Duncan, Alfred and Nolan, Charles (2024) *Adam Smith and the bankers: retrospect and prospect.* National Institute Economic Review, 265 . ISSN 0027-9501.

Adam Smith and the Bankers: Retrospect and Prospect[1,2]

Alfred Duncan*

Charles Nolan**

June 2023

Abstract

Adam Smith promoted free banking—private, competitive, convertible banknotes. He also supported restrictions on banks. We study Smith's views and the era in which they developed, suggesting his 'regulations' were a backstop against banks' risks to depositors but primarily monetary stability. In modern parlance, Smith supported macroprudential regulations to underpin monetary stability, as did Friedman and Schwartz (1963) the US FDIC. We discuss why Smith's vision for banking went unrealised. Bank regulation became microprudential and ran aground in 2008/9. The prominence of macroprudential regulation now provides a chance to reorientate regulation to support monetary stability. Early signs are not promising.

JEL: B12, B31, E58, G21
Keywords: Adam Smith; Banking Regulation; Economic History; Monetary Policy.

* University of Kent
a.j.m.duncan@kent.ac.uk
** University of Glasgow
charles.nolan@glasgow.ac.uk

[1] This paper was written to mark the tercentenary of Adam Smith's birth. The title of the paper recognises the singular contribution of Professor Sydney G. Checkland (1916-1986), of the University of Glasgow, to Scottish banking history and 'Adam Smith and the Bankers', (Checkland, 1975a, b). His insights will be apparent in what follows.

[2] We thank for helpful comments participants at a workshop on Adam Smith at the National Institute for Economic and Social Research. We are also indebted to Mark Billings, Sayantan Ghosal and Alex Trew for very helpful insights. The usual disclaimer applies.

Introduction

For a long time, Adam Smith's views on money, banking and credit seemed to enthuse economists and historians of economic thought much less than his other economic insights. As far as money and banking go, and related topics such as the balance of trade, some have argued that he said little of any real novelty, perhaps committed serious gaffs, and even inexplicably passed over Hume's insights with which undoubtedly he was familiar. On banking, his clear advocacy of restrictions on laissez faire remain a disappointment to some. And what came to be known as the real bills doctrine has been, in turn, exploited by policymakers and denounced as a fallacy by economists[3]. More recently, some scholars have urged a rethink on Smith's contributions in these areas, arguing his monetary insights were sound, his banking restrictions sensible and his real bills 'advocacy' misunderstood.[4]

In this paper we try to do four things. First, we provide an historical overview of banking for the period in which Smith was writing. Outlining its salient features, and following Checkland (1975a) and Rockoff (2011), two events are emphasised that appear especially influential in the evolution of Smith's thinking; specifically an episode of overissuing of currency and a systemic bank collapse. Second, Smith's vision is then described of what a 'good' banking system should look like, as well as good monetary policy (and central banking[5]). We view Smith's manifesto for good banking through a set of five *internal rules* and five *external rules*. Smith endorsed what we call 'internal rules' or what might be labelled 'good banking practice' but otherwise championed free entry into banking and competitive note issue. But his support for 'free banking' was qualified by 'external rules.' The reasons he thought the law should intervene are analysed as are the trade-offs that Smith seems to have identified as important. Third, we show that while Smith recommended that banks should be numerous, joint stock, competitive, unlimited liability, with a right of issue, but legally restrained in other ways, by the middle of the nineteenth century that vision was effectively dead. Some of the reasons

[3] His critics got in early. Sen (2013) notes that Jeremy Bentham criticised some banking regulations Smith supported (in particular, usury laws directed at 'prodigals and projectors'). Bentham and Smith disagreed on much else, as is well known. Modern critics of Smith on money and banking would include: Charles Rist (1940) who questions a number of aspects of Smith's monetary economics including whether he understood the difference between money and credit; indeed Rist is nigh on dismissive in places; Viner (1937) famously remarked that Smith's failure to acknowledge Hume's prices-specie-flow mechanism is one of the great mysteries in the history of economic thought; Taylor (1965) is almost as dismissive in places as Rist (op. cit); Selgin (2018) criticises the restrictions on small notes that Smith strongly advocated; Mints (1945) coined the expression 'the real bills doctrine', with some help from Smith (see later) and denounced it as a fallacy, a view that monetary economists widely accepted; Thornton (1802) also questioned Smith's real bills advocacy, though for different reasons to Mints'; and few economists perhaps would instinctively support usury laws for the reasons as did Smith, see Hollander (1999). Overall, Schumpeter (1954) thinks the Wealth of Nations was a singular contribution but, as a matter of fact, contained no new ideas.

[4] Those urging a rethink specifically on his monetary and banking views include Laidler (1981, 1991) who explains why the real bills fallacy is not to be pinned on Smith, Wennerlind (2000), and Rockoff (2011), see below. Arnon (2011) a recent historian of monetary thought is also more positive as to Smith's contribution. See also Sen (2013).

[5] He had, of course, little to say directly about what we would now think of as central banking and monetary policy, although he had a sense of the possible role of the Bank of England as lender of last resort as we discuss below.

why that was the case are discussed. Finally, an attempt is made to deduce how Smith might have approached some of the current issues in banking regulation. The rules he thought banks should follow and have imposed on them by Government reveal, it is argued, a concern for smaller, less sophisticated borrowers but primarily a desire to ensure that the medium (media) of exchange is (are) reliable. Bank's may well be inclined both to overexpansion and excessive riskiness, Smith noted. Thus, Friedman's (1959) observation around "the peculiar difficulty of enforcing contracts involving promises to pay but that serve as a medium of exchange and of preventing fraud in respect of them" is something that likely would have resonated with Smith. As would the issue of contagion of fear driving bank collapses emphasized by Friedman and Schwartz (1963); indeed, Smith made exactly this point in justifying legal restrictions on banks. Ultimately, it is suggested that for Smith the rules that banks should follow in their management of their balance sheets, or have imposed on them by the law, are driven by one overriding issue: to ensure the integrity of money and monetary exchange. Arguably that perspective has been obscured in modern approaches to financial regulation.

Overview of Banking and Money in Scotland in the Eighteenth Century

Smith appears to have thought that the Scottish banking system had boosted Scottish growth and had much else to commend it. Hence, this section outlines some of the key features of the Scottish banking system at the time of Smith's writing of the Wealth of Nations which was published in 1776.

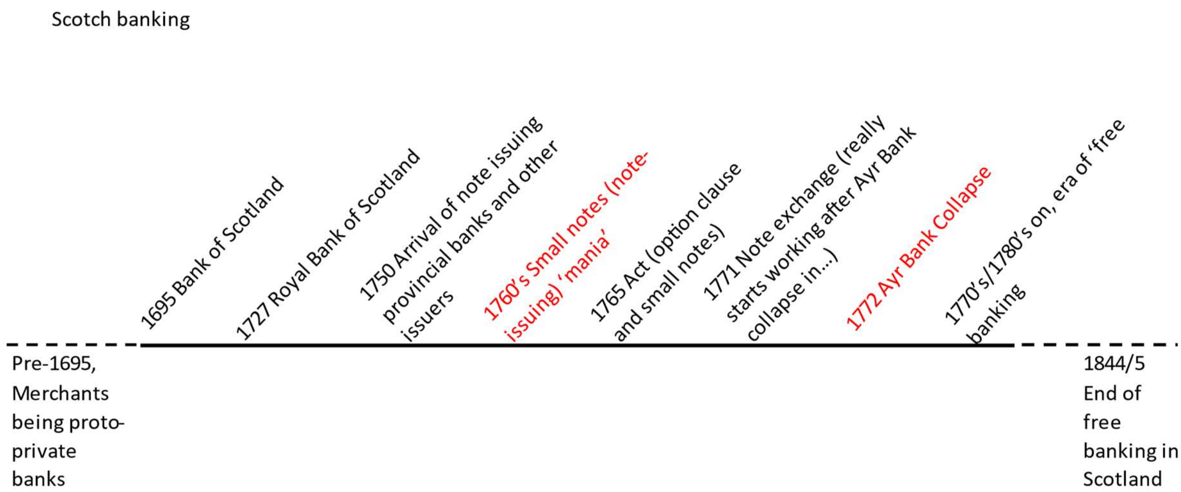
Towards the end of the eighteenth century, Scottish banks were perceived by some as more robust than the banks south of the border, in part because they (the Scottish banks) were less constrained. Nevertheless, we suggest that the Scottish banking system faced challenges just as did the English system. Significantly, the clearing of notes in Edinburgh, whilst an important element in the purported emergent stability of the Scottish version of free banking, had no clear rules on overissue. That could, and did, lead to frictions that appear never really to have been resolved, until 1844/45 when free banking in effect came to an end. Smith believed that banks, completely unconstrained, were prone to instability. Indeed, the Scottish experience led him to promote clear guidelines as to how banks should be run internally[6] and to support external rules that the state should also impose.

The timeline below gives a brief overview of Scottish banking from its emergence towards the end of the early modern period, with the Bank of Scotland's incorporation in 1695, one year after the Bank of England's, up until Scottish free banking's final demise in 1844/5 with the Peel Acts which all but ended the Scottish banks' ability to decide on their issues and placed the Bank of England as the marginal supplier of new base money henceforth.

Two entries on the timeline are highlighted in red, the small notes mania in the 1760's, and the collapse in 1772 of Ayr Bank. Both events, particularly the failure of Ayr Bank, are

[6] Checkland (1975b) suggests that some of what Smith promotes was already regarded as good practice within the banking community although he does not give any sources for this.

discussed in the Wealth of Nations and Smith scholars suggest these events were formative in Smith's understanding as to the stability of banking. Both are mentioned presently, and discussed in more detail later.



By the time of the Wealth of Nations, Scottish banks were of three major types. First, there were the chartered banks. The two largest were Bank of Scotland, chartered by an Act of Parliament in 1695, and The Royal Bank of Scotland, which obtained a Royal charter in 1727.[7] These are sometimes referred to colloquially as the 'Edinburgh banks' or the 'public banks'. Up until around 1750 the only other class of banks were the non-issuing Edinburgh private banks who discounted bills (which the two public banks then did not) and who used the two banks' issues (i.e., paper money) in their discounts. Both the public banks enjoyed limited liability, something not generally available to banks in Scotland or the rest of Britain until 1858[8]. From around 1750 on, there arose a third class of banks—the provincial note-issuing banks (Munn, 1981). Until the 1770's this amounted to four banks in Glasgow and another five in other towns (Checkland, 1975b).

After 1750 a period of direct intense banking rivalry between the two big Edinburgh banks subsided as they faced a common perceived threat in the provincial banks. These banks were free to issue notes—as were other 'firms'—and they proved, it would seem, to be effective competition to some degree at least. Checkland (1975a, b) describes the small-notes mania

[7] A third chartered bank, The British Linen Company, was granted a Royal charter in 1746. This bank was somewhat smaller than the two Edinburgh chartered banks, although its authorised capital was on a par with the other two. See Checkland (1975a).

[8] Banks in Britain were permitted to limit their liability from 1858. However, very few established banks in England immediately sought to limit liability. In the decade or so after the legislation was enacted most of the limited liability banks in England were new banks, and there were no limited liability banks in Scotland (save the three chartered banks that had always been limited liability). That would change in both England and Scotland after the failure of the City of Glasgow Bank in 1878.

of the 1750's and 1760's as one in which notes for as little as 1 shilling (and even smaller) were circulating having been issued by 'a great many parties' who thus became de facto quasi-banks of issue and with the resulting notes often taken up in an economy chronically short of small change and in particular silver coin. Outright frauds were also widely reported although no detailed data on the extent either of 'mania' itself or its fraudulent component exist[9].

The two big Edinburgh banks were worried about these competitor issues and wanted to supress them legally. They argued that suppression would support monetary stability. No doubt their motives were mixed. Smith (1776) would also worry, looking back at this period, that the most vulnerable, small traders and people of modest means, would be the most exposed in such a situation of widespread monetary mistrust. More recently, as noted below, free-banking scholars have suggested Smith rather badly misread this situation.

After the short life and spectacular collapse of Ayr Bank in 1772 (again, see below), the Edinburgh private banks initially diminished in importance; only four remained of the perhaps more than twenty that had been in operation[10]. There were a number of other non-issuing banks elsewhere in the country. However, as the note-issuers rose in number, the non-issuers declined although they remained numerically significant until the end of the century (see the Table below from Cameron, 1967). So, towards the latter part of the eighteenth century in Scotland, in addition to the two dominant limited liability banks, and the British Linen Company (also joint stock and limited liability), there were many smaller provincial banks and the Edinburgh private banks. The note exchange, based in Edinburgh, was in regular operation by 1774. And by 1780 the Bank of Scotland had also begun to construct a substantial branch presence. The Royal Bank would be slower in doing so. From this time on, the banking system seems overall to have been relatively stable and the Scottish system of banking gained a reputation among some (contemporary and future) observers for stability and effectiveness, at least compared to the system south of the border.

According to data from Cameron (an extract of his 1967, Table 3.1, p. 66 is reproduced below), Scotland went from having one bank of issue at the start of the eighteenth century, rising to 12 at the time of the Ayr Bank collapse, and rising to 18 by the turn of the nineteenth century. It would rise still further and for the next 30 years there would be 30 banks of issue. The branch network would grow similarly: From zero at the start of the eighteenth century, to 93 by the start of the next century. And it would grow much more strongly thereafter.

Several features of the Scotch banking system distinguished it from that south of the border. First and foremost was the size of banks in Scotland; banks of issue in Scotland were not restricted as to the number of partners as they were in England. They had as a result more capital resources and were quicker to develop branch networks. Next, by the 1770's the various note issues of reputable banks in Scotland were all accepted at par by other banks in Scotland and cleared in Edinburgh. Third, the population appeared to be very comfortable with paper money and small notes. The flip side of this, is that there may have been less cash

[9] See Kerr (1908) and Rockoff (2011) for some of the anecdotes and details.

[10] Estimates of the precise number of private banks in Edinburgh seem to vary a little. Cameron in the table in the text suggests that in 1770 non-issuing banks numbered 17. The suggestion of twenty or so banks mentioned in the text comes from Checkland (1975a,b). There is no disagreement amongst historians that there was widespread failure of Edinburgh private banks following the collapse of Ayr Bank.

(i.e., specie) in the economy. Fourth, Scotland had more than one dominant bank although none of them had the monopoly privileges of the Bank of England nor had they the close links to Government that the Bank of England had.

	1704	1750	1760	1770	1780	1790	1800	1815	1825	1836	1845	1865
Banks of Issue	1	5	5	12	12	17	18	30	31	29	19	12
Other banks	0	8	18	17	14	1	1	10	7	3	1	0
Branches		0	2	11	17	54	64	84	141	229	368	682
Total Offices	1	13	25	40	43	84	93	124	179	261	388	694
Circulation (£1,000's)	51	163	393	700	1,200	2,100	3,500	3,164	4,058	3,281	3,351	5,003
Deposits (£1,000's)	0	140	290	600	1,000	2,000	7,000	15,000	21,000	25,000	33,192	57,180

Adapted from Cameron, 1967, Table III.I, page 66.

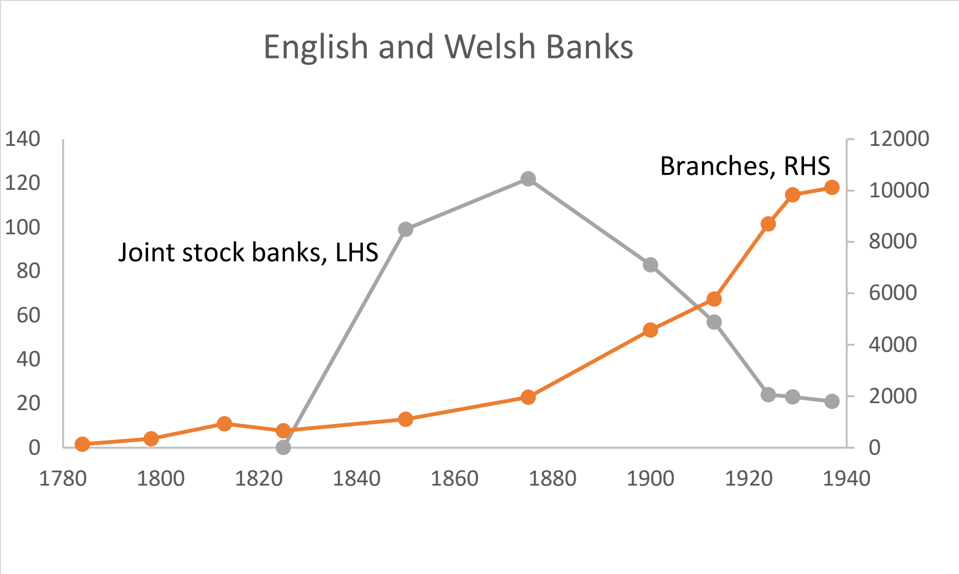
Size and Governance of the Scottish Banks

The provincial Scottish banks were partnerships with unlimited liability. There were no restrictions, as in England, as to the numbers of partners. Moreover, as described below, these partnerships were, in effect, joint stock in that the shares could be relatively easily passed on/sold and so there was an effective mechanism to separate ownership from control (Acheson et al., 2011, Turner, 2014).

