Hysteresis in Banking and Financial Industry

Theory, Evidence & Policy

Author: Jozef Patrnciak
Department of Economics, Erasmus School of Economics, Erasmus University Rotterdam

August 25, 2017

First supervisor: Professor Dr. Bas Jacobs
Second reader: Professor Dr. Casper De Vries

Abstract

Hysteresis, the historic memory or path dependency, is often informally mentioned as an important determinant of banking concentration. However, until now in the area of general banking competition, hysteresis did not receive proper formal treatment either theoretically or empirically. The main objective of this research is to fill this gap. Theoretical foundations for introducing hysteresis into the banking industry will be laid by generalising the popular Monti-Klein model of banking competition, to allow for both fixed entry and per-period cost and extending it to quasi-dynamic setting. As will be shown, such a model implies multiple path dependent equilibria of banks' activities. The implications of the theory of hysteresis developed herein are tested by applying unit root tests to the data. These generally find some evidence in support of the presented theoretical model. Implications of these findings for policy are briefly discussed as well.

Key words: hysteresis, Monti-Klein model, microeconomics of banking, unit root testing.
JEL classification: G21; D20; L10.

Acknowledgements: I would like to thank Professor Dr. Bas Jacobs and Professor Dr. Casper de Vries for their useful remarks and comments. All errors herein are the sole responsibility of the author.

Contents

1 Introduction

The role of financial intermediation in the Great Recession sparked renewed interest in the study of the determinants of concentration and competition in the financial industry. Hysteresis, the historic memory or path dependency, is often informally mentioned as an important determinant of banking concentration. However, until now in this area, hysteresis did not receive proper formal treatment either theoretically or empirically[1]. The main objective of this research is to fill this gap. Theoretical foundations for introducing hysteresis into the banking industry[2] will be laid by generalising the popular Monti-Klein model of banking competition, to allow for both fixed entry and per-period cost and extending it to quasi-dynamic setting. As will be shown, such a model implies multiple path dependent equilibria of banks' activities. Subsequently, implications of the theory of hysteresis developed herein are tested by applying unit root tests to the data. These generally find some evidence in the support of the presented theoretical model. Implications of these findings for policy are discussed as well. To be more specific, this model suggests that policy makers should adopt more dynamic regulatory approach.

This inquiry is relevant because banks and more broadly financial intermediaries are one of the most important entities in the modern economy. Empirical studies show that financial sector has first order impact on economic growth (see Levine, 1997, 2005), that they disproportionately contribute to the raising standards of living for poor (see Beck et al., 2007, Honohan, 2004), and that small firms get the largest share of positive benefits brought about by financial development (see Beck and Demirguc-Kunt, 2006). The current literature was perhaps best summed up by Merton (1993) who argues that "a well developed smoothly functioning financial system facilitates the efficient life-cycle allocation of household consumption and the efficient allocation of physical capital to its most productive use in the business sector."

Thus, there is no doubt that a healthy banking sector is necessary for sustainable economic progress. However, what constitutes a healthy banking sector is still a matter of ongoing controversy (Berger et al., 2004). Allen and Gale (2000) argue

---

[1] While some scattered remarks can be found across the wider literature, to the best knowledge of the author, no concise treatment can be found. For example, popular handbook on microeconomics of banking by Freixas and Rochet (2008) does not provide treatment of this topic whereas comparable handbook of applied international trade theory by Bowen et al. (2012) devotes a whole section to the concept. This is not a criticism of the handbook of Freixas and Rochet, rather this helps to illustrate the gap and lack of previous treatment of this topic.

[2] Where banking industry and banks are broadly considered as all financial intermediaries that turn deposits (of various forms) into loans and investments.

that higher concentration of banking industry enhances financial stability. According to proponents of this view, higher concentration leads to higher market power and profits. The latter can serve as a buffer during negative shocks. Moreover, they argue that it is easier to regulate few banks rather than many. On the opposing side, Boyd and De Nicolo (2005) argue that increased concentration and market power leads to higher commercial interest rates under which only riskier projects apply for funding. Furthermore, it might not necessarily be true that regulating smaller number of banks is easier as it might be disproportionally more complex. In addition, numerous scholars point out that more concentrated banks are protected by 'too big to fail' implicit guarantees leading to higher-risk taking incentives and moral hazard (see Boyd and Runkle, 1993, Labonte, 2015, Mishkin, 1999, O'hara and Shaw, 1990, Stern and Feldman, 2004).

The empirical evidence is mixed as well (Berger et al., 2004). Beck (2007) shows that more concentrated banking sectors are less prone to systemic crises, especially if a country has developed and competition promoting institutions. But the author points out that this seems contradictory and should be interpreted with caution. Empirical study of Nicoló et al. (2004), using different proxy for systemic risk, finds that the probability of fail actually increases with bank concentration. Furthermore, Boyd and De Nicolo (2005) argue that the literature in favour of the view that concentration improves financial stability is quite weak. For example, Yeyati and Micco (2007) find that in the wake of the Great Recession, the banking industry in the U.S. was so concentrated that even a failure of one 'megabank' would overwhelm the resources available to the Federal Deposit Insurance Corporation (FDIC).

Despite the mixed evidence, one thing that both the concentration-stability and the concentration-instability camp can agree on is that, concentration and competition play a crucial role in determining the systemic stability. Given this, it is surprising how little attention is paid to systematic analysis of impacts that hysteresis has on the concentration and competitiveness of banking industry. Most authors attribute the increased concentration either to regulation or technological change (see Barth et al., 2008, Freixas and Rochet, 2008, Yeyati and Micco, 2007). These assertions are not necessarily wrong, but they do not capture the whole story.

The aim of this research is to fill this gap by examining the exit and entry decisions of financial intermediaries. This work will show that both the concentration and competitiveness in the banking industry are not dependent only on its present characteristic but are also path dependent. This result will be derived using generalised Monti-Klein model of banking industry, that includes both per-period fixed costs as well as entry costs. Moreover, implications of the aforementioned model will be tested using unit root tests. Since these tests generally find evidence in support of the hysteresis hypothesis, policy implications are discussed as well.

This work is organized as follows. The second chapter provides a brief overview of hysteresis and previous literature, theoretical and empirical alike. Third chapter proves few useful propositions which show that hysteresis is present in the banking industry. The fourth chapter provides empirical evidence supporting the theory developed in this paper. The fifth chapter discuses the policy implications of hysteresis. Finally, the sixth chapter offers conclusive remarks.

2 Hysteresis in Past and Present

This section provides a brief summary of the past research on hysteresis. Unfortunately, as mentioned earlier, past literature on hysteresis in banking is virtually non-existent. Therefore, this section will draw upon previous research from related fields such as international or labour economics.

The concept of hysteresis was originally developed in physical sciences, to describe systems with historic memory, and it appeared in economics only in late 1960s. According to Franz (1990), its explicit use in the economic science can be traced back to Samuelson (1965) and other forerunners including Georgescu-Rogen (1971), Phelps (1972) and Kemp and Wan (1974). However, Cross and Allan (1988) argue that the idea of hysteresis was already at least implicitly present in the thought of Frisch, Kaldor and Schumpeter. Arguably the concept was already implicit even in the writings of Alfred Marshall[3]. The theory of hysteresis was further developed mainly in the field of international economics and labour economics[4]. International economics especially offers very rigorous theories of hysteresis, and this is mainly thanks to the work of Baldwin, Dixit and Krugman (Franz, 1990).

The usual narrative of these models is that the entrant firms face firm-specific and non-recoverable entry costs[5]. This set up makes the entry 'investments' sunk ex-post. Because of this, entry is worth while only if market prices sufficiently exceed the unit costs. Therefore, a temporary increase in the price may lead to an entry but a subsequent return to 'normal' will not lead to exit as long as variable costs are covered. Thus there are multiple equilibria of economic activity dependent on the past shocks, or in other words hysteresis (Belke et al., 2014).

For example, Dixit (1989) examines how optimal entry and exit decisions of a firm under uncertainty, especially price uncertainty, can lead to hysteresis in the particular industry. He finds that even relatively small 'sunk cost' can have large hysteresis effects. Later, Dixit (1992) showed that there is inverse relationship between price volatility and probability of hysteresis. The rationale for this is that in a more volatile environment, agents have more incentive to try to wait for more favourable outcome.

Baldwin (1988) developed a model where temporary exchange rate fluctuations can have hysteresis effects on quantities, prices and also on the number of firms active

---

[3] See Marshall (1890). Principles of Economics, pp. 425-428. Marshall writes about tastes that are acquired during disturbances, and then never forgotten, nor implications of those for productivity. These ideas implicitly presuppose hysteresis.

[4] In recent years the concept of hysteresis was also extended to macro economics (see DeLong and Summers, 2012).

[5] These are often market specific. Some examples include creating brand awareness (or advertising more broadly) and the set up costs of distribution networks.

in a particular trade. Baldwin goes even further and tries to provide some empirical support for his model. Specifically, he formulates several implications of his model, such as that hysteresis should cause structural break which would be accompanied by increased absolute value of elasticity in the aggregate import demand. Afterwards he tests these implications using two-stages least square (2SLS) approach. Unfortunately, his empirical findings provide only very tentative support for the presence of hysteresis. Later this model was further extended by Krugman and Baldwin (1989).

