Socially Useless? The crucial contribution of finance to economic life

Philip Booth, Professor of Finance, Public Policy and Ethics, St. Mary's University, Twickenham and Diego Zuluaga, Head of Financial Services and Tech Policy, Institute of Economic Affairs, London

Introduction

Public animosity towards the financial sector is a longstanding phenomenon. Because finance involves the creation of intangible value, many observers have struggled to comprehend its social contribution. Furthermore, ignorance about fundamental features of the financial system, such as risk, uncertainty and the opportunity cost of money, led monotheistic religions to proscribe or severely constrain the charging of interest. This in turn fueled prejudice against people who, pushed out of other occupations and less constrained by their own creeds, dedicated themselves to finance.

We have advanced much since the times of Shakespeare's Shylock, yet the suspicion of finance remains and is progressively stoked by financial crises. As we will show, there are important problems with the present system of financial regulation and risk management. But claims that financial activity is largely self-serving and, in the words of a prominent former financial regulator, "socially useless" are unfounded and dangerous.

The claims are unfounded because a closer look at the type and volume of financial activity in developed countries reveals that little of it is unrelated to customer needs, such as transfers of resources across time and space, risk diversification and insurance, hedging and the mitigation of uncertainty, and income-smoothing across people's lifetimes. They are dangerous because the doctrine that finance is useless beyond an arbitrary threshold may lead to crude policy measures that will restrict people's ability to lend, borrow and insure themselves.

Furthermore, reliance on statutory regulation to address the perceived problems of finance may heighten not mitigate the exposure of taxpayers and households, especially poor ones, to adverse events such as recessions and speculative bubbles.

What is "socially useless"?

The idea of the financial sector being "socially useless" was popularized by (Lord) Adair Turner in 2009, who suggested that the UK financial sector had become larger than was optimal. Turner pointed to "fixed income securities, derivatives, trading and hedging, and possibly also asset management and share trading" as areas that had grown too big. Though he emphasised that it was extremely difficult for a regulator to know in advance which activities it would be beneficial to restrict, he did suggest a tax on financial activities to reduce the size of the financial sector to its optimal level (Turner 2009).

In a speech to the United Nations in 2015, the Pope said: "beginning in 2008 the trend of food prices has changed: doubled, then stabilized, but always with higher figures in comparison to the preceding period. […] we cannot overlook financial speculation: for example the high prices of wheat, rice, corn, soy, which fluctuate on the stock market, perhaps they are linked to profits and, therefore, the higher the price the greater the profit."[1] We will look specifically at speculation further below.

[1] See: https://w2.vatican.va/content/francesco/en/speeches/2015/june/documents/papa-francesco_20150611_fao.html

Turner, for his part, was making two distinct claims. The first is that the financial sector is too big. Economic theory suggests that an activity will reach larger-than-optimal size when the marginal cost faced by firms is not the full marginal cost of undertaking the activity. In other words, there is an external cost borne by third parties which the firm does not take into account, so that the firm will produce more than it would if it faced all relevant costs.

There are a number of ways in which the financial sector may fit this description. Firstly, bank deposits can carry a state guarantee which is effectively unlimited. This lowers banks' cost of capital and will lead them to extend more credit than they would without state deposit insurance. Secondly, there is an implicit subsidy whereby large financial institutions, whose significance is deemed to make their potential failure an event to be avoided at all costs, will be rescued by the state. This lowers the required return from creditors and investors, since the probability that they will lose their capital is reduced by the implicit bailout guarantee. This implicit subsidy will, in turn, encourage banks to grow bigger. Haldane (2012), speaking at the IEA's Beesley Lectures, estimated the implicit subsidy provided by governments to the world's largest banks at $70bn annually between 2002 and 2007.[2]

It is important to note that this issue of externalities is a separate problem from that of modern financial instruments not having any social value in the first place. Moreover, the diagnosis that explicit and implicit capital subsidies have deleterious effects may be widely shared, but the policy prescriptions vary considerably. Turner argued in favour of taxation and regulation, but as we discuss below, the second-order effects from these interventions might outweigh any beneficial effects, and it is not clear that they would resolve the underlying problem. Instead, it is worth considering more direct ways of addressing the discrepancy by removing state guarantees.

The purpose of this article is to defend the position that the financial sector is intrinsically useful. Below we review the functions of financial markets and the contribution of the sector to the wider economy. We discuss the argument, made in a series of empirical papers, that financial-sector growth beyond a threshold is associated with poor economic outcomes. We also discuss the economic role of speculation.

The role of finance

Financial markets broadly fulfill the following functions:

- Maturity transformation: borrowing short and lending long
- Facilitate money transfers and payments
- Reduce the cost of matching lenders and borrowers
- Transfer capital across time
- Diversify risk
- Lower information and monitoring costs for lenders, and lower search costs for borrowers.

Levine (2005) highlights five key functions of the financial system:

- The production of information about possible investments, and capital allocation on that basis;

[2] To estimate the implicit subsidy, Haldane used the difference between "standalone" and "support" credit ratings.

- Monitoring of investments and corporate governance after finance has been provided (cf. Diamond 1984);
- Risk management, trading and diversification
- Mobilisation and pooling of savings
- Facilitation of payments and exchange.

The traditional description of the function of banks and the financial system more generally is that they resolve a 'constitutional weakness' in the economy (Hicks 1939). Households wish to have a secure method of saving, and other households and companies need a secure source of funds. In general, it is often suggested that households wish to lend short so that their assets can be easily liquidated. Firms need to finance their activities through longer-term borrowing. The financial liabilities created by firms and which become the financial assets of intermediaries enable firms to borrow long term. These liabilities include bank loans and securities and encapsulate the intermediaries' role of 'maturity transformation.'

Debt and equity instruments are mechanisms through which companies raise funds. Those securities provide liquidity for the ultimate saver because they can be traded on secondary markets. When companies borrow money from banks, it creates liquidity for the ultimate saver because banks are able to manage their business based on the assumption that not all households will want to liquidate their savings at the same time: an application of the law of large numbers.

Financial institutions serve to reduce transactions costs

Financial intermediation can be seen as the process through which the savings of households are transformed into physical capital. It can be understood as a chain. At one end of the chain, we have households giving up consumption and saving part of their income. They then save these funds through financial institutions or intermediaries such as banks, pension funds and insurance companies.

Financial institutions then either lend directly to corporations or purchase securities in corporations, thus buying assets that offer a financial return. Corporations then use the capital raised from the issue of securities to purchase physical capital. The returns from capital are then passed back down the chain, through paying returns to the holders of securities (or paying interest on bank loans) and the institutions that hold securities then pay returns to their savers on their saving products.

The financial system exists to reduce the cost to businesses of raising capital and the costs to individuals of postponing consumption. The institutions that are involved in this process include pension funds, banks, investment trusts, mutual funds, investment banks, etc. Innovation in the financial system should be seen in the context of its potential usefulness in performing these functions. Financial institutions specialise in reducing transactions costs such as the cost of screening the risk of the ventures in which people are investing, the cost of finding entities that want capital, the costs of transferring ownership of investment interests and the costs of diversification.

Thanks to the existence of financial markets, households can spread risk across a range of institutions and also across of range of potential borrowers. Risks can be pooled so that investors are less exposed to the failure of individual investment projects. If investment takes place through institutions, all the investor has to do is analyse the soundness of the institution and not of the underlying investments. There is no need for all the pairwise assessments of all borrowers by all households. Financial intermediaries can also monitor risk on a continual basis. Banks can monitor companies that have borrowed from them when

deciding whether to continue lending (Diamond 1984). Purchasers of securities (particularly of equities) can monitor by exercising voting rights or by selling shares.