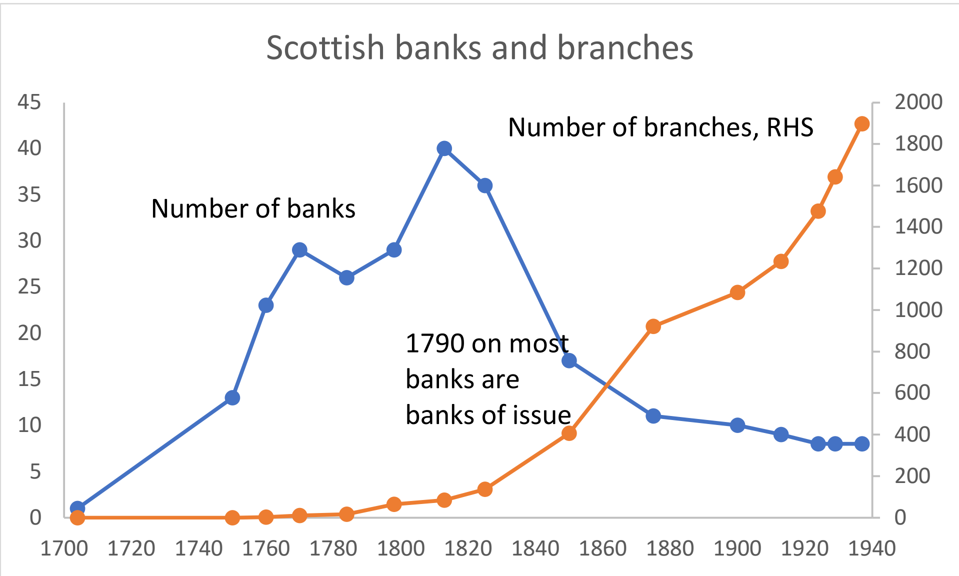
The restriction on banks of issue in England dated from 1708 and the recharter of the Bank of England. Feaveryear (1963, p154/5) reproduces the relevant part of the Act:

> No partnership…exceeding the number of six persons, in that part of Great Britain called England, to borrow, owe, or take up any sum or sums of money on their bills or notes, payable at demand, or at any less time than six months from the borrowing thereof.

Whether or not this left room for joint stock banks to exist so long as their liabilities were not 'on demand' is dismissed by Feaveryear; he argues the intention of the Act was clear, repeated in various Bank of England (Bank) charters and not seriously doubted at the time. Corroborating Feaveryear, Grossman (2010) notes that the 1742 chartering act made crystal clear the Bank's "privilege of exclusive banking". That aside, note the application of this restriction to "that part of Great Britain called England."



Data in chart from Collins (1988) based on underlying sources of Cameron (1967), Nishimura, (1971) and Collins' own research. Looking at the underlying data we note that there is some disagreement between Pressnell (1956) and Cameron (op.cit.), p. 33. Pressnell (op. cit.) thinks there are more branches and fewer banks compared with Cameron.



In any case, from 1694 until 1826 the Bank of England was the only bank in England, note-issuing or not, with more than six partners.

The data in the charts above for Scotland draw on Cameron (1956). The contrasting profiles of 'joint-stock' banking in Scotland and England and Wales are apparent. The number of banks in Scotland had peaked ahead of joint stock banking being permitted in England and Wales (following the 1825/6 crisis). The peak in the number of banks in Scotland coincides with the peak in private banking in England (not shown) which declined secularly and had all but disappeared by the end of the eighteenth century. The rise of the branch network was quicker in Scotland. Nishimura (1971) argues that the rise of branch networks in England was later than many earlier scholars had suggested.[11] In any case, the branch network hits its inflection point in the last 1880s on this data (which is from Collins (op. cit.) but based on Cameron (op. cit.), Checkland (1975a) and Nishimura (op. cit.) and Pressnell (op. cit.)). Also indicated in these charts, is an increasing concentration in the banking sectors across all of Great Britain. We return to that below.

That said, as documented by Pressnell (1956), it is interesting to note that most issuing banks in England had somewhat fewer partners than the permitted maximum of six. Pressnell (1956) reports that note issuing banks in May 1822 totalled 552, of which 375 had three or fewer partners. He reports that 26 issuing banks had 6 partners. The average number of partners per bank was 3. Acheson et al. (2011) point out[12] the equivalent institution in Scotland had over 43 partners. Indeed, 11 Scottish banks had more than 50 partners and one bank had 424 (see the data reported in Munn, 1981, p. 152ff).

Acheson et al. (2011) and Turner (2014) argue that Scottish partnership law allowed Scottish banks to be, in effect, joint-stock institutions with transferable shares—unlimited liability notwithstanding—an effective separation of ownership from control, and separate legal personality. They argue that this latter facet short-circuited a hold-up problem that bedevilled the English banks; and it was this hold-up problem, these authors suggest, that explains the small number of partners in the typical English bank, smaller typically than the legal permissible. In other words, the binding constraint on bank size in England was not the "obnoxious clause", as Joplin (1822) labelled it, that gave the Bank of England the monopoly of joint-stock banking and limited every other bank of issue to a maximum of six partners; it was the nature of English partnership law and the consequent inability to separate ownership from control.

Turner (2014) suggests that note-issuing banks, hobbled by hold-up problems and inefficient dissolution costs, and with a small number of partners were prone to a form of risk-shifting. That is, if partners' wealth was diminished for whatever reason, then the partners faced incentives to undertake more risky projects; with less wealth on the line, attitudes towards risk softened. However, that still may not explain why expanding the number of partners was not also a way of effectively limiting the liability of all partners. There seem to have been some interesting trade-offs here and why they were resolved in favour of fewer partners in many instances is not entirely clear. In any event, if Turner (2014) is correct, then the limit placed on note-issuing banks may have been unnecessary as partnership law was the real binding constraint on the stability of the English banking system. And so one might ask: What effect did the limit of six really have if partnership law was the real binding constraint?[13]

---

[11] To be more specific, Nishimura (1971) is referring to what he calls "balanced" branch networks spanning net saving and net borrowing regions of the economy. He is specifically looking to understand the decline of the inland bill market and seeking to counter arguments in King (1936).

[12] See also Turner (2014, p103ff).

[13] It may be possible to investigate this issue further by a closer examination of the regional dispersion of the banks and the number of partners, perhaps by comparing the number of partners in agricultural-area banks and manufacturing area banks, the latter being typically in areas that were net borrowers and the former net lenders. There are also empirical regularities with respect to note issuance and use across these different areas which may also be useful in that investigation. For example, see Wood (1939).

In any event, as well as squashing any effective competition to the Bank of England, this paucity of partners was held by many at the time to account for the widespread fragility of banks in England compared with their Scots counterparts. And others would argue that private banks of issue had a propensity to overissue, or otherwise pursue disordered patterns of issue and retirement[14].

One well-known contemporary commentator who blamed the 1708 Charter for the fragility of English country banks put it thus:

> I believe that trade is pretty much the same in both nations, or if there is any difference, that the merchants of Scotland are the most speculative, and least stable of the two. But the true cause of the difference is to be found in the nature of their respective banking establishments; the Scotch banking being joint stock companies, while the English banks are private concerns." Joplin (1822)[15].

Joplin (1822) and others of his viewpoint made little headway in their arguments against what he called "that obnoxious clause" cited above by Feaveryear (1963). However, following the crisis of 1825/6, Scottish banks appeared to do much better and certainly did not experience the widespread failures seen in England (where about 10% of all banks failed). That crisis led to fundamental reform.

The upshot was the 1826 Banking Co-Partnerships Act ("An Act for the Better Regulation of Co-Partnerships of Certain Bankers in England"). This was a major piece of legislation that lead eventually to growth in both the number and size of banks. As Grossman (2010) notes, this was the beginning of the end of the Bank of England's monopoly on joint stock banking, although it would remain dominant in that market for some time afterwards (see King, 1936). There was no need to apply it to Scotland as the Act was consciously trying to emulate a perceived strength of the Scottish system.

The 1826 Act did not amend the requirement of unlimited liability which the Bank of England possessed, and the Bank retained other privileges too as banker to the Government and monopoly supplier of notes in the capital.

The 1826 Act gave English banks in effect the same legal possibilities as Scottish banks; banks could be joint-stock companies with shareholders required to have joint and several unlimited liability. Stock could be more easily transferred—a major difference to the set-up under partnerships. Although not formally part of the Act, Directors of a bank had to approve share transfers. This was to make sure shares were transferred to individuals of sufficient wealth such that the unlimited liability obligation remained evenly spread across the owners.

---

[14] And that basic division of opinion would develop into a defining aspect of the debate between the so-called Banking and Currency Schools which would culminate in, but not end with, the Peel Acts.

[15] Interestingly, Joplin (1822) also makes a number of points that may echo in spirit those made by Acheson et al. (2011) with respect to the governance of the Scottish banks.

In sum, as England pondered how to reform its banking sector following the crisis in the years after the Resumption, it looked to certain features of the Scotch system, namely its de facto joint stock structure, to restore a semblance of stability. However, the systems were still somewhat different. Scotland had developed an effective note-clearing system with notes accepted at par and returned rapidly to the participating banks. Scotland did not have a central bank although it was increasingly linked to the London market as a financing source. Moreover, Scotland would develop, ahead of England, a branch structure whilst English country banks would remain essentially unit banks for some time. As discussed below, Scotland had smaller note denominations and fought to keep these despite objections from some in the British Parliament.

Issues with the Scottish Banking System

Detailed data on the Scottish macroeconomy of the time do not exist. Cameron (1967, p94) suggests that a reasonable estimate is that Scotland's per capita income in 1750 was "no more than half that of England's, but that by 1845 it very nearly equalled England's". Many observers, contemporary and more recent, and indeed Smith himself, have been inclined to put this robust growth in part down to the banking sector.[16] Whilst Smith seems to see many strengths in the Scottish banking system at the time of the publication of the Wealth of Nations, there were issues with the system. Some of these issues he recognised explicitly and thought banks themselves should address for their own good; others he saw as weaknesses of the system as a whole, needing Government legislation as a remedy. And there were other issues that he seemed not to identify or to care about.

The first issue that Smith seems not to discuss is the fact that despite urging free entry, banking in Scotland over his lifetime was dominated by two big banks: Bank of Scotland and The Royal Bank of Scotland. These banks were much larger than any of the other banks until Ayr Bank was established in 1769. Clearly, free entry into banking did not necessarily mean a competitive banking system—even absent the two public banks. Moreover, these large banks' actions or inactions were of economy-wide importance. And that was despite the fact that neither of the public Edinburgh banks (nor Ayr Bank) were linked to Government as was the Bank of England.[17]

The second issue pertains to a key feature of free banking. It appears that very few details remain of the how the Edinburgh note exchange system actually worked (Munn, 1981). The Edinburgh exchange, as Munn recounts, had been preceded by more local agreements between banks often on a bilateral basis to accept and exchange one another's notes. Such an agreement had been in place between the two Edinburgh banks since 1751.

[16] This is the conventional wisdom so far as Scotland is concerned. Smith (1776) as noted appears to endorse that view as do Kerr (1908), Cameron (1967), Checkland (1975a) and many others. For an important paper looking at the English experience in the nineteenth century and concluding that banking did indeed have a stimulative, that is causal, effect on industry see Heblich and Trew (2019).

[17] The Royal Bank of Scotland had an agency role in the remitting of customs receipts to London; see Checkland (1975a, p.104).

Nonacceptance of notes of other banks was inconvenient to traders, limited a bank's circulation, could lead to frictions between banks, and became more costly as the economy grew and trading networks expanded internally.[18] Ayr Bank made the first decisive moves in November 1769 to end this source of friction between the non-Edinburgh banks by accepting all 'reputable' notes at par and inviting reciprocal acceptance. This move was widely accepted (although not by the public banks). This move was still, it would seem, initially a sequence of bilateral deals. But the agreements were struck that regular (weekly) clearing of notes would take place in Edinburgh and the initial bilateral arrangements became multi-lateral. It was not long before the public banks became parties to the clearing (May, 1771); the Edinburgh banks would accept the provincial bank notes. Their customers apparently put pressure on them for acceptance of other notes, and the banks themselves were aware that the new banks were pushing their circulation more widely. However, as Munn (op. cit., p26) observes there was also a sense that this clearing would also help limit overissue by provincial banks. The Aberdeen Banking Company refused to join this nascent clearing system and indeed urged the other banks to decline its notes. It felt its business would be stymied by the clearing of notes and as it felt its business was largely local nonparticipation should leave its notes in circulation for longer. This bank was not the only bank to resist the attraction of clearing sensing the discipline that the note exchange imparted on banks of issue exceeded the benefits of wider acceptance (Munn op. cit., p27).

However, there are indications that the rules were never really settled and the judgement that a bank was over-issuing could be contentious (e.g., Checkland, 1975a). Munn (op. cit.) also notes that it was not long before the public banks tried to strong-arm non-participants into joining. It would seem that the big banks did start to take a view on over-issue of participating banks; or, were urged to discipline banks perceived of overissuing. Some indications to this effect are contained in Checkland (1975a). In time this oversight, and any lender of last resort role, would need to have been formalised. Gorton and Tallman (2018) discuss in great detail similar issues during the National Banking Era in the United States, and how private banks and the clearing houses to which they belonged managed with little official intervention during crises.

There were several key issues that he did acknowledge. Smith argued that banks, including the Bank of England, did not always act prudently or in their own interests. He indicates that while the Scottish economy had benefitted from its ability to operate with little in the way of hard cash, that efficiency gain was attended by potentially serious risks. And these risks needed to be handled in part by banks behaving prudently and in part by government regulation. First the issue of too little specie is addressed. Then the small notes mania of the 1760's and the collapse of Ayr Bank in 1772 are considered, these being the two events

[18] A feature of the so-called 'bank wars' was collecting up of rival's issue and demanding specie. It also led to sharp practices such as 'note picking', underhand attempts to replace a rival's notes with one's own notes. See Munn (op. cit.), p.23ff.

highlighted above on the timeline[19]. From these events emerge Smith's concerns about banks as do his arguments as to how those risks should be managed.

<u>Too little specie?</u>

Smith argued strongly that the economizing of gold and silver by the substitution of paper money that was nevertheless redeemable for specie at sight, was a beneficial development. The core advantage of such backed paper money was that it increased the real output of the economy. The gold and silver replaced by paper could be used to purchase abroad goods for consumption or could otherwise be invested in productive projects yielding future consumption. Smith was careful to emphasise that such 'liberated' specie had to be employed abroad since the paper money which was its replacement could not of itself boost domestic output; in more modern terminology, Smith was essentially assuming that domestic output was at its full employment level; adding paper circulation in addition to the gold and silver would simply be inflationary (see the Wealth of Nations, p326: The whole paper money of every kind that…)

But the risk was that the system may have been prone to operate with too little specie. Ironically, the efficiency of note clearing may have exacerbated that risk. Munn (op. cit.) notes that as the clearing system became established, the public enjoyed greater confidence in notes and sharp practices, such as note picking, ended. The result was that banks were operating on 'fractional specie reserves which were often as low as 2 or 3 per cent of demand liabilities'.[20] A feature of the system that underpinned business expansion could however be taken too far —possibly to a potentially dangerous level if convertibility was to be a founding principle of issue (Kerr, 1908). And this lack of convertibility may especially hit the least vulnerable. Moreover, banks may not always act in their, or society's, best interests.

