J Bus Ethics (2016) 139:287–297
DOI 10.1007/s10551-015-2611-7

In Which Context is the Option Clause Desirable?

Mathieu Bédard[1,2]

Received: 12 September 2014 / Accepted: 10 March 2015 / Published online: 28 March 2015

Abstract The option clause is a contractual device from free banking experiences meant to prevent banknote redemption duels. It has been used within the Diamond and Dybvig (J Pol Econ 91: 401–419, 1983) framework to suggest that very simple contractual solutions can act as an alternative to deposit insurance. This literature has, however, been ambiguous on whether the option clause can replace deposit insurance outside of those two contexts. It will be argued that the *theoretical* clause does not generally affect the likelihood that a solvent bank goes bankrupt *because* of a bank run, as empirical evidence suggests it is already near null, and that the exercise of the clause will have the effect of *diminishing* the size of creditor claims on bank assets because it exacerbates the agency problem of bank debt. It will therefore be argued that the clause is only desirable in (a) free banking systems that are historically devoid of bank runs in the first place and have other means of managing debt-related agency problems and (b) under the unrealistic assumption that bank runs are self-fulfilling prophecies. It will be argued that the agency problem of bank debt makes the option clause undesirable outside of free banking systems.

Keywords Financial stability · Bank runs · Free banking · Option clause · Insolvency

JEL Classification N23 · G21 · G33

✉ Mathieu Bédard
mathieu.bedard@tse-fr.eu

[1] Toulouse School of Economics, Manufacture des Tabacs, Aile Jean-Jacques Laffont, 21 Allée de Brienne, 31015 Toulouse Cedex 6, France

[2] Aix-Marseille School of Economics, CNRS and EHESS, Centre de la Vieille Charité, 2, rue de la Charité, 13002 Marseille, France

Introduction

The 80s and 90s saw the Diamond and Dybvig (1983) model become the canonical model of bank runs and a centerpiece of financial stability theory. Free bankers would respond to this challenge by developing a full-fledged theory around a contractual device out of subset of the Scottish free banking era, going from 1730 to 1765, that economic history literature had recently highlighted (Checkland 1975; Vaubel 1984; White 1995).[1] The Scottish banking system of that era can be characterized as having a free entry, and where the right to issue currency was universal (White 1992, p. 157; White 1995). There was therefore no possible monetary policy, or even a central bank to conduct it. The use of deposits as we know them today was relatively uncommon, the clause rather applied to the conversion of banknotes to gold and silver. In general, banking regulation was minimal and benign, the check on banknote emission was its secondary market and clearing mechanism through reflux. To prevent inimical abuses of these reflux mechanisms, Scottish banks would use the so-called option clause.[2] Through the clause a bank could, at its discretion, pay the bearer of its notes within 6 months rather than immediately, compensating him by paying what was then the maximum interest rate allowed

[1] This theoretical literature essentially comprised Dowd (1988), Dowd (1991, 1993, 1996), Cowen and Kroszner (1989), Schuler (1992b), Selgin (1993, 1988), White (1999), Selgin and White (1987, 1994), and Selgin and White (1997).

[2] Henceforth cited as "the clause" for the sake of brevity.

by usury laws. These theoretical developments would present the clause as an alternative to deposit insurance.

The Scottish historical anecdote illustrates some stabilizing properties of the clause in case of "duels."[3] During those predatory "duels," a competing bank would suddenly present a large quantity of bank notes for conversion into specie. If the bank has been holding fractional reserves of specie, and these assets are less than perfectly liquid, a critical mass of note bearers demanding immediate conversion to base money can cause insolvency either through firesale losses or an inability to pay debts as they come due. The predator-bank would then presumably acquire its market share and/or its assets at a discount (White 1995, pp. 23–27). These duels were rarely successful, and rarely even triggered the prey-banks' option clause, but when it did, it allowed more time to orderly liquidate assets and satisfy note bearers.

Bank runs as described by Diamond and Dybvig (1983) are very different from those duels. They are self-fulfilling prophecies, where depositors rush en masse to the bank because they expect it will be insolvent, thereby causing its insolvency. Deposit insurance solves this problem by making sure that depositors do not lose anything in the event of a bank run, so they do not actually have an incentive to run in the first place. The clause is presented as an alternative to deposit insurance. By giving the bank time to liquidate its assets before it has to satisfy customers, it will not face firesale losses as a result of a bank run. Just like in the case of deposit insurance, the cause's *ex post* properties make it *ex ante* efficient, and depositors have no incentives to run.

While free bankers explicitly and unambiguously recognize that free banking era Scotland never experienced this kind of self-fulfilling bank runs (Selgin and White 1997, p. 271),[4] they use the clause's example to suggest that very simple contractual solutions could emerge and be used by banks in case of Diamond and Dybvig bank runs, and seemingly constitute a private sector alternative to government-provided deposit insurance.

Although the clause is often discussed by theorists of free banking in the context of a broader discussion on free banking which includes features such as multiple banks of issue and sometimes extended liability for shareholders, they also argue that the clause is desirable outside that context. Discussing modern banks, Dowd (1991, p. 332) suggests that the "option clauses could also be inserted into the contracts governing deposits as well," in Selgin and White (1996, p. 91) "[a]n option clause in note and demand deposit contracts gives the bank the option to suspend payments in the event of a run." White (1999, Chap. 6) discusses the clause and its application to runs on deposits in discussion that sometimes addresses the modern-day banking systems. It is common for these discussions to underline the similarity to modern-day Notice of Withdrawal clauses on checking accounts, even though they carry no penalty. Evans (2013) discusses the clause and gives a modern example of "quasi option clause" and "quasi withdrawal notice" (the equivalent of the clause for deposits) to prove that they are explicit and not hidden or deceptive features of banking. The question of whether the clause can constitute an alternative to deposit insurance outside of the context of free banking can be ambiguous, and the terms of the debate are not clearly explained. This research clarifies this point.