All the above functions involve the reduction of transactions costs of various types. The financial system also reduces transactions costs thought of in the narrow sense of the cost of transferring a financial interest from one person to another. Through the use of standardised, and often tax-efficient, securities sold on exchange (or 'second hand' markets), financial institutions allow people to access their capital at any time. The bank deposit system allows the transfer of financial interests in a similarly efficient way. Without the ability to transfer ownership in this way, the cost of capital to companies would be much greater and households would be reluctant to save.

Banks, in particular, organise their business in such a way that they facilitate money transmission for the purpose of the exchange of goods, services and financial interests. Banks could just involve themselves in money transmission (narrow banks) without being involved in lending money to businesses and other households. However, in most banking systems, the roles of money transmission, bank lending and financial intermediation go hand-in-hand. The cost of alternatives to money transmission in this way would be prohibitive. Money transmission does not have to rely on the transfer of bank deposits. Money market mutual funds have developed which allow the transfer of units from one person to another with a cheque book or equivalent mechanism and thus can be used for money transmission purposes.

Insurance companies, and some other financial institutions, are involved in what is often described as "asset transformation". They both pool and transfer insurable risks so that the financial liability held by the institution is of a different financial form from the asset held (for example, insurance company liabilities are contingent due to the insurance services they provide yet the assets are not subject to the same type of contingency).

Financial institutions, like institutions in other markets, are continually evolving and subject to innovation. Recently, peer-to-peer lending networks have emerged, enabling households to diversify their risk whilst omitting the intermediary of a formal bank. Mobile money transfers have developed in Africa as a way of promoting money transfer functions without using the banking sector which has been difficult for households and businesses to access efficiently (cf. Mbiti and Weil 2011). In China, 2,000 platforms intermediate £100bn of peer-to-peer lending (Weinland 2017). Exchange traded funds allow extraordinary degrees of diversification of investment portfolios with minimal cost to the investor. Most of the instruments of which people are suspicious in the modern financial sector and which are described as 'socially useless' perform some useful function in terms of reducing transactions costs or facilitating diversification.

A thought experiment

The best analogy between the financial sector and another sector of the economy is probably between the financial sector and supermarkets. It would be possible to prepare dinner by visiting a chicken farmer, a market gardener to buy a cabbage, another farmer to buy some potatoes etc. However, such a process would be incredibly time consuming and involve the sacrifice of a huge amount of real economic activity. Similarly, the financial sector brings together savers and businesses in a way that reduces costs for both parties and thus allows both to achieve their objectives efficiently.

A reductio in absurdum thought experiment is often useful in communicating the benefit of the financial system. Imagine an unintermediated system in which a household wished to save for retirement and to protect against the risk of the main earner dying. The latter is almost impossible to envisage. The former would be extremely costly.[3] The household would have to analyse different investment prospects, taking a lot on trust. The household would only be able to invest in two or three projects. In order to liquidate an investment, the household would have to undertake complex legal processes. Saving for retirement would be much more risky and probably impossible for all but the very wealthy.

Are financial markets short-term-oriented?

Because securities are turned over many times during their lives and may be owned for short periods of time, investment markets are often thought to be "short-termist". It is difficult to calculate a figure for the average holding period of a share. However, a few months is often suggested. However, many of the trades that lead to such a low average are undertaken by particular investors who may add liquidity to markets and reduce anomalies within the markets, but whose function is not to invest the capital of savers. On the other hand, passive funds hold 13 per cent of the market and hold shares indefinitely.[4] The average holding period of a share is not a good indicator of "short-termism".

Indeed, there is a great deal of confusion in discussions of short-termism. Two quite different forms of short-term orientation in financial markets tend to be conflated. The first involves short-termism on the part of investors and shareholders, who discount future cash flows at high rates and thus value future cash flows a lot less than they value cash flows closer to the present. This has been discussed by Haldane (2011) – Haldane is Chief Economist at the Bank of England.[5] On the other hand, there is the short-term orientation of managers, driven by quarterly performance measurement, which has nothing to do with discount rates and all to do with reporting requirements and their own remuneration.[6] The policy causes, if any, and implications of the two types of short-termism are also different.

Haldane (2011) suggests that investor short-termism is driven by myopia, but that is only one possible cause of the apparent high rate of discounting companies apply when deciding whether to invest in new projects. Investors may have become more risk-averse, and therefore value future cash flows less, since all investment projects carry some risk. Or their preference for today's consumption against tomorrow's may be high or have increased in recent years. Alternatively, increased policy uncertainty relating to the regulatory environment in which firms are operating might increase risk premiums. Indeed, the work of Baker et al (2016) suggests that the finger of short-termism should be pointed elsewhere, possibly at central banks such as the Bank of England. Baker et al (2016, page 1) write: "Using firm-level data, we find that policy uncertainty is associated with greater stock price volatility and reduced investment and employment in policy-sensitive sectors like defense, healthcare, finance and infrastructure construction. At the macro level, innovations in policy uncertainty foreshadow declines in investment, output, and

[3] This is why, in the old days, for poor households the only – imperfect – way of providing for old age and the early death of the breadwinner was to have lots of children and to have them early.

[4] See: https://www.economist.com/news/business/21717069-firms-are-increasingly-accused-failing-look-ahead-misdiagnosis-corporate

[5] There may be, in the language of behavioural economics, 'hyperbolic discounting' by investors, or their rates of time preference may have changed.

[6] Investor short-termism will, in an efficient market where agents act in the interest of principals, also be reflected in short-termism by managers. But manager short-termism in this case is only a symptom.

employment in the United States and[…] 12 major economies. Extending our US index back to 1900, [economic policy uncertainty] rose dramatically in the 1930s (from late 1931) and has drifted upwards since the 1960s."[7]

If the driver were myopia, there might be scope for policy intervention in the form of a choice architecture that enables individuals to better understand future versus present cash flows (Thaler and Sunstein 2003). But if rational investor preferences have simply shifted, then there are neither grounds for nor an obvious form of policy intervention. If the cause is government policy uncertainty raising risk premiums, then the solution is to reform and reduce intervention by government – or at least ensure that the framework in which policy is determined is coherent[8].

The second form of short-termism, by managers of public firms who are required to report quarterly and whose compensation is set on that basis, can be addressed with policy reforms which give greater flexibility of supervision. A combination of EU and national regulator rules have required quarterly reporting in various jurisdictions. These have now been relaxed in the UK and in the EU more generally. Whether quarterly reporting really does make a difference to investment is a moot point. Nalarreddy et al (2017), for example, suggest otherwise[9]. However, as discussed in Sternberg (date) the purpose of corporate governance regulation is to ensure that the agents (managers) fulfill the objectives of the principals (owners). If the objectives of the owners are long-term, owners should develop corporate governance arrangements that would reflect that. In turn, stock exchanges are incentivised to develop rules for quoted companies that lower the cost of capital thus providing an environment which is conducive to companies issuing shares and having them traded on the exchange (see, Arthur and Booth, 2010). The objective should not simply be to provide the maximum amount of information possible, nor should it be to provide the kind of information that would encourage management to be short-termist.

Indeed, short-termism is not an inevitable feature of financial markets, as is demonstrated by two of the most admired types of investing currently in operation: value investing of the sort undertaken by Warren Buffett, and venture capitalism which focuses on innovative firms. Neither can be described as short-term-oriented. Buffett's Berkshire Hathaway has an average holding period of one year, with 20 per cent of its stocks held for longer than two years (Hughes, Liu and Zhang 2010). Venture capitalists, on the other hand, take on average five years to exit their investment via mergers and acquisitions, and seven years for initial public offerings (IPOs) (CB Insights 2013). Twitter, for instance, took 77 months from first funding to listing on the public market. In 2015, the average tenure of S&P 500 chief executives had been 11 years. Furthermore, 75 per cent of Amazon's market value is justified by profits anticipated more than a decade hence.[10] US tech companies often have huge valuations, way above the level that could be justified by current profits. Indeed, there is a whole sector of investment markets described as "unicorns" which contains start-up companies with a market value of over $1bn. There are around 250 such companies with

[7] See:

[8] Recent examples of how government action might have raised the cost of capital include the UK government's policy of imposing price caps on energy companies and a whole array of actions in relation to the residential rented housing market such as increases in transactions taxes, limit on the deductibility of interest from financing costs, and a whole range of regulations implemented in the name of tenant protection.