In this vein, Smith noted: "…the business of the country is almost entirely caried out by means of the paper of those different banking companies…silver very seldom appears…and gold still seldomer." The extent of the use of paper issue in place of specie marks Scotland apart. Smith famously explains the efficiency of paper money but adds a caution:

> The gold and silver money which circulates in any country may very properly be compared to a highway, which, while it circulates and carries to market all the grass and corn of the country, produces itself not a single pile of either. The judicious operations of banking, by providing, if I may be allowed

[19] An additional issue is around the so-called optional clause. As we explain below, the optional clause allowed the issuer to delay redemption. Both the small notes issue and the optional clause were addressed in legislation passed 1765. The Law outlawed the optional clause. The Law also limited the minimum note in Scotland to be £1 and the guinea, although in England the £5 was the minimum denomination. Smith favoured a £5 minimum for Scotland.

[20] These numbers appear consistent with those reported in Cameron, op. cit., p. 88. Indeed, he reports rations somewhat lower than these based on analysis of banks' accounts (e.g., Glasgow Ship bank).

> so violent a metaphor, a sort of waggon-way through the air, enable the country to convert, as it were, a great part of its highways into good pastures and cornfields, and thereby to increase very considerably the annual produce of its land and labour. The commerce and industry of the country, however, it must be acknowledged, though they may be somewhat augmented, cannot be altogether so secure when they are thus, as it were, suspended upon the Daedalian wings of paper money as when they travel about upon the solid ground of gold and silver. Over and above the accidents to which they are exposed from the unskillfulness of the conductors of this paper money, they are liable to several others, from which no prudence or skill of those conductors can guard them. WN, p. 349

Smith was not naïve and was aware that fractionally backed paper could be dangerous. As Rockoff (2011) shows, Smith's worries grew as he developed his ideas on money and banking ahead of the final version of the Wealth of Nations. In earlier drafts of the above passage, Smith extols the virtues of paper money as a great efficiency, but the latter concern as to its fragility is absent. Rockoff (2011) argues, as indicated also by Checkland (1975a, b), that the failure of the Ayr Bank had the decisive impact on Smith's thinking. In the normal course of things, there were powerful correctives to counter over-issue by any single banking institution, of promissory notes or bills or via the Scotch system of "cash accounts" (WN, p324). First of all, notes in excess of that required by the locality in which the bank is operating will experience an increase in demand for money (that is, cash/specie) as its notes are returned more often (or later, sent for clearing in Edinburgh). It is costly to store and replenish the stock of specie and so banks will be aware of this cost. And scholars of the Scottish system have suggested that the Edinburgh note exchange was an effective mechanism to ensure this happened, although this was likely only fully in place after 1774 (Munn, 1981) once some of the lessons of Ayr bank had been learned[21].

However, Smith argued that experience taught that banks had not always understood this. And he names the Bank of England and the Scotch banks as examples of banks who have overissued. He also discusses at length, although not by name, the misdeeds and missteps of the Ayr Bank (WN, p341). And as discussed later, lending limited amounts and for short-term only against 'real' bills was part of the correct remedy for this tendency (along with a limit on low denomination notes and an end to the optional clause). And a like caution was to be observed with respect to cash accounts by requiring "frequent and regular repayments from all their customers…" (WN, p332). But bill market abuses could and did occur and these could be extensive sometimes, (WN, p339); indeed, so extensive that banks upon discovering the fraud may try to "withdraw gradually" in order to avoid making bankrupts of the projectors and the bank itself. The idea was to make life difficult for the discounters in the hope that they

[21] Although clearly Ayr Bank had been a prime mover in setting up the clearing system, it had still been able to overexpand.

would discount with other banks. The situation more widely could be made worse as a result of such behaviour.

In short, Smith believed paper money to be a material efficiency gain that was helping Scotland grow strongly and catch up on England. But he was equally clear that there were attendant risks. He understood the risks of fractional reserve banking and was aware that banks of all stripes had from time to time in the past overissued; sometimes that was due to unavoidable error and sometimes not.

Small notes mania in the 1760's

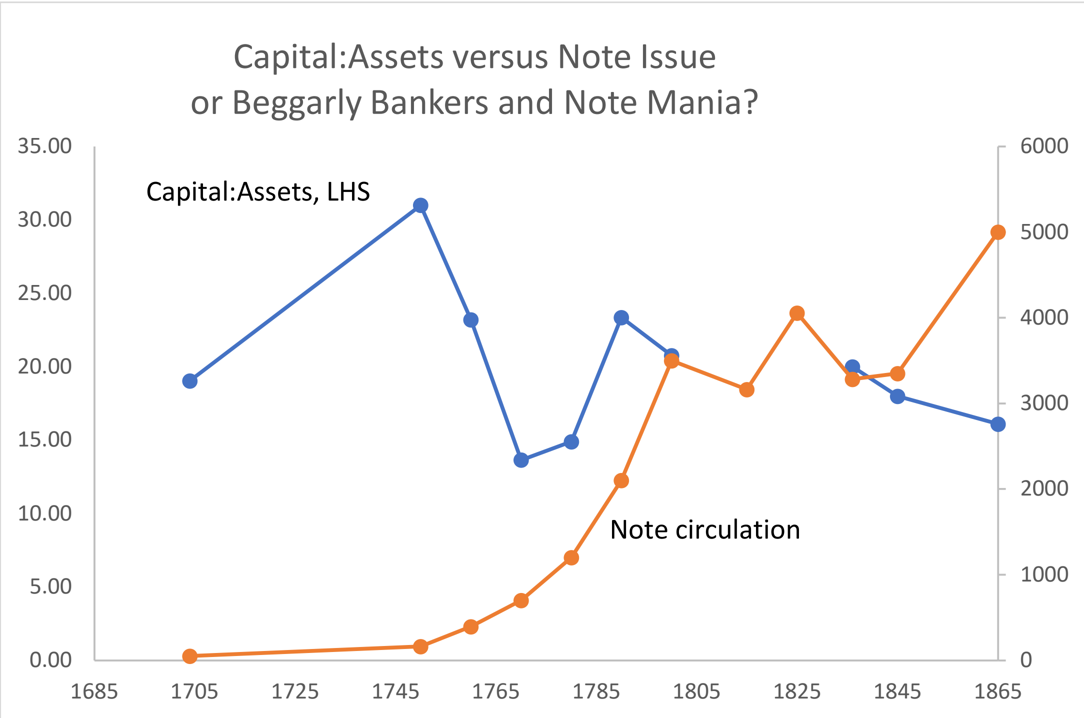
Smith was appointed as a Professor at Glasgow University in 1751 and he remained there until the beginning of I764. It is thought that the early drafts of the Wealth of Nations were began in the early 1760's. And it was during this time that Smith reported that something like a small notes mania emerged. The details surrounding this mania are not entirely clear (not least because there is little data on note issue). Moreover, the precise dating of the mania is somewhat contested as is the size of the problem. Subsequent writers, particularly those of the free banking school, have suggested that Smith, no doubt inadvertently, exaggerated the extent of the problem and indeed may have misunderstood the economics of the issue (Goodspeed, 2016; Selgin, 2018). Others, such as Kerr (1908, p844ff), Clapham (1944, vol. I p.160), Hamilton (1956) and Checkland (1975a, 104ff), acknowledge the contemporary perception that such a mania existed and report some salient facts. They do not contest that such a mania existed[22] and appear to concur that it was harmful to some degree. Checkland suggests that the mania covered the period from the 1750's to the 1760's and that in some measure it grew out of banking wars between the two public banks in Edinburgh and their attempts to gain a foothold, each at the expense of the other, in Glasgow near to the sites of many international traders.

Shortly after these two institutions set up what in effect were almost like satellite banks, Checkland (op. cit.) reports that there was a proliferation of small notes not just in Glasgow but throughout Scotland. These notes, he reports, were for even as little as one shilling. Checkland (op. cit.) accords with Smith that many impecunious ("beggarly" in Smith's language) bankers emerged. But he suggests it was not just 'banks' as that issued many small notes. There were no restrictions in Scotland at this time as to who could issue notes; if a debtor could persuade a creditor to accept a promissory note, of any value, then he or she was free to do so. Checkland (op. cit.) also notes that for various reasons there was a shortage of silver coin during this period. In that case, small notes could make a serviceable substitute for absent coin; that was especially so for wage payments. The two main public banks, at least initially in this period, did not supply small notes in sufficient measure to fulfil these needs. Many writers have noted that dubious notes also appeared all over the country not only emerging from the larger cities such as Glasgow, Edinburgh, Dundee, and Aberdeen—cities with a number of banks—but also somewhat smaller conurbations such as Perth, Dumfries, Dunkeld, Montrose, Falkirk, Kirkliston and Auchtermuchty (Checkland, op. cit.).

[22] Clapham (op. cit.), for example, refers to "that short and curious phase of Scottish banking history" that nevertheless established legal precedent in the statutory regulation of note issue, albeit as yet only in one part of Britain.

Perhaps these notes did not exhibit any great systemic risk. Nevertheless, so far as the banks were concerned there was a danger that mistrust might infect small notes more generally. Since those small notes were an important source of revenue for banks, especially the large public banks who wanted to dominate issue, concern amongst the bankers appears to have been quite widespread. It is also worth pointing out that the small note mania, such as it was, was an issue for Scotland and not England since in England small notes below £5 were illegal. In any case, at the start of the 1760's, and four years in to the Seven Years' War (1756–1763), note issue was increasing in part as the two public banks became more active. By 1762, Checkland reports that the small notes mania was reaching its peak.

The data below from Cameron (1967) provides a small amount of evidence perhaps consistent with some of Smith's concerns around beggarly bankers. It shows that the ratio of banks' capital to assets declined from a peak in 1750 through 1770 while notes issued increased over the same horizon, although deposits at that time were still relatively modest. This data only refer to partnerships that were clearly banks whilst the small notes mania was also reflective of a wider tendency to issue promissory notes and indeed notes of a more obviously fraudulent nature. As indicated, no data on 'non-bank' issue or fraudulent issue is available although much anecdotal discussion exists. Nevertheless, what data we have on the banks seem consistent with a wider potential concern that within the banking sector rising note issue was accompanied by a trend rise in bank gearing (the ratio of capital to assets).[23]



Data from Cameron, 1967.

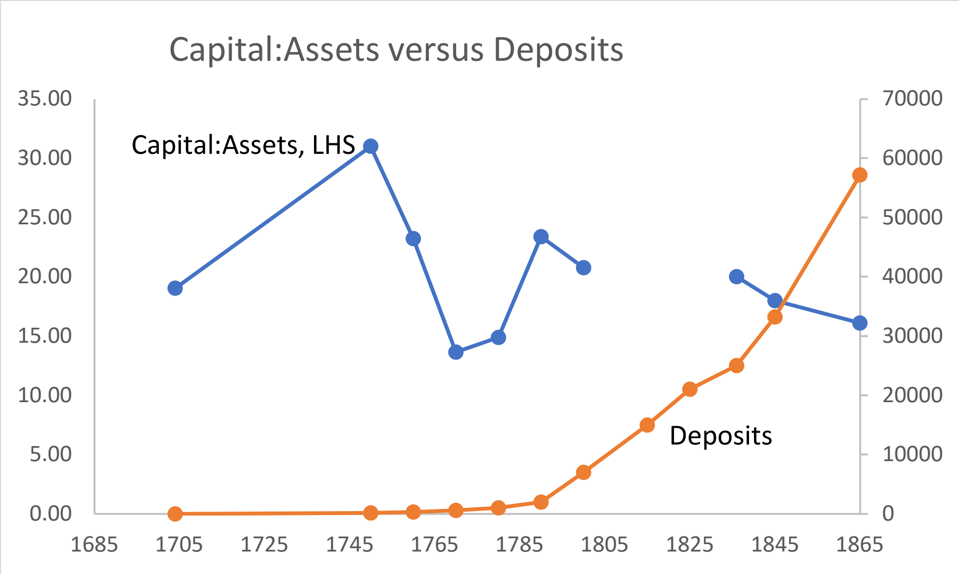
[23] Cameron (1967) does not describe in detail the source of the data used here and in the table above on page 5. It is a mixture, he says, of published data, data from some primary/archival investigation and data culled from Parliamentary papers.

Checkland seems to suggest that in fact the situation was quite tense. Scotland for a variety of reasons had lost much specie. And indeed Cameron (1967) using fragmentary pieces of data from some of the banks' contemporary accounts confirms that specie levels were very low (see footnote 20). Partly this seems to have been war-related: taxes had to be paid and the revenue did not return in the form of Government military expenditure as it had previously. The war was expensive more generally and the balance of payments also struggled on account of poor harvests. Speculative flows headed for London. All of this put pressure on the (already meagre) gold stock and circulating specie in Scotland. As did the usual wartime inflation. Rising prices tends to push up the value of silver and gold coin as bullion. As a result of all this, not much gold was to be had from England as the Bank of England was also short. Moreover, the small notes mania did not evaporate quickly according to Checkland as new banks emerged outside the main banking centres in Scotland even as the war came to an end in 1763.

Smith no doubt was well aware of these issues and the resulting tensions. And although Britain was not forced off gold—and neither did the Scotch banks suspend payments—it would have been apparent, one imagines, to Smith that such an outcome had been all too possible.

In summary, then, Smith viewed Scotland's banks' ability to economise on specie as a blessing, though a mixed one. The system in practice might tend to operate with altogether too little specie to weather much in the way of a shock. The obvious incentives under free banking to provide small notes, and the ready demand for them, might induce 'beggarly' bankers to enter the competitive fray. The victims of banks who run out of specie would be, Smith argued, those least able to bear it and the least financially sophisticated. It might also cause people to question paper money more generally. He alludes to such contagion as we show later. Moreover, Smith no doubt worried that banks which economised too much on cash reserves might not only inconvenience the everyday man and women; they could even, in extremis, imperil the system itself.

The small notes mania and competitive pressures on specie might afflict many banks at once and pose a risk to the system. However, single institutions could also pose a risk. The failure of Ayr Bank was a key event for Smith.



Data from Cameron, 1967.

Ayr Bank (Douglas, Heron, and Company) and the Crisis of 1772

As noted, the late 1760's was a difficult period in Scottish banking with the two main banks struggling to cope with the new banking firms that were being set up and themselves coming in for criticism for under-lending. Difficulties at those public banks and a non-cooperative approach to the clearing of the "challenger banks'" issues led to a perceived dearth of credit (Checkland, 1975a, pp121-124). That was despite the number of banks increasing fairly rapidly, from five in 1740 to 32 in 1769[24]. The Act of 1765 had also complicated things, outlawing notes of less than £1 and outlawing optional clauses on notes (White, 1984)[25]. For some, these measures were an egregious error the full import of which would be clear during the Suspension, some thirty years in the future (Selgin, 2018; see also Goodspeed, 2016, for a detailed analysis). For others, Smith included, these measures were needed to ensure the banks held enough specie in part to protect traders and poorer people from 'beggarly' banks. At any rate, into these stringent times was born the Ayr Bank (more properly known as Douglas, Heron and Company). It was founded in 1769. It was very strongly capitalised and its list of shareholders included many extremely wealthy landowners and prominent individuals of substantial means. It was immediately a major rival to the Edinburgh public banks. By the time of its collapse in 1772, the bank supplied 25% of the notes circulated by Scottish banks, 25% of total deposits, and 40% of total bank assets. The public banks, the Bank of Scotland and the Royal Bank of Scotland, accounted for only 21% of total assets (Rockoff, 2011; Checkland, 1975a).

Initially, the Ayr Bank's impact appears to have been beneficial. As noted earlier, the Ayr Bank accepted at par the notes of private banks so helping to end the banking wars. Hitherto, the two public banks accepted only each other's notes. The Ayr Bank's example forced them to change and ushered in the Scottish note circle (interbank note clearing) which became a key mechanism to avoid overissues as well as make notes acceptable in a wider geographic area

[24] These numbers from Checkland (1975a) appear very similar to those estimates in Cameron (1967). Cameron suggests that in 1770 there were 29 banks, of which 12 were issuing banks.
[25] The 'optional clause' is discussed presently but essentially allowed issuers to delay redeeming notes.

