION BASED ON THE INDUCTIVE METHODOLOGY

by

RICHARD WERNER
University of Southampton

1 Empirical Challenges

Alan Greenspan confessed in 2008 to recognizing a ‘flaw’ in mainstream approaches to how financial markets work (Congress, 2010). Donald Kohn (2009), as Vice-Chairman of the Federal Reserve, admitted:

It is fair to say . . . that the core macroeconomic modelling framework used at the Federal Reserve and other central banks around the world has included, at best, only a limited role for . . . credit provision, and financial intermediation. . . . asset price movements and the feedback among those movements, credit supply, and economic activity were not well captured by the models used at most central banks.

The research agenda which culminated in macromodels without banks (Walsh, 2003), without monetary aggregates (Woodford, 2003), or without a financial sector (most real business cycle and dynamic stochastic general equilibrium models), has not been successful.

Macroeconomics has proceeded down the wrong path. In order to find the right one, we need to return to the crossroads at which it was taken. In the 1980s, the prevailing approaches faced deep problems. At the time, classical, many neoclassical, Keynesian, monetarist, post-Keynesian and many eclectic approaches shared one fundamental pillar linking money \((M)\) to the economy: the ‘equation of exchange’ or ‘quantity equation’:

\[
MV = PY
\]

For monetary policy purposes, to explain and forecast GDP \((Y)\) and prices \((P)\), as well as to estimate the money demand function, a stable velocity \(V\) was required. This stability was the thread on which macroeconomics hung. It held up for a while, but increasingly a ‘velocity decline’ was observed: nominal GDP did not grow as much as the money supply. Also known as the ‘breakdown of the money demand function’ or the ‘mystery of the missing money’, it spawned a large literature, pointing out the dire consequences for macroeconomics (the quantity equation had ‘came apart at the seams during the course of the 1980s’, Goodhart, 1989; a worldwide ‘puzzling’ anomaly, Belongia and Chalfant, 1990; money was not in any predictable relationship to the tangible economy, and could no longer be measured accurately; once ‘viewed as a pillar of macroeconomic models’, the quantity equation ‘is
now . . . one of the weakest stones in the foundation’, Boughton, 1991). Some likely reasons were contemplated (deregulation, liberalization, financial innovation, development of more and rapidly changing financial instruments), although they are not convincing (should financial innovation not have increased the velocity of money?). While the velocity decline has accelerated in many countries since, economists have turned away from the problem—not because a solution had been found, but because they had given up searching for an answer.

This was precisely the time when the moneyless economic models became increasingly attractive to economists. They were a form of escapism: instead of rising to the empirical challenge to improve on their understanding of the monetary sector, economists chose to pretend that money and banks did not exist and had no influence on the economy. This was possible thanks to the prevalence of the hypothetico-deductive approach, which places little emphasis on empirical facts.

The new moneyless economics however merely produced further empirical puzzles and ‘anomalies’ which by today have also discredited it: ‘The notion that there is something about banks that makes them “special” is a recurrent theme’ (Blanchard and Fischer, 1989, p. 478), one that is empirically well supported (e.g. Peek and Rosengren, 2000; Ashcraft, 2005; Werner, 2005; see also Fama, 1985). However, economists failed to identify what makes banks special. So the recurring banking crises have remained unexplained as well—a highly visible and embarrassing refutation of moneyless and bank-less macroeconomics.

Meanwhile, the focus in these recent models had shifted from the quantity of money to its price (interest), in the hope that this would behave in a more predictable way (inversely related with economic activity and Granger-causing it). But interest rates tend to be positively correlated with economic activity and there is at least as much evidence that Granger causality runs from economic growth to interest rates. Furthermore, interest rates and models centred on them could not explain asset prices (especially during the recurring episodes of asset bubbles and busts), bank lending or international capital flows.