[9] https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2817120

[10] See: https://www.economist.com/news/business/21717069-firms-are-increasingly-accused-failing-look-ahead-misdiagnosis-corporate

a market value of around $750 billion. In other words, investors are being patient and attaching a high value to long-term profits to which there is considerable uncertainty attached.

It is difficult not to conclude that the debate about short-termism conflates too many phenomena to be a contribution to our understanding of how financial markets operate in practice. Examples and anecdotes alone do not disprove the hypothesis but, at the same time, the evidence of those who suggest there is short-termism does not seem to be robust.

Is the financial sector 'duping' consumers and investors?

Zingales (2015) has argued that financial firms are harming consumers and investors through misselling. He suggests that such 'duping' can take the form of predatory sales, whereby customers are sold products they do not understand and would not buy if they did, and of bundle sales in which one of the components in the bundle is harmful but overlooked by unsophisticated buyers.

Consider the case of predatory lending in the market for payday loans. Because of its political saliency, there are abundant empirical studies of the welfare effects of this form of credit. Yet, the evidence is mixed. Payday loans appear to make it easier for people to meet unexpected expenses (see Morse 2011). There is also evidence of lower bankruptcy rates, fewer bounced cheques and fewer borrower complaints when payday lending is authorised (Morgan and Strain 2007). On the other hand, payday loans appear to increase bankruptcy rates among marginal borrowers (Skiba and Tobacman 2011).

More interestingly with regard to Zingales' argument, Morgan (2007) builds a model in which predatory lending is profit-maximising for the lender. However, he finds no evidence that payday lending in the United States – where the market for such loans is most developed – fits the definition.

The case that customers are being 'duped' is also questionable in asset markets. It is well-known that actively managed mutual funds tend to underperform their benchmarks over long periods (Malkiel 1995). Because they charge hefty fees which have a strong detrimental impact on compound rates of return, there is a powerful case that putting one's savings into a mutual fund is a fool's errand if there are cheaper alternatives (cf. Bogle 2007).

Yet, the emergence of cheaper alternatives is precisely what has occurred as evidence of the sub-optimality of active investing has mounted. As of 2016, index tracker funds accounted for 28.5 per cent of assets under management in the United States (Moody's 2017, cit. in Wadhwa 2017). This is up from less than five per cent twenty years ago. Note that the creation of index funds was a direct consequence of the development of portfolio theory and the efficient markets hypothesis (Markowitz 1952; Fama 1970).[11] It is a transparent case of free-market financial theory shaping real-world outcomes for the better.

How about the bundling of bad products – presumably in some concealed fashion – together with good ones that customers wish to purchase? Zingales gives the example of expensive credit cards attached to current accounts, but the case can be generalised to any market that features informed and uninformed consumers. This is a long-standing discussion in economics since Varian's (1980) seminal article on why businesses have sales.

[11] Bogle (2007) gives an account of how studying under Paul Samuelson motivated him to consider the potential of a tracker fund.

It is not in dispute that, under plausible circumstances, the uninformed will pay higher prices (Gabaix and Laibson 2006). But, by charging them more, the provider can lower prices for informed consumers. Total welfare would rise if all customers were informed and the price was uniform, but given that there are uninformed consumers the best outcome may be to have differentiated pricing which results in the uninformed subsidising the informed (Armstrong and Vickers 2012).[12]

In fact, what spurs individuals to inform themselves, thereby putting downwards pressure on prices, is the fact that the uninformed pay more. Ryanair travellers can thank those who check in luggage and buy food on board for at least part of their consumer surplus. The same is true for current accounts and add-ons such as unauthorised overdrafts.

Of course, this is not to suggest that a situation where consumers are uninformed is better than one in which they are better informed. However, it is a failing of neo-classical economics and not of markets to think of markets as being institutions in which people trade with perfect information. There might be particular regulatory interventions that will increase welfare in consumer financial markets beset by information asymmetries (though we are doubtful), however the starting point should be a realisation that markets are institutional settings in which information is discovered and communicated.

There is certainly no shortage of regulations in the UK context to try to deal with the alleged failings of markets. However, the fact that such are still felt to persist might lead us to question whether regulation is inhibiting the development of well-functioning markets. Just to provide some context, in 2011 alone the UK financial regulator brought in regulation or issued guidance, advice, discussion documents or consultations, totalling 4.3 million words. This is more than five times the number of words in the Bible. This included a 585-page consultation on the regulation of the mortgage market, which then led to a 312-page document on regulations relating to the sale of mortgages in 2012.

Does financial sector growth harm economic growth?

For many years, the leading study of the relationship between finance and growth was King and Levine (1993), who examined 80 countries in the 1960-1989 period. Their analysis yielded a strong positive relationship between measures of financial development and economic development.[13] Their study also found financial development to be a good predictor of *future* economic growth. Subsequent papers further bolstered the argument that both bank credit and stock market depth positively affected growth.[14]

But a few more recent studies have contradicted the consensus. Cecchetti and Kharroubi (2012) find a significant negative relationship between average credit-to-GDP ratios and economic growth for a panel of developed and emerging countries between 1980 and 2009. In a subsequent paper (Cecchetti and Kharroubi 2015), this one looking only at mature markets, they find a negative relationship between credit *growth* and income growth.

[12] Whether some customers are uninformed because it is rational, i.e. their cost of becoming informed is too high, or because they overestimate their probability of avoiding extra charges, or because they hyperbolically discount future payoffs, may affect the policy implications but does not affect the argument above.

[13] King and Levine (1993) used private credit to GDP, financial liabilities to GDP and private bank assets as indicators of financial development; and per capita GDP growth, growth in the capital stock, investment and allocative efficiency as indicators of economic development.

[14] Cf. Levine (2005) for a review of those studies.

An obvious reason for the discrepancy are the different time periods of the studies. Panel (a) in Figure 1 shows credit as a share of GDP in the countries studied by Cecchetti and Kharroubi (2015) for the period 1960-1989, the time frame of King and Levine (1993). Panel (b) shows the same indicator for the period 1980-2009, which is the one Cecchetti and Kharroubi (2015) consider.

Figure 1 (a). Domestic private credit as a percentage of GDP, selected countries, 1960-1989