(since acceptors of notes could clear them all together in Edinburgh and not have to go various banks' provincial offices). The Ayr Bank expanded rapidly opening branch offices and lending vigorously. Its notes expanded rapidly. "It was not long before it was said that the Ayr Bank notes represented some two-thirds of those in the Circle. The two Edinburgh public banks, now confronted with a rival whose power and behaviour were unpredictable, further contracted their own lending, thus increasing the demands upon the Ayr Bank." Checkland, 1975. P.127.

The overexpansion of credit by the Ayr Bank led, it would seem, to much speculative activity; its borrowers were unable to invest in "real" projects in such short order while the Ayr Bank was keen to get its notes into circulation. Eventually, the speculative activity led to strains as notes were returned for redemption and the company kept afloat by borrowing heavily on bills. By the end of 1771, the two Edinburgh banks were suspicious of Ayr Bank notes. The London market too was suspicious of Scottish bills. Some of the directors of the Ayr Bank too were having misgivings about their loan book (Hamilton, 1956). In June of the following year Neale, James, Fordyce and Downe, a major London-based bank collapsed. This bank had extensive Scottish connections especially to the Ayr Bank. Scottish paper in London fell in value and many banks with Scottish connections failed. And indeed, there were international dimensions to the failure of Neale and Co as Fordyce had been speculating in British East India stock. Many Scottish banks failed—notably most of the Edinburgh private banks—and the Ayr Bank suspended payments on 25 June 1772. The bank reopened in September to redeem notes but finally closed in August 1773. The Ayr Bank had tried but failed to negotiate a loan from the Bank of England[26] and the Edinburgh public banks, although the Bank of England did extend a loan to an important export house. Hamilton (1956) cites a writer to the London Chronicle in 1772 who noted that the Ayr Bank

> …will find it impossible to carry on [its] business as a Banking Company independent of the Bank of England, that being the great source of the British funds and credit, without whose countenance and occasional aid, no banker nor merchant even in London, can do business with safety and profit.

What followed the collapse of the Ayr Bank is interesting. Checkland (1975a) notes:

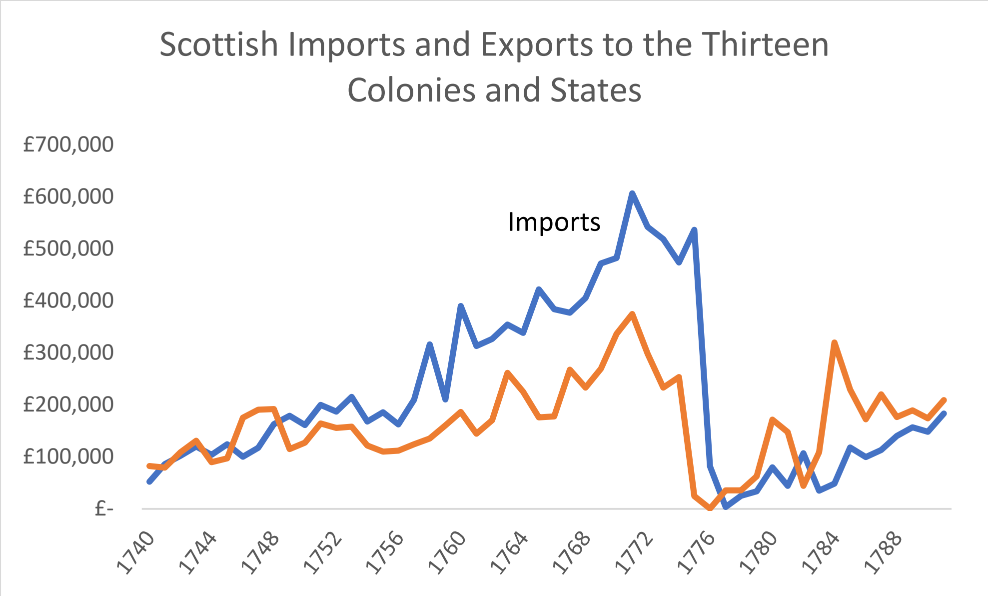
> The most important task was to ensure that the notes of the Ayr Banking Company were redeemed, so that the acceptability of notes in general, in Scotland, might not be irreparably damaged. The Bank of Scotland and the Royal Bank decided to accept Ayr Bank notes to the extent of the good security they could provide. This meant relying ultimately upon the revenues of the landed estates of the principal proprietors. Large though the Company's obligations were, the security, because of unlimited liability, was good. It was estimated that the capital value of the land and estates of the proprietors was some six to seven millions sterling. With the principal

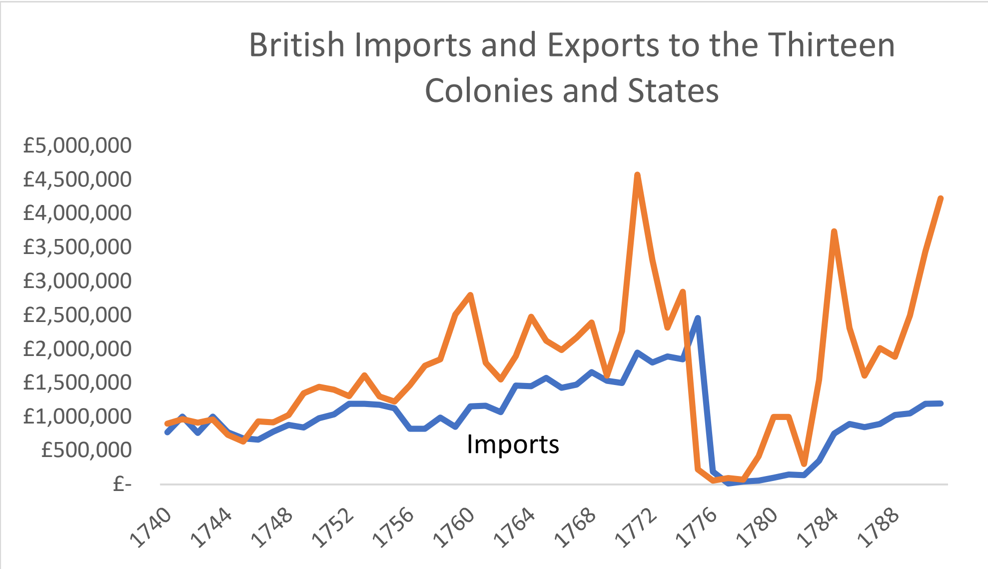
[26] The Bank of England's terms were rejected by the Ayr Bank directors.

partners securely pledged, the two banks ran no risk. But there was no hope of resuscitating Douglas, Heron and Co." p131-132

It seems that the Ayr bank was not simply a 'bubble,' but itself came on the back of a long upswing in trade and growth (Hamilton, 1956). Various key sectors had been growing strongly, notably linen, and there was, according to Hamilton (op. cit.) and Checkland (1975a), a sense that the banks, especially the Edinburgh banks, were not lending sufficiently to support commerce. Hamilton (op. cit.) seems to argue that Ayr Bank's misdemeanours notwithstanding the long boom was due to come to an end. Other writers have been more censorious of the role the Ayr bank played in the slowdown of the economy and a post-collapse report into the bank recites a litany of failings (Rockoff, 2011; Munn, 1981). The lack of data makes it difficult to judge whether Ayr Bank was more of a cause than effect of a wider slowdown. However, the Scottish economy appears to have had been growing for some time ahead of the founding of Ayr Bank. Some data constructed by Price (1975)[27] may provide a modicum of evidence concerning the 'boom' and its ending. The charts below reproduce Price's (op. cit.) data. The data construct individual and aggregate series for imports and exports between both Britain and Scotland on the one hand, and the thirteen colonies and state of America, on the other. Scottish trade with the colonies was overwhelmingly (c. 94% for imports, c. 90% for exports) concentrated on Virginia and Maryland. A large part of these exports was undoubtedly linen (see Durie (1973) and the data therein presented) and much of the imports were tobacco (Devine, 1975; Olson, 1983). Price (op. cit.) reports that by 1770-1774 Scottish imports from the colonies accounted for about 30% of the total British import value (exports were about 10% of the British total). The major features of the data are apparent. First, there is a strong upward trend from earlier in the eighteenth century. That seems to be consistent with Hamilton's (op. cit.) argument that the Scottish economy had been growing strongly for some time. Second, there is clear evidence of the impact of the onset of hostilities between America and Britain; the major contraction across all series begins in 1775. However, ahead of that contraction, there is peak in Scottish imports and exports in 1771. That is the year ahead of the Ayr Bank's demise, although concerns around the bank had been rising ahead of that, as described earlier. Indeed, imports contracted through 1774 before rising slightly in 1775. However, it is interesting to observe that British imports more or less continued growing through 1771 (when Scots imports had peaked) until 1775 and the contraction caused by the outbreak of hostilities. Thus, the data as they stand corroborate Hamilton's view that a long period of robust growth looks to have been underway. However, the data also may be consistent with the Ayr Bank possibly having caused a sharper slowdown in imports (and perhaps exports) than might otherwise have been the case.

[27] As Price (op. cit.) explains, the data were part of an effort by the British Government to assess economic affairs as tensions with America escalated. There are a number of possible issues with this data which Price (op. cit.) discusses. Probably most significant are those related to the price data. In any event, Price's data seem strikingly close to some data cited by Hamilton (1956) in footnote 4, page 409.





By August 1775 Hamilton (1965) notes that only 112 of the 226 partners of the bank remained solvent but no creditor lost out. The collapse of so big a bank with apparently limited consequences is striking some argue. White (1984) notes that this was one of a number of crises that would confront the Scottish banks up through 1844. He argues that they weathered these surprisingly well and with very few failures, and better than banks in England when the crises were experienced in common. White (1984) and many others have put this down to the existence of 'free banking' in Scotland compared with England. That is not to say that the crisis had no 'real' effects or that they were necessarily short lived. Hamilton's (op. cit.) and Price's (op. cit.) data suggest that Scotland's international trade may have suffered for several years. And certainly, some at the time were not quite so sanguine. Famously in late June of 1772 Hume wrote to Smith:

> We are here in a very melancholy Situation: Continual Bankruptcies, universal Loss of Credit, and endless Suspicions . . . even the Bank of England

> is not entirely free from Suspicion. Those of Newcastle, Norwich, and Bristol are said to be stopp'd: The Thistle Bank has been reported to be in the same Condition: The Carron Company [an iron works, and pioneer of the industrial revolution] is reeling, which is one of the greatest Calamities of the whole; as they gave Employment to near 10,000 people. (Smith, Correspondence 1987 [1740-1790], p. 131, quoted in Rockoff, 2009)

Hume went on to ask whether these events had affected Smith's theory that he was developing. Some scholars argue that indeed it did affect Smith. Checkland (1975b) mentions both the small notes mania and the Ayr Bank collapse. Hugh Rockoff has argued that the crisis of 1772 was decisive in convincing Smith that banking should be as free as possible but not completely so (e.g., Rockoff, 2009, 2011). As was noted previously, early drafts of the book (The Wealth of Nations and the lectures upon which some of it is based) indicate, Rockoff argues, a change in his views on bank risk. At any rate, in the final published version of the Wealth of Nations Smith lists all the ways the Ayr Bank (although he never names the bank explicitly) went wrong. First, the bank lent long term. The borrowers were raising capital for long term investments, whereas for Smith banks should not be in that business; that was one part of the underlying rationale of the real bills doctrine. Second, the bank also engaged in what would now be called 'connected lending.' In other words, the bank lent to its shareholders in effect reducing the capital in the Bank as well as calling into question the probity of the bank. Third, the bank had expanded too quickly. In particular, the bank had made unwise acquisitions of other banks. Fourth, underlying these problems was a dash for growth. Ayr Bank was massively capitalised (connected lending notwithstanding) and there was much demand at the time for its notes, but the supply of issue outstripped the supply of legitimate investments. The bank found its notes returning at such a rate as to put a strain on its specie reserve. At the note clearing in Edinburgh, specie as well as bills on London could be used; increasingly the Ayr Bank was redrawing drafts as they came due. As a result, its short-term debt was piling up.

It would seem that the Scottish experience of banking, both its positives and negatives, influenced greatly Smith's thoughts on how banking ought to be conducted.

What did Smith believe comprised a 'good' banking system?

Most of Smith's views on banking are contained in Book II, Chapter II of the Wealth of Nations entitled "Of Money"[28]. The chapter is a wide-ranging discussion on banking matters covering historical and contemporary developments (nationally and internationally) as well as theoretical and policy developments. The Chapter covers such topics as how Smith sees money somewhat as a factor of production. However, physical currency is costly and can be economised substantially by the use of paper money boosting appreciably the economy's performance. This is in fact is what Smith thought had happened in Scotland in the 25 or so

[28] Its full title is: Of Money Considered as a Particular Branch of the General Stock of the Society, or the Expence of Maintaining the National Capital.

years up to the publication of the Wealth of Nation with the widespread proliferation of banks in Scotland (WN p. 322) and their ability to issue paper. However, having established what he saw as the sizeable benefits of paper money and competitive banks of issue, he turned to the risks with the system. He argues that while the total stock of paper money cannot exceed the gold and silver that would have served had the paper money not existed, and that individual banks face clear pressures not to overissue, some individual banks (and even the Bank of England) had, despite their own best interests, at times overissued. Smith observes that issues are typically put into circulation via discounting bills or via cash accounts. Banks, he argued, should only lend against real bills; that is bank loans should be collateralised and self-liquidating (in other words, the loan/debt is extinguished upon final sale of goods). The value of bills should be in the amount that the merchants and traders would otherwise have had to hold themselves; a core business of banks, then, is to facilitate the economising of cash balances. Moreover, such balances should be "small." This should help guard against overissue which had, at its root, "overtrading," the result, in turn of a combination of "bold projectors" and the "imprudence and inattention" of the banks.

In the round, the Chapter "Of Money" (along with remarks on the desirability on joint stock banking in Book V to which we turn later) may be regarded as an analysis of the benefits and risks associated with banking and paper money, and what actions are required to mitigate these risks. Smith concludes "Of Money" by saying that the only restrictions required on banks are that notes for small amounts should be prohibited and that all notes should be repayable on demand. Subject to these, the bankers' trade "may, with safety to the public, be rendered in all other respects perfectly free." (p. 358). However, in is a sense that conclusion seems to sit oddly with the facts and with Smith's own analysis. It seems to ignore some of the issues just noted—the potential for large systemic banks to dominate the system and how the rules of the note exchange ought to be formulated and enforced. It also ignores the fact that in reality Smith was actually endorsing quite a comprehensive set internal and external rules, not just the two he noted. In effect, Smith recommends a series of five *internal rules* that banks should adhere to, as well as five *external rules* that need to be imposed of banks. Together these would deliver, his argument seems to be, an efficient banking system and, crucially, a stable monetary system.

<u>Smith's Internal Rules</u>

The main internal rules or guidelines to which banks should adhere are: First, that banks should lend against real bills. Second, banks should only lend short-term. What exactly that might mean may be open to interpretation. But for inland bills, Smith mentions the preponderance of dated bills at two and three months. He does not appear to discuss anywhere the effective increase in usance necessary to cover international trade, something of much importance to the Scottish economy.[29] Third, there should be no lending for fixed investment purposes or property loans. Fourth, banks should lend only "small" amounts. Again, what exactly that means is open to debate. But Smith states that banks should never lend more than what the trader "might otherwise be obliged to keep by him unemployed." Finally, cash credits (introduced by The Royal Bank of Scotland in 1728) should be governed by the same rules as discounting bills: there should be no long-term finance (rolling over of balances endlessly); accounts should regularly be settled; and all lending should be against high quality security/guarantees.

It seems that "real bills" were central to these internal rules and that Smith thought of these as akin to a voluntary code of conduct and as best practice. Importantly, however, Smith does not recommend any direct rules from the legislature to enforce them, much less is there any indication that these rules should be implemented by any kind of regulator. The real bills doctrine is perhaps the most controversial—perhaps most misinterpreted—of Smith's banking doctrines. For some it came to be seen as a monetary doctrine rather than a banking doctrine. It is worth looking at what Smith actually wrote.

Real Bills Doctrine

Although The Royal Bank of Scotland invented the cash-credit (an overdraft backed by two guarantors and perhaps security) and this seemed to be adopted more widely as the eighteenth century wore on, it was note issue and bill discounting that formed the core of many earlier banks' business. Smith had clear views on both notes and lending.