Japan posed a particular challenge to macroeconomics, old and new. First its success was difficult to explain. Then a recession and banking crisis occurred that had not been predicted, and that resisted standard cures. The main policy recommendation, to lower interest rates, failed to end a slump that is now entering its third decade. The more traditional Keynesian advice to boost fiscal expenditure produced record national debt (triggering other problems), but no lasting recovery. Some think that the ‘credit view’ approach or the ‘liquidity trap’ argument have provided answers. Not so. The credit view consists of the bank lending and balance sheet approaches, and the credit rationing argument (Jaffee and Russell, 1976; Stiglitz and Weiss, 1981). The latter is undoubtedly correct: it maintains that banks
keep interest rates below the hypothetical equilibrium level, since demand
for money and credit is always large, then ration and allocate credit, to
minimize non-performing loans. But it is a microeconomic theory in search
of macroeconomic consequences: lack of bank credit, we are told, can be
easily compensated by credit from non-bank financial intermediaries or
‘direct finance’ in the form of public debt, or equity. However, empirically
this does not happen. When banks are damaged, go bust, are closed or
even restructured, the economy tends to suffer (Peek and Rosengren, 2000;
Ashcraft, 2005; Leary, 2009; Voutsinas and Werner, 2011). Likewise,
the lending or balance sheet channel views’ predictions were not supported
by the Japanese evidence: they argue that monetary policy, working
through interest rates, should be more effective thanks to this additional
channel. But in Japan’s bank-centred economy, monetary policy was sin-
gularly ineffective. Instead of explaining the puzzle, they merely made it
bigger.

Meanwhile, the liquidity trap argument has not addressed, let alone
answered, the question why over a decade of significant interest rate reduc-
tions have failed to stimulate the economy: the argument is merely about the
moment in time when interest rates have hit the lowest possible level (March
2003, in Japan’s case, after over a decade of ineffective rate reductions), and
it merely says that at this lowest level interest rate policy will not be effective.
It turns out the liquidity trap argument is a tautology stating that when
interest rates have fallen so low that they cannot fall lower, then they cannot
be lowered! On the questions at hand, why over a dozen rate reductions over
a decade have failed to stimulate the economy, how Japan got into this mess
in the first place or how it can be avoided in other countries—the theory
maintains a detached silence.

The lack of success of demand management policies initially boosted the
supply siders, whose moneyless real business cycle theories were used to
justify policy advice to improve economic performance through deregulation,
liberalization and privatization. Economies with freer markets should be
more efficient, while more controlled economies, employing non-market
mechanisms, should be less successful, we are told. This flies in the face of the
economic performance of the German economy from 1933 to the 1960s, the
Japanese, Korean, Taiwanese economies from about the mid-1930s onwards,
and the Chinese economy since 1982—all of which were based on non-market
mechanisms and government intervention in the form of clever institutional
design and direction of credit (more on this below). But also the argument
that structural reforms lead to better performance was disproven: when
Japan increased the number of cartels in the 1960s to over 1000, economic
performance improved; when it reduced them, deregulated and liberalized the
economy, resulting in the free market economy of the late 1990s with zero
cartels (with financial markets more liberalized than in the USA), economic
growth also dropped to zero (Werner, 2003).
There is a future for macroeconomics, if it can explain these empirical puzzles. Below I suggest a solution that has performed well empirically. It also has the advantage of being simpler: the principle of parsimony holds that a simpler explanation, relying on fewer assumptions, is preferable to a more complex one. Finally, it is not based on the deductive methodology, which was instrumental in leading macroeconomics into a cul-de-sac.