In the field of labour economics, hysteresis was used by Blanchard and Summers (1986) as an explanation for the persistent high unemployment levels in Europe. They focused especially on hysteresis caused by insider-outsider bargaining between workers in particular industry, and showed that insiders have incentive to maintain the level of unemployment from previous periods. Moreover, their empirical findings indirectly seem to be supporting their story.

Later DeLong and Summers (2012) popularized the concept in the field of macroeconomics too. These authors developed a very influential macroeconomic model which showed that if there is a considerable amount of hysteresis in the economy, if multipliers are higher than during normal times and if economy is at or near zero lower bound, then debt financed expansionary policy may be in fact self-financing. Insofar this paper shows the importance of studying and understanding the phenomenon of hysteresis in designing optimal macroeconomic policy. However, some contributions to this topic were already made earlier by Buiter (1987) and others.

The theory of hysteresis in banking industry developed in this work is mainly built upon the previous research from international economics. This work is especially influenced by the works of Baldwin (1988, 1989) and Dixit (1989, 1992) who provided foundations for the modern literature on hysteresis in general. Nevertheless, the rest of the aforementioned literature provided valuable lessons as well. However, this work diverges from the Dixit modelling approach by nesting the idea of hysteresis in the popular Monti-Klein model. This is done to make the idea more compatible with the industrial organization approach to banking competition.

During the late 1990s and early 2000s hysteresis literature shifted its focus disproportionally more on empirics. The reason for this intellectual shift is that while the concept of hysteresis is theoretically attractive, empirical research on it prior to late 1990s was very tentative. This was by no means due to lack of effort or ingenuity on the part of the above mentioned scholars. Rather at the time there were virtually no tools that would make it possible to directly test for the presence of hysteresis (Belke et al., 2014).

As a matter of fact, to the present day, literature has not settled on a single way of testing the hysteresis hypothesis (Belke et al., 2014). This is because such testing is not as straightforward as it might seem. Although hysteresis is, simply put, just dependence of an economic variable on its initial conditions and past realisations, it is generally not possible to test these by standard 'naive' ordinary least squared (OLS) regression. There are several reasons for this.

To start with, most economic variables exhibit certain short run dynamics or persistence that shows as autocorrelation (Verbeek, 2008). However, hysteresis is not, nor should be confused with, a simple autocorrelation. As long as autocorrelation declines over time, we can talk of persistence but not hysteresis as in such cases past events only influence present ones but not determine them. Hysteresis, on the other hand, by its definition implies that the autocorrelation function contains a unit root[6].

The difference can perhaps be best seen by looking at the mathematical representations of these different processes. Following Verbeek (2008) and Hamilton (1994), consider an $AR(1)$ model with a deterministic linear trend of a variable N_t

$$N_t = aN_{t-1} + \delta + \psi t + \epsilon_t,\ t = 1, 2, ..., T, \tag{1}$$

where $|a| < 1$, and N_0 is the initial value. The moving average (MA) representation of the solution for N_t is

$$N_t = a^t N_0 + \mu + \mu_1 t + \epsilon_t + a\epsilon_{t-1} + a^2\epsilon_{t-2} + ... + a^3\epsilon_{t-T} \tag{2}$$

with the mean being constant around the trend,

$$E[N_t] = a^t N_0 + \mu + \mu_1 t \to \mu + \mu_1 t \text{ as } T \to \infty. \tag{3}$$

Similarly, the variance is constant as well

$$V[N_t] = V[\epsilon_t + a\epsilon_{t-1} + a^2\epsilon_{t-2} + ... + a^3\epsilon_{t-T}] = \sigma^2 + a^2\sigma^2 + a^4\sigma^4 + ... = \frac{\sigma^2}{1-a^2}. \tag{4}$$

The deviations from the mean of this process $N_t - E[N_t] = N_t - \mu - \mu_1 t$ are a trend stationary mean reverting process. Note that while the past events influence

---

[6] Some authors use the term 'partial hysteresis' for persistent processes, and 'pure hysteresis' for the unit root processes (see O'Shaughnessy, 2000). Others prefer to confine the term hysteresis only to non-linear systems (see Amable et al., 1995, Archibald, 1995, Cross, 1995, Cross and Allan, 1988, Cross, 1987). Moreover, these authors would prefer to apply the term persistence to unit roots. However, such terminology is quite confusing because the term persistence is already commonly applied in economics to autocorrelated processes, where past realisations do not determine the present ones (Belke et al., 2014). To avoid further confusion, this paper uses terms 'persistence' and 'hysteresis' following Franz (1987), to denote autocorrelated and near unit root processes or unit root processes respectively.

current ones they do not determine them. By the Frisch-Waugh-Lovell (1933, 2008) theorem, such relationship can be estimated by OLS with a trend term but as it does not describe hysteresis it is of no interest to us.

However, in the case where $a = 1$ the AR(1) model takes a different form

$$N_t = N_{t-1} + \delta + \epsilon_t, \; t = 1, 2, ..., T. \tag{5}$$

In such case the autoregressive polynomial $a(L) = (1 - L)$ contains a unit root, making $a(L)$ not invertible. The solution for N_t is then given by

$$N_t = N_0 + \sum_{i=1}^{t} \Delta N_i = N_0 + \sum_{i=1}^{t} (\delta + \epsilon_i) = N_0 + \delta t + \sum_{i=1}^{t} \epsilon_i. \tag{6}$$

where N_0 is again the initial value. Also, the moments are given by (Verbeek, 2008)

$$E[N_t] = N_0 + \delta t \text{ and } V[N_t] = t\sigma^2. \tag{7}$$

In this case, broadly speaking, the *AR* process actually exhibits hysteresis (Franz, 1987). As the equation 6 shows, the present values are determined by the initial conditions and by the sum of past shocks. As can be seen from the expressions for mean and variance, standard OLS cannot be applied in this case. These show that, in this case, both mean and variances are not constant but rather time dependent and ill defined. Not to mention that errors of such regression would be non standard and dependent on the exact specification, save for special cases. Therefore OLS cannot be generally applied at all as the major Gauss-Markov assumptions are violated (Verbeek, 2008). In this case the OLS is inherently biased and cannot be as easily salvaged as when only heteroskedasticity or autocorrelation is present[7].

An important caveat here is that the unit root processes, where all past shocks determine the present value, are a special case of hysteresis. Because of this, some authors such as Blanchard and Summers (1986) do not shy away from extending the term hysteresis to include near unit root processes as well (i.e. AR processes with the (sum of) coefficient(s) having value less but close to unity). Hysteresis was originally developed to describe systems with 'selective memory', where only certain shocks or certain types of shocks have permanent effect. This broader family of non-linear systems with hysteresis can be analysed using techniques developed by Mayergoyz (1991), Krasnosel'skii and Pokrovskii (1989). In these models the current state of an economic variable depends on non-dominated extremum values of past

---

[7] An interested reader can find more detailed discussion and proofs in Verbeek (2008) and Hamilton (1994) (or any other standard econometric handbook).

shocks. However, the sharp dichotomy based on stationarity and non-stationarity of a process makes exposition easier and more accessible[8]. Moreover, due to their highly non linear nature, these models cannot be generally estimated using the standard ‘naive’ OLS techniques as well.

The methods for dealing with these problems did not enter mainstream until recently and that is why prior to 2000s empirical research on hysteresis was scarce. The ‘traditional’ way of testing the hysteresis hypothesis prior 2000s was very indirect. For example Baldwin showed that his model postulates that hysteresis would cause structural break accompanied by increased absolute value of elasticity in the aggregate import demand. Afterwards, these implications were tested using 2SLS. Nonetheless, as the author himself points out there are no reasons to believe that hysteresis implies an one-time structural break. Moreover, the elasticity in the aggregate import demand could also change due to hysteresis unrelated reasons. Thus, by its own nature, this was also very tentative approach without a wider applicability. Thanks to recent advances in the time series analysis, this way of testing for hysteresis is on the decline.

Currently, the empirical literature on hysteresis can be divided into two major competing approaches. The ‘structural approach’ adopts the more nuanced view of hysteresis, and focuses on estimating non-linear models that are built upon the field specific theories of hysteresis. On the other hand, unit root approach adopts the sharp dichotomy described earlier in this section and simply applies unit root tests to the series.

The structural approach is not easy to summarize as it encompasses diverse range of empirical strategies. This makes an exhaustive review of these methods impossible here but interested readers can find more details in the sources cited herein. The simpler variants of this approach use a non-linear extension of more sophisticated OLS models. Non-linear modelling is necessary as hysteresis, by its definition, implies that the underlying relationship is non-linear or asymmetric. This is because in systems with historic memory, the effect of explanatory variables vary depending on both the magnitude of current changes as well as on the past realisations of the variables.