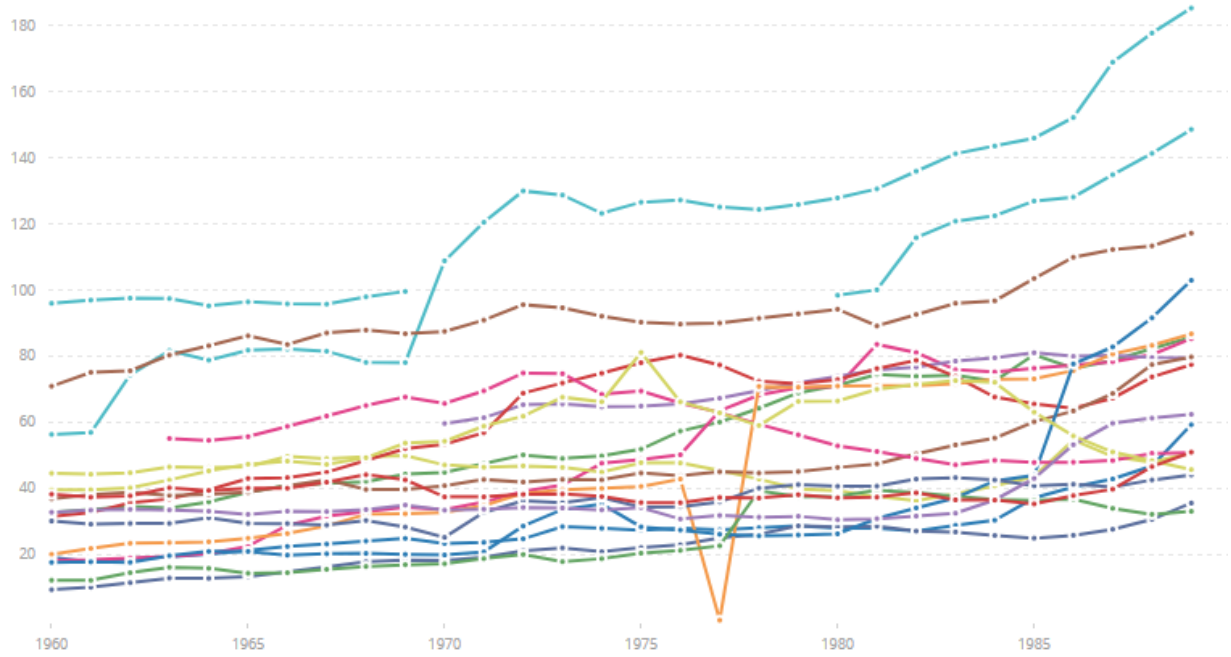
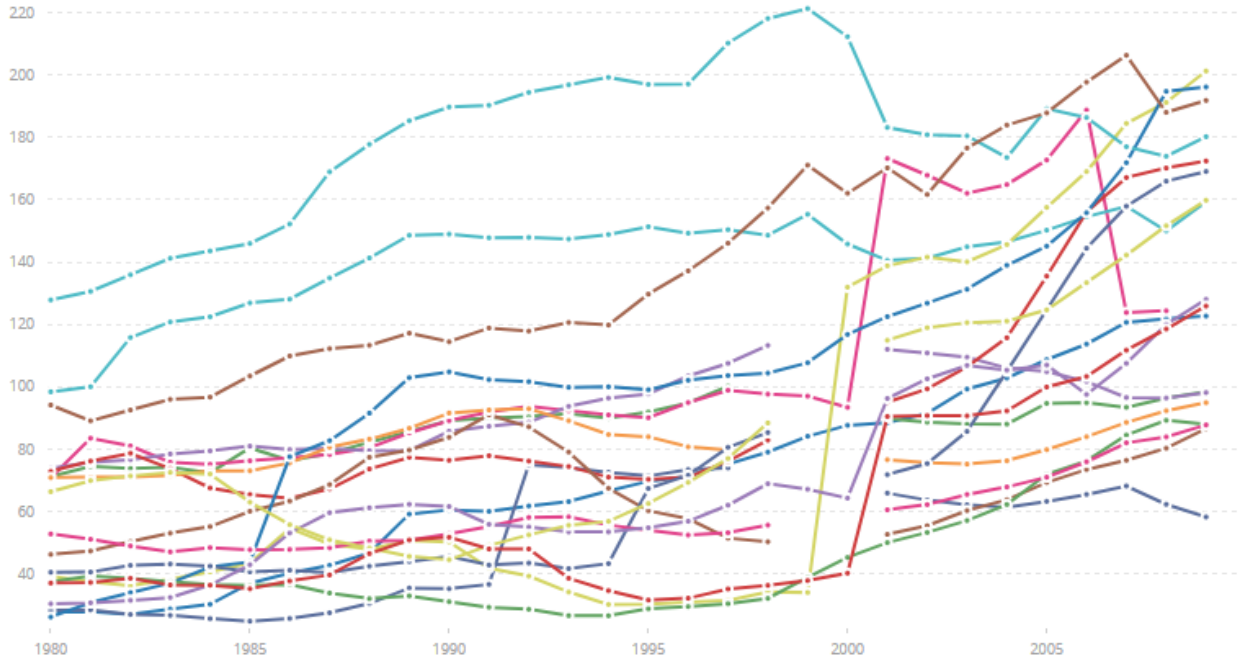


Figure 1 (b). Domestic private credit as a share of GDP, selected countries, 1980-2009



The 1990s and 2000s were a period of substantial expansion of private credit. Despite finding themselves at quite different levels of credit-to-GDP, all the developed countries in Cecchetti and Kharroubi's (2015) study experienced strong credit growth in the 20 years to 2009 and ended the period at a higher level. This is in contrast to the stability in credit-to-GDP ratios observed in countries during the period analysed by King and Levine (1993).

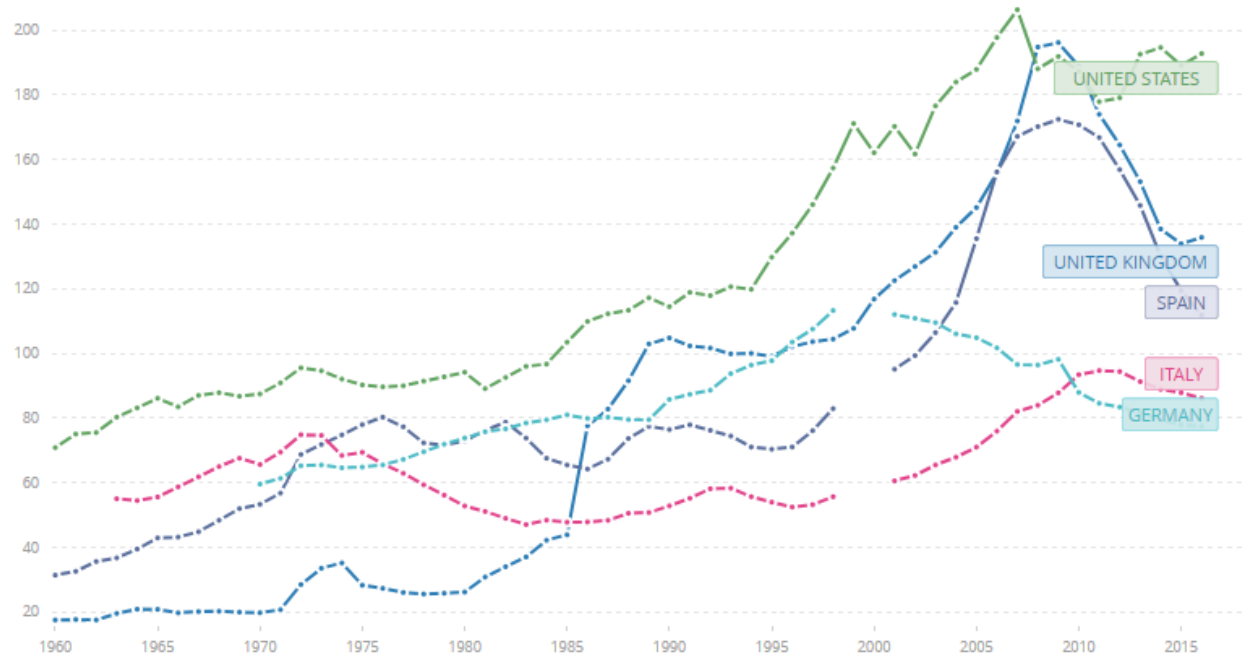
Why is it significant? The later period saw capital markets liberalisation around the world, banking deregulation in many jurisdictions (cf. Martín-Aceña 2013; Guiso et al. 2004) and, most notably, the advent of the single currency in many EU countries. Eleven out of the twenty countries studied in Cecchetti and Kharroubi (2015) are euro members, with all bar Greece having joined at the outset.

It is well known that euro membership altered capital flows between the member countries, notably spurring lending from the core – Germany, France and the Netherlands – to the periphery – Spain, Greece, Italy, Ireland (Lane 2013). The effect of euro membership is thus potentially significant, yet the drivers of credit flows are complex and multiple, and they may include mispricing due to declining nominal interest rates in the periphery, misallocation by politically influenced banks, and a belief by core country banks that periphery treasuries would not let borrowers fail. As these are distinct factors from credit growth, it is surprising that the impact of the euro was omitted as a variable in Cecchetti and Kharroubi's (2015) regressions.