> When a bank discounts to a merchant a real bill of exchange drawn by a real creditor upon a real debtor, and which, as soon as it becomes due, is really paid by that debtor; it only advances to him a part of the value which he would otherwise be obliged to keep by him unemployed, and in ready money for answering occasional demands. The payment of the bill, when it becomes due, replaces to the bank the value of what it had advanced, together with interest. The coffers of the bank, so far as its dealings are confined to such customers, resemble a water pond, from which, though a stream is continually running out, yet another is continually running in, fully equal to that which runs out; so that, without any further care or attention, the pond keeps always equally, or very nearly equally full. (WN, P331.)

And Smith is equally clear on what the bank should be lending for. It should not be for what we might term fixed capital, long term investment, property, and mortgages. These

[29] It may be that the consignment structure of trade and the funding of different legs of the export/import could be effected by a sequence of, say, three month (mostly internal) bills. But that raises another set of issues alluded to by Thornton (1802). See below.

restrictions are the clear corollary of the real bills doctrine and clearly indicate that Smith did not think of real bills as a theory of the aggregate money supply but as bank "regulation").

> …the capital which the undertaker of an iron forge, for example, employs in erecting his forge and smelting-house, his workhouses and warehouses, the dwelling-houses of his workmen, &c.; of the capital which the undertaker of a mine employs in sinking his shafts, in erecting engines for drawing out the water, in making roads and waggon-ways, &c.; of the capital which the person who undertakes to improve land employs in clearing, draining, enclosing, manuring, and ploughing waste and uncultivated fields, in building farm-houses, with all their necessary appendages of stables, granaries, &c. The returns of the fixed capital are in almost all cases much slower than those of the circulating capital; and such expences, even when laid out with the greatest prudence and judgment, very seldom return to the undertaker till after a period of many years, a period by far too distant to suit the conveniency of a bank. (WN, p334.)

It is clear enough from the language used by Smith why it came to be called the 'Real Bills' doctrine. He argues that lending should be short term and against security. The amount so lent, should be 'small' and should replace specie balances that the borrower would otherwise be required to store—unproductively—had the bank not existed. From these guidelines Smith may not even be endorsing loans for what we might now call 'working capital.' He is certainly arguing that banks should not be funding long term loans or fixed capital or property as the quote above makes abundantly clear. And he applied the same stricture to the cash credit, an innovation he praised (as did his friend David Hume, although Hume worried it could lead to overissue).

Smith follows this discussion with some observations on bill market abuses, no doubt motivated by Ayr bank's collapse. Such abuses would also play a significant role in the great financial crises of the nineteenth century and would be part of the focus of various Parliamentary inquiries into these "derangements of credit." Smith therefore was clearly not naïve to the risk that bills could be abused and that they could easily become accommodation or finance bills[30]. But he did not go so far as to suggest any external or legal prohibitions in this area. But the possibility of such abuses, suggests he did believe that banks could be complicit in over issue.

[30] Although here views would evolve, and some (e.g., Thornton, op. cit.) would defend such bills as legitimate business.

Smith's real bills doctrine was criticised by a leading protagonist in the 'bullionist controversy.' Henry Thornton, a preeminent monetary and banking scholar, argued that Smith over-simplified the nature of the relation between bill finance and issue, implying he was much less sanguine than was Smith on the restriction to real bills in constraining credit/issue.[31] An argument based on real bills was used by the Bank of England to defend itself against the charge that inflation during the Suspension—in effect the Bank argued that it had not overissued paper currency once it had been freed from the legal requirement to redeem notes in specie. The Bank argued that, on account of lending only against real bills, they could as a matter of logic not overissue. Any inflation, or gold premium observed in the market, must have been due to other factors.[32] However, the real issue appears to centre on where the Real Bills Doctrine is applied. Was it a theory of the *aggregate* money supply? Or did it apply to banks individually (but somehow not in the aggregate). It seems clear, as indicated above, that Smith was referring to banks individually. It was essentially a normative statement as to how banks *should* lend. Historically, however, it became identified with fallacious reasoning about the aggregate money supply, or even inflation. Moreover, Smith's version of the real bills doctrine applied to small, competitive banks of issue who were legally required to redeem in specie on demand. His discussion clearly did not apply to a limited liability, financial giant which Smith recognised as an arm of the state and whom he had explicitly asserted as having overissued in the past and who was now (during the Suspension) legally exempt from paying out on its notes. Nevertheless, the Bank had employed a real bills defence of its actions, and would not be the only central bank to do so. In defending its conduct of monetary policy during the Great Depression in the 1930's, the US Fed also invoked a real bills defence.[33]

However, even as a theory of individual bank behaviour Thornton (1802) argued that Smith's discussion of issue against real bills was incomplete and that a multiple of notes might be issued against the same transaction depending on the 'supply chain' of sale and repurchase before the goods were indeed finally sold at market. This qualification or refinement of the real bills doctrine may have traction even on Smith's assumptions.

In short, over issue could occur for three reasons, two of which appear to be clearly acknowledged by Smith. First, banks might not act in their own long-run self-interest; they might act irrationally and imperil their ability to redeem in specie. Second, bill market abuses could and did occur and Smith knew it. Sometimes the banks were unaware of the abuse, sometimes not. But abuses happened and overissue could result. And third, Thornton

[31] See Thornton (1802). He makes two key criticisms of the real bills argument of Smith. The first criticisms is discussed beginning on p. 85. Moreover, he goes on (p. 87) to argue that accommodation bills may not be simply classified as "fictitious". The second is on p. 253; he suggests that the real bills doctrine does not provide a secure basis for supposing issue will be held in control. Mints (1945) would pick up on these points.
[32] See Viner (1937), pages 136ff. Newby (2012) offers a new perspective on British monetary management around the Suspension period.
[33] Lloyd Mints coined the phrase "Real Bills Doctrine" noting that "[Adam Smith] is the first of a long succession of writers, extending to the present day, who have integrated into a systematic exposition certain ideas in regard to control of the quantity of bank credit, the kinds of assets banks should hold, the provision of elasticity in the currency by means of bank credit, and, finally, the provision for liquidity. See Humphrey and Timberlake (2019) for an analysis of the Fed's "use" of the doctrine to defend its actions during the Great Depression..

demonstrated that a number of bills—and a corresponding larger issue of paper money—might be erected on a single real transaction. Smith argued that competition would go a long way to alleviate some of the risks. But he argued for further banking regulations and so was unconvinced that the market could be entirely trusted. The debate on whether competitive banks of issue were conducive to monetary stability and external adjustment became a central issue between the Currency and the Banking Schools.[34]

Smith's External Rules

What we label as Smith's five external rules are: First, banks should be constrained as to the maximum interest rate they can charge on their loans (usury). Second, bank owners/partners should face unlimited liability[35]. Third, banking (i.e., banks of issue) should be competitive—there should be with free entry. Fourth, all bank notes should be convertible on demand. Fifth, the law should specify the minimum denomination of notes. The first three 'rules' were part of the legal landscape in Scotland, but it is worth emphasising their significance since, as subsequent developments would have it, they were not immutable and they did play a significant role on Smith's banking 'model.' The final two rules were introduced by the 1765 Act and it was these "two rules" that Smith emphasised at the conclusion of Book II, Chapter II needed to be applied to an otherwise free banking.

Smith recommended a £5 minimum, although the 1765 Act settled on £1. His assessment of the small notes mania and the collapse of the Ayr Bank—especially the latter—seem to have convinced him that these restrictions were required (even though they obviously did not avert that failure). However, there are a few things worth observing. The rules that Smith endorsed and which were introduced into Scotland in 1765, were not the only external restrictions in place on banks. The first was the Usury Law, limiting at that time the maximum interest rate to 5%. And the second was that, along with free-entry, all non-chartered banks, indeed all firms (partnerships), carried unlimited liability. That was probably more significant in minimising the fallout from Ayr Bank, as we suggested above. Smith took unlimited liability for granted but more controversially supported the restrictions on banks' maximum lending rate.

Smith's overriding concern was for the stability of the currency. And it was this that drove him to support the usury laws as we now explain.

---

[34] That debate around whether private banks of issue were dangerous to monetary stability or not is covered in detail in Wood (1939).

[35] That unlimited liability was an 'external rule' can be seen from other times and places that experimented with free banking and where prior legislative norm was not pre-imposed on corporate structure. In those instances, unlimited liability did not invariably emerge as the 'equilibrium choice' of contracting parties. See Friedman and Schwartz (1983).

Usury Restrictions

In England and Scotland, as elsewhere in Europe, there had been long-standing prohibitions on charging any interest on lending. After the Reformation, Calvin permitted interest on commercial loans, according to Munro (2011), but not for loan to the needy. In England, Henry VIII's Parliament permitted interest of up to 10% (on all loans); any higher rates were then defined as constituting usury. This was revoked in 1552 following a religious backlash but revived under Elizabeth I, in 1571. The maximum rate was reduced in stages to 8% in 1624, to 6% in 1651 and to 5% in 1714, the effective rate for all of Adam Smith's life. That remained in place until 1833 when the Bank of England was allowed to charge a higher rate, although with some restrictions initially (See King, 1935). That relaxation was in part driven by the need to control credit expansion, a policy concern that was probably absent in earlier centuries. The abolition of the usury laws was finally completed in 1854.

Smith notes:

> The legal rate, it is to be observed, though it ought to be above, ought not to be much above the lowest market rate. If the legal rate of interest in Great Britain, for example, was fixed so high as eight or ten per cent, the greater part of the money which was to be lent, would be lent to prodigals and projectors, who alone would be willing to give this high interest. Sober people, who will give for the use of money no more than part of what they are likely to make by the use of it, would not venture into the competition. A great part of the capital of the country would thus be kept out of the hands which were most likely to make a profitable and advantageous use of it and thrown into those which were most likely to waste and destroy it. Where the legal rate of interest, on the contrary, is fixed but a very little above the lowest market rate, sober people are universally preferred as borrowers, to prodigals and projectors. (WN II.iv15: 357)

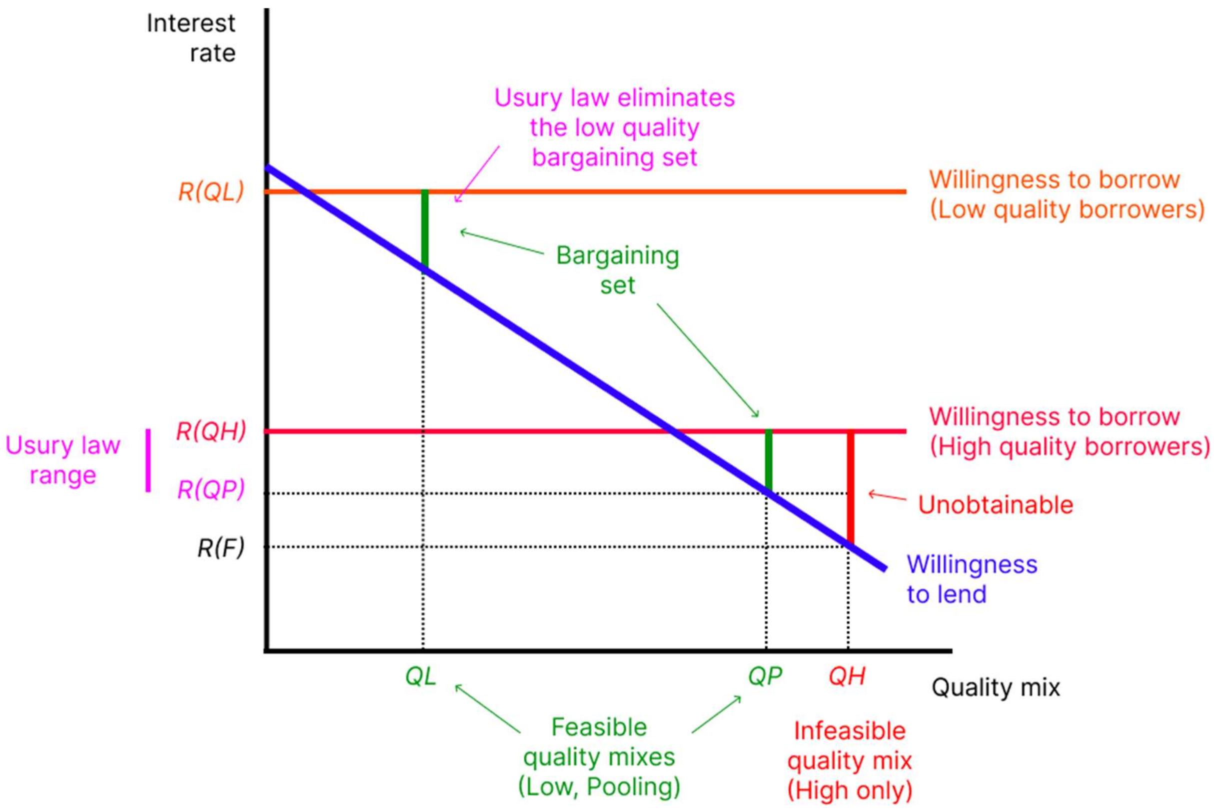
We can analyse the effects of Smith's support for the usury law using a simple 'Market for Lemons' model. Consider a bank who is looking to lend to a borrower and cannot immediately differentiate between high quality borrowers (sobers) and low quality borrowers (prodigals). Prodigals are willing to borrow at high interest rates (up to *R(QL)*) but sobers are only willing to borrow at interest rates up to *R(QH)* < *R(QL)*. The lender's required interest rate depends on the probability that a borrower is a prodigal, demanding a higher interest rate if all the borrowers are prodigals. The lender is more willing to lend at a lower interest rate if the probability of the borrower being a prodigal is low.

As we have drawn the diagram, the bargaining set, where borrowers are willing to borrow, and lenders are willing to lend, has two spaces (in green). In the low interest space, there is a mix of sober and prodigal borrowers. In the high interest space, there are only prodigals. There is an unobtainable, but ideal space where lenders only lend to sober borrowers (in red). Unfortunately, lenders have no way of excluding prodigals from asking for loans. The famous Akerlof market breakdown result can be found by shifting the willingness to lend schedule upwards, until it still crosses the red space, but no longer crosses the green spaces.

A usury law interest rate between *R(QP)* and *R(QH)*, as Smith supported, removes the high interest space from the bargaining set, leaving only the low interest space, where loans are extended to a mix of sober and prodigal borrowers.

Smith's final line in the quote above states "sober people are universally preferred as borrowers, to prodigals and projectors" but the usury law does not (on its own) eliminate prodigals as borrowers, rather it generates a pooling equilibrium where prodigals and sober borrowers alike are extended credit. If sober borrowers are sufficient in number, then this will be a good outcome.

The best (or constrained efficient) outcome would be a separating contract, that somehow allows the lender to distinguish between sober and prodigal borrowers, charging low interest rates to sober borrowers and high interest rates to prodigals. Smith's usury law on its own does not generate that outcome, but Smith's other proposals may do so. Specifically, Smith argued that credit should be short term, and any overdrafts frequently cleared, as explained above. There is theoretical support for Smith's frequent clearing proposal in helping to identify borrower types; borrowers expecting brighter future prospects are more likely to select into short term borrowing, subjecting themselves to the risk of future refinancing, than borrowers expecting future challenges (see Flannery, 1986; Diamond, 1991). In our view, the two proposals should be treated as a single proposal with two complementary branches, with the frequent clearing helping to support the revelation of borrower types, ultimately supporting an outcome that is close to the separating equilibrium---the area marked *unobtainable* in the diagram.



In sum, Smith's support for usury is a sophisticated analysis of the problem of adverse selection in credit markets. It also bears some similarity with modern US policymaking, particularly in health insurance and in peril crop insurance markets, albeit these are markets where policymakers explicitly want to achieve a pooling equilibrium, which is not the intended outcome for Smith.

For Smith, the usury law was primarily a way to stop banks becoming excessively risky, although he also appeared to see some merit in the traditional case for usury limits in supressing extortive practices. As the model sketched suggests, there are circumstances in which Smith's intuition along this dimension is valid even though it is a blunt tool and is Pareto-dominated may by other policies (or combinations of policies).