Prominent examples from this category are the non-linear autoregressive distributed lag (NARDL) models (Belke et al., 2014, Shin et al., 2011). The standard autoregressive distributed lag (ARDL) models are an extension of OLS that includes

---

[8] Such a sharp division admittedly comes at a cost of losing some nuance. However, as will be seen later, it is not unusual in applied literature as the more nuanced view of hysteresis can often be too impractical. Especially the more common unit root approach for testing hysteresis is based on this dichotomy.

lags of the dependent variable and NARDL also allows for non-linearities or asymmetries in the effect of explanatory variables over time. These get around the problem of non-stationarity by transforming all variables into their difference-stationary form and examining hysteresis through testing implications of a theoretical hysteresis model from their field. These can also be extended to test for hysteresis more directly through cointegration and error correction. More specifically, the bounds test approach developed by Pesaran et al. (2001) allows for inclusion of non-stationary terms alongside of their difference-stationary representations. Thanks to this, both short and long run dynamics can be examined consistently. For example, Verheyen (2013) applied the NARDL model to export demand focusing on the potential non-linear hysteresis effects of exchange rate that should be present in the data. Under hysteresis exporters should not adjust their prices as a response to every change in the exchange rates. The non-linearity is introduced by first expressing the original exchange rate as three partial sums, a sum that captures positive exchange rate movements above selected threshold, a sum that captures the negative movements under another threshold and a sum that allows for both negative and positive ('small') movements in between these thresholds. However, Verheyen fails to find a robust and significant hysteresis effect present, which might be due to the fact that the exchange rate changes he studied were too small for inducing hysteresis.

Some authors try to examine hysteresis by extending the 'classical' linear cointegration methods such as the Engle and Granger (1987) or Johansen (2002) models. These models take advantage of the fact that even non-stationary variables may revert to common stochastic trend or equilibrium. This 'long-run' cointegrated relationship can be consistently estimated by the Fully Modified OLS (FMOLS) proposed by Phillips and Hansen (1990), Park's 1992 Cointegrating Regression (CCR) or Dynamic OLS (DOLS) advocated by Saikkonen (1991) and Stock and Watson (1993). Lagged errors from such regressions can be used in a standard time series OLS to estimate the rate of adjustment towards the long-run equilibrium. However, in the standard cointegration models, an error correction mechanism is modelled linearly but, as explained earlier, the hysteresis theory suggests a highly non-linear effects and error correction mechanics (Belke et al., 2014). As a response to this problem, Balke and Fomby (1997) introduced a model that allows for sharp jumps in the rate of adjustment towards the equilibrium. This model can be further extended to allow for smooth transition in the rate of adjustment (Teräsvirta, 1994).

Other techniques are more related to the hysteresis models developed by Mayergoyz (1991), Krasnosel'skii and Pokrovskii (1989). These are one of the most direct ways of testing for hysteresis as, in these models, current states of an economic variable depend on non-dominated extremum values of past shocks. As a result, these models have a ‘selective memory’ because, although all disturbances affect the variable, only the non-dominated extremum values of disturbances are retained in the ‘memory’ of the system. The most popular of these, are the models built on Preisach dynamics which can be used to model such systems. For example, in their application, Piscitelli et al. (2000, 1999) introduce an algorithm for transforming variables in a Preisach-aggregation consistent way. This approach relies on an identification strategy that first determines the weight of the area of active firms, based on specific assumptions on the distribution of entry and exit triggers of firms and their weights. However, the main drawback is that due to complexity, only the most simple uniform weight distributions are assumed (Belke et al., 2014).

An interesting structural technique for testing hysteresis was developed by Göcke (1994, 2001) and Belke and Göcke (2001, 2005). Their method approximates macroeconomic hysteresis loops using generalised play dynamics which allow for existence of partial linear relationships with different slopes. This is done by introducing the ‘play’ and ‘spurt‘ sections across which the relationship can differ. The ‘play’ section allows for weaker response of the dependent variable while the ‘spurt’ section allows for a stronger response of the dependent variable to the ‘forcing variable’. The ‘forcing variable’ is a term used for variables whose change lead to a change in the economic behaviour of the observed units. In this model, permanent (hysteresis) effects occur when the movements take place on the spurt line. Based on the movements on the spurt line, authors use their algorithm to create an artificial ‘shift’ variable. The ‘shift’ variable integrates the movements in the ‘spurt’ area whereas movements in the ‘play’ area are filtered out. The filtering and consequently the ‘shift’ variable depends on the width of the play area. Such a model can be estimated by the FMOLS.

An advantage of the non-linear structural models is that, as mentioned previously, they take more nuanced view of hysteresis (Krasnosel’skii and Pokrovskii, 1989, Mayergoyz, 1991). That is, these models do not necessarily require all past shocks to affect the present realisations of an economic series. Rather, in contrast to the unit root approach, they allow for a selective memory where only a certain kind of shocks (usually the large ones) will lead to hysteresis. Thanks to this, these models properly reflect the highly non-linear nature of hysteresis effect that is often suggested by theoretical literature (O’Shaughnessy, 2000). Further advantage of these models is that they can be used to obtain a wider range of policy relevant parameters as they are not restricted to examining one economic variable at a time like the unit root approach. Thus these models can be used not only to test for hysteresis but also to estimate the effect of various other variables controlling for the hysteresis effects.

Nevertheless, these models also have their drawbacks. Their highly nuanced view of hysteresis makes proper modelling challenging. The evidence provided by these models can be quite sensitive to the exact specification. Moreover, because these models try to closely follow the underlying theory, the application varies across different fields. This allows for a great deal of discretion on the side of the modeller which is not necessarily desirable, especially in areas where previous research is lacking. In addition, these can be in practice often applied to test hysteresis only indirectly (Belke et al., 2014). However, their main drawback vis–à–vis this research is that these models are usually quite demanding in regard to the quantity of data required. This might not be an issue in the fields that focus on examining hysteresis caused by the exchange rate movements or hysteresis in (un)employment. Data for these and related variables are, in most cases, readily available even on higher frequencies. This does not generally hold for all data that would be required for estimation of the model developed in the next section. For example, the Federal Deposit Inusurance Corporation's (FDIC) Historical Statistics on Banking (HSOB) dataset , one of the most prominent sources of data on the U.S. commercial banking sector from 1934 to present, keeps track of many important variables only for approximately last 30-40 years. Moreover, the data are collected only at a yearly frequency. Some of these years would have to be further sacrificed due to necessary differencing and inclusion of lagged variables. This might not be a problem for estimation of more traditional models as data are also collected on state level and thus it is possible to utilize larger samples using various panel techniques. But hysteresis is a long-run phenomena and thus having a large number of observations in a temporal dimension is far more important than the number of observations across the cross-section.

Another strain of research builds on the sharp dichotomy outlined earlier and examines the hysteresis hypothesis using the unit root tests. This is still the more common approach to testing the hysteresis hypothesis and it is based on examining the order of integration or more specifically the (non)stationarity of an economic variable(s). While the individual techniques for unit root testing vary, they are all at their heart based on different ways of distinguishing the autocorrelated processes where past realisations determine the present one (see the equation 6) from those where they do not (as described by the equation 3). Like the structural approach, the unit root literature offers a rich variety of models that can be applied. But in a stark contrast to literature on the structural approach, individual unit root tests are closely related to each other making an overview of wide range of methods, as in the structural approach, unnecessary[9]. Rather we will focus only on the more recent and advanced applications.

---

[9] Interested readers can find such overview in Hamilton (1994), Stock and Watson (2007), Verbeek (2008) and Wooldridge (2015).

Thanks to the increased availability of dated panels, most authors prefer to test hysteresis using panel unit root tests. These are usually based on the extension of the univariate augmented Dickey–Fuller (1979) (ADF) test that tests the null hypothesis of unit root against the stationary or trend stationary alternative (depending on exact specification). This is commonly done by regressing the past lags of a variable on its first difference representation (Verbeek, 2008). The lags are included to get rid of the autocorrelation that would otherwise lead to a bias[10]. The most popular panel extension of these are the Levin–Lin–Chu (2002) test, Im, Pesaran, and Shin (2003) (IPS) test and the Fisher-Maddala, Wu (1999) test. These tests first apply individual unit root tests to each member of a panel and subsequently test for the joint significance of the result. The main difference between these comes from different assumptions on the joint significance and differences in producing the panel test statistics. The Levin–Lin–Chu (2002) is quite restrictive as it assumes homogeneity of coefficients, both in the null unit root and alternative stationary hypothesis. The Im, Pesaran, and Shin (2003) (IPS) and Fisher-Maddala, Wu (1999) tests allow for heterogeneity of coefficients in the alternative hypothesis, the difference being the methodology for getting the overall test statistics. To be more specific, the latter collects natural logs of p-values derived from the individual τ statistics whereas the former collects the τ statistics first to calculate the overall p-value. The advantage of Maddala, Wu (1999) test is that it avoids issues of sequential or joint T, N asymptotic.

However, one problem of these tests is that they have hard time distinguishing a unit root process from a stationary process with structural breaks. This can be problematic as one-time structural level shifts that are not related to hysteresis, such as those caused by technological change, can lead to erroneous non-rejection of the unit root null hypothesis. Thus it is imperative to explicitly control for such shifts as failure to do so would bias the test toward the type II error. This is, to a certain degree, analogous to the omitted variable bias problem.