2 An Alternative Approach

2.1 The Role of Banks

Leading economic models and theories consider banks to be mere ‘financial intermediaries’, helping to allocate resources by collecting savings and distributing them to investors in need of funds (see, for example, Walsh, 2003; Mishkin and Eakins, 2006). Based on an analysis of empirical data and institutional facts (all of which require a study of history) it can be shown, however, that banks are not mere financial intermediaries. They perform an important macroeconomic role that makes them special, and that needs to be reflected in models: they are the creators of the money supply (Werner, 2005). In most countries, including the UK, they create about 97 per cent of the money supply through the process of credit creation. A tiny proportion only is created by the central bank. As Werner (2005) shows, the banks’ power to individually create money through the process of credit creation is based on the regulatory and accounting regime banks are subjected to. Currently, banks are allowed to individually create new purchasing power by simultaneously booking an asset and a liability when a new credit (‘loan’) is granted (a fact not mentioned in macrotextbooks or macromodels). By virtue of a signed ‘loan’ contract, banks are allowed to add the amount of the loan to the asset side of their balance sheet, while the borrower’s current account is credited with the same amount. In this way, banks can create new deposits ‘out of nothing’, without transferring money away from elsewhere in the economy. Effectively banks pretend that borrowers have deposited money, treating them as if they had done so. Since banks are the accountants of the economy, others cannot tell the difference. The fictitious deposits become real. The fact that the vast majority of money circulating in economies is created by profit-oriented private sector enterprises (the commercial banks) is not widely known by economists or the general public (and those who knew about it chose to ignore its potential implications). But it is acknowledged in less well-known passages of reports by central banks.4

4The actual process of money creation takes place primarily in banks... Of course, they do not really pay out loans from the money they receive as deposits. If they did this, no additional money would be created’ (Federal Reserve Bank of Chicago, 1994, p. 3; p. 6); ‘... when banks loan money to people and businesses, the overall money supply increases’ (Federal Reserve Bank of Kansas City, undated, p. 4); ‘Over time... Banknotes and commercial
2.2 The Link between Bank Credit Creation and the Economy

That banks ration and allocate credit is recognized in the literature. But this fact takes on a whole new dimension of importance when it is combined with the recognition that banks are the creators of the money supply. This is the missing link that causes credit rationing to have macroeconomic implications (Werner, 1992). Since the credit market is rationed and determines the money supply, the quantity and the quality of credit creation are key factors shaping the economy.

We can now revisit the quantity equation. The original formulation by Irving Fisher (Fisher and Brown, 1911) was:

\[ MV = PQ \]

Fisher said that the ‘effective’ money \( MV \) (assumed to circulate and be used for transactions) is equal to the value of transactions \( PQ \) (the sum of all pairs of prices times quantities transacted).\(^5\) In other words:

The total value of transactions during any time period must be the same as the amount of money used to pay for these transactions (Werner, 2005).

Due to a lack of data on transactions, but increasing availability of GNP data, first Cambridge, then all economists argued that the stock of money should be proportional to ‘total nominal expenditures’ (equation 1), apparently on the principle that at night we should search under the lamp post for our lost keys). But this is accurate only if

\[ PY = PQ \]

or, in other words, if nominal GDP is a robust proxy for the value of total transactions in the economy for which money is changing hands. Asset transactions, such as of the financial and property variety, are of substantial volume in most economies, yet are not included in GDP. It is therefore necessary to distinguish between GDP-based transactions and non-GDP-based transactions (see also Werner, 1992, 1997). Werner (1992, 1997) argues and shows in the Japanese case that the widely observed velocity decline is not due to ‘disintermediation’, financial innovations or deregulation, as the literature has argued, but simply due to an increase in money used for transactions that are not part of GDP (asset transactions). Theore-
tically, we can disaggregate the transactions data any way we wish. It is an empirical issue whether we can find suitable statistical data to proxy the theoretical breakdown. Following Werner (1992, 1997), we break both sides of (2) down by distinguishing money used for transactions that are part of GDP and money used for transactions that are not. A similar distinction seems to have been contemplated by the likes of Fisher, Keynes or Friedman. They failed to implement it however due to their focus on monetary aggregates (M1, M2, M3)—which cannot be disaggregated by their use. More fundamentally, these have no business in the quantity equation, since they measure money out of circulation (they are really ‘savings aggregates’). Circulating money that is used for transactions can however be measured by credit creation, which can also be disaggregated (Werner, 1992, 1997). We can thus improve on the traditional equation of exchange by substituting credit (C) for money (M) and formulating two equations, differentiating the use of credit money:

\[ CV = P Q Y \]  
\[ CV = P Q \]  
with \( V = \text{const.} \)

\[ CV = P Q \]  
\[ CV = P Q \]  
with \( V = \text{const.} \)

For growth:

\[ \Delta(CR) = \Delta(PQ) = \Delta(Y) = \Delta(\text{nom. GDP}) \]  
\[ \Delta(CF) = \Delta(PF) \]  