The arguments in favor of the clause are summarized by Dowd (1991, p. 763) as follows: (1) using the clause increases the size of the creditor's claim on bank's assets and (2) it eliminates the possibility that a solvent bank goes bankrupt because of a bank run. Therefore, (3.a) the clause reduces the incentives to run, and (3.b) makes runs less damaging when they do occur.

It will be argued that the clause does not generally affect the likelihood that a solvent bank goes bankrupt *because* of a bank run, as empirical evidence suggests it is already near null, and that the exercise of the clause will have the effect of *diminishing* the size of creditor claims on bank assets. It will therefore be argued that the clause is desirable only (a) in free banking systems that are historically devoid of bank runs anyway and (b) under the unrealistic assumption that bank runs are self-fulfilling prophecies as in Diamond and Dybvig. Both run contrary to the empirical evidence. This research will therefore respond to points (1), (2), and (3.b) in Dowd (1991), according to which the clause eliminates the possibility that a bank goes bankrupt due to a bank run, and that it makes runs less damaging.[5] The clause, it will be argued, is not desirable outside the institutional context from which it emerged,[6] and that even within this context the clause is not necessary beyond its historical role in preventing "duels."

This paper also emphasizes the differences between the *theoretical* clause and the *historical* clause underlined by

[3] For other examples of option clauses in free banking systems, see Jonung (2000) for 1864–1903 Sweden and Schuler (1992a) for nineteenth century Canada. See also Hammond (1957, p. 174) for one known case in 1804–1809 Boston.

[4] A contemporary anonymous comment cited in White (1995, p. 143) is particularly revealing "A run upon any bank, such as happens in England sometimes or a panic, are terms the meaning of which is hardly understood in Scotland."

[5] For a critique of point (3.a), namely that the clause reduces the incentives to run, see Yeager (1993), Gherity (1995), and Shah (1997). However, note that it would seem like simple modifications in the payoff and interest accumulation, not involving additional transaction costs, could solve this problem.

[6] More specifically, in a context of competing money issuance and extended liability of shareholders.

Goodspeed (2014) and Gherity (1995). The *theoretical* clause is a suspension of the convertibility of all of a single bank's banknotes, while the *historical* clause was a mean to selectively deny convertibility and discriminate between classes of note bearers.

To do this, we will first give a quick overlook of the Diamond and Dybvig (1983) model and the *theoretical* clause. We will then turn to the *historical* clause in Sect. "Discrepancies Between the *Theoretical* and the *Historical* Clause," and explain the context of its emergence, and how it differs from the *theoretical* clause. Section "The limits of Diamond and Dybvig's significance" explores the difference between the predictions of Diamond and Dybvig and the empirical literatures findings. In Sect. "Desirability of the clause," we ask whether either clauses are desirable. To do this, we offer an alternative model for bank runs, more consistent with stylized facts, introducing an agency problem of bank debt. Section "Conclusion" will conclude.

The Theoretical Framework of Diamond and Dybvig and the Clause

To better appreciate the theoretical arguments about the clause and how it provides an answer to the Diamond and Dybvig (1983) challenge, it is necessary to revisit this model and modifications by Selgin (1993).

The Diamond and Dybvig model is a prisoner's dilemma where investment opportunities are strictly long term and costly to interrupt. It is eventually revealed to agents whether they have immediate or distant (long term) needs, but they find out only after having invested, and the urgency (or lack thereof) of their needs is strictly private. To pool the risk of having needs before investments mature, a bank offers a "deposit" contract that improves the interim payment in case of investment interruption, at the cost of a lesser long-term payoff. The bank is essentially offering an insurance against liquidity needs that would appear before investment maturity by reducing the residual payment after project completion.

The model can lead to a rational bank run when one adds the assumption that depositors face a sequential service constraint (clients are served on a first-come-first-served basis). Ignorant of the true consumption needs of each individual agent demanding early interruption of the investment, the "bank" compensates all clients the same. If all agents decide to withdraw at an interim date, like it is the case in a bank run, the bank will be insolvent before the lasts in line receive anything. Agents with genuine intermediate needs will seemingly "starve."

The model therefore has two Nash equilibria: (1) the first where everyone acts according to their needs and (2) a second where long-term agents self-fulfillingly believe, for some reason left unexplained,[7] that other long-term agents want to withdraw early.

The bank may take action to remain in the good equilibrium by announcing that "deposit" convertibility will be suspended if there are too many withdrawals. This solution is considered *ex ante* efficient in the benchmark model because agents with long-term needs will not withdraw early. There is, however, a problem of *ex post* efficiency when one assumes a random distribution of consumption profiles. Agents still find out their own consumption pattern, but not the proportion of short-termists, and might believe there are too many early withdrawals, and run on the bank even with a forthcoming suspension. Some short-term agents may not exercise their rights if they are at the back of the line. Since the role of the Diamond and Dybvig bank is specifically to provide insurance against the risk of being a short-termist, suspension of convertibility is strictly dominated by a Government intervention that would tax early withdrawals if they are too numerous. This redistribution scheme is curiously labeled "deposit insurance."

Much has been written in terms of criticism of this model,[8] Of particular interest is Selgin (1993), who questions the suspension of convertibility as portrayed in Diamond and Dybvig.

Indeed, since the Diamond-Dybvig bank invests and dispenses the economy's sole homogenous consumption good there are no media of exchange, if the bank suspends convertibility, all of the economy's goods are locked up in the bank. But modifying the original model to assume that it is possible to trade bank debt, and to include some other vendor of consumption good to make for more complete markets, allows for short-term agents to sell their claim to the vendor, and the vendor to act as a long-term agent with those claims. A bad 'bank run' equilibrium can also exist in this framework, when one or both types of agents lose confidence in the bank for any reason, and ask for redemption instead of using their claims as a media of exchange.