The authors highlight changes in credit growth and productivity growth in two countries, Spain and Ireland, and show a negative correlation. But both countries experienced large housing credit bubbles from the late 1990s, which in the case of Spain was largely driven by state-owned banks under political direction (Martín-Aceña 2013; IMF 2012). Even if one controls for real estate sector employment, as Cecchetti and Kharroubi do, we are left with a financial sector which was to a large extent not driven by market signals but political allocation. But the drivers of the relationship between finance and growth in Spain's case are likely to be distinct from those in more market-driven banking systems such as Britain's or Germany's. Thus credit growth may be more a symptom than a cause of growth-depressing afflictions.

To take a different example, in Greece there was strong private sector credit growth in the 2000s, but the level of borrowing was lower than that from the public sector. Between 1995 and 2009, private credit grew from 29 to 88 per cent of GDP, whilst public debt-to-GDP climbed from 98 to 135 per cent (World Bank 2017; OECD 2017). Public expenditure as a share of GDP increased from 47 per cent when Greece joined the euro to 54 per cent in 2009. In the context of a debt-fuelled boom, it is not surprising that credit growth would correlate with productivity growth declines, but again the circumstances of the relationship are Greece-specific.

Figure 2. Private domestic credit as a percentage of GDP, selected countries, 1960-2016



Source: World Bank

The complexity of drawing policy lessons from the relationship between credit growth and national income growth is borne out in Figure 2. The first striking feature of the chart is the disparity in trajectories of different countries. They are at quite different levels of credit-to-GDP – compare Germany's 80 per cent with 190 per cent in the US – and do not follow similar patterns.

There is also no obvious relationship between credit growth and GDP growth. Germany was underperforming its historical average in the 1990s, when credit was growing, whilst Italy, Spain and the UK were booming in the years when credit grew at its strongest in the 2000s. The big spike in Britain's credit-to-GDP ratio in 1985 can be linked to deregulatory measures, whilst reductions in the ratio after the crisis mask differences in whether the reduction came from reduced credit – as is almost certainly the case in Britain and Spain – or from a decline in both credit and income growth.

The importance of idiosyncratic factors in driving the relationship between finance and growth is supported by the finding in Guiso, Sapienza and Zingales (2004) that local financial development has a significant effect on local GDP growth even in the context of an integrated financial market. The finding in Alfaro et al. (2004) that well-developed local financial markets gain from FDI more than less developed ones points in the same direction.

There are other problems with the literature on finance and growth. It is not just that the components of the dataset are heterogeneous, but also that sample sizes tend to be small. Cecchetti and Kharroubi's (2015) sample of 20 countries over 30 years, during which there were several instances which one might describe as incidences of regime change,[15] mean that there is potential to find inconsistent statistical relationships.

[15] The entry of China into the global economy, trade liberalisation, the fall of the Berlin Wall, and the introduction of the euro are the most salient in the period 19800-2009.

Additionally, there is a troubling disconnect between theory and evidence (Levine 2005). Specifically, theory suggests ways in which specific financial institutions and products – retail banks, investment funds, futures, peer-to-peer lending platforms – reduce transaction costs and raise welfare, and the circumstances – asymmetric information, behavioural biases – in which financial development might harm welfare (Armour et al. 2016). Empirical studies, on the other hand, tend to use aggregate measures of credit growth and relate them to aggregate measures of welfare such as national income. No distinction is thus made between forms of finance, so the policy lessons that may be drawn are inevitably limited.

Does the financial sector create inequality?

A key criticism of all those who have argued that there is a problem with modern finance is that the financial sector is too big. This is frequently linked to an alleged rise in inequality[16]. In a book that was acclaimed widely on the left, political scientist David Marquand strongly criticised the financial sector and its growth. Marquand (2015) blamed the financial sector for inequality and rent-seeking, citing the US as being at the extreme end of a financialised economy. Marquand's work was intended to be polemical, but more sophisticated analyses have made similar points to those of Marquand about the link between finance and inequality. For example, Courene et al (2015), writing for the OECD, argue that the growth of the financial sector was a contributor to inequality. Turbeville (2015) has produced a very detailed analysis of the growth of finance in the US and concluded that the sector has been important in promoting inequality. Whilst the same cannot be said of the OECD study, Turbeville's analysis suffers from a focus on one era in one country.

Looking at international data, prima facie, there does not seem to be any obvious relationship between the size of the financial sector and inequality. If Luxembourg is excluded as being a special case of a tiny country with a huge financial sector, the largest financial sectors in the world as a proportion of national income are to be found in Switzerland, Australia, the Netherlands, the US, the UK, Iceland, Ireland, Belgium and Denmark (in that order)[17] [18]. In all these countries, the share of the financial sector in the economy by gross value added is between 5 per cent and 10 per cent. The US is not at the extreme, as Marquand likes to suggest. Furthermore, these countries are a mix of high, medium and low income inequality countries. Indeed, three of the OECD's eight most equal countries are in this top ten list of financial sectors.

The OECD analysis of financialisation suggests that, in certain circumstances, the growth of the financial sector can raise inequality. One of its policy suggestions is that the sector should be required to hold high capital buffers, to which we would object for various reasons[19]. However, its other two policy recommendations that implicit subsidies should not be given to the financial sector and that tax systems should not artificially discourage the use of equity capital would be supported by anybody who believes in a genuinely liberalised financial sector. If the financial sector is encouraged by implicit subsidies (for example, government guarantees and activities of central banks that effectively prevent bankruptcy and protect insiders) then we would expect the financial sector to have grown beyond the socially optimal

[16] Though, globally, inequality is falling, as it is in the UK. In about two-thirds of OECD countries, inequality is rising.
[17] With the addition of Luxembourg, these countries made up the ten largest financial sectors as measured by gross valued added as a proportion of national income in 2015.
[18] See: http://researchbriefings.parliament.uk/ResearchBriefing/Summary/SN06193#fullreport
[19] For example, it would reduce the probability of failure, thus entrenching large established players within the sector. And it would impair lending, thereby reducing the trend growth rate.

size. However, the response to this should not be to attack the financial sector but to remove the subsidies so that the financial sector is genuinely liberalised.[20]

Overall, the criticisms of the financial sector that are commonly made are, if not spurious, extremely unconvincing. Merely to argue that the financial sector has important functions does not, alone, make the case for a bigger and more sophisticated financial sector. However, it sometimes seems that we focus with a microscope on the problems and thus miss the wood for the trees and sometimes miss the government interventions which encourage the financial sector to grow in an undesirable way. In the next sections, we consider some of the important functions of finance and examine in more detail the social function of speculation which is often seen as a 'useless' activity.

The value of the financial sector

When we drill down into the financial sector, its social usefulness becomes clearer. Of the total value added by the UK financial sector (£124bn in 2016)[21], approximately 50 per cent is exported. Although there are also imports of financial services, the UK has a huge trade surplus in the sector of £44bn or 3 per cent of national income (cf. Tyler 2016). Even if the financial services output of the UK economy is socially useless the fact that much of the output of financial services is exported, thus providing the income to buy other services and consumer and investment goods, renders the argument largely irrelevant. Whatever their social utility, exported financial services have no detrimental effect on the economy of the country that is producing them. Indeed, it is notable that financial services provides just over 3 per cent of all the jobs in the economy despite producing 7.1 per cent of gross value added in the economy.

The financial services sector is also diverse. Since the financial crisis, images have been conjured up of an industry that produces opaque and complex products, leading to critiques such as those of Zingales, discussed above. Approximately one-quarter of the UK financial sector is made up of insurance services which are important for protecting against household and business risks and for providing income protection in old age.