Luckily, Zivot and Andrews (1992) developed an extended version of the ADF test that controls for such shifts by including an endogenously determined breakpoint in the level or level and trend. Nonetheless, the Zivot and Andrews test allows only for one structural break. Also panel extensions of the model are too impractical to be worthwhile as this would require expected values and variances of the ADF t-statistics for all possible break locations in the sample. To deal with these shortcomings, Lee and Strazicich (2003), Lee et al. (2004) propose a minimum Lagrange Multiplier (LM) unit root test which allows for more than one endogenous structural break.

---

[10] Selecting the right amount of lags is crucial as over-selection leads to loss of power (see Verbeek, 2008).

The minimum LM unit root test also has very good asymptotic properties. For instance, in general, the critical values do not asymptotically depend on structural shifts under the null. Using this property Im et al. (2005) propose the panel extension of the minimum LM unit root test. Thanks to these advantages, this method is increasingly becoming a more popular way of testing for hysteresis. Examples of applying the minimum LM unit root to the unemployment include Lee et al. (2009), Gomes and da Silva (2009), Lee et al. (2010) or Romero-Avila and Usabiaga (2007).

In general the unit root approach to hysteresis has several advantages vis-à-vis the structural approach. First, the unit root approach can be (and is) applied more uniformly across the literature. Of course, the exact specification differs to reflect the particularities of data, such as the differences in the level of autocorrelation of the series and similar differences, but otherwise they are not large. This makes the unit root approach much more comparable and standardized across fields, leaving less space for discretion. This is advantageous as wide room for discretion may easily lead to unintended and unconscious rationalisations of choices that confirm our preconceived notions, despite the best efforts of the modeller. This is even more problematic if there is virtually no prior empirical research on the topic.

Second, the unit root approach is less data intensive than the structural approach. While unit root tests also suffer from low power in datasets with short temporal dimension (see Verbeek, 2008), they require data only on the variable of interest itself. This is precisely because unit root tests are generally not based on structural models. Rather these tests only look at whether the present realisations of an economic variable are determined by the past ones or not. This task generally does not require addition of control variables, except for the lags of variable itself or controls for intercept, trend and structural breaks of these where appropriate. In contrast structural tests can require data on multiple variables, some of which may only be indirectly observable, depending on the model which they are built on.

The disadvantage of unit root tests is, as mentioned earlier, that they are a special case of hysteresis where all past events determine the present state of a variable. Hysteresis allows for dynamics with a selective memory where only certain types of shocks determine the present state of an economic variable (Belke et al., 2014). Unit root dynamics is non-selective and every past shock affects the present realisations of an economic variable. As a result of this, in a hysteresis, two opposing but equally large shocks generally result in a new equilibrium, while in a *linear* unit root process such shocks will leave the equilibrium unchanged (Amable et al., 1994, Cross, 1994, Piscitelli et al., 2000). The main practical implication of this is that the rejection of a unit root hypothesis does not necessarily imply the absence of hysteresis in general. Rather this only implies that the series does not follow a special case of hysteresis.

That is, the hysteresis where all past shocks determine the present realisations of an economic variable. Because of this some scholars also consider an evidence for near unit root process as an evidence for hysteresis (Blanchard and Summers, 1986).

Another drawback of unit root tests is that they are generally devoid of a deeper economic meaning. These tests only look at whether the past realisations of an economic variable determine the present ones. As such they generally cannot be used to answer deeper questions about the underlying economic relationships or to provide estimates of policy relevant coefficients (other than the impact of past lags on present realisations). However, this comes as a consequence of the advantages presented previously. Thus this is a feature of the test, not a ‘bug’.

The main takeaway of this overview of methods used to test for hysteresis is that despite all these advances in econometric methods made during last three decades, testing for hysteresis is still a very peculiar and challenging exercise. This is because so far literature did not settle on some standard approach for testing the hysteresis hypothesis, as there are no explicit tests to do so (Belke et al., 2014, Hallett and Piscitelli, 2002). Because of this it is not easy to settle on a proper approach for testing the hysteresis hypothesis implied by the theoretical model developed in the next section. On one hand, such a model would be perfect for the structural approach that must be built on solid theoretical foundations. On the other hand, data limitations in this case undermine the credibility of such an approach. The unit root approach only requires data on the number of banks, and these are easier to observe. Thus, it is not surprising that the FDIC reports such data consistently for the whole span of their dataset (1934-2015). The unit root approach may be more suited for the task, due to its fairly standardized application, especially if one takes into consideration the lack of previous empirical examinations of such model in this field. Furthermore, as will be shown in the next chapter, there are good reasons to suspect that shocks leading to hysteresis are likely. In such cases the line between unit roots and ‘genuine’ hysteresis gets blurred as a unit root process can be viewed as special case of hysteresis where every shock leads to hysteresis effect.

3 Hysteresis in the Monti-Klein Framework

As mentioned previously, this paper nests the idea of hysteresis in a generalised version of the Monti-Klein model of banking industry. The Monti-Klein model is uniquely suited for studying hysteresis as it is the main workhouse of the industrial organization strain of research on the banking competition. Moreover, while the Monti-Klein framework is relatively simple and accessible, it provides a rich set of insights and testable predictions (Freixas and Rochet, 2008). These attributes make the Monti-Klein model an excellent tool for introducing the idea of hysteresis into the banking industry in a clear and rigorous yet accessible manner.

Hysteresis is introduced into this model by allowing for the existence of both fixed entry and per-period ('maintenance') costs, as well as extending the model into a quasi-dynamic setting. For the sake of transparency, let us begin with the textbook version of the model and the aforementioned extensions will be added throughout the chapter as necessary.

The use of the entry and exit barriers and triggers is inspired by the previous work on hysteresis in the field of international and general economics. The works of Baldwin (1988) and Dixit (1989, 1992), who where amongst the first authors to consider the consequences of such barriers, were highly influential. However, this work nests the hysteresis in the Monti-Klein model making it somewhat distinct from those models. Thus the novelty of this model does not only lie in application of hysteresis to banking industry but also in its approach.

Following Freixas and Rochet (2008) consider the case of N identical banks, indexed by $n = 1, ..., N$. Each bank faces downward sloping demand for loans $L(r_L)$ and upward sloping supply of deposits $D(r_D)$. However, let us work with the inverse demand $r_L(L)$ and supply $r_D(D)$ functions, as this is more convenient. Banks' choice variables are the volume of loans L_n and deposits D_n. Banks also face the interbank market rate r (which can also be interpreted as the central bank's rate). The net position of a bank on the interbank market will be given by $M = (1 - \alpha)D - L$, where α is the reserve coefficient. Banks will face cost function $C(D, L)$ which is assumed to be twice differentiable and satisfies the standard convexity assumptions. To ensure that the model can be solved analytically, the cost function $C(D, L)$ will be assumed to be linear (thus $C(D, L) = \gamma_D D + \gamma_L L$). The point where we diverge from the textbook example of Monti-Klein model is the addition of fixed firm-specific and non-recoverable entry costs e, which makes them sunk ex-post, and per-period fixed 'maintenance' costs f. While these additions were inspired by the models of hysteresis developed in the context of international trade (See Baldwin, 1988, 1990), they are completely innocuous in the context of banking as well.

There is no doubt that a potential entrant in the banking industry faces high entry or set-up costs. These come from various sources. Some of these are natural. For example, Dell'Ariccia et al. (1999) show that in the presence of uncertainty about the borrowers' creditworthiness, adverse selection and information asymmetries between entrant and incumbent bank can erect high entry barriers. In fact they show that in standard Bertrand competition with two incumbent banks, the entry of a third bank will be blocked. Of course, in practice we can see new banks entering the industry, but the point is that these entrants face high entry costs just due to the nature of banking industry. But not all entry barriers are natural. Others are man-made as most developed countries highly regulate the entry into the banking and financial industry (Barth et al., 2006, 2008, Besanko and Thakor, 1992, Freixas and Rochet, 2008). These costs can be empirically quite high and effective. For example, according to data gathered by Barth et al. (2008), who measured entry barrier on scale ranging from 0 to 8, 132 out of 153 countries scored 7 or higher (with the U.S. scoring 8). Nevertheless, this is not a criticism of such regulations as it could be argued that they are welfare improving[11]. The point is that, for better or worse, they exist and can be substantial. Thus, it is reasonable to include them in a model of banking industry.

Similarly, there is plenty evidence of existence of non-trivial per-period fixed ('maintenance') costs in the banking industry. Again some of these occur naturally as the typical banking activities entail high fixed costs (Pulley and Humphrey, 1993). These can come from various sources such as the costs of advertising, deposit management or others. Here as well, some fixed per-period costs are introduced by the fiat. These include various regulatory requirements which have to be satisfied regardless of the volume of the services provided by banks (Barth et al., 2006, Freixas and Rochet, 2008). Because of this, it is quite reasonable to include fixed per-period costs in the model as well.