As a response to this kind of bank run, the Selgin (1993) bank can resort to two types of suspension of convertibility. The first is a 'bank holiday' where all bank activity would stop until investment maturity, much like in the original Diamond and Dybvig model. If there are more than one bank, the clearing mechanism comes to a halt, and claims on each bank would stop being generally accepted. The

[7] Diamond and Dybvig suggest 'sunspots' as a reason for this instant change in expectations. It is a way of saying that the source for these beliefs is exogenous to the model, and can be any belief, right, or wrong, about anything that might affect banking even in the most remote and indirect way, further emphasizing the bank's fragility. This is why it is sometimes called the random withdrawal theory of bank runs.

[8] For a review of the critical literature on Diamond and Dybvig see White (1999, pp. 127–133).

second is a more partial restriction where the bank suspends convertibility of deposits to specie, but still allows for its claims to be traded through banknotes and checks. These do not affect the bank's solvency. The bank would choose to offer deposit contracts with the latter type of suspension, because short-term agents would not invest in it otherwise.[9]

To make this contract incentive-compatible, Selgin and White (1994, p. 1729) suggest it could feature an indemnification like the option clause historically did. Dowd (1988, pp. 322–323) investigates the mechanisms governing the optimal clause. The bank chooses a deferment period and compensatory interest rate such that the depositor is indifferent between the use of the clause or not, and which maximize bank profit.

From the bank's point of view, the clause is not only beneficial *ex post* by allowing banks to find specie at more favorable prices than through rapid firesale of assets, but also *ex ante* by reducing the likelihood that bank runs occur. This compels Dowd (1991, p. 763) to suggest that depositors would *prefer* contracts with an optional suspension clause, not only because of the two above-mentioned benefits, but also because it increases the size of their claims on bank assets by eliminating the possibility that a solvent bank will default. The bank has an incentive to invoke this clause as soon as it allows for a return to liquidity less expensively than short-term borrowing. This may be in the case of bank runs, as well as in cases where the bank can only borrow at an unfavorable interest rate, either because the market interest rate are temporarily too high or because its credit rating has gone sour (Dowd 1988, 1993).

Discrepancies Between the *Theoretical* and the *Historical* Clause

Goodspeed (2014) and Gherity (1995) emphasize that the literature on the *theoretical* clause is very different from the *historical* clause. A first important difference mentioned in the introduction, is that the clause was not intended to protect banks against runs, but against "duels." A second difference is that, because it is set in the Diamond and Dybvig framework with its privacy of information assumption, when triggered, the *theoretical* clause seems to cover all notes of the bank, in a similar fashion to suspensions orchestrated by clearinghouses during the *National Banking Era* in the US. The clause as it was practiced historically was, however, triggered individually, so as to discriminate between different classes of depositors. For the clause to be indeed triggered, and the accumulation of interest to begin, the bearer had to effectively been refused redemption and have had each of his notes signed and stamped. To understand the differences between the two versions of the clause, we will now turn to the historical context in which it emerged.

Fatherhood of the option clause is attributed to the *Bank of Scotland*, who used it extralegally three times in 1704, 1715, and 1728 (White 1995, p. 23; Gherity 1995, p. 718). It was not the only bank to use it, but it was the first to make it an explicit clause written on high denomination banknotes in 1730.[10] The *Bank of Scotland* was trying to protect itself against redemption runs orchestrated by its rival, the *Royal Bank of Scotland* rather than trying to fix an inherent instability in the banking system (Dowd 1991, pp. 769–770). The notes were issued with the optional clause that the bank could, at its discretion, pay the bearer within 6 months, given a compensation. In its official, explicit version it was used for the first time in 1762, when the Seven Years' War funding had created a drain of Scottish gold to London (Goodspeed 2014, p. 37).

The clause was adopted by more and more Scottish banks until it was made illegal in 1765, when a banking law required total and immediate convertibility of banknotes (White 1995, p. 26), favoring the three chartered banks who needed this clause less than their small competitors.

When first introduced, the clause was accepted by the customers of the *Bank of Scotland* despite the fact that the *Royal* refused to provide it, going as far as to boast their immediate convertibility on their banknotes. Selgin (1993, p. 357) notes that despite the fact that some Scottish banks banknotes circulated at a discount, the banknotes of the *Bank of Scotland* were accepted at face value during these periods of suspension, which he interprets as evidence that it brings net benefits. Goodspeed (2014, pp. 37–38) notes that inclusion of the clause by two smaller banks in 1756 was met "with considerable public support," and that over the 1762–1765, when the clause would effectively be invoked, observers would suggest themselves that more Scottish banks adopt the option clause (Goodspeed 2014, pp. 39–40). It seems to have been widely recognized that the clause provided benefits in terms of protection against predatory note redemption orchestrated by other banks.

However, discussing an option *clause* suggests that it is always explicit and contractual. This was not always the case. Indeed, the use in 1704 for a period of 6 months, and in 1715 and 1728 for periods of 8 months had been made

[9] Selgin mentions a third type of bank runs, simply suspending the activities of the clearing house. There is, however, no room for a clearing house in the Diamond and Dybvig framework, as there is only one bank.

[10] The *Bank of Scotland* was forced to do so because it was the only one to include in its bylaws the right to summary diligence upon convertibility of its notes (Gherity 1995, pp. 717–718), in contradiction with the informal ad hoc use of the clause.

without the prior approval of the bearers. Also, after the clause was made illegal in 1765, some banks continued to practice a version of it illegally and selectively toward ill-intentioned note bearers.

The Scottish free banking system was in a constant state of selective suspension of convertibility toward abusive redeeming. This selective suspension was, prior to 1762, de facto practiced by borderline dishonest "legally impermissible limitation of convertibility" maneuvers (Checkland 1975, p. 186). It seems to have been common knowledge that it was impossible for rivals and arbitrage-driven "specie lifters" to convert a significant amount of banknotes instantly, but that converting a small amount of notes in the normal course of business would be possible in most cases (Checkland 1975, p. 185).