As time wore on, there was increasing scepticism around the role of usury laws. However, this scepticism revolved not around 'prodigals and projectors' but around the control of money and credit at the aggregate level, and in particular in response to external drains. Hayek (1939) notes comments by Tooke[36] describing the difficulty the Bank had in 1795 retaining gold in Britain faced with Government demands for war finance, a French return to the gold standard and continued demand at home for loans. The laws on usury resulted in the bank rationing credit, with deleterious effects on the London money market[37]. The dismantling of the usury laws begin systematically in 1833 as the Bank struggles with emerging policy pressures; see King (1935).

Optional clause

In the early period of Scottish banking, the optional clause was to limit malicious competition between banks where banks might accumulate each other's notes and present all at once for redemption. Checkland (1975b) also reports that banks could threaten to employ the clause if their own borrowers were overly keen on getting their hands on gold. Smith, however, had supported the outlawing of these clauses not, it would seem, to protect banks but to protect customers of banks, to restore the equilibrium exchange rate between Scotland and England, and to retain specie in circulation. Smith also referenced critically in Book II, Chapter II, practices by governmental authorities in the American colonies who issued their own currency, declared it legal tender but made the notes redeemable only several years after issue. Smith does not discuss at length his support for the prohibition of the optional clause, but overissue/lack of specie seem to be behind much of his concern; the issue of paper money could go too far and threaten convertibility. Although banks had every incentive not to overissue, nevertheless it could happen, Smith believed, and did happen:

[36] See Hayek's introduction to Thornton (1802), p. 39ff.

[37] Thornton (1802) argues later (p. 254ff) in his book that usury ceilings cause borrowers to demand too much in the way of credit. But again the "microeconomic" distortion seems to be a second-order concern to the macro/monetary dimension.

> Had every particular banking company always understood and attended to its own particular interest, the circulation never could have been overstocked with paper money. But every particular banking company has not always understood or attended to its own particular interest, and the circulation has frequently been overstocked with paper money.
>
> By issuing too great a quantity of paper, of which the excess was continually returning, in order to be exchanged for gold and silver, the bank of England was for many years together obliged to coin gold….
>
> Over and above the accidents to which they are exposed from the unskillfulness of the conductors of this paper money, they are liable to several others, from which no prudence or skill of those conductors can guard them.

<u>Small Notes Restriction</u>

Along with the optional clause, Smith wanted a prohibition on small notes—those under £5. In the event, as observed earlier, the £1 note became the lower limit. Smith saw that a minimum note size would leave paper money in the hands of those capable of monitoring the banks. And it would leave the smaller trader and merchant and labourer in possession of hard cash. In this way, the latter would be more protected from any banking bumps on the road. Were smaller notes made available, as occurred for example during the mania of the early 1760's, there was a real risk of the banking system de-monetising to a dangerous degree. Indeed, in extremis, with too many small notes in circulation, a collapse in confidence could imperil convertibility.

Smith thus concluded:

> To restrain private people, it may be said, from receiving in payment the promissory notes of a banker, for any sum whether great or small, when they themselves are willing to receive them, or to restrain a banker from issuing such notes, when all his neighbours are willing to accept of them, is a manifest violation of that natural liberty which it is the proper business of law not to infringe, but to support. Such regulations may, no doubt, be considered as in some respects a violation of natural liberty. But those exertions of the natural liberty of a few individuals, which might endanger the security of the whole society, are, and ought to be, restrained by the laws of all governments, of the most free as well as of the most despotical. The obligation of building party walls, in order to prevent the communication of fire, is a violation of natural liberty exactly of the same kind with the regulations of the banking trade which are here proposed.

The party wall analogy and its clear allusion to "contagion" is reminiscent of the Friedman and Schwartz (1963) narrative of the waves of bank failures during the Great Contraction in the US between 1929 and 1933. And Smith's nervousness around this fragility is reinforced by another of his famous metaphors cited above concerning the Daedalian wings.

In short, then, Smith's external rules may have been necessary to underpin "free banking" in Scotland. Friedman and Schwartz (1983) summarise various free banking episodes in the US between 1791 and 1836 and conclude that: "Various degrees of laissez-faire prevailed in the several states, but nowhere did it lead to unlimited liability, freely interconvertible bank notes, security of both note holders and depositors from loss, and the other favorable characteristics of the Scottish banking system." (p.50)

What was the Point of the Internal and External Rules?

Adam Smith thus recommended a banking sector that followed certain internal rules and was subject to some external rules. The five internal rules were guidelines on how banks should manage their business. To recall, they were: First, banks should lend against real bills. Second, banks should lend short-term. Third, banks should not finance fixed investment or property lending. Fourth, banks should lend "small" amounts. Fifth, cash credits should be governed by the same rules as discounting bills: There should be no long-term finance, accounts should regularly be repaid, and all lending should be against good security/guarantees. The five external rules were: First, banks should be limited as to the maximum interest rate they could charge (usury). Second, banks should have unlimited liability. Third, banking (i.e., banks of issue) should be competitive with free entry. Fourth, all notes should be convertible on demand. Fifth, there should be a limit of notes of low denomination. These final two rules were introduced by the 1765 Act, and it was these "two rules" that Smith concluded needed to be applied to an otherwise free banking system (although as we note, there are other "external rules" that he took for granted, especially the first and second just mentioned).

For Smith, following these internal and external rules would deliver monetary integrity with what he regarded as the minimum of intervention. In particular Smith's rules would appear to result in no need for a central bank of issue or central reserve of specie, something on which Bagehot (1873) would later opine. Smith argued that the free entry of banks and their proliferation would be a desirable development. Whilst individual banks might fail in the normal course of things, none individually would be dominant and so could easily be let go without affecting the wider economy. This would result in widespread confidence in banking firms and their issues, and there would no need regulate banks (save for the two restrictions above on small notes and redeemability). Finally, banks could safely be allowed joint stock status, although still with unlimited liability. We explain this last point presently. And as a result of all these rules, note issue could never be excessive *for long*. That insight is what would later be labelled by Viner (1937) as *The* Banking Principle.

Smith and Joint Stock Banking

As described earlier, the Scotch banking system had many features of joint stock firms (excluding limited liability). Smith was typically sceptical of the joint stock firm as he regarded it as often less efficient than a partnership and able to survive only due to state-sanctioned privileges such as monopoly rights and limited liability. Smith argued that joint stock companies, even when not protected by the State, typically performed poorly when faced with competition. That is to say, they had a poor survival rate (see the discussion in Anderson and Tollison, 1982).

Smith says:

> The only trades which it seems possible for a joint stock company to carry on successfully, without an exclusive privilege, are those, of which all the operations are capable of being reduced to what is called a Routine, or to such a uniformity of method as admits of little or no variation. Of this kind is, first, the banking trade; secondly, the trade of insurance from fire, and from sea risk and capture in time of war; thirdly, the trade of making and maintaining a navigable cut or canal; and, fourthly, the similar trade of bringing water for the supply of a great city. [p815][38]

Anderson and Tollison (1982) are intrigued by Smith's endorsement of joint stock banking:

> Basically, Smith found that although the joint-stock form had a poor survivorship record, there were a limited range of activities where it seemed to compete effectively with alternative forms of organization. This observed range of activities was small but quite diverse, appearing superficially to be a somewhat peculiar collection. The common characteristic Smith noted was that all of these lines of business-banking, insurance, and waterworks-seemed to involve problems conducive to routine decision making…. In effect, it was easy to see if a manager was shirking in these employments by answering the question, is he following

[38] Smith, earlier in the discussion, argues that a key difference between the co-partnery structure and the joint stock company is the unlimited liability entailed in the co-partnership. We assume that he is envisaging joint-stock unlimited liability banks although he does not appear to be completely explicit on this point.

> the rule? ...Smith's analysis is interesting, but it is hard to sustain a characterization of banking and insurance as routine undertakings…
>
> While the daily operations of a bank may superficially appear to be repetitive, a bank's success depends upon the critical acumen of its managers, especially with respect to large loans. (p. 1244/5)

The puzzle seems to be that Smith is recommending a role for joint stock banks on the basis that they are simple enterprises with, in modern parlance, little by way of serious principal agent problems. However, it is important to note that Smith’s observations come in Book V of the Wealth of Nations after Smith has set out in Book II his “model” of banking and the various internal and external guidelines that he envisages being in place. Perhaps the puzzle is not so much that Smith endorsed joint stock banking but that he may have thought the law of partnership as it applied is Scotland was deficient as regards banking? Perhaps by not including any discussion of this in Book II implies that he did not think this much of an issue at all. In any case, the prevailing view is that Scotland was able weather the 1825/6 crisis because Scottish partnership law provided Scots banks with a “flexibility” deprived the English and Irish banks. Turner (2014) notes. [39]

> However, unlike in England and Ireland, other Scottish banks did not face a six-partner restriction. Furthermore, and more fundamentally, because the Scottish commercial-law system closely resembled those of its civil-law Continental trading partners, the partnership banks in Scotland enjoyed the privilege of a separate legal personality. This allowed these banks to separate ownership from control, enabling them to develop a managerial hierarchy, which in turn facilitated transferable stock. As a result, these Scottish partnership banks were effectively quasi-joint-stock companies with unlimited liability, which is why Scotland did not need to pass the equivalent of the 1826 Banking Co-partnership Act. (p.37-p.38)

Did Smith’s vision for banking prevail?

By the end of the nineteenth century almost nothing of real substance that Smith recommended for banking and monetary control would prevail. Smith’s ideas on monetary and banking control would be criticised, by the likes of Henry Thornton (1802). However, as White (1984) suggests, and before him Joplin (1822), many of the criticisms of competitive banks of issue, and hence indirectly of Smith, overlooked the fragility of small banks that was as likely the result of the monopoly privileges accorded the Bank of England as they were of country banks of issue as such. Nevertheless, whilst banks would be almost completely joint

[39] See also Acheson et al (2011).

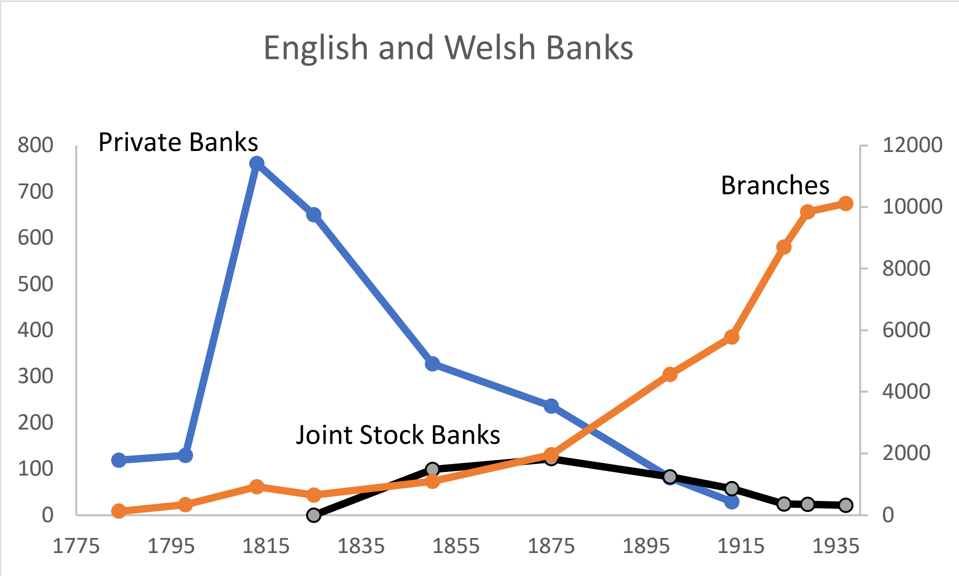
stock enterprises and convertibility would be retained, Britain had passed through a number of substantial banking crises. These often revolved around banks in lending practices and bill market abuses that Smith cautioned against. This led to intense debates around money and banking policy. The upshot was that the usury laws were dismantled, beginning in 1833, and by 1844 the Bank of England was the monopoly supplier of legal tender notes and the nation's guardian of specie. This one gold reserve legally backing the issue was meant to impart monetary stability; be that as it may, it did not guarantee banking stability. And three times in the ensuing decades the Act had to be suspended to allow the Bank of England to play the role of lender of last resort (and potentially supply notes in excess of the legal maximum).

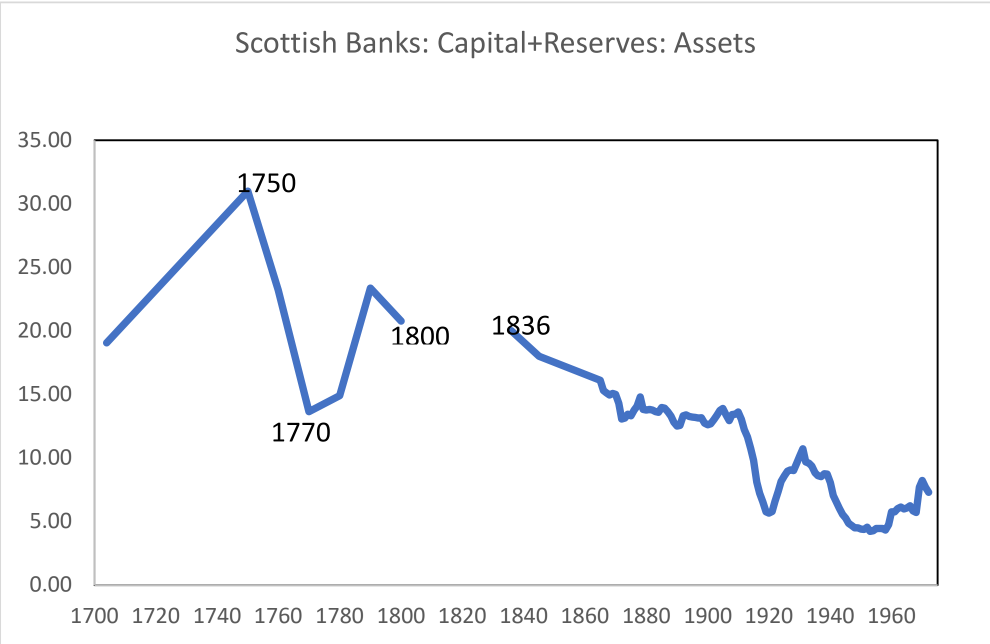
However, the Bank of England's role as lender of last resort imparted huge moral hazard into the financial system. Lord Overstone (a key influencer and defender of the Peel Acts and the preeminent leader of the Currency School and himself a former banker) observed in a letter to the Chancellor of the Exchequer after the 1844 Act had been suspended for the second time:

> The misfortune, as it appears to me, is this. The vicious system of Credit and Banking which is the source of the evil – will derive additional strength, and the upholders of that system will feel encreased confidence – from the assurance that in all future emergencies the Law will be relaxed for their assistance and protection....
>
> For this danger I do not see the remedy. The vicious system of Credit, and all the dangerous practices dependant upon it, will extend themselves. Never mind the apparent restraints of the Law, there is no reality in them, dash forward with boldness – the reins will certainly break.
>
> This consideration leads me to anticipate future convulsions, increasing in magnitude, and more formidable in their consequences. Bill brokers and Joint Stock Banks have now ascertained, or at least have convinced themselves that they can terrify the Govt can stop the Bank of England, can convulse the country – unless they are protected by Govt interference whenever their own extravagance brings about or threatens public inconvenience. Twice we have refused to face the danger, of course we must make up our minds to be trampled upon by the bully. (Quoted in Eltis. 2001).