This model also introduces certain assumptions about the entry and exit behaviour of competing banks. Entry occurs as long as the profits are positive, thus the entry condition can be expressed as $\pi_n > 0$. This makes sense as a rational bank would enter the industry only if it expects to recoup the costs of entry into the industry. The exit is assumed to happen under the following condition $0 - e > \pi_n$ or $0 > \pi_n + e$. The entry costs e are being added back to the profit because they are sunk ex-post and thus do not affect the decision anymore. Therefore, after the entry, exit will occur only if profits drop below zero excluding the non-recoverable firm specific fixed costs.

---

[11] For an overview of arguments in favour of and against entry barriers, see Freixas and Rochet (2008), Barth et al. (2008) and sources cited therein.

After describing the set up of the model introduced here, this section proceeds to show that there is hysteresis in the number of active banks in a market. The starting place is the derivation of optimum profit and choice variables. Putting all the above mentioned parts of the set up together gives the following profit function of a representative bank n:

$$\pi_n(L_n, D_n) = r_L\left(L_n + \sum_{m\neq n} L_m\right)L_n + rM - r_D\left(D_n + \sum_{m\neq n} D_m\right)D_n - C(D_n, L_n) - f - e. \tag{8}$$

Substituting for the net position on the interbank market $M = (1-\alpha)D_n - L_n$ and by factoring the common terms produces:

$$\begin{aligned}\pi_n(L_n, D_n) = &\left(r_L\left(L_n + \sum_{m\neq n} L_m\right) - r\right)L_n + \left(r(1-\alpha) - r_D\left(D_n + \sum_{m\neq n} D_m\right)\right)D_n \\ &- C_n(D_n, L_n) - f - e.\end{aligned} \tag{9}$$

Note that this is simply a generalized version of the Monti-Klein model. In a special case where $f = e = 0$, the equation 9 collapses into the standard ‘textbook’ Monti-Klein model. To show how hysteresis occurs in this generalized Monti-Klein framework, let us start by proving that this set up allows for multiple equilibria where the equilibrium number of firms differs while holding everything else constant. This is an important step as hysteresis is fundamentally a concept of multiple equilibria and equilibria selection. Thus, the natural starting point is the proposition 1 which proves that multiple equilibria do exist in this framework.

Proposition 1 *Let the entry costs be higher than fixed per-period costs $e > f$, then for a given equilibrium profit π^* such that $e > \pi^* > f > 0$ there will be multiple possible equilibrium values of the number of active firms in the market N^*.*

Proof 1

In this environment, representative bank faces the following optimization problem:

$$\max_{D_n L_n} \pi_n(L_n, D_n) = \left(r_L\left(L_n + \sum_{m\neq n} L^*_m\right) - r\right) L_n + \left(r(1-\alpha) - r_D\left(D_n + \sum_{m\neq n} D^*_m\right)\right) D_n - C_n(D_n, L_n) - f - e. \tag{10}$$

The representative bank tries to optimize the intermediation margins on loans and deposits minus the costs. Assuming that π is concave, the optimum π^*_n can be found with respect to the choice variables L_n and D_n using the following first order conditions (FOCs):

$$\frac{\partial \pi_n}{\partial L_n} = r'_L\left(L^*_n + \sum_{m\neq n} L^*_m\right) L^* - r_L\left(L^*_n + \sum_{m\neq n} L^*_m\right) - r - \gamma_L = 0 \tag{11}$$

$$\frac{\partial \pi_n}{\partial D_n} = -r'_D\left(D^*_n + \sum_{m\neq n} D^*_m\right) D^* - r_D\left(D^*_n + \sum_{m\neq n} D^*_m\right) + r(1-\alpha) - \gamma_D = 0. \tag{12}$$

In a Nash equilibrium, each bank n chooses strategy profile $\mathcal{S}_n(L_n, D_n) \in (0, \bar{\mathcal{S}})xR_+$ which is the best response to strategy profiles of other banks described by the FOCs. The conditions for interior equilibrium are given by the FOCs and $D_n, L_n > 0$. In a symmetric equilibrium arguments maximizing the π_n are:

$$\arg\max_{L_n, D_n} \pi_n = \left\{L^*_n = \frac{L^*}{N}, D^*_n = \frac{D^*}{N}\right\}. \tag{13}$$

Substituting this result for L^*_n and D^*_n back into the first order conditions and the profit function gives:

$$\frac{\partial \pi_n}{\partial L_n} = r'_L\left(L^*\right)\frac{L^*}{N} - r_L\left(L^*\right) - r - \gamma_L = 0 \tag{14}$$

$$\frac{\partial \pi_n}{\partial D_n} = -r'_D\left(D^*\right)\frac{D^*}{N} - r_D\left(D^*\right) + r(1-\alpha) - \gamma_D = 0, \tag{15}$$

and

$$\pi^*_n(L^*_n, D^*_n) = \left(r_L\left(L^*\right) - r - \gamma_L\right)\frac{L^*}{N} + \left(r(1-\alpha) - r_D\left(D^*\right) - \gamma_D\right)\frac{D^*}{N} - f - e. \tag{16}$$

Equation 16 gives the optimum equilibrium profits of a representative bank. Expressions 14, 15 and 16 again highlight that the banks' profits depend on the width of the intermediation margins on loans and deposits respectively, and the costs of managing these. This is similar to the solution to standard Monti-Klein model. The difference, in this case, is that banks face the fixed per-period and bank specific non-recoverable entry costs as well.

Also, rearranging the equation 16 derives an equilibrium number of banks as a function of profit.

$$N^* = \left(\left(r_L\left(L^*\right) - r - \gamma_L\right) L^* + \left(r(1-\alpha) - r_D\left(D^*\right) - \gamma_D\right) D^*\right) \frac{1}{\pi^* + f + e} \tag{17}$$

Using the entry and exit conditions, it is also possible to solve for a band within which the equilibrium number of banks stays constant. Such 'band of inaction' features commonly in the hysteresis literature (see Baldwin, 1989, Bowen et al., 2012, Dixit, 1992). Let us start by deriving the upper entry bound denoted by $\bar{N}^*$. To find an expression for this, start with the entry condition

$$\pi^* > 0. \tag{18}$$

Substituting for the optimum profit (π^*) from the equation 16 gives:

$$\left(r_L\left(L^*\right) - r - \gamma_L\right) \frac{L^*}{N} + \left(r(1-\alpha) - r_D\left(D^*\right) - \gamma_D\right) \frac{D^*}{N} - f - e > 0 \tag{19}$$

Now, because our interest is in a boundary at which a bank is indifferent between entering the industry and staying inactive, the inequality can be replaced with equality. Solving for N gives us the 'entry boundary' $\bar{N}$.

$$\bar{N} = \left(\left(r_L\left(L^*\right) - r - \gamma_L\right) L^* + \left(r(1-\alpha) - r_D\left(D^*\right) - \gamma_D\right) D^*\right) \frac{1}{f + e} \tag{20}$$

The exit boundary can be derived using similar steps. This boundary describes the points of indifference between staying in the industry and exiting. Starting with the exit condition, substitute the optimal profit from equation 16, let the relationship hold with equality and solve for $\underline{N}$. The exit condition is defined as:

$$0 > \pi^* + e. \tag{21}$$

Substituting the expression for the equilibrium profits (π^*) gives

$$0 > \left(r_L\left(L^*\right) - r - \gamma_L\right)\frac{L^*}{N} + \left(r(1-\alpha) - r_D\left(D^*\right) - \gamma_D\right)\frac{D^*}{N} - f - e + e. \qquad (22)$$

Because we are looking for the points where banks are indifferent between exiting and staying active, the inequality can be replaced by equality again. Rearranging and solving for N produces the $\underline{N}$:

$$\underline{N} = \left(\left(r_L\left(L^*\right) - r - \gamma_L\right)L^* + \left(r(1-\alpha) - r_D\left(D^*\right) - \gamma_D\right)D^*\right)\frac{1}{f}. \qquad (23)$$

An important corollary that follows from the solutions to the upper entry and lower exit bounds is that $\bar{N} > \underline{N}$ (since $\frac{1}{f} > \frac{1}{f+e}$) regardless of optimum values and model's parameters as long as aggregate variable profits (the expression in the brackets) and fixed per-period costs are greater than zero. That is as long as:

$$\left(r_L\left(L^*\right) - r - \gamma_L\right)L^* + \left(r(1-\alpha) - r_D\left(D^*\right) - \gamma_D\right)D^* > 0 \text{ and } e, f > 0. \qquad (24)$$

Positive aggregate variable profits are necessary for market to exist in the first place. Thus this corollary shows that as long as there is a market, upper entry bound will always be higher than the lower exit bound. This holds regardless of the chosen parameters as long as $e, f > 0$.

To prove the proposition 1, it is enough to just consider whether there can exist a whole set of optimum numbers of banks, instead of just one unique equilibrium. That is:

$$\exists \boldsymbol{N}^* = \{N_1^*, N_2^*, ...N_k^* | k > 1\}.$$

Thanks to the entry and exit bounds (described by equations 20 and 23), all possible values of N can be split into the following three ranges:

1. The range above the entry bound: $(\bar{N}, \infty)$.
2. The range below the exit bound: $[0, \underline{N})$.
3. The range between the entry and exit bounds: $[\underline{N}, \bar{N}]$.