Indeed, before the clause era, and also during and after, the Scottish banks used ingenious stalling tactics to deny the convertibility of their notes to predatory, high-volume redeemers, such as agents of rival banks and "specie lifters" looking to benefits from the premium on bills of exchange on London banks.[11] If the note bearer presenting an unusually large amount of notes for redemption "made serious trouble, the matter would be noted and they would find the obtaining of credit more difficult in future," "[a]t best they would get a little specie and perhaps bills on London" (Checkland 1975, p. 185), in an attempt to seemingly export the problem to another place of business. They would question their customer's loyalty toward their bank, or threaten to tighten their credit in the future. If the "specie lifter" would refuse, the banks would count deliberately slowly, carefully inspect each note, take very frequent breaks to go on imaginary errands (Meulen 1917, p. 103) or close shop early (Goodspeed 2014, pp. 36–37). As a last measure resort from 1762 to 1765, the clause would formally be invoked.

The *historical* clause, instead of allowing for a suspension of all conversions to specie, was mainly used to discriminate between classes of note bearers. From the introduction of the official version of the clause in 1730 until its illegalization, Goodspeed (2014, pp. 37–38) demonstrates that it was used mainly from 1762 against arbitrageurs trying to export specie and benefit from foreign opportunities. There are no evidence of banks actually using it even in the case of duels, as the threat of its use seems to have been sufficient.

This kind of discrimination between classes of depositors is impossible in the Diamond and Dybvig (1983) model. Indeed, unlike the case of duels where it is easy to identify inimical redeemers, if only because of the unusually large volume of notes presented for conversion compared to the regular customer, it is impossible to discover the true hidden consumption needs of Diamond and Dybvig depositors. If this kind of discrimination had been available, it would have allowed the bank to discriminate between depositors exercising their right of withdrawal for immediate consumption, and depositors withdrawing in fear for their bank's solvency, suspending convertibility only for the latter and thus avoiding the bad equilibrium with self-fulfilling bank runs. Additionally, in the context of the Diamond and Dybvig framework, once the tipping point has been reached, all depositors automatically run of the bank. It is not necessary to specify that claims have to be effectively redeemed for the clause to be triggered, hence the "announcement" version of the clause. The literature on the stability of free banking, because it responded directly to Diamond and Dybvig, had to internalize the terms of the debate and adopt these assumptions too, abandoning the selective and discriminatory features of the clause.

The Limits of Diamond and Dybvig's Significance

The literature on the clause, if interpreted as an alternative proposal to deposit insurance,[12] is problematic insofar as it is intended primarily for Diamond and Dybvig runs, which do not correspond to bank runs stylized facts.

First, unlike the model would suggest, bank runs are generally not responsible for the initial shock (Selgin 1992). Gorton (1988) studies the National Banking Era in the US, and finds that for each of the 7 crises he identifies, bank runs were rather the result of a previous event announcing a possible depreciation of banking assets.[13] Likewise, Calomiris and Gorton (1991) find that over 1875–1913 all banking panics (generalized run on all banks) happened within the quarter following an abrupt increase in business failures. Mishkin (1991) studies bank panics from 1857 to 1988, and finds that for all but that of 1873, panics occur well after the recession has started.

Second, banks that do go bankrupt because of a bank run are those that are pre-run insolvent. Banks that are solvent can generally borrow from other banks and other institutions, historically clearinghouses, have a large repertoire of possible solutions to help banks in crisis. While bank runs and associated liquidity problems can be aggravating factors, even in the worst bank panic episodes they are causes of bank failure only in exceptional circumstances (Kaufman 1987, 1988). Even in the most fruitful historical era in

[11] Goodspeed (2014, pp. 38–39) explains the mechanism through which individuals could write bills of exchange on a dummy in London, change it for bank notes including a premium, redeem the bank notes and quickly send the specie to the dummy in London before the bills' term. He likens the actual use of the clause to a private capital control measure to sterilize "hot money" flows.

[12] Either Diamond and Dybvig (1983) "deposit insurance" or genuine deposit insurance.

[13] This does not mean that problems on the liability side cannot accelerate these difficulties.

terms of bank panics and runs, the American *National Banking Era*, runs were a primary cause of failure in only one case out of 594 bank bankruptcies (Calomiris and Gorton 1991, p. 154). Calomiris and Mason (1997) study the banking panic of June 1932 in Chicago and find that no pre-run solvent banks failed. Reviewing this literature, Benston and Kaufman (1995, p. 225) conclude that "the policy implications of the Diamond and Dybvig model are not very useful for understanding the workings of the extant banking and payments system."

A third reason is that, most runs have in fact been partial "verification" runs. Depositors eventually figure out that the bank will likely survive the crisis, and runs stop. This is impossible in the Diamond and Dybvig framework; once initiated, the run must always go through and make the bank fail. Ó Gráda and White (2003) study a single bank from the 1850s in New York State. They investigate depositor behavior through individual account data, and particularly through the panics of 1854 and 1857. The bank survived both. They find that runs are not sudden, but involve a learning mechanism where random beliefs are progressively dropped, while behavior motivated by legitimate signals become more important over time. Panic does not displace learning in the market process.

Finally, if Diamond and Dybvig are correct, it should apply to all fractional-reserve banking systems without deposit insurance. But, as evidence by the US-centric literature cited, bank runs are much more common in U.S. history than elsewhere, and bank panics are specific to the American National Banking Era and attributable to bank regulation of that era, such as the ban on branch banking that made mergers with insolvent banks impossible, and the bond deposit system that limited emission at a critical time (Smith 1991). Bordo (1990, p. 24) compares bank panics internationally and comments that "the difference in the incidence of panics is striking." While over the 1870–1933, the US had four panics, there were none in Britain, France, Germany, Sweden, and Canada despite the fact that "in all four countries, the quantitative variables move similarly during severe recessions to those displayed here for the U. S." Table 2-1 in Schwartz (1988, pp. 38–39) report that from 1790 to 1927 the U.S. experienced 14 panics, while the Britain, the only other country with as many observations, experienced eight, all of them before 1867.[14]

Not only does Diamond and Dybvig imply that bank runs are much more costly than empirical evidence does suggest,[15] but the model also shroud their benefits. The next section explores those benefits, and what becomes of the clause if bank runs are motivated by bad news under uncertainty, rather than irrational random withdrawals.