It is difficult to measure the recorded value of the set of services to which commentators such as Turner refer when they discuss the social utility of the sector. However, recent analysis from the Bank of England looking at the contribution of finance to GDP shows that financial intermediation undertaken by non-deposit-taking financial institutions – a proxy for the short-term speculative activity often targeted by critics – accounts for just 9 per cent of all financial sector output, equivalent to 0.7 per cent of national income (Burgess 2011). Attempts to argue that this activity overwhelms or distorts the economy are therefore enormous exaggerations. The vast majority of land, labour and capital in the UK are employed in activities that have nothing to with modern financial instruments.

But even this activity is demonstrably useful. Securitisations, which were implicated in the financial crisis, provide opportunities for the diversification of risk and for the movement of risk outside the banking system. They thereby facilitate increased lending by banks. This was widely welcomed by central banks before the financial crisis. Certainly, there were particular problems which led to the financial crisis and

[20] See Lilico (2011??) on deposit insurance reform, plus references on ending too-big-to-fail and the consequences of below-equilibrium interest rates.

[21] See: Tyler (2017). http://researchbriefings.parliament.uk/ResearchBriefing/Summary/SN06193#fullreport

securitised mortgages appeared to be at the scene of the crime. However, this does not make them the culprit nor does it provide evidence that, in their proper context and when not encouraged by regulation, complex financial instruments such as CDOs are a problem in and of themselves. If regulatory and tax systems artificially encourage complexity, this is something that can be dealt with at source.

The reality is that the financial sector is productive, diverse and socially valuable and additional regulation and taxes, which tend to be imposed indiscriminately, simply serve to reduce the productivity of the sector and the economy as a whole. But, what about the more exotic parts of the system that are widely criticised by finance academics, moral philosophers, regulators and practitioners? In the next section, we specifically examine the value of speculation. It is this kind of economic activity which it is often argued is self-serving rather than supporting the "real economy".

The added value of speculation

Speculators have long been heaped with opprobrium. Pope Francis' attack on commodity speculation noted above is one in a long line of critiques. Speculation is often associated with the use of derivative instruments such as futures, options and swaps and there has been a huge growth in the use of such instruments in recent years. The total nominal outstanding in derivatives markets currently stand at over $1.2 quadrillion (though this vastly overstates the economic exposure).

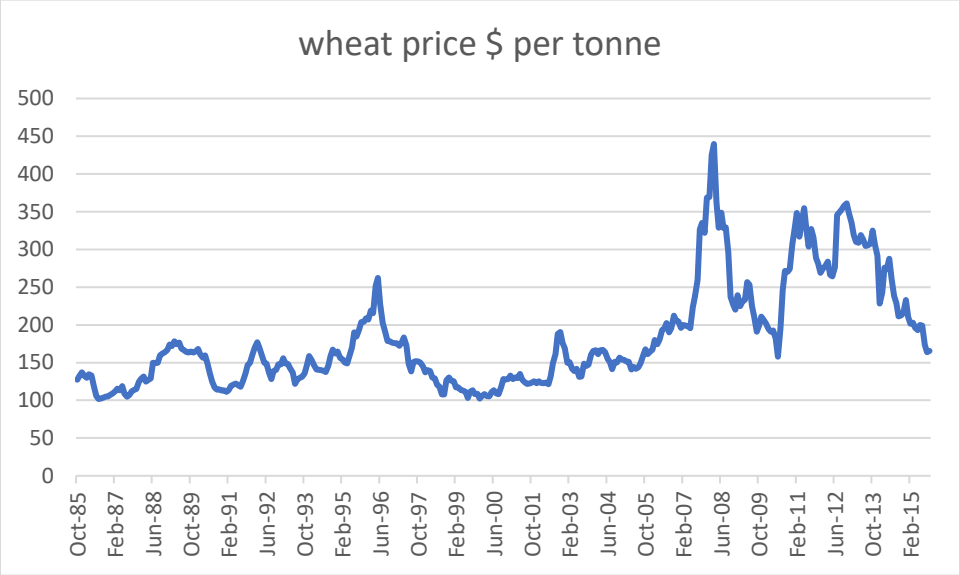
Critiques of speculation do not only come in polemical form. Nobel prize winning economist, Robert Shiller (2017) describes common-place discussions of speculation in the US in the 1920s and 1930s. Nakrosis (2013) examines objections to speculation in moral theology. Discussion of the issues in moral theology goes back at least 800 years and probably further. Nakrosis' conclusions are interesting, because he notes that speculation can have economically valuable effects and, in those cases, it should not be assumed to be morally wrong. However, even in circumstances where it is morally wrong, it does not follow that it should be made illegal or even made more difficult via regulation or taxes.

Following the 2009 financial crash, the Archbishop of York had strong words to say about speculators, commenting: 'To a bystander like me, those who made £190 million deliberately underselling the shares of HBOS, in spite of its very strong capital base, and drove it into the bosom of Lloyds TSB Bank, are clearly bank robbers and asset strippers.'

Commentators often relate speculation to problems in the real economy that are faced by people trying to make a living from the production of non-financial goods and services. For example, Pope Francis argued that speculation in financial markets, by people disconnected from the real economy, leads to devastating consequences for people who need to buy food. In particular, he argued that speculation required higher prices to deliver profits to the speculators and so speculation led to higher food prices.

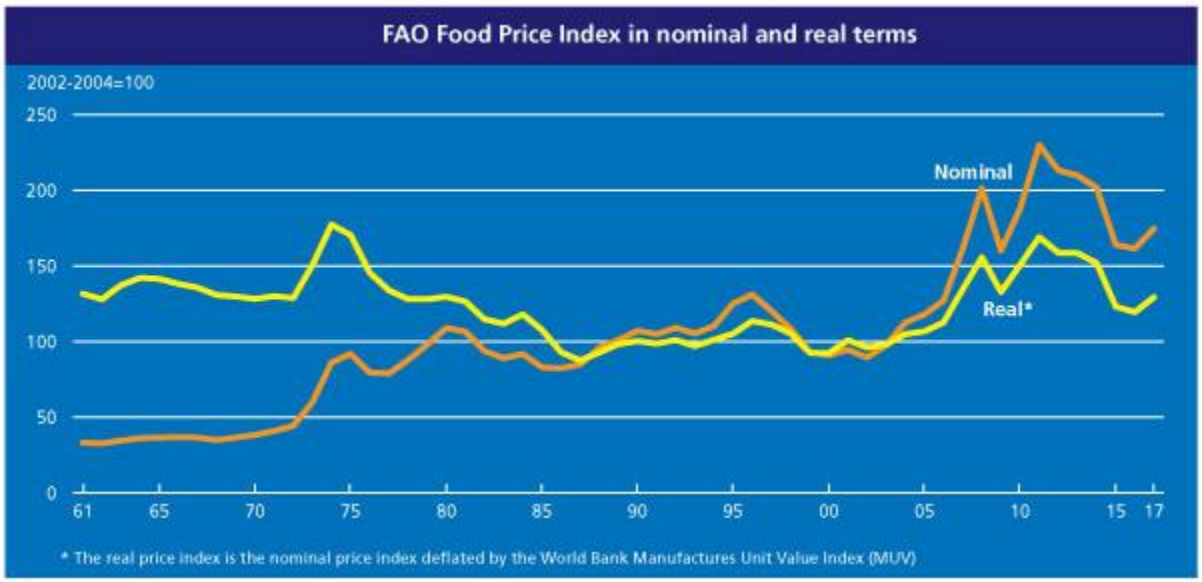
It is, in fact, difficult to think of any plausible mechanism by which prices can be driven ever-higher by speculation motivated by the desire for more profit, though the possibility that prices might be driven higher temporarily will be discussed below. When commodity derivatives are traded, there are two sides to the trade. One side benefits from the price going up and the other side benefits from the price going down. There is no net gain from prices going up as far as the speculators are concerned, just a different distribution of gains between the different types of speculators. Indeed, the Archbishop of York argued that speculation must drive prices down (in share markets) whilst Pope Francis argued that speculation must drive them up (in food markets). Both cannot be correct, though neither might be.