From equation (6) it follows that bank credit creation will boost nominal GDP growth, if it is extended for transactions that contribute to GDP. Credit creation will boost the value of asset transactions (thus often their prices), if used in asset transactions. This explains why the apparent velocity decline coincides with periods of asset booms. Bank credit creation in the form of consumer credit adds new purchasing power that will boost consumption demand, but does nothing to increase the amount of goods and services available: *ceteris paribus*, consumer price inflation ensues (consumptive credit creation). Credit creation for investments in the production of goods and services or productivity enhancement is less likely to create inflation and more likely to boost real GDP (productive credit creation). Only the production of goods and services can over time and in aggregate deliver the income streams to service

---

6This was recognized by Fisher and Brown (1911), Fisher (1926), Keynes (1930) and Milton Friedman (1968), who pointed out that ‘Each side of this equation can be broken into subcategories: the right-hand side into different categories of transactions and the left-hand side into payments in different form’ (p. 435).
bank credit and repay the principal. Thus only the latter type of credit creation is sustainable and not crisis-prone.

2.3 Asset Boom/Bust Cycles and the Role of Banks

Banks are the creators and allocators of the money supply. Their decisions concerning the aggregate quantity of credit created and its quality (i.e. type) shape the economic landscape. Despite being awarded this public privilege, banks have not been required to ensure a desirable outcome for the economy. They have been able to focus on profit maximization alone. As a result, they have tended to create and allocate credit for financial transactions. These deliver potentially high returns in the short term (capital gains being a function of aggregate financial credit creation) and hence high bonuses for bankers, but also systemic problems. They produce asset inflation and ultimately banking crises: as soon as bank credit creation for non-GDP transactions slows, the asset bubble bursts, taking down the banks with it, if it has been large enough. In the UK, for instance, the proportion of bank credit allocated for unproductive, asset purchases has multiplied in recent years, amounting to a multiple of GDP. Since such unproductive credit creation cannot in aggregate be paid back, it must turn into non-performing loans that will cripple the banking sector: bank equity is quickly insufficient to cover the size of non-performing loans. Hence banks are soon insolvent, at which time it is usually suggested that tax money be used to rescue them. Non-performing assets also make them more risk-averse, hence reducing credit further, producing a downward spiral. Banking activity is always procyclical, since banks create the credit that enables the majority of economic transactions (further analysis and empirical evidence in Werner, 1997, 2005).

2.4 Conclusions

This simple model resolves the empirical puzzles of the old and the new macroeconomics: Banks are special, because they create the money supply. Credit rationing thus has macroeconomic implications. The velocity decline (and money demand function breakdown etc.) was due to the misspecification of the quantity equation and unrecognized growth of non-GDP transactions. Fiscal policy could not boost growth in Japan, because it was bond-financed (i.e. not linked to credit creation; equation (6) makes clear that nominal GDP growth is determined by credit creation for GDP transactions; increased government expenditure merely raises the government share of the unchanged income pie and thus crowds out private demand). Asset bubbles and busts are due to credit creation for non-GDP transactions and end when this dries up, creating banking crises. Capital flows can result from credit creation, irrespective of interest rates (Werner, 1994,
1997). Interest rate policy per se cannot stimulate growth, as it does not feature in equation (4) or (6). This is not surprising in a world of rationed markets and disequilibrium (a good description of our world, since mainstream economics has proven that the requirements for equilibrium are so stringent that we can be sure it will never exist), which are determined by the short-side principle: quantities determine outcomes, not prices, in this world. The model can also explain recent empirical research on the role of credit in a number of contexts (e.g. Buyukkarabacak and Krause, 2009). Even the supply-side puzzles are explained: Germany 1933–1960s was so successful, because its banks allocated credit largely to productive purposes, while speculative and consumptive credit remained repressed. Likewise, the factor at the core of the ‘East Asian Economic Miracle’ in Japan, Taiwan, Korea and China has been their direct intervention in the credit market through central bank ‘window guidance’ of bank credit to higher value added use and repression of the macroeconomically potentially harmful credit for consumption or speculation (until the 1980s, in the Japanese case; Werner, 1998, 2002, 2003). An incentive-compatible and efficient regime of ‘credit guidance’ can be designed and replicated in other countries (the hurdle for such intervention to be welfare-enhancing is much lower than in the theoretical dream worlds of mainstream economics: in non-fiction economics no market is in equilibrium). There are also many other policy implications, not only concerning how to boost growth in periods of recession or banking crises (use the central bank to purchase non-performing loans at zero cost to the taxpayer; boost bank credit for productive purposes) and how to avoid banking crises once and for all (ban or severely restrict credit creation for transactions that do not contribute to GDP—a rule that could be implemented easily as all transactions can be classified thus and loan officers already spend much time to ascertain the use of the credit), but how to engineer stable, sustainable growth with high employment and low inflation. For instance, developing countries do not need to borrow from abroad to boost domestic growth (foreign credit cannot be used domestically). Instead, they merely need to focus on productive credit creation at home. But on these and other issues, see Werner (2005).