Desirability of the Clause

Desirability of the *Theoretical* Clause

To understand the benefits of bank runs, and how the clause might interact with them, it is necessary to consider bank runs beyond the canonical model and focus on the origins of bank runs as asset side problems, as the stylized facts suggest, rather than one strictly related to the liability side of the balance sheet.

But first it should be recognized that in a static world with perfect information, but with more realistic assumptions about bank liabilities,[16] bank runs would not be a problem. Depositors would run when they anticipate a bank failure, and would force the closure of inefficient banks. Runs would then be optimal because banks would close before they have had time to realize losses on the capital of deposits (Gorton 1985).

Bank runs can only become a problem when information is less than perfect, and depositors can make errors. Indeed, the opacity of bank portfolios means that depositors are not fully informed about what the bank does with their deposits. They monitor their bank through imperfect and noisy signals. Bank runs result from a revision of expectations following bad news. Depositors ignore not only if their bank is concerned with the bad news, but also the severity of the impact. In Gorton (1988), for example, it is bad news about a macroeconomic indicator of business results, but we can also imagine the default of another bank having invested in similar assets, etc. It may not only be direct bad news about the bank, such as downgrade from a ratings agency, but also more distant bad news relating to third party risks to which the bank might be exposed.

In such a setting, not all bank runs are net costs but only bank runs on a solvent bank (false positive), and also the absence of runs on an insolvent bank (false negative), because these phenomena suggest Pareto dominated

[14] Note that in Schwartz's table over the period studied by Bordo (1990), the U.S. had three panics, while France, Germany, and Sweden had one. Canada and Britain had none.

[15] For an argumentation that costs of banking crises are in fact attributable to the prudential regulation supposed to limit these crises, see Kaufman (1987), Selgin (1994) and Calomiris (2010). See Demirgüç-Kunt and Detragiache (2002) for empirical evidence that it makes banking crises more likely.

[16] In the Diamond and Dybvig (1983) model, the deposit contract relates to a debt/equity hybrid which is very different from traditional banking contracts (Dowd 1992, pp. 111–112). The absence of a clear and distinctive class of equity supposes that the bank has no capital reserve. Long-term agents are therefore never protected by equity. Also consider that even in the good equilibrium the bank's liabilities always exceed its assets in the interim period. As this section will attempt to demonstrate, there is nothing mysterious or harmful in an insolvent bank experiencing bank runs.

situations. Depositors then err with respect to the "nirvāṇa" situation of omniscience. Even modern bank run models in the Diamond and Dybvig tradition, such as Goldstein and Pauzner (2005) or Gu (2011), think of bank run triggers in terms of a mechanism to cope with asymmetries of information rather than the unexplained 'sunspots' of the canonical model. Reasoning in terms of an information problem also opens the door to the possibility that depositors learn, and that mistaken running depositors can eventually be reassured about the capacity of their bank to repay depositors before they cause the bank to be insolvent.

On the other hand, "unmistaken" bank runs have important properties. Demandable deposits provide customers with a possible vote of no confidence *vis-à-vis* their bank (Calomiris and Kahn 1991) . It solves the agency problem of bank debt. The apparent fragility of demandable deposits contracts is a way for the bank to voluntarily submit to depositor veto (Diamond and Rajan 2001). By precipitating the failure of the bank, it prevents the bank from substituting assets for riskier but more rewarding equivalents.

This asset substitution can occur when banks are sheltered from loss. They then transfer more risk on creditors than initially agreed. In the case of "unmistaken" runs under a limited liability regime, asset substitution can be aggravated by the exercise of the clause. Because the bank is insolvent, shareholders and managers have nothing to lose anymore, equity having been depleted and the receivership being imminent. They have lost their status of residual claimants, and when liquidation comes they will not receive anything. If assets substitution takes place, risks that they would have supported themselves under normal solvency conditions are now shifted unto captive depositors.

These risks are more acute for banks than other firms because they encounter much more opportunities in the course of their normal business, involving daily refinancing and new loans (Flannery 1994). Financial institutions' assets are difficult to describe and contract on because of acute informational asymmetries. Contractual mechanisms usually employed in these circumstances such as covenants, where creditors describe the course of action the debtor must follow when approaching insolvency, generally do not work for them.

Unhampered bank runs can therefore be thought of as a "dirty" mechanism to "crowdsource" bank supervision, and force insolvent banks into insolvency resolution procedures. The relatively poorly informed actions of depositors force other financial institutions, with much greater analysis capabilities, to reassess the bank's robustness. In the case of false-positive runs, they will lend to these banks, and in cases of genuine insolvency let them fail, preventing asset substitution.

Several theorists of free banking have assumed that the clause is a net benefit because it allows the bank to avoid the false-positive runs that would precipitate insolvency. But not all runs are false positives, and in cases of true positives, there are no *a priori* reasons that the bank would not go bankrupt anyway, even exercising the clause. The clause does reduce the costs associated with the liquidation of assets necessary to satisfy depositors as they come, but does not affect losses associated with depreciation of assets.

That is to say that, first, it is not the case that the clause is sufficient to avoid insolvency and, second, it opens up the possibility for the clause to aggravate the bank's financial distress by sheltering it from the fail-safe of demandable debt. Even if the clause creates good *ex ante* incentives for depositors not to run on solvent banks, it is not *ex post* efficient in that insolvent banks can use the suspension of convertibility to engage in riskier activities and bet themselves out of insolvency unsanctioned.

The literature on the clause recognizes the possibility of moral hazard and proposes various solutions to align incentives. Gorton (1985) suggests a contract where the suspension of payments is contingent upon an audit by an external source. The bank would then have no incentive to invoke the clause to save time, since it would then not only repay depositors but also pay the costs of control. Elsewhere, it recognizes the problem of abuse of this clause outside of bank run contexts, and Selgin and White (1994, p. 1729) suggest that the clause could include a compensation, as the clause historically did in Scotland. For a sufficiently high interest rates, they suggest that it would not be in the interest of the bank to use the clause in non-emergency situations.