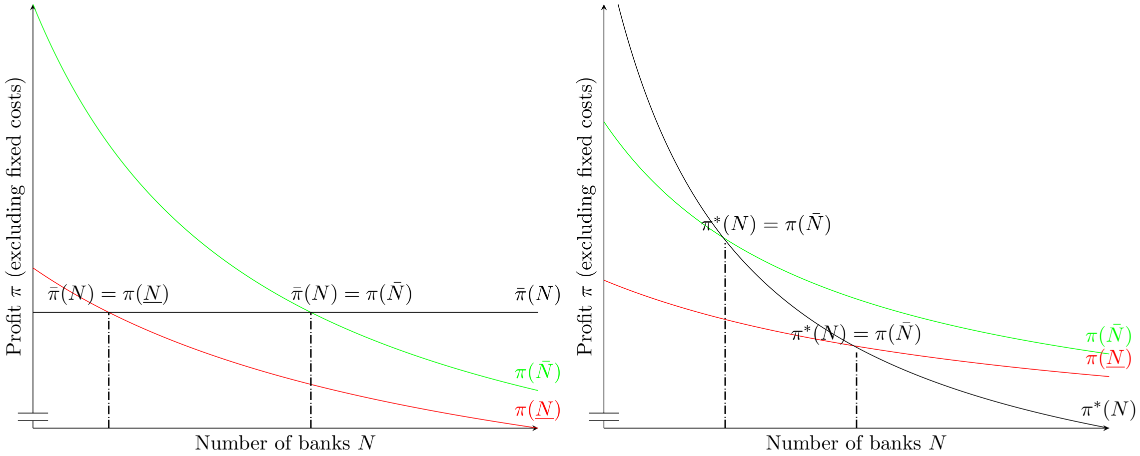
Now let us conduct the following thought experiment. Consider a rate of profit π which is constant for different values of the number of banks N, that is $\pi(N) = \bar{\pi}$. Such a rate of profit is possible because the choice variables L^*, D^* can adjust to make the constant profit optimal for different number of active banks. The first range is clearly not an equilibrium since above the entry boundary profits will be sufficiently high to attract new entrants thereby forcing N to adjust until $N = \bar{N}$. The second range, $[0, \underline{N})$ is again a disequilibrium. In this case the profits are too low to make this number of active banks consistent with the equilibrium. Thus, in this case N will be pushed back to $N = \underline{N}$. Inasmuch, no equilibrium result in this range is possible as well. In the third and last case, equilibrium can be sustained as profits are not high enough to trigger entry, but are not too low to force some incumbents to leave. Therefore, any N that lies within the $[\underline{N}, \bar{N}]$ range will represent one of the multiple potential equilibrium numbers of banks as long as $\underline{N}$ and $\bar{N}$ are not equal. However, this was already proven by comparing the equations 20 and 23.

Analogously, it can be proven that there are also multiple equilibria in the rate of profits at a given fixed number of active banks. This can be done by the same thought experiment but now by holding the number of firms constant and considering possible equilibria for the ranges of profits. Generally any combination of π and N will be an equilibrium as long as this combination lies within area under the $\pi(\underline{N})$ and above the $\pi(\bar{N})$. The area of such equilibrium space is defined by:

$$\int_0^\infty \pi(\bar{N})d\bar{N} - \int_0^\infty \pi(\underline{N})d\underline{N}. \tag{25}$$

The function $\pi(\bar{N})$ is the profit at any arbitrary number of $\bar{N}$ where firms are indifferent between entering and staying out of the market. Similarly, the function $\pi(\underline{N})$ is the profit at any arbitrary number of $\underline{N}$ where firms are indifferent between continuing to operate and shutting down. This is the so called 'band of inaction' that is an important part of the hysteresis models (Baldwin, 1990, Bowen et al., 2012, Dixit, 1989, 1992). The band of inaction is the area that enables multiple equilibria to exist and thus this proves the proposition 1.

This process is visualised by the figure 1a. The figure 1a especially highlights how a certain fixed rate of profit can be consistent with different number of banks. Similarly, it can be seen that the different rates of profit could be an equilibrium holding the number of banks constant by drawing a vertical line passing through the $\pi(\bar{N})$ and $\pi(\underline{N})$. Moreover, it is also worth noting that it is not necessary to hold either the profits or the number of firms constant to show that there is a range of optimum profit where multiple equilibria can exists. As long as the equilibrium profit function ($\pi^*(N)$) is monotonically decreasing there will be a range of multiple equilibria between the points where the equilibrium profits cross the entry and exit bound (i.e. in range given by $[\pi^*(N) = \pi(\bar{N}), \pi^*(N) = \pi(\underline{N})]$). This is shown on figure 1b. Nonetheless, the thought experiment with constant profit is still important



(a) Multiple equilibria holding the rate of profit constant.

(b) Multiple equilibria at a variable rate of profit

Figure 1

as an important feature of hysteresis is that even shocks that fully reverse themselves may lead to different rates of profit.

Also, some may wonder why the bounds are curved. The reason for this is that as the number of banks in industry decreases the market power of incumbent increases, and this can be used to deter entry. However, as the number of incumbent increases the bounds converge to constants determined by the model's parameters, fixed per-period costs and entry costs. This is not a novel result as this is one of the arguments used by the proponents of the concentration stability view mentioned in the introduction (Allen and Gale, 2000).

Proving the proposition 1 is an important step towards getting the hysteresis result, as existence of multiple equilibria is a necessary condition for hysteresis. However, this is not a sufficient condition. By definition, a process exhibits hysteresis if current equilibria are dependent on the past events (shocks). Therefore, next part will prove that:

Proposition 2 *The optimum number of banks at a given time $N(t)$ is determined by past shocks and the initial number of banks $N(0)$.*

Proof 2

To prove the proposition 2, let us start by a comparative static exercise. Consider the changes to the equilibrium volume of loans (L) and deposits (D) and the number of banks (N) as a result of small changes in interest rates on loans (r_L) and deposits (r_D) and the unit costs (C). The effects of these small changes can be shown using the chain rule which gives us the following system of equations:

$$dL = \frac{\partial L}{\partial \pi}\frac{\partial \pi}{\partial r_L} dr_L + \frac{\partial L}{\partial \pi}\frac{\partial \pi}{\partial C} dC, \tag{26}$$

$$dD = \frac{\partial D}{\partial \pi}\frac{\partial \pi}{\partial r_D} dr_D + \frac{\partial D}{\partial \pi}\frac{\partial \pi}{\partial C} dC \tag{27}$$

and

$$dN = \frac{dN}{d\pi}\frac{d\pi}{dr_L} dr_L + \frac{dN}{d\pi}\frac{d\pi}{dr_D} dr_D + \frac{dN}{d\pi}\frac{d\pi}{dC} dC. \tag{28}$$

The equation 28 is most important for this proof. This relationship says that the change in the number of banks depends on how the number of banks change with respect to profit, which in turn depends on changes in r_L, r_D and C. The equation 28 can also be rewritten in a more interesting form. Begin by multiplying the first part of the first term by $\frac{\pi N}{N \pi}$ and second part by $\frac{r_L \pi}{\pi r_L}$. Next, multiply the first part of the second term by $\frac{\pi N}{N \pi}$ and second part by $\frac{r_L \pi}{\pi r_L}$. Finally, multiply the first part of the third term by $\frac{\pi N}{N \pi}$ and second part by $\frac{C \pi}{\pi C}$. None of these operations change the expression as they are all equal to one. All these adjustments give the following:

$$dN = \frac{dN}{d\pi}\frac{\pi N}{N\pi}\frac{d\pi}{dr_L}\frac{r_L \pi}{\pi r_L} dr_L + \frac{dN}{d\pi}\frac{\pi N}{N\pi}\frac{d\pi}{dr_D}\frac{r_L \pi}{\pi r_L} dr_D + \frac{dN}{d\pi}\frac{\pi N}{N\pi}\frac{d\pi}{dC}\frac{C \pi}{\pi C} dC. \tag{29}$$

Let us define the elasticity of the number of banks active with respect to profit as $\varepsilon_{N\pi} = \frac{dN}{d\pi}\frac{\pi}{N}$. The elasticity of the number of banks with respect to profit is positive as the number of banks cannot be negative by definition. The profit will always be non-negative in the equilibrium and thus the derivative of number of banks with respect to profit will be positive.

$$\frac{dN}{d\pi} = \left(\left(r_L\left(L^*\right) - r - \gamma_L\right) L^* + \left(r(1-\alpha) - r_D\left(D^*\right) - \gamma_D\right) D^*\right)\left(\frac{1}{(\pi^* + f + e)^2}\right)\left(-\frac{d\pi^*}{dN}\right). \tag{30}$$

In addition, define the elasticity of profit with respect to the lending and deposit rate as $\varepsilon_{\pi r_L} = \frac{d\pi}{dr_L}\frac{r_L}{\pi}$ and $-\varepsilon_{\pi r_D} = \frac{d\pi}{dr_D}\frac{r_D}{\pi}$ respectively. In this case it is trivial to see from the equation 10 that $\varepsilon_{\pi r_L}$ is positive and $\varepsilon_{\pi r_D}$ negative. Finally the elasticity of profit with respect to costs can be defined as $-\varepsilon_{\pi C} = \frac{d\pi}{dC}\frac{C}{\pi}$. Again the negative sign is clear from the equation 10. Substituting these definitions back to the equation 29 and rearranging produces the following expression:

$$dN = \left(\varepsilon_{\pi r_L}\frac{dr_L}{r_L} - \varepsilon_{\pi r_D}\frac{dr_D}{r_D} - \varepsilon_{\pi C}\frac{dC}{C}\right)\varepsilon_{N\pi}N. \tag{31}$$

This is a very important result as the expression above shows that the change in the number of banks depends on a number of important relationships. First, the flow of banks depends on the elasticity of profit with respect to the interest rate on loans ($\varepsilon_{\pi r_L}$) and deposits ($\varepsilon_{\pi r_L}$) respectively, as well as the relative change in the rates of interest on loans ($\frac{dr_L}{r_L}$) and deposits ($\frac{dr_L}{r_L}$). Second, the change depends on the elasticity of profit with respect to unit costs $\varepsilon_{\pi C}$ and their relative change $\frac{dC}{C}$. Third, the change depends on the elasticity of number of banks with respect to profit $\varepsilon_{N\pi}$ and currently operating number of banks N.