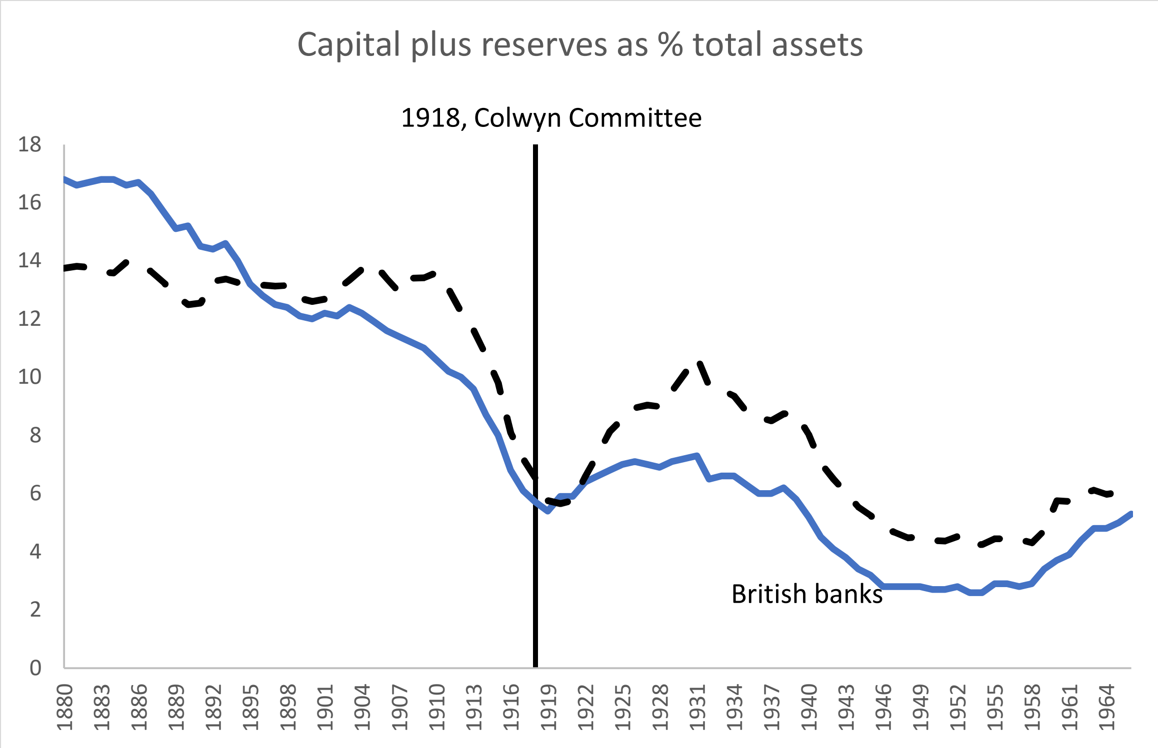
Moreover, the banking sector itself after about 1850/60 or so started to become very concentrated and its gearing rose, as the first chart below suggests. In 1878 the City of Glasgow Bank failed, collapsing due to a, by then, familiar mix of malfeasance and poor lending. That failure led to widespread adoption of limited liability by banks which, although it had been available for some years, had been largely resisted by banks. The concentration of the sector intensified so much so that by 1918 an official inquiry, known as The Colwyn

Committee (more formally, the Treasury Committee on Banking Amalgamations), was set up to investigate the sector (Billings et al., 2021). And the advantage of joint stock banking did not appear to have led to a highly capitalised sector.[40]





[40] Turner (2014) notes that even though limited liability was widely adopted by the end of the nineteenth century, the extended liability structure of bank shares lasted until well into the twentieth century providing an additional level of security. Billings and Capie (2007) show that bank capital for much of the twentieth century up until 1970 was probably somewhat higher than implied by published accounts because of hidden reserves. Turner (2014, p. 131) argues, on the other hand, that these ratios also need to be adjusted in the other direction for loan losses. The broad trends, at any rate, are agreed upon. The comments of Herbert Foxwell reported below suggest that there were by the time of the Colwyn Committee concerns about bank capitalisation, extended liability notwithstanding.



The decrease in the capital resources of the banks looks to have been broadly similar in Scotland and the UK. The first chart above for Scottish banks combines the data up to 1800 from Cameron (1967) with data from Checkland (1975a) from 1836 on[41]. The second chart above compares the Scottish experience with Britain as a whole over the shorter sample for which Sheppard (1971) provides data. The secular decline in capital ratios, or rise in gearing, is apparent in both charts, and particularly steep ahead of the inquiry into bank amalgamations. The Colwyn Committee did not recommend radical reform such as breaking up banks, nor did it recommend legal prohibitions on bank mergers. However, it did in effect put an informal halt to further immediate concentration in the sector. Mergers amongst the big five were in effect prohibited and further amalgamations involving the largest banks with smaller banks required Treasury and the Bank approval. Gorton and Tallman (2018, p. 72) suggest that a US congressman coined the phrase 'too-big-to-fail' following the rescue of Continental Illinois Bank in 1984. However, the too-big-to-fail problem had been clearly identified almost 70 years earlier. The only academic economist on the Colwyn Committee was Herbert Foxwell. He is quoted in Billings et al. (2021) observing

> … institutions of a certain size are "too large to let go". It inspires the customers with confidence when a bank reaches a certain size; they say that whatever the state of things this bank will not be allowed to fall: a crash would carry the market away. I confess I think this rather an unfortunate presumption, though it may be a sound one, because I believe it is partly responsible for the very slender reserves which our banks carry, and the very small proportionate capital.

[41] It goes without saying that these two data sources are not consistent. Our interest here is simply in broad trends.

The plots above show banks had indeed been increasing their gearing for some time. Turner (2014, p. 159-160) suggests that by the time of the Colwyn Committee, or shortly thereafter, it was widely (by the Bank and senor politicians) accepted that the big banks could not be allowed to fail. Various factors, including two world wars, government financial repression and a Great Depression would ensure that the "too large to let go" problem would lie dormant for some decades but by the 1970's it would return with a vengeance.

Why did Smith's views not prevail?

The proximate reason why Smith's views did not prevail is largely as a result of the victory of the Currency School over the Banking School and the Acts of 1844/45 (Peel Acts). One of the arguments of the Currency School was that the English and Welsh country (provincial) banks had a tendency to overissue their paper. Derided tendentiously as 'paper mints'[42], and often having their competence questioned, Peel became convinced that the country banks could operate independently of the Bank of England and that this was potentially dangerous to the country. For example, he worried that an attempt to address an external drain might be made more difficult or worse by the provincial banks.[43]

That the country banks were a danger to monetary and credit control was *perhaps* the majority opinion, both then and now (see Feaveryear, 1963, p. 239ff for the case against the country banks). That said, important voices have questioned that view, arguing in effect that the Bank of England did ultimately control (see e.g., Viner p 154ff). This debate had rumbled on since the Bullionist controversies but was given even more vigour following the English country bank failures of 1825/6.

Related to, though distinct from, this argument around the country banks, the Currency School argued that in a mixed currency system (that is one where paper and specie circulated) to ensure convertibility and stable adjustment as in Hume's prices-specie-flow the paper component of the circulation had to *be made to* behave like gold. That is to say, the circulation had to be controlled. This, Viner (1937) would label *The* Currency Principle. Into this picture, the notion of many competitive banks of issue seemed to fit less than easily.

[42] The phrase appears to be William Cobbett's who railed against country banks; see Fetter (1965), page 109. Irving Fisher (1936) was more restrained but equally critical of banks in the aftermath of the Great Depression in the US in 1930's. He referred to banks, as having acted "like so many private mints". See the Preface to the first edition.

[43] Viner (1937) is a classic reference for the monetary debates. Wood (1939) is a very detailed analysis of the debates and looks in some detail at the data. See also the very clear overview in Fetter (1965).

White (1984) argues that free banking has been airbrushed out of the great monetary debates of the early 19th century. He argues that Henry Thornton was in some ways very influential in setting the terms of the debate; namely, that free banking would be a dangerous route for England to take. He observes that Thornton (1802) was critical of the country banks and their susceptibility to runs. He also notes that Thornton believed that the Bank of England, at least in principle, was a stabilising force not just in the capital but in the wider country. He believed that any competitor in note issuing in the metropolis would be highly destabilising and would rob the Bank of the ability to control the general level of credit in the country both in stable and, more importantly, in unstable times. However, White is very critical of Thornton's failure to recognise that the country banks fragility, as he sees it, was due to a prior restriction. In other words, it was due to the Bank of England's legal monopoly on joint stock banking which, he argues, had been secured at the expense of wider banking stability. That was despite the fact that Thornton himself notes in passing that the Scottish banks had been free of bank runs in 1797. For White (op. cit.) and the free banking school, the combination of competitive issue, no limitations on the number of partners, note clearing, and no central bank reserve, meant the free banking system in Scotland was bound to be substantially more stable than the English system.

In short, the free banking school argues that Scotland's apparently more robust banking structure was the result of two features. First, a greater number of partners enabled banks to be larger, more diversified, and much better capitalised. Second, it was composed largely of banks of issue (at least in the latter part of eighteenth century and early part of the nineteenth, see Cameron's 1967 data reported earlier) and not dependent on a central bank of issue. The competitive nature of these banks and the rapid clearing of notes meant that no bank dominated the others and problems were quickly uncovered. And as noted earlier, some contemporary writers such as Joplin (1822) essentially made that same case.

The ultimate reason why Smith's vision for money and banking was unrealised was perhaps that as the British economy developed, became more integrated and moved increasingly centre stage in the world economy, two very different banking systems might find it hard to coexist. As Smith recognised, the Bank of England was already at the time of the Wealth of Nations "a great engine of state" (Kosmetatos, 2018)[44], and its role was only set to increase. While he typically opposed joint stock companies with privileges granted by government, he was restrained in overtly criticising the Bank of England[45]; he even seemed to acknowledge what we might now call lender of last resort interventions and (perhaps?) with a degree of approval. Moreover, the Scottish banks were also reliant on the London markets, and hence indirectly on the Bank of England, for liquidity, something the free bank analysis can overlook.

[44] Kosmetatos (2018) argues that the crisis in particular in 1772 did elicit action by the Bank that was more distinctly tending in the direction of lender of last resort (LoLR). He concludes: "Neither of these mid-eighteenth century (1763 and 1772) crises can be included among those seminal upheavals, like 1825-1826, 1857 or 1866 that punctuated the discourse on central bank action, and which even led to changes in the legal and monetary framework governing banking practice. But this earlier experience demonstrates that many of the intervention techniques used in those more famous episodes were already being developed by the second half of the eighteenth century, even if they were governed by practicality rather than theoretical conviction." No doubt Henry Thornton (1802) drew on these interventions in his recommendations for LoLR interventions, and upon which Bagehot (1873) built.

[45] See West (1997) for a critical/puzzled take on Smith's apparent reticence to critically appraise the position of the Bank of England.

Perhaps for this reason alone—the central importance of the Bank of England[46]—the Scotch system could not come to dominance in Britain, or even endure. Indeed, both the Currency School and the Banking School supported having a "central bank" even though they disagreed on how that bank should operate (and how dangerous to monetary stability were country banks of issue). The free bankers—who would have done away with a central bank—were a minority view. And so, England being much the larger country and economy and London the dominant financial centre, not merely in Britain but also increasingly the world, the Scottish system's influence on wider questions was increasingly side-lined.[47]

Smith and banking today

In this section, an attempt is made to understand why Smith endorsed the internal and external regulations that were identified earlier, in order to try to understand how Smith might assess the current state of bank regulation. That is: What were the concerns that led him to endorse qualifications to free banking in his time? Might those concerns still be legitimate today? Or perhaps the concerns of the past have been superseded by others, or the balance amongst the concerns may have shifted? In particular, following the earlier discussion we look at his defence of usury, the banning of the optional clause and small note restrictions. We then turn to joint stock banking and his attitude to regulation more widely.

In discussing what is assumed to be the Ayr Bank, Smith recognises very clearly the manifest failures of what 'his' internal rules. As well as lending to inappropriate clients (whose aims were essentially dishonest) and bill market abuses, the bank itself had lent inappropriately, breaching all of Smith's guidelines including those on real bills. The upshot was that the total circulation was larger under a mixed (paper plus specie) currency than it would be under a pure specie circulation: That is, it breached The Currency Principle (Viner, 1937). Recall, that the Ayr Bank was far from being a small unit bank, supplying as it did most of the paper issue in Scotland and indeed its circulation being used to buy bills on London. Smith notes:

> The paper which was issued upon those circulating bills of exchange, amounted, upon many occasions, to the whole fund destined for carrying on some vast and extensive project of agriculture, commerce, or manufactures; and not merely to that part of it which, had there been no paper money, the projector would have been obliged to keep by him, unemployed and in ready money for answering occasional demands. The

[46] The Bank always had its critics, some vociferous, that demanded an end to its "privileged" position. Even David Ricardo proposed plans to reduce substantially the Bank's importance in monetary affairs as it was run, he believed, by "ignoramuses" in matters of monetary issues. Nevertheless, nothing that was happening north of the border appears to have had much bearing on the Bank's critics.

[47] There was of course some accommodation by Government in London to the specifics of the Scotch system, such as the legality of the £1 and the special provisions in 1845 for Scotland and Ireland as part of the Peel Acts. But in the sweep of monetary history these are footnotes. See Checkland (1975a) and Fetter (1965) for details.

> greater part of this paper was, consequently, over and above the value of the gold and silver which would have circulated in the country, had there been no paper money. It was over and above, therefore, what the circulation of the country could easily absorb and employ, and, upon that account, immediately returned upon the banks in order to be exchanged for gold and silver, which they were to find as they could. It was a capital which those projectors had very artfully contrived to draw from those banks, not only without their knowledge or deliberate consent, but for some time, perhaps, without their having the most distant suspicion that they had really advanced it.

In other words, Smith's rules, internal and external, on banks were all about underpinning monetary stability. They were not microprudential as we might now call them. Their justification, or perhaps significance, is largely that which Friedman and Schwartz (1963) accorded the FDIC in underpinning monetary stability following the contraction in the US between 1929 and 1933. Smith's rules, if anything, were macroprudential in nature, just as was deposit protection in the US[48].

So, while his arguments as regards usury, the optional clause, small notes restrictions, joint stock and unlimited liability banks, and what we label his "internal rules" for banks, all tend toward protecting depositors and businesses and, to an extent, protecting banks from themselves, ultimately they are all about ensuring the integrity of circulating notes. It seems clear that Smith had no desire to create a formal banking code, much less design a banking regulator. His "external rules" along with free entry into banking would ensure that banking would be sufficiently stable and risk free and that the inevitable bank failures would, unlike The Ayr Bank, be of minor significance.

It may also be worth noting that even the pure free banking model would find it hard to avoid some "regulatory oversight," especially where nothing stops "large" banks from emerging. In particular, note-clearing is a double-edged sword. Whilst it may act as a disciplining device and work to supress overissue by individual banks, eventually there will need to be rules of the game or at least rules for clearing up after crises. Who would devise those rules and who would implement them? And can that be done in a way that maintains support of participants in clearing? Moreover, there will inevitably be times when decisions are contentious and consequential; who rules when another bank's issues are excessive, or collateral is below par? And how is it to be decided when support is both necessary and right, or even more controversially when such support is not to be extended? How might any losses be apportioned?

And so, no-one had a desire to introduce a banking code (neither the Currency School nor the Banking School argued for this). One was briefly introduced as part of the 1844 Act but was

[48] Around 90% of commercial banks were part of the FDIC scheme at its inception.

soon all but discarded[49]. It would be 1979 before Britain moved to legislate on banks in any meaningful way. Whether Smith might have endorsed more rules as time went on is hard to say. On the one hand, as Stigler (1956) notes: "Natural liberty seems to have been little more than a working rule, and Smith proposes numerous departures from natural liberty because the participants are incompetent or fail to consider external effects of their behaviour……"

On the other hand, however, Smith famously remarked,

> "People of the same trade seldom meet together, even for merriment and diversion, but the conversation ends in a conspiracy against the publick, or in some contrivance to raise prices.

Less observed, is that after stating this he goes on to say:

> It is impossible indeed to prevent such meetings, by any law which either could be executed, or would be consistent with liberty and justice. But though the law cannot hinder people of the same trade from sometimes assembling together, it ought to do nothing to facilitate such assemblies; much less to render them necessary.
>
> A regulation which obliges all those of the same trade in a particular town to enter their names and places of abode in a publick register, facilitates such assemblies. It connects individuals who might never otherwise be known to one another, and gives every man of the trade a direction where to find every other man of it.

In other words, Smith senses cartelised behaviour being a risk from identifying a group for specific regulation. Moreover, the link between the British state and the great chartered companies of his day alerted him to the risks of Government involvement in business. He appears to have sensed that regulation might become a "Game of Bank Bargains," as Calomiris and Haber (2014) have recently argued. And indeed, elements of a UK banking cartel were safely established by the start of the twentieth century with the full knowledge and support of the authorities under the guise of 'oversight.' The cartel would be an enduring feature for much of the twentieth century; see Griffiths (1973) and Turner (2014) and the references therein.