Both of these solutions are, however, only satisfactory within the Diamond and Dybvig framework where insolvency is an exclusive consequence of the bank run, and once the bank run is avoided, the bank has no reason to become insolvent. However, in the case of a bank that is pre-run insolvent, or about to be insolvent, the shareholders would have nothing to lose and could either pass these costs over to the depositors in case of bankruptcy, or capture the benefits in the case of a return to solvency. These solutions are further complicated by the existence of a manager whose incentives might be misaligned with both shareholders and creditors.

Dowd (1991, pp. 767–768) and Dowd (1988, p. 325) suggests that, in addition to compensation, the liability of the shareholders could be extended when the bank triggers the clause, imitating unlimited liability. This would indeed solve the problem of moral hazard, but we are not aware of the existence of any such system of modulating shareholder liability regime. Extended liability is a much more ambitious and challenging policy that has the potential to suppress bank runs with or without an option clause.

Outside of the Diamond and Dybvig framework, it is therefore not obvious at all that the clause is desirable.

Presumably, depositors would not accept a deposit contract including this clause.

Desirability of the *Historical* Clause

We, however, go further and argue that it is also the case that the *historical* clause is not desirable outside of the context of its emergence.

It is true that the clause was accepted in the Scottish free banking system, as well as other systems of free banking or almost-free banking systems,[17] while according to our above analysis it would be against their own interest. When these clauses were adopted, there was no systematic abuse of the clause to prevent *bona fide* depositors from redeeming their notes in the normal course of business. In fact, there even seems to be no anecdotal abuses of the clause. There are (at least) two reasons for this, resulting from the institutional context of free banking, limiting the abuse of the clause and making it desirable.

In its historical context

First, there was no abusive use of the clause because there were numerous banks of issue. Banks were competing for the use of their banknotes as money, and this was the most important source of bank profit.[18] Abusing the clause would have affected the reputation of the bank and the acceptability of its notes.

Second, in most free banking systems, shareholders were subject to extended or unlimited liability. This was the case in Scotland, and while the three chartered banks had the privilege to depart from this rule, Goodspeed (2014, pp. 41–42) suggests (1) they did not invoke it against *bona fide* customers and restricted its use to high denomination notes and (2) the chartered banks made up for it by substantially lower leverage levels (Goodspeed 2014, p. 86). This means that risk-shifting unto depositors was made impossible by liability rules, and shareholders had to face the consequences of their actions on their own private wealth. They did not have an incentive to engage in asset substitution when nearing insolvency. For the few banks for which this was not the case, competition with their unlimited liability counterpart pushed them to hold more equity. Moreover, the risk of being liable beyond your equity in the company is internalized by the shareholders, and induces a less risky bank behavior, and indeed there were relatively few bank failures in free banking systems.[19]

This institutional environment therefore limited the moral hazard associated with the clause ensuring it is not abused when it is not needed, making managers and shareholders liable even when the bank has exhausted its equity. The *historical* clause was therefore desirable in this context.

In a modern context

The clause does not appear, however, to be desirable under central banking as a substitute for deposit insurance.

Under free banking, competition in currency atoned the moral hazard created by the clause. By definition, competition for banknote usage is absent from modern systems with monopolies of emission. Bank runs and the clause, in this context and in the absence of deposit insurance, would rather take the form of deposits and notices of withdrawal.[20] Or, as Evans (2013) puts it, "quasi notices of withdrawal," to underline that in their current iteration they do not include a financial compensation.

However, the comparison stops here. Deposits in modern banking are not a primary source of bank profits. Banks offer a menu of other services, such as payment systems, which are the focus of competitive rivalry and the biggest source of profit. The reputation with regard to deposit services is relatively not as important as reputation was to bank notes in free banking systems. The consequences of abusing the Notice of Withdrawal would not be as directly threatening to the bank's livelihood as abuse of the clause would have been. A serious concern for moral hazard remains. This means that demandable/runnable claims in modern banking are not as discipline-inducing as demandable/runnable claims were in free banking systems.

The second feature of free banking systems to make the clause acceptable to depositors was extended liability, and/or high capitalization levels. Both are absent under modern central banking systems. This prompts us to conclude that

[17] See *supra* note 3.

[18] While gold and/or silver were the units of account, individual banks issued bank notes payable to bearer. The banking profit on this type of activity was based on the acceptability and use of its banknotes. The more they were accepted and used, the more banknotes took time to come back to the issuing bank for conversion to base currency. The bank then had a vested interest that its notes be widely and generally accepted, as it allowed it to allocate assets to more lucrative venues than reserves.

[19] See Goodspeed (2014, pp. 99–135) and White (1995, pp. 27–29) for the Ayr Bank failure, the "exception that confirms the rule" in the loose rhetorical sense.

[20] Under a regime of multiple banks of issue, a run on deposits is generally not a problem. The bank can respond by transforming deposits into bank notes without affecting its solvency or even requiring to sell assets, and there are no *a priori* reasons that a run on bank money should transform to a run on the base currency (see Selgin 1988, p. 134). In comparison, monopolies of issue like our modern central banking systems eliminate the possibility of offering banknotes to depositors because it requires central bank cash on hand. So while modern central banking systems have no use for the option clause on *banknotes*, the case has been made that option clauses might be applicable to *deposits* in those systems.

notices of withdrawals, as an alternative to deposit insurance for modern banks, are much less desirable as a run-proofing mechanism than the clause was to free banks. If they were used in the event of a run, moral hazard could have the effect of diminishing the creditors' claim on the bank by creating a venue for risk-shifting.

That is not to say that "quasi option clauses" and "quasi withdrawal notices" might not be desirable for other reasons than run-proofing. Their function is to provide a way for modern bank branches to respond to unusually large withdrawal demands, and in that respect they might very well be desirable.