As it happens, as if shown by Figure 3, food prices have certainly not systematically risen as the availability and use of speculative instruments has increased. The figure shows wheat prices since 1985.



The wheat price today was more or less the same in 2015 in nominal dollars as it was in 2009, 2005, 1995 and 1985. And this makes no allowance for general price inflation. In real terms food prices have been falling for a long time despite rising populations. The Food and Agriculture Organisation (FAO) of the United Nations, in its index of food price changes, shows no increase in price volatility between the 1960s and 1970s and the present (Fig.4).

Fig. 4. FAO Food Price Index, 1961-present



Prices do not always stabilise at higher levels as Pope Francis suggested. If anything, they have been stabilising at lower levels. Of course, spikes occasionally occur due to natural or policy-related factors or a combination of both. In 2008, for example, the rush to bio-fuels reduced food supply.

There is, though, a separate and important question of whether speculation might lead to more volatility in commodity prices. This, in turn, might create pinch points where people, in particular situations are unable to access food. This reasoning applies to all commodities, but we will use food to illustrate the issues.

Speculation, derivatives and hedging – the example of food

Financial instruments can be used to manage risk by those operating in the real economy. This is not speculation as such, though the creation of the instruments may lead to or require speculators. For example, a farmer or a food processor might wish to buy or sell commodities they do not currently have in order to protect themselves against price changes that can be disruptive to their business and they can do that using derivatives. A farmer producing cocoa could sell produce at a price that is fixed today before his harvest is gathered in because all the costs have to be borne before the crop is produced. This is simply a form of insurance (hedging). In the same way, a processor might wish to buy cocoa at a price that is fixed in advance of the crop being produced in order that they can develop their marketing strategy fix price lists, and so on.

For these reasons, simple futures or forward contracts can develop. They exist in all sorts of markets and have existed since ancient times. However, if we leave it to farmers and processors to simply make arrangements between themselves, there are several disadvantages. There will be relatively few buyers and sellers of the contracts; both will be locked into contracts which might not turn out to be appropriate as situations change; the contracts will probably be expensive; one party will be at the risk of the other party failing financially; and so on. As a result of these risks and other obstacles, financial futures markets develop that allow buying and selling of forward interests thus making the markets more efficient and prices more stable for producers. Furthermore, the use of financial instruments allows speculators to assume risks, and to pass them on, which might otherwise have been borne by vulnerable manufacturers and farmers. In various ways, speculators and others in financial markets can then diversify those risks.

From hedging to trading

Once such futures are developed, it is difficult to prevent people buying and selling them even if they are not farmers, processors or others who have an important interest in insuring themselves against volatility in food prices. Indeed, the general availability of such instruments is important in ensuring that the market is liquid. It is at this point that there becomes a perceived problem of investors, traders and speculators buying commodity futures.

So, the question arises, does such trading in commodity futures cause problems in food markets? It is difficult to see how this would be so.

The benefits of such instruments can be illustrated by an example. Assume that people expect a rise in the price of wheat by 50 per cent over six months and that the current price is 100 currency units per tonne. Let's also assume that there is a futures contract available that allows a trader to agree today a price at which they will buy a tonne of wheat in six months' time. Given that the traders expect the wheat price to be 150 in six months' time, it is highly likely that the futures contract will settle on a price of around 150. In six months' time, the buyer of the future will have to buy a tonne of wheat from the seller

of the future for 150.[22] Assume a hedge fund buys that contract. In practice in six months' time the hedge fund pays 150 and receives an amount of money equal to the price of a tonne of wheat. In fact, the contract is designed, of course, so that the holder receives the difference between the price of a tonne of wheat and 150. The buyer – or speculator – gains from any increase in the price of wheat. At this level, the arrangement is not connected to the market for wheat: there is no mechanism that links futures speculation and the actual price of wheat in this example. The hedge fund has just taken a bet on the price of wheat. We may approve or disapprove of such bets, but the wheat market is unaffected.

Now assume there is speculation in the futures contract and everybody wants to buy it for some reason. The futures contract might now rise in price to 200. This means that, in six months, the buyer receives the difference between the price of a tonne of wheat and 200. If the price of wheat does not rise to 200 within six months those who have bought the future will make a loss and those who have sold it will make a profit. Perhaps this is what the Pope has in mind, though it should be noted, again, that there are two sides to the transaction. The question is whether this type of speculation can affect the wheat price.

There are various ways in which the price of wheat can react to and follow the price of the future. In other words, there are ways in which speculation can drive up prices, at least in theory. For example, traders could somehow corner the market in wheat and withdraw it from the market in order to keep the price up so that they make profits on their futures contracts. In some circumstances this would be illegal, but it is, in fact, impossible in relation to the vast majority of commodities and there is no real evidence of it ever happening. In general, futures traders never have access to the underlying commodity and it would be very difficult to access sufficient to affect the price. OPEC countries' recent struggles to raise the price of oil by curtailing their own production is just the latest example of the failure to corner a market over any extended period.

Secondly, the speculation in the futures market might somehow affect decisions to produce wheat, thus driving up the price of the underlying commodity. However, the opposite is likely to happen in this case. If farmers see that the futures price is high, it gives them an opportunity to lock in a high price for their produce today. They might be able to plant today and sell wheat futures to guarantee them the high price. But, this should increase food production more quickly than would otherwise happen and therefore bring down food prices and reduce volatility in prices.

Lastly, farmers might store their wheat when the futures price is higher than the current price instead of bringing it to market immediately.[23] They might do this because they could effectively lock in a price of 200 to be received in six months' time by selling the wheat future. Most futures markets are in storable commodities and some – such as metals - can be stored indefinitely. So, if a farmer sees the wheat futures price at 200, he might sell the future and put the wheat in storage , thereby locking in the futures price.

If there are people going hungry at the time, this might be thought a bad thing. And it is just possible to envisage a mechanism where it might cause problems in some circumstances. However, in general, the effect is benign. In a sense, the farmer is doing exactly what happens in the Bible – storing in years of relative plenty in order to sell it when there is a shortage as, presumably, the price rises to 200 in six

[22] Certain practicalities have been omitted such as the treatment of interest and the fact that at the end of the contract settlement does not take place in such a "clunky" way.

[23] This phenomenon is widespread enough in the oil market to have its own name: contango. See http://www.investopedia.com/terms/c/contango.asp.

months' time because there is a relative shortage. The futures market facilitates the storage of commodities so that they can be sold at a time when there are shortages and higher prices. Although the person producing commodities is taking the wheat off the market now, he is bringing it back on to the market when the price is even higher. The futures market helps him to do that.

Financial speculation in commodities is likely to have the following effects. It will:

- Make it easier for relatively poor farmers to get information about prices
- Make it easier and cheaper for relatively poor farmers to manage their own risk
- Leave relatively poor farmers less exposed to price changes

And it could have the following effects:

- Bring forward increases in production in times of shortage
- Facilitate storage in times of surplus

The argument that speculation and derivatives in food and other commodities might reduce supply in times of shortage and increase volatility seems possible but implausible and the evidence seems to back that up. As one study puts it: "In theory, feedback effects are possible; in practice, no empirical proof has been found. Studies show that in recent years, storage levels for soft commodities did not increase as prices rose. The levels stayed steady, or even dropped."[24] Price volatility in commodities where there are no futures markets (such as tungsten) has, in recent years, increased more than volatility in commodities where there are futures markets.

Whilst it might be the case that volatility of commodity prices in general does not increase, it could be the case that financialisation of commodities markets and the creation of derivative markets related to them could lead to price spikes and increases in volatility at specific times. This might happen if information were unevenly distributed or if markets were operating inefficiently. However, the evidence for this is weak[25]. Similar conclusions were drawn in a major study of oil market speculation. Though the possibility was left open that some forms of speculation were more damaging than others, no evidence was found for a causal link between financial markets and oil price rises (Fatthouh et al 2012).