3 Methodology Matters

The credit guidance principles deployed by Germany and the East Asian ‘miracle economies’ are not based on the (failed) principles of a socialist planned economy. Instead, they are derived from capitalist economic theory that relies on private ownership, but does not adopt unrealistic assumptions about the behaviour of agents or markets. It was developed following the inductive methodology. The main reason why macroeconomics has for so many decades been developing largely unperturbed in the wrong direction is the ‘Ricardian vice’ of the deductive methodology, which
is supremely suited to manipulation or political abuse, as axioms and assumptions can be designed to deliver any desired result (such as that government intervention is bad, big corporations and banks must not be restricted, free trade is best, etc.). The sciences use the inductive methodology, and there is no good reason why economics should not also. Deductive economics ignored the pivotal role of banks in the economy, which is an institutional quirk that can only be detected by careful empirical observation. That we rely on private enterprises to create and allocate the money supply indeed does not follow from axioms, standard assumptions or rational argument. Indeed, it is a feature that needs to be reconsidered: surely, thanks to 21st century technology, monetary systems can be designed in a more transparent, equitable and stable fashion than by allowing private profit-oriented enterprises to create and allocate our money, without any attempt to guide their decisions towards the greater good (for concrete regulatory proposals, see Dyson et al., 2010; Werner, 2010).

REFERENCES


The question of whether the 2007/8 financial crisis was a symptom of a problem with macroeconomics, or rather a problem of finance theory and poor market regulation was taken up by several speakers. Alistair Milne (Cass Business School) differed from the views expressed by Guido Ascari that the financial crisis was just about financial markets and risk-taking. He argued that surely there was a macroeconomic dimension, perhaps associated with the current account imbalances and the build-ups of stocks of debt and credit. He drew a contrast between the view that the global imbalances were merely a reflection of the more efficient financial sector in the Anglo-Saxon west compared with the surplus countries, and that therefore they are the efficient destination for international capital flows, and were not a source of concern. This view was expressed by Caballero et al. (2008). But since then Caballero has changed his views quite radically, and now expresses almost the opposite view that the inefficiency of financial intermediation was one of the ingredients of the crisis. Milne inferred from this that we really do need models of heterogeneous agents, models that allow for the possibility of capital flows going to the ‘wrong’ places. Milne noted also that discussion of macro-prudential regulation had not featured in the presentations. He argued that it is important to integrate analysis of macroeconomic and regulatory aspects of the crisis; and for macroeconomics to pay attention to financial developments and financial aggregates, so as to be able to spot where crises might be building up; and not deal with financial supervision as a separate matter.