Now recall that due to the entry and exit conditions, the number of banks does not change within the 'band of inaction'. Thus, as long as changes to the interest rates on loans (dr_L) or deposits (dr_D), or changes in the unit costs (dC) are not large enough to move the profits out of the 'band of inaction' $\left(\pi(\bar{N}) \geq \pi(N) \geq \pi(\underline{N})\right)$, the implied value of the elasticity of number of banks with respect to profit is zero ($\varepsilon_{N\pi} = 0$). Thus the change in the number of banks can be described by the following piecewise function:

$$dN = \begin{cases} \left(\varepsilon_{\pi r_L}\frac{dr_L}{r_L} - \varepsilon_{\pi r_D}\frac{dr_D}{r_D} - \varepsilon_{\pi C}\frac{dC}{C}\right)\varepsilon_{N\pi}N & \text{if } \pi(N) > \pi(\bar{N}) \text{ or } \pi(\underline{N}) > \pi(N) \\ 0 & \text{if } \left(\pi(\bar{N}) \geq \pi(N) \geq \pi(\underline{N})\right), \end{cases}$$

This is just another representation of the 'band of inaction'.

Now consider the following thought experiment. First, assume that the banks' profits and the number of banks are at an equilibrium. By proposition 1, such point of equilibrium can only lie within the 'band of inaction'. Now suppose that the equilibrium rate of profit gets disturbed by a random shock. This can happen through either interest rate on loans, interest rate on deposits or banks' unit costs. Furthermore, let us consider the case of a negative shock, sufficiently large to move the profits out of the 'band of inaction'. In such a case, banks will start exiting the market as profits fall below the exit bound. However, each time a bank exits the market, remaining banks are able to raise their interest rates. This is because, in the Monti-Klein model, market power is inversely related to the number of banks (this can be seen from the FOCs given by equations 14 and 15). The higher market power enables banks to charge more for their loans and pay less for their deposits. Moreover, as banks leave the market, the supply of loans and demand for deposits diminishes, leading to higher interest rate on loans and lower one on the deposits, assuming constant demand for loans and supply of deposits. Once the profits are high enough as to not trigger the exit, the number of banks will reach its new equilibrium at a point right above the exit bound. If now the shock fully reverses itself and the rate of profit returns to its original equilibrium, the number of banks will not change as the original equilibrium rate of profit was within the 'band of inaction' where number of banks cannot change due to exit and entry conditions. Thus while profits can return to their original equilibrium, the number of banks stays in their new equilibrium. For the number of banks to get to their original equilibrium, market would have to experience a sufficiently high positive shock. This is the essence of hysteresis as this implies that the equilibrium number of banks depends only on certain (large) shocks, and the effect of these shocks persists even after they have died out.

This result can also be extended to a dynamic setting. Assume that the interest rate on loans and deposits, costs and consequently the number of banks in the market depend also on time t. These changes are assumed to be fully expected by the banks, although in principle an uncertainty could be introduced into the model, and they can be thought of as small changes that stem from shifts in preferences and technology across the time. Now, considering small deviations from the original solution with respect to time and by denoting the time derivatives of variables by $\frac{dx}{dt} = \dot{x}$ the equation 31 can be rewritten as a following differential equation:

$$\dot{N} = \left(\varepsilon_{\pi r_L} \frac{\dot{r_L}}{r_L(t)} - \varepsilon_{\pi r_D} \frac{\dot{r_D}}{r_D(t)} - \varepsilon_{\pi C} \frac{\dot{C}}{C(t)} \right) \varepsilon_{N\pi} N(t). \tag{32}$$

This is only further generalisation of the comparative static exercise.

Assuming that elasticities stay constant (which is reasonable at least for the short to medium run), the solution to the differential equation 32 can be obtained by dividing the equation by N and integrating both sides.

$$\int_0^t \frac{\dot{N}}{N} dt = \left(\varepsilon_{\pi r_L} \int \frac{\dot{r_L}}{r_L(t)} dt - \varepsilon_{\pi r_D} \int \frac{\dot{r_D}}{r_D(t)} dt - \varepsilon_{\pi C} \int \frac{\dot{C}}{C(t)} dt \right) \varepsilon_{N\pi}. \tag{33}$$

Denoting the initial number of banks in the market by $N(0) = N_0$ implies the following solution:

$$|N(t)| = \left(|r_L(t)|^{\varepsilon_{\pi r_L}} \cdot |r_D(t)|^{-\varepsilon_{\pi r_D}} \cdot |C(t)|^{-\varepsilon_{\pi C}} \right)^{\varepsilon_{N\pi}} N_0, \tag{34}$$

where the absolute values can be gotten rid off as the number of banks and costs must be non-negative by definition. The interest rates r_L and r_D could be negative in principle, but that would imply that bank is paying interest on its loans to firms and is charging depositors interest on their deposits. This is unlikely, although not impossible as a recent empirical evidence on negative interest rates shows (Jobst and Lin, 2016). Getting rid of the absolute values gives

$$N(t) = \left(r_L(t)^{\varepsilon_{\pi r_L}} \cdot r_D(t)^{-\varepsilon_{\pi r_D}} \cdot C(t)^{-\varepsilon_{\pi C}}\right)^{\varepsilon_{N\pi}} N_0. \tag{35}$$

The relationships described by equations 35 proves that there is hysteresis in the number of active banks in this model. The equation 35 shows that the current number of active banks depends on the initial number of banks N_0. Moreover, the equation 35 also highlights that the number of banks remains constant unless the elasticity of number of banks with respect to profit is non-zero ($\varepsilon_{N\pi} \neq 0$). As proposition 1 shows, this holds only outside the 'band of inaction', where $\varepsilon_{N\pi} \neq 0$ holds. As a result, only a shock sufficiently large to push profits outside the band, will result in new number of banks given by 35. Therefore, the current number of banks is determined by past shocks and the initial conditions. This whole process is visualised on the figure 2.

The figure focuses only on the case of a negative shock considered here but a positive shock works through the same mechanism in a mirrored way. This clearly demonstrates that the number of banks, in this model, is determined by the initial conditions ($N(0)$) and by past shocks that are sufficiently large to trigger the hysteresis effect. Therefore, expression 35 proves the proposition 2. Propositions 1 and 2 taken together prove that there is hysteresis in the number of active banks in the market.

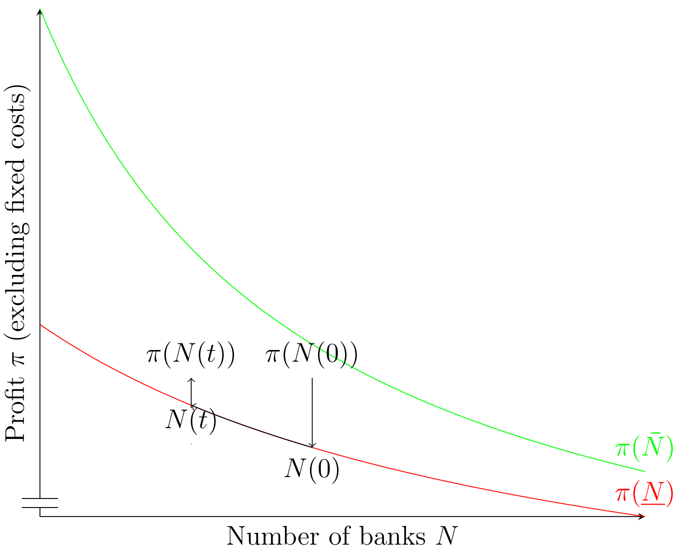


Figure 2: Visualisation of the hysteresis effect.

Nonetheless, a careful reader might wonder how likely it is that shocks from all three channels (r_L, r_D, C) combine in a way to create the hysteresis effect. After all in principle these shocks could be offsetting each other rather than reinforcing each other. However, as it turns out, there are strong theoretical reasons to suspect that these shocks are reinforcing and that unfavourable shocks are more likely than favourable ones.