Where they have been used, they have not been a significant source of moral hazard or even invoked much during banking crises. This is not inconsistent with our claim that if used as a run-proofing device in the absence of deposit insurance, it would lead to significant moral hazard. In the current context, runs are already being suppressed by deposit insurance. There is no need for banks to shift risks upon depositors, as these risks are already being supported by taxpayers through deposit insurance funds. Ironically, it is deposit insurance that makes these quasi-clauses acceptable under modern central banking systems.

Conclusion

Table 1 summarizes the contexts where the clauses are desirable. We have argued that self-fulfilling prophecy runs are a theoretical curiosity with little consequence for the analysis of banking policies, and that the institutional framework where duels appeared no longer exists. Remains the possibility of bank runs through asymmetry of information, where either clauses are not desirable due to asset substitution.

Discrepancies between the clause in its theoretical incarnation and its historical evidence are many. The *theoretical* clause concerns self-fulfilling prophetical runs, while the *historical* concerns note redemption duels. The *theoretical* applies uniformly to all note bearers or depositors through an announcement, while the *historical* was used to discriminate between predatory and *bona fide* note bearers. Either kind of clauses are not desirable in the current institutional context because it dispenses with the informational and incentive benefits of runs, and aggravates the agency problem of bank debt.

Table 1 Contexts in which option clauses are desirable

	Clauses	
	Historical	Theoretical
Contexts		
Diamond and Dybvig framework	Non-applicable	Yes
Free banking duels	Yes	No
Central banking w/ltd-liability	No	No

Therefore, it should not be considered as a private alternative to other prudential policies "as is" outside of its original institutional context, as it faces several challenges. If bank runs are asymmetry of information driven, rather than a problem of self-fulfilling prophecies, and runs alone virtually never lead banks to failure, the option clause immediately becomes much less interesting, or even necessary. Bank runs would be made more harmful by the exercise of the clause.

Finally, in the absence of deposit insurance, but in an otherwise modern institutional context with a central bank, other institutions could manage these situations, such as the automatic stay of corporate bankruptcy laws. Once initiated, it prevents creditors from exercising their claims on the bank's assets, much like the option clause. But unlike the clause, bankruptcy's automatic stay limits the possibility of asset substitution by submitting the bank's decisions to *ex ante* judicial review, and prevents all creditors (not just depositors) from 'running' on the bank. This would allow time for the bank to find a viable solution out of illiquidity and/or insolvency, and a forum where creditors are represented.

In light of these problems, it is not clear that the option clause has a practical application to modern banking stability problems. In a modern institutional context with a central bank, it is not in the interest of depositors to rely on the terms of a deposit contract including an option clause to suspend the convertibility for financial stability.

We, however, insist on the fact that this does not mean that the case for the clause is without merit, as it demonstrates that even in the flawed framework of Diamond and Dybvig, very simple institutions can solve financial instability problems. The argument in favor of the clause as a private prudential measure is consistent if it is a simple illustration suggesting that learning and innovation can take place, and that relatively simple contractual solutions may emerge and solve market failures. We also contend that the clause is perfectly fit to respond to the literature opposed to fractional reserve banking, when these oppositions are based on an inherent financial instability argument. Often the framework is, consciously or not, reminiscent of Diamond and Dybvig self-fulfilling prophecies.[21]

This article has also implicitly suggested, without addressing the question directly, that it might not even be necessary to replace deposit insurance with a specific alternative institution.

[21] For responses to this literature using the clause, see Selgin and White (1996), Cachanosky (2011), Evans (2013), and Evans (2014).

Acknowledgments The author would like to acknowledge that this research was partially conducted with the support of a Bradley Foundation Doctoral Fellowship for a visit at New York University.

References

Benston, G. J., & Kaufman, G. G. (1995). Is the banking and payments system fragile? *Journal of Financial Services Research*, *9*(3–4), 209–240.

Bordo, M. D. (1990). The lender of last resort: Alternative views and historical experience. *FRB Richmond Economic Review*, *76*(1), 18–29.

Cachanosky, N. (2011). A comment on Barnett and Block on time deposit and Bagus and Howden on loan maturity mismatching. *Journal of Business Ethics*, *104*(2), 219–221.

Calomiris, C. W. (2010). The great depression and other 'contagious' events. In A. N. Berger, P. Molyneux, & J. O. S. Wilson (Eds.), *The Oxford handbook of banking* (pp. 693–710). Oxford: Oxford University Press.

Calomiris, C. W., & Gorton, G. (1991a). The origins of banking panics: Models, facts, and bank regulation. In R. G. Hubbard (Ed.), *Financial markets and financial crises* (pp. 109–174). Chicago: University of Chicago Press.

Calomiris, C. W., & Kahn, C. M. (1991b). The role of demandable debt in structuring optimal banking arrangements. *The American Economic Review*, *81*(3), 497–513.

Calomiris, C. W., & Mason, J. R. (1997). Contagion and bank failures during the great depression: The June 1932 Chicago banking panic. *The American Economic Review*, *87*(5), 863–883.

Checkland, S. G. (1975). *Scottish banking: A history, 1695–1973*. Glasgow: Collins.

Cowen, T., & Kroszner, R. (1989). Scottish banking before 1845: A model for Laissez-Faire? *Journal of Money, Credit and Banking*, *21*(2), 221–231.

Demirgüç-Kunt, A., & Detragiache, E. (2002). Does deposit insurance increase banking system stability? An empirical investigation. *Journal of Monetary Economics*, *49*(7), 1373–1406.

Diamond, D. W., & Dybvig, P. H. (1983). Bank runs, deposit insurance, and liquidity. *Journal of Political Economy*, *91*(3), 401–419.

Diamond, D. W., & Rajan, R. G. (2001). Liquidity risk, liquidity creation, and financial fragility: A theory of banking. *Journal of Political Economy*, *109*(2), 287–327.

Dowd, K. (1988). Option clauses and the stability of a Laisser Faire monetary system. *Journal of Financial Services Research*, *1*(4), 319–333.