[49] With some provisos, banks were part of company legislation from 1857 with the enactment of the Joint-Stock Companies Banking Act.

How would Smith Approach banking regulation?

Our argument above is that Smith endorsed rules on banking in so far as they contributed in important ways to monetary stability—the integrity of note issue and the avoidance of contagion. To that end, he wanted banks to observe an internal code of conduct. But that code was underpinned by a law to limit balance sheet risk (usury), measures to ensure sufficient specie was in circulation (the 1765 Act mandated convertibility on demand and a lower limit on note denomination so that some specie would be in circulation), and measures to ensure banks could fail safely (unlimited liability and free entry into banking). These rules applied to all banks and no rules were applied to one bank and not to another (the chartered banks were the exceptions). In other words, rules were, in modern terminology, macroprudential and intended to underpin monetary policy and stability. And it was for much the same reason that Friedman and Schwartz (1963) 'endorsed' the establishment of deposit protection in the US in 1933.

If the above argument is correct, then presumably Smith would devise rules that would support the currency and so the first question would be: Who should supply the medium of exchange? There are two basic options. The first option would be to follow Friedman (1959) and endorse the State as the monopoly supplier of a fiat currency[50]. A second option might be competing private sector fiat (fiduciary?) currencies as suggested by Hayek.[51] Of course, the first option has come to dominate and there seems little likelihood of option 2 emerging. So, how did Friedman propose to reconcile monetary stability and bank regulation?

Friedman (1956) argued that the state should be the monopoly supplier of paper notes and that banks should hold 100% reserves against deposits (the so-called Chicago plan). The plan was formulated originally by Frank Knight, Lloyd Mints, Henry Simons, and other well-known Chicago economists[52].

He endorsed government control of the money supply on the basis of four main arguments. First, there is a resource cost of running a pure commodity currency. That led to a tendency, he argued, for unbacked paper currency or fiduciary issue to emerge. Next, as noted earlier, he observed the "peculiar difficulty of enforcing contracts which involve promises to pay but that also serve as a transactions medium and the consequent tendency to their fraudulent

[50] A possible variant might be a commodity standard of some sort. Hayek thought this would be desirable if the State were the monopoly supplier. Hayek thought a gold standard would be better than nothing, but that a better standard could be devised.
[51] Again, an obvious variant here would be competing private commodity standards.
[52] More specifically, the plan was to abolish the fractional reserve system and impose 100% reserves on demand deposits. Phillips (1992) explains how the plan was formulated, written and passed to FDR, and what happened subsequently.

use." Third, Friedman argued that the "technical monopoly character" of pure fiat currency meant that it was necessary to set an external limit on its amount. And finally, "the pervasive character of money which means that the issuance of money has important effects on parties other than those directly involved and gives special importance to the preceding features." Friedman (op. cit.) concludes that "[I]t is dubious that the market can by itself provide such a framework."

Of course, the US in the end did not adopt the Chicago Plan. Instead, the Banking Acts of 1933 (also known as Glass-Steagall) and 1935 formed the core of the financial reforms. Amongst other things, these separated commercial and investment banking, created the FDIC, and strengthened the decision-making at the Fed, providing the Board with more power to regulate banks and set interest rates.

In the intervening years that perspective—regulating banks only in so far as doing so is necessary for monetary stability—has perhaps been lost. Certainly, in the post-WWII the focus has been on microprudential regulation of banks. As the Bretton Woods system broke down, international banks and capital markets began to flourish and central banks around the world were confronted by "new" risks of increasingly large, innovative (not always in a good way) and internationally active banks. For example, in the UK the secondary bank crisis materialised more or less at the same time as Herstatt Bank failed in Germany and Franklin National Bank failed in the US. Bank regulation became an increasingly important task for national authorities; but that regulation was largely microprudential. As it developed, it was focussed mostly on individual bank risks, depositor protection and in minimising the cost of bank failure (Goodhart, 2011). The notion that there were risks common across banks and that the banking sector had the potential seriously to disrupt monetary policy receded somewhat.[53]

It is difficult to pin down the degree to which microprudential regulation was either a success or a failure for most of the post-WWII. In the UK, regulation developed at the same time as banking and financial services were liberalised. The periodic bank failures and credit booms were for some a price worth paying along a transition path to a more efficient economy; for others they reflected an enduring Achilles heel in financial markets; for still others they were a predictable outcome of a political game over income and wealth distribution. The global financial crisis in 2008/9 and its aftermath may have changed the terms of the debate. There has in the UK been a clear pivot towards a macroprudential perspective on financial regulation. As well as global moves towards increased capital and liquidity requirements for banks and other financial intermediaries, in the UK the enactment of the Vickers Commission proposals on bank ringfencing, and the establishment of the Financial Policy Committee (FPC) within the Bank of England have been amongst the most striking institutional developments anywhere. The Bank of England is now responsible for

[53] That is not to say, as regulations developed, there had been no recognition of these more macro issues. See for example Crockett (2000) and Borio (2003).

operational monetary policy (via the Monetary Policy Committee, MPC), microprudential regulation (via the Prudential Regulation Authority, PRA) and macroprudential regulation (via the FPC). The Bank is also the resolution authority for UK regulated financial intermediaries. The FPC's statutory responsibility is to assist the Bank in ensuring the stability of the UK financial system. The FPC can request or require additional capital and liquidity requirements to be imposed on all regulated intermediaries if it deems financial stability is at risk. It has also been granted additional, far-reaching powers, such as those to limit aggregate bank exposure to the property market.

It could be argued that since the global financial crisis the UK has moved in the direction of Smithian/Friedmanite regulation in embracing a macroprudential perspective on regulation. Monetary policy was neither able to avert the crisis nor on its own to help the economy recover from its aftermath. So, perhaps macroprudential measures will help restore the efficacy of monetary policy.

However, the system of rule setting is very complex here in the UK and elsewhere. Microprudential regulation remains a key focus of policymakers and is becoming ever more complex. Moreover, macroprudential is also very heterogenous across countries. And even in the UK, where macroprudential policies and institutions are amongst the most well-formulated anywhere, there are serious issues of accountability both as to objectives and accountability. It is not clear what policymakers are ultimately trying to achieve with these myriad policies.[54] On this, Smith had the advantage.

[54] See Duncan and Nolan (2019, 2023a,b) for discussions of these issues.

References

Graeme G. Acheson, Charles R. Hickson & John D. Turner (2011) Organisational flexibility and governance in a civil-law regime: Scottish partnership banks during the Industrial Revolution, *Business History*, 53:4, 505-529

Gary M. Anderson and Robert Tollison (1982) Adam Smith's Analysis of Joint-Stock Companies. *Journal of Political Economy*. vol. 90, issue 6, 1237-56.

Arie Arnon (2011) Monetary Theory and Policy from Hume to Wicksell: Money Credit and the Economy. Cambridge University Press.

Mark Billings and Forrest Capie (2007) Capital in British Banking, 1920-1970. *Business History* Vol. 49. 139-162. 10.1080/00076790601170231.

Mark Billings, Simon Mollan, and Philip Garnett (2021) Debating banking in Britain: The Colwyn committee, 1918. *Business History* Vol. 63, no. 6, 944–965 https://doi.org/10.1080/00076791.2019.1593374.

Claudio Borio (2003) Towards a macroprudential framework for financial supervision and regulation? BIS Working Papers, No.128.

Andrew Crockett (2000) Marrying the micro- and macroprudential dimensions of financial stability. BIS Speech. September.

Charles W. Calomiris and Stephen H. Haber (2014) *Fragile by Design*: *The Political Origins of Banking Crises and Scarce Credit*. Princeton University Press.

Rondo Cameron (1967) Chapter 3: Scotland 1750-1845 in *Banking in the Early Stages of Industrialization: A Comparative Study*. New York: Oxford University Press.

Sydney G. Checkland (1975a) *Scottish Banking: A History, 1695–1973*, Glasgow. Collins.

Sydney G. Checkland (1975b) Adam Smith and the Bankers in A.S. Skinner and T. Wilson (eds), *Essays on Adam Smith*, Oxford: Clarendon Press, pages 504–23.

John H. Clapham, (1944, 1970) *The Bank of England: A History*, 2 vols Cambridge University Press.

Michael Collins (1990) *Money and Banking in the UK: A History*, London: Routledge.

Thomas M. Devine (I975) *The Tobacco Lords: A Study of the Tobacco Merchants of Glasgow and Their Trading Activities, c. I 740-90* Edinburgh University Press

Douglas W. Diamond (1991). Debt Maturity Structure and Liquidity Risk. The Quarterly Journal of Economics, 106(3), 709–737.

Alfred Duncan and Charles Nolan (2020) Reform of the UK Financial Policy Committee. *Scottish Journal of Political Economy*. Volume 67, Issue 1. February 2020 Pages 1-30

Alfred Duncan and Charles Nolan (2023a) Aligning the ring-fencing and resolution regimes. Evidence submitted to House of Lords Inquiry into UK ring fencing and resolution regimes.

Alfred Duncan and Charles Nolan (2023b) Bank of England: How is Independence Working? Evidence submitted to Houses of Commons Inquiry in Bank of England Independence.

Alastair J. Durie (1973) The Markets for Scottish Linen, 1730-1775. *The Scottish Historical Review*, Vol. 52, No. 153, Part 1 April, pp. 30-49 (20 pages)

Walter Eltis (2001) Lord Overstone and the Establishment of British Nineteenth-Century Monetary Orthodoxy. *Discussion Papers in Economic and Social History* Number 42. University of Oxford.

Albert Feavearyear (1963) The Pound Sterling: A History of English Money. Second Edition with E. Victor Morgan. Oxford, Clarendon Press.

Frank W. Fetter (1965) *Development of British Monetary Orthodoxy, 1797–1875*. Cambridge: Harvard University Press.

Irving Fisher (1936) *100% Money*. Adelphi Company. New York.

Mark J. Flannery (1986), Asymmetric Information and Risky Debt Maturity Choice. The Journal of Finance, 41: 19-37.

Milton Friedman (1959) *A Program for Monetary Stability* Fordham University Press.

Milton Friedman and Anna J. Schwartz (1963) *A Monetary History of the United States, 1867–1960*. Princeton University Press.

Milton Friedman and Anna J. Schwartz (1983) Has government any role in money? *Journal of Monetary Economics*. 17, pp 37-62.

Charles Goodhart (2011). *The Basel Committee on Banking Supervision: A History of the Early Years, 1974–1997*. Cambridge University Press

Tyler B. Goodspeed (2016) *Legislating Instability: Adam Smith, Free Banking, and the Financial Crisis of 1772.* Harvard University Press.

Gary B. Gorton and Ellis Tallman (2018) *Fighting Financial Crises – Learning from the Past***.** University of Chicago Press.

Brian Griffiths (1973) The development of restrictive practices in the UK monetary system, *The Manchester School of Economic & Social Studies*. vol. 41, issue 1, 3-18

Richard S. Grossman (2010) *Unsettled Account: The Evolution of Banking in the Industrialized World since 1800*, Princeton, NJ: Princeton University Press.

Henry Hamilton (1956) The Failure of the Ayr Bank. *The Economic History Review*, Vol. 8, No. 3 pp405-417.

Andrew G. Haldane (2015) Who Owns a Company. Speech at University of Edinburgh Corporate Finance Conference 22 May 2015. Available at Bank of England webpages.

Stephan Heblich and Alex Trew (2019) Banking and Industrialization. *Journal of the European Economic Association*, Volume 17, Issue 6, December, Pages 1753–1796.

Samuel Hollander (1999) Jeremy Bentham and Adam Smith on the usury laws: a 'Smithian' reply to Bentham and a new problem, *Journal of the History of Economic Thought*, 6:4, 523-551, DOI: 10.1080/10427719900000042.

Thomas Joplin (1822) An Essay on the General Principles and Present Practices of Banking, in England and Scotland. Classic Reprint Series.

Thomas M. Humphrey and Richard H. Timberlake (2019) *Gold, the Real Bills Doctrine, and the Fed: Sources of Monetary Disorder, 1922-1938*. Cato Institute.

Wilfred T. C. King (1935) *A History of the London Discount Market*. Routledge.

Andrew W. Kerr (1908) History of Banking in Scotland. Adam and Charles Black. London.

Paul Kosmetatos (2018) Last Resort Lending before Henry Thornton? The Bank of England's Role in Containing the 1763 and 1772-1773 British Credit Crises. *European Review of Economic History*, 23, 299–328.

David Laidler (1981) Adam Smith As a Monetary Economist. *The Canadian Journal of Economics* Vol. 14, No. 2 May, pp. 185-200

David Laidler (1991) *The Golden Age of the Quantity Theory* Princeton University Press

Lloyd W. Mints (1945) *A History of Banking Theory in Great Britain and the United States.* The University of Chicago Press

Charles W. Munn (1981) The Scottish Provincial Banking Companies 1747-1864. John Donald Publishers Ltd. Edinburgh.

John H. Munro (2011) Usury, Calvinism, and Credit in Protestant England: from the Sixteenth Century to the Industrial Revolution. Working Paper No. 439, University of Toronto, Department of Economics.

Elisa Newby (2012) The suspension of the gold standard as sustainable monetary policy, *Journal of Economic Dynamics and Control* Elsevier, vol. 36(10), pp. 1498-1519.

Shizuya Nishimura (1971) *The Decline of Inland Bills of Exchange in the London Money Market 1855-1913*. Cambridge University Press.

Alison G. Olson (1983) The Virginia Merchants of London: A Study in Eighteenth-Century Interest-Group Politics *The William and Mary Quarterly*, Vol. 40, No. 3 July, pp. 363-388.

Ronnie J. Phillip (1992) The "Chicago Plan" and New Deal Banking Reform. Working Paper No. 76. Levy Economics Institute.

Jacob M Price (1975) New Time Series for Scotland's and Britain's Trade with the Thirteen Colonies and States, 1740 to 1791. *The William and Mary Quarterly*, Vol. 32, No. 2, pp. 307-325

Leslie S. Pressnell (1956) *Country Banking in the Industrial Revolution*. Oxford, U.K.: Clarendon Press.

Charles Rist (1940) *History of Monetary and Credit Theory: From John Law to the Present Day*. Routledge.

Hugh Rockoff (2009) Upon Daedalian Wings Of Paper Money: Adam Smith And The Crisis OF 1772 NBER Working Paper 15594. December.

Hugh Rockoff (2011). Parallel Journeys: Adam Smith and Milton Friedman on the Regulation of Banking. *Journal of Cultural Economy*, 4:3, 255-283, DOI: 10.1080/17530350.2011.586849

Joseph A. Schumpeter (1954) *History of Economic Analysis*. Allen and Unwin.

George Selgin (2018) Scottish Banks and the bank Restriction, 1797-1821: Parts 1-3. Published on-line at The Cato Institute.

Amartya Sen (2013) The Contemporary Relevance of Adam Smith, in *The Oxford Handbook of Adam Smith* Christopher J. Berry (ed.) et al.

David K. Sheppard (1971) The Growth and Role of UK Financial Institutions 1880-1962. Methuen and Co Ltd. London.

Adam Smith (1776) An Inquiry into the Nature and Causes of the Wealth of Nations. Edited by Edwin Cannan. The Modern Library Edition, 1994.

William L. Taylor (1965) *Francis Hutcheson and David Hume as Predecessors of Adam Smith*, Durham: Duke University.

Henry Thornton (1802, 1939) An Inquiry into the Nature and Effects of the Paper Credit of Great Britain. Edited by F. A. von Hayek. Routledge.

John D. Turner (2014) Banking in Crisis: The Rise and Fall of British Banking Stability, 1800 to the Present. Cambridge University Press.

Jacob Viner (1937) Studies in the Theory of International Trade. Routledge.

Carl C. Wennerlind (2000) The Humean Paternity to Adam Smith's Theory of Money. *History of Economic Ideas* Vol. 8, No. 1, pp. 77-97

Edwin G. West (1997) Adam Smith's Support for Money and Banking Regulation: A Case of Inconsistency. *Journal of Money, Credit and Banking*, Feb., Vol. 29, No. 1, pp. 127-134

Lawrence White (1984) *Free Banking in Britain*. Cambridge University press.

Elmer Wood (1939) *English Theories of Central Banking Control 1819-1858*. Cambridge. Harvard University Press