Peter Spencer (University of York) concurred with the view that the financial crisis was not just a problem that arose from the financial markets and finance theory. It was clear to many people that the housing boom was unsustainable. But he and others failed to foresee the speed and severity of the financial crisis, in which many financial markets literally broke down. And while there are well-known models of credit rationing (Stiglitz and Weiss) and quantity-constrained market disequilibria, such as Muellbauer and Portes, which predict the consequences of such a breakdown, these ideas were not incorporated in the models in general use at the time. Their inclusion leads to non-linear models with multiple solutions. To linearize such models to get a unique solution is to throw away their key insights. Agreeing with Richard Werner, he asserted that it will be essential to examine disequilibria and short-side constraints, and it will not be easy.
David Cobham (Heriot-Watt University) brought the discussion of the causes of the financial crisis back to monetary policy. The authorities had not been willing to react to asset prices. The conventional view, articulated by the Chairman of the Federal Reserve Board, Ben Bernanke, was that they would clean up after a bubble burst, and all would be well. Strikingly, while the Central Banks have adopted radically new policies with Quantitative Easing, they remain unwilling to change their non-interventionist view on asset prices. The idea that monetary policy should lean against the wind needs to be brought back into the debate. Pursuing this line of argument further, Marcus Miller observed that Central Banks have surely learned not only that they need to take account of more indicators, but also that they need more instruments. Charles Bean, Deputy Governor of the Bank of England, had recently commented, in a reference to Jan Tinbergen, that they needed as many instruments as they had targets. This goes beyond leaning against the wind. But even this does not deal with the endogeneity of the crisis. Andrew Haldane of the Bank of England has described it as an event in a ‘doom loop’ where a crisis is followed by a bailout and then another even bigger crisis. Can we be sure that the bailouts that have worked for now have not induced changes in behaviour that will trigger more crises in future, unless there is serious change? Marx may be more relevant than Keynes in understanding all this.

Another theme that emerged in the general discussion was the question of what criteria and benchmarks might be used to judge the usefulness of a model. Hashem Pesaran re-iterated the three criteria developed in his presentation: relevance, consistency and adequacy. Consistency does not require a dynamic stochastic general equilibrium (DSGE) model; it might refer to a broader notion of consistency, or to consistent accounting. The adequacy of a model needs to be judged against its main predictions and purposes. Models might have different purposes, such as predicting financial crises, designing macroeconomic stabilization policy, or teaching; and different models will be adequate for different purposes. While researchers might fail to agree on the adequacy of a model, because there are multiple criteria for it, discussion of it is bound to bring about at least partial convergence; and there needs to be more discussion of adequacy of models.

Guido Ascari agreed that models should be adequate for their intended purpose, e.g. for particular policy questions, and did not need to be adequate for everything. Hence the popular macroeconomic models (such as those of Christiano, Eichenbaum and Evans, or Smets and Wouters) cannot be criticized for failing to predict the financial crisis. They were not intended, and their assumptions do not enable them, to do that. A model intended to explain or predict the crisis would need different assumptions.

In concluding remarks, Hashem Pesaran acknowledged that there are many recent research papers dealing with heterogeneous agents, market disequilibrium and other pathologies, but felt that they are not prevalent in the
discipline, and that it will be hard to assimilate them into the mainstream, by which economics is judged. He also remarked that merely having heterogeneous agents in a model does not enable it to deal with occasional outbreaks of correlated behaviour among agents, the importance of which he had developed in his presentation. Guido Ascari disagreed. He noted that many younger economists who have recently become tenured in the top university economics departments in the USA are researching the issues of heterogeneous agents and market imperfections. While much of their research is not yet applicable in quantitative analysis of policy, he felt sure it would be soon.

Hashem Pesaran argued that, when looking for parsimony in models, it was important not to create a straitjacket, and DSGE could easily become one, which would restrict research on heterogeneous agents, imperfect information and so on, if researchers continued within the DSGE framework and attempted to add these features to it. He further wondered how such models could be estimated and how checks that they are identified could be made. He argued in favour of a more flexible approach to relating data to the theory and to data; and for diversity in empirical approaches to reduce risk of failure.

Richard Werner appealed for greater observance of the principle of Ockham’s Razor in macroeconomics. Other things being equal, the simpler explanation is the better one. He argued that economics was too much in love with complexity. On a different note he felt that Central Banks had a heavy responsibility for the crisis; and although they had performed abysmally and failed catastrophically, they had avoided fierce criticism. Academics, he felt, had not criticized them robustly enough. He singled out the Chinese central bank as an exception, whose successful use of credit controls recognized the absence of equilibrium in the system.

Reference