Dowd, K. (1991). Option clauses and bank suspension. *Cato Journal*, *10*(3), 761–773.

Dowd, K. (1992). Models of banking instability: A partial review of the literature. *Journal of Economic Surveys*, *6*(2), 107–132.

Dowd, K. (1993). *Laissez-faire banking*. New York: Routledge.

Dowd, K. (1996). *Competition and finance : A reinterpretation of financial and monetary economics*. New York: St. Martin's Press.

Evans, A. J. (2013). In defence of 'Demand' deposits: Contractual solutions to the Barnett and Block, and Bagus and Howden Debate. *Journal of Business Ethics* (in press).

Evans, A. J. (2014). What is the Latin for 'Mayonnaise'? A response to Bagus, Howden and Gabriel. *Journal of Business Ethics* (in press).

Flannery, M. J. (1994). Debt maturity and the deadweight cost of leverage: Optimally financing banking firms. *The American Economic Review*, *84*(1), 320–331.

Gherity, J. A. (1995). The option clause in Scottish banking, 1730–65: A reappraisal. *Journal of Money, Credit and Banking*, *27*(3), 713–726.

Goldstein, I., & Pauzner, A. (2005). Demand-deposit contracts and the probability of bank runs. *The Journal of Finance*, *60*(3), 1293–1327.

Goodspeed, T. B. (2014). *Upon Daedalian wings of paper money: Adam Smith, free banking, and the financial crisis of 1772*. PhD thesis. Harvard University. http://scholar.harvard.edu/files/goodspeed/files/goodspeed_upon_daedalian_wings_-_chapters_1-5_march_2014.pdf (visited on 09/07/2014).

Gorton, G. (1985). Bank suspension of convertibility. *Journal of Monetary Economics*, *15*(2), 177–193.

Gorton, G. (1988). Banking panics and business cycles. *Oxford Economic Papers*, *40*(4), 751–781.

Gu, C. (2011). Herding and bank runs. *Journal of Economic Theory*, *146*(1), 163–188.

Hammond, B. (1957). *Bank and politics in America. From the revolution to the Civil War*. Princeton: Princeton University Press.

Jonung, L. (2000). *The economics of private money: Private bank notes in Sweden 1831–1902*. Unpublished manuscript (in preparation).

Kaufman, G. G. (1987). The truth about bank runs. In C. England & T. Huertas (Eds.), *The financial services revolution. Policy directions for the future* (pp. 9–40). Boston: Kluwer Academic Publishers and Cato Institute.

Kaufman, G. G. (1988). Bank runs: Causes, benefits, and costs. *Cato Journal*, *7*(3), 559–587.

Meulen, H. (1917). *Industrial justice through banking reform; An outline of a policy of individualism*. London: R.J. James.

Mishkin, F. S. (1991). Asymmetric information and financial crises: A historical perspective. In R. G. Hubbard (Ed.), *Financial markets and financial crises* (pp. 69–108). Chicago: University of Chicago Press.

Ó Gráda, C., & White, E. N. (2003). The panics of 1854 and 1857: A view from the emigrant industrial savings bank. *The Journal of Economic History*, *63*(1), 213–240.

Schuler, K. (1992a). Free banking in Canada. In K. Dowd (Ed.), *The experience of free banking* (pp. 79–92). New York: Routledge.

Schuler, K. (1992b). The world history of free banking: An overview. In K. Dowd (Ed.), *The experience of free banking* (pp. 7–47). New York: Routledge.

Schwartz, A. J. (1988). Financial stability and the Federal Safety Act. In W. S. Haraf & R. M. Kushmeider (Eds.), *Restructuring banking and financial services in America* (pp. 34–62). Washington: American Enterprise Institute.

Selgin, G. A. (1988). *The theory of free banking*. Totowa, NJ: Rowan & Littlefield.

Selgin, G. A. (1992). Bank lending 'Manias' in theory and history. *Journal of Financial Services Research*, *6*(2), 169–186.

Selgin, G. A. (1993). In defense of bank suspension. *Journal of Financial Services Research*, *7*(4), 347–364.

Selgin, G. A. (1994). Are banking crises free-market phenomena? *Critical Review*, *8*(4), 591–608.

Selgin, G. A., & White, L. H. (1987). The evolution of a free banking system. *Economic Inquiry*, *25*(3), 439–457.

Selgin, G. A., & White, L. H. (1994). How would the invisible hand handle money? *Journal of Economic Literature*, *32*(4), 1718–1749.

Selgin, G. A., & White, L. H. (1996). In defense of Fiduciary media–or, We are *Not* Devo(lutionists), We are Misesians!. *The Review of Austrian Economics*, *9*(2), 83–107.

Selgin, G. A., & White, L. H. (1997). The option clause in Scottish banking: A comment. *Journal of Money, Credit and Banking*, *29*(2), 270–273.

Shah, P. J. (1997). The option clause in free-banking theory and history: A reappraisal. *The Review of Austrian Economics*, *10*(2), 1–25.

Smith, B. D. (1991). Bank panics, suspensions, and geography: Some notes on the "Contagion of Fear" in banking. *Economic Inquiry*, *29*(2), 230–248.

Vaubel, R. (1984). Private competitive note issue in monetary history. In P. Salin (Ed.), *Currency Competition and Monetary Union* (pp. 59–73). The Hague: Martinus Nijhoff.

White, L. H. (1992). Free banking in Scotland before 1844. In K. Dowd (Ed.), *The Experience of Free Banking* (pp. 157–186). London: Routledge.

White, L. H. (1995). *Free banking in Britain: Theory, experience, and debate, 1800–1845* (2nd ed.). London: The Institute of Economics Affairs.

White, L. H. (1999). *The theory of Monetary institutions*. Malden, MA: Blackwell Publishers.

Yeager, L. B. (1993). Rev. of Laissez-Faire Banking, by Kevin Dowd. *The Manchester School of Economic & Social Studies*, *61*(3), 321–323.