Benefits from speculation

Thus, when it comes to speculation linked to commodity markets, it is difficult to see systematic and serious problems arising from financial markets and easy to see advantages to those involved in commodity markets from linked financial markets. However, we can go further: there are many potential benefits from speculation.

Returning to the 'robber barons' who short-sold bank shares, this is one of the more obvious situations where speculation can bring significant benefits. If information is discovered that suggests a company is over-valued (for example in relation to the credit risk on a bank loan book), in general, it is better that such information is reflected in the share price as quickly as possible so that action is taken. The main beneficiaries of short-selling not taking place are those in board rooms whose mistakes go unpunished by

[24] https://www.db.com/cr/en/concrete-Agricultural-speculation-and-commodity-prices---is-there-a-link.htm
[25] See https://www.db.com/cr/en/docs/zef_dp_161.pdf

markets for longer. As Copeland and Booth (ed) 2009 points out, if short-sellers had been more active when RBS announced plans to takeover ABN Amro, taxpayers might have been spared some of the costs of rescuing the bank in 2008.

The situations in which short-sellers can distort markets by forcing prices down to the extent that a firm may face bankruptcy are extremely limited and could only occur in specialist and illiquid markets. In liquid markets, if short-selling leads to lower share prices, other investors who disagree with the short-sellers are likely to buy the stock. Short-sellers also provide greater liquidity in markets by ensuring that a greater volume of stock is available for those who wish to buy. There is widespread evidence for this, including during a period when the FSA banned short-selling during the financial crisis[26]. Of course, increased liquidity makes market manipulation more difficult and this is perhaps a side-benefit of the development of derivatives and speculation.

Perhaps one of the greatest benefits of short-selling is to bring an end to government policy that is causing damage by distorting markets. The most obvious example of this is the Exchange Rate Mechanism (ERM) crisis in 1992 when the short-selling of the pound allowed the over-valuation of sterling in the ERM to be exposed and thus brought about an end to the extreme monetary policy that was designed to prevent sterling from finding its proper market value. Following sterling's exit, there were 61 consecutive quarters of economic growth.

George Soros was one well-known example of a beneficiary of short-selling in the ERM crisis. On 16th September, the president of the Bundesbank let it be known that he believed a re-alignment of the ERM was possible. It was clear that interest rates in the UK were higher than deemed reasonable by domestic conditions. Either there would have to be internal deflation or sterling's value would have to be allowed to fall and interest rates adjust. George Soros' hedge fund took out a $10billion short position against sterling. This was a safe bet because there was only one way sterling could go – the value of sterling would either stay the same (if there were no re-alignment) or would go down (if there were a re-alignment). It can be argued that the creation of exchange rate regimes that are susceptible to such one-way bets is undesirable per se and that action taken by markets to undermine them is beneficial. However, at the very least, in these cases, markets act to restore equilibrium exchange rates in a system where governments have taken actions that attempt to impose inconsistent policy environments that damage the real economy.

Conclusion

The value of financial markets is under-appreciated. Financial markets perform fundamental functions which are vital in reducing transactions costs in the economy for businesses and households. Without well-functioning financial markets, business would find if much more costly to raise capital and households would find retirement, protection against everyday risks and day-to-day transactions impossible.

Of course, those who criticise financial markets do not deny this. However, they do argue that particular sectors within financial markets are problematic or socially useless. They also argue that financial markets can grow too big or can be short-termist.

[26] See: https://www.lseg.com/sites/default/files/content/documents/short-selling-restriction-market-quality-december-2008.pdf

These anti-finance claims come from a number of sources. Most of the claims lack credibility. However, some of the more evidence-driven claims from economists are not without credibility and do demand scrutiny. The evidence that large parts of the financial sector undermine the real economy or cause social problems appears more circumstantial when put under closer scrutiny. At the very least, the claims apply to a portion of the financial sector which is a vanishingly small proportion of domestically produced and consumed national income.

However, this does not mean that government incentives that promote the growth of particular parts of the financial sector should not be criticised. There is a potential "do no harm" reform agenda on which there can be wide agreement. For example, the US government should not underwrite securitized housing debt and, more generally, should play no part in the housing finance market. We should remove the tax discrimination against equity finance. More controversially, all financial reform efforts should be put into ensuring that we have a legal framework so that insolvent banks can be wound up safely. This would end the implicit subsidy that larger banks receive from the state. More controversially still, if this is done, arguably we can dispense with international bank capital regulation that promotes gaming including through the creation of complex financial instruments in order to reduce regulatory capital charges. One industry which is still growing is that which produces financial regulation along with the related compliance industry[27]. We might question its social utility and also whether it encourages the growth of those parts of the financial sector which are opaque and, according to some, not beneficial to society at large.

References

Armour, J., et al. (2016) Principles of financial regulation. Oxford University Press.

Armstrong, M., and J. Vickers. (2012) Consumer protection and contingent charges. *Journal of Economic Literature* 50.2 (June): 477-493.

Burgess, S. (2011) Measuring financial sector output and its contribution to GDP. Bank of England Quarterly Bulletin 2011 Q3.

Cecchetti, S.G., and E. Kharroubi. (2015) Why does financial sector growth crowd out real economic growth? BIS Working Paper No. 490 (revised). Bank for International Settlements.

--------. (2012) Reassessing the impact of finance on growth. BIS Working Paper No. 381.

Cour

Diamond, D.W. (1984) Financial intermediation and delegated monitoring. *Review of Economic Studies* 51.3 (July): 393-414.

Fama, E. (1970) Efficient capital markets: a review of theory and empirical work. *Journal of Finance* 25.2 (May): 383-417.

[27] Projecting from figures produced by Andy Haldane, if the number of financial regulators and number of people in finance continue to grow in the next 60 years at the rate achieved in the last 30 years, there will be more regulators than people working in finance by 2070. This excludes compliance officers.

Gabaix, X., and D. Laibson. (2006) Shrouded attributes, consumer myopia, and information suppression in competitive markets. *Quarterly Journal of Economics* 121.2 (May): 505-540.

Guiso, L., P. Sapienza, and L. Zingales (2004). Does local financial development matter? *Quarterly Journal of Economics* 119.3 (August): 929-969.

Haldane, A. (2011)

--------. (2012) On being the right size.

Hicks, J.R. (1939) Value and capital.

International Monetary Fund. (2012) Spain: the reform of Spanish savings banks technical notes. IMF Country Report No. 12/141 (June).

King, R.G., and R. Levine. (1993) Finance and growth: Schumpeter might be right. *Quarterly Journal of Economics* 108.3 (August): 717-737.

Lane, P.R. (2013) Capital flows in the euro area. Economic Papers 497 (April). European Commission.

Levine, R. (2005)

Malkiel, B. (1995) Returns from investing in equity mutual funds 1971 to 1991. *Journal of Finance* 50.2 (June): 549-572.

Markowitz, H. (1952) Portfolio selection. *Journal of Finance* 7.1 (March): 77-91.

Martín-Aceña, P. (2013) The savings banks crisis in Spain: when and how? World Savings and Retail Banking Institute.

Mbiti, I., and D.N. Weil. (2011) Mobile banking: the impact of M-Pesa in Kenya. NBER Working Paper 17129.

Moody's Investor Services. (2017) http://uk.businessinsider.com/moodys-investor-services-report-on-active-and-passive-investment-2017-2?r=UK&IR=T

Thaler, R., and C. Sunstein. (2003) Libertarian paternalism is not an oxymoron. University of Chicago Public Law and Legal Theory Working Paper No. 43.

Turner, A. (2009) How to tame global finance. *Prospect*, 27 August 2009.

Varian, H. (1980) A model of sales. *American Economic Review* 70.4 (September): 651-659.