Let us start by looking back at the differential equation 32. This expression shows that the change in the number of banks depends on three channels. These are the growth rate of interest rate on loans, the growth rate of interest rate on deposits and the growth rate of unit costs.

$$\dot{N} = \left(\varepsilon_{\pi r_L} \frac{\dot{r_L}}{r_L(t)} - \varepsilon_{\pi r_D} \frac{\dot{r_D}}{r_D(t)} - \varepsilon_{\pi C} \frac{\dot{C}}{C(t)} \right) \varepsilon_{N\pi} N(t). \tag{32}$$

Starting with the growth rate of interest on loans, there is a strong reason to suspect that this rate is more flexible downwards than upwards. Stiglitz and Weis 1981 proved that when banks cannot perfectly and costlessly monitor the lender behaviour, they have an incentive "to ration credit rather that raise the interest rate when there is an excess demand for loanable funds." This is bad both for lenders and borrowers. The former has to resort to less efficient and consequently less profitable method of supplying credit than it could in a situation with no market failures, while the latter finds it increasingly hard to get credit. In principle, lenders could still try to raise the debt-equity ratio of borrowers (the collateral), but Stiglitz and Weiss prove that this will not happen within their model. The reason why this does not occur is that raising collateral requirements disproportionally decreases the demand from less wealthy and less risky individuals, assuming that there is decreasing absolute risk aversion. Driving the less risky individuals from the market is the precise opposite of what the financial institutions want, so they will avoid that. The implication of this result for this paper is that the lending rates should be quite inflexible upwards. Therefore, shocks that decrease the interest rate on lending should be far more common.

Moreover, it is reasonable to assume that the directions of shocks to these channels are correlated with each other. To see this, just consider the usual course of events during recessions and more generally in times of financial distress. During recessions the value of collateral erodes relative to the debt burdens, which in turn increases the cost of distinguishing between 'good' and 'bad' borrowers (Bernanke, 1983). During the bust, borrower defaults and falling asset prices which lower the value of collateral, make lending more riskier and costlier enterprise than it is during the booming periods. Also, widespread bank failures destroy borrower-specific knowledge. It is more costly for other banks to lend to those borrowers.

At the same time, these periods of distress can lead to run on deposits (Diamond and Dybvig, 1983, Shin, 2009), with large depositors usually being the first to run (Huang and Ratnovski, 2011). This forces the banks to increase their deposit rates to compensate depositors for higher risk of default. The deposit insurance goes a long way to mitigate this, but it is no silver bullet. This is because many countries restrict deposit insurance only to a certain level and to certain institutions, the standard commercial banks. However, insights from this research broadly apply to all financial intermediaries[12].

Therefore, this ‘unholy coincidence’ of market failures and attributes of banking sector make large and adverse shocks likelier than they otherwise would. By extension, this also makes hysteresis more likely as banks cannot easily accommodate negative shocks to deposit rate or costs. This also suggests that there is an asymmetry in likelihood of hysteresis shocks that reduce the number of banks and those that increase it. This is mainly due to upward inflexibility of lending rate.

---

[12] As was mentioned earlier, ‘bank’ is broadly used as any financial intermediary in this work.

4 Empirical Application to the U.S. Banking Sector

The previous section shows that in a generalised Monti-Klein model, hysteresis implies that the number of active banks in the market is dependent on the initial conditions and sufficiently large shocks which cause a hysteresis effect. Moreover, it was also argued that the 'unholy coincidence' of the sector's traits makes adverse hysteresis shocks relatively likely. This section tests these predictions using the unit root approach.

The unit root approach is both simple and popular way of testing for hysteresis. However, as already explained in the second section, a drawback of this approach is that in unit root processes all shocks are 'remembered' by the system whereas genuine hysteresis allows for a 'selective memory' where the influence of only some shocks is retained by the system. Thus the unit root tests test a special case of hysteresis where all past shocks are retained by a system. Nonetheless, as the likelihood of hysteresis inducing shocks increases, the lines between unit root and hysteresis get blurred. This makes the unit root tests more applicable in general. Yet this important distinction means that the rejection of the unit root hypothesis can also just indicate that the hysteresis inducing shocks are not very likely. Nevertheless, failing to reject the unit root hypothesis would provide an evidence in favour of hysteresis in the number of active banks.

The unit root tests will be applied to the banking sector in the United States. The U.S. is perfect for such empirical application as it has fairly developed banking sector that seems to satisfy the assumptions made in this theoretical model. For example, there is an evidence for both high entry and per-period fixed costs in the U.S. banking sector (see Barth et al., 2008, Pulley and Humphrey, 1993). Moreover, the U.S. banking sector has a sufficient data availability for the empirical application.

This research uses data from the Federal Deposit Insurance Corporation's (FDIC) Historical Statistics on Banking (HSOB) dataset. The figure 3 shows how the number of FDIC insured commercial banks changed throughout the last century. The red lines mark all great and widely recognized financial panics during the observed period. The first red line shows the collapse of Breton-Woods system and the oil embargo of 1973-74. The second red line marks the 'Black Monday', a great market crash of 1987. The third red line represents the 1997 Asian financial crisis. The fourth red line depicts the 2001 dotcom crash. Finally, the fifth red line marks the sub-prime crisis and Great Recession of 2008.

The figure 3 shows that the number of U.S. banks was quite stable in the 1934-1985 period and started to decline rapidly in the period after. Interestingly, as it can be seen, the 1934-1985 period experienced only one large and widespread financial panic while the second period experienced four of them. After each of these periods the number of banks seem to drop further and further. Moreover, the FDIC dataset does not include data from the Great Depression as it was actually set up as a response to it. If it did, it would reveal a similar pattern. According to Walter (2005), in 1921 there were about 31,000 active banks in the U.S. Yet as the FDIC dataset shows, toward the end of Great Depression in 1934, there were only 14,146 banking institutions left. Furthermore, we can see that this decline continued until the end of Great Recession as in 1940 there were only 13,442 institutions. This somewhat supports the narrative described by the generalized Monti-Klein model derived in section 3.

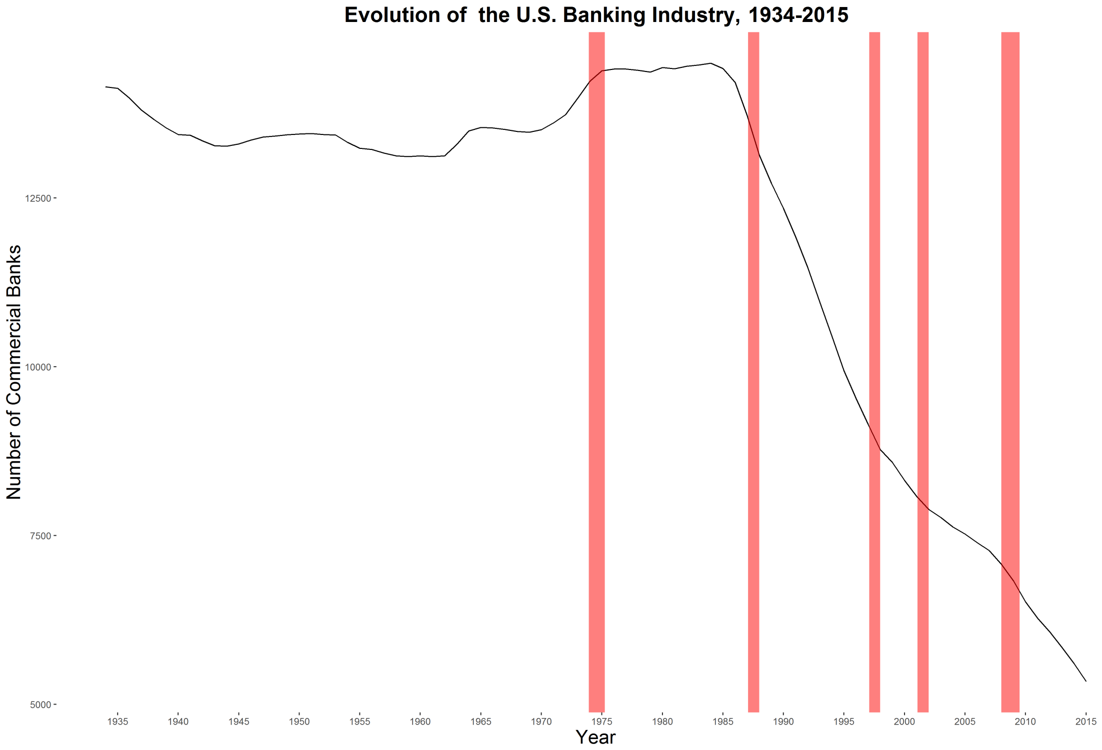


Figure 3: The number of commercial banks (Source: Federal Deposit Insurance Corporation (2017)). Highlighted areas mark the fall of the Breton-Woods system and the oil embargo of 1973-74, 'Black Monday', a great market crash of 1987, Asian financial crisis of 1997, dot-com crash of 2001 and Great Recession of 2008.

However, one must be very careful in interpreting the aggregate data. To start with, the timing of the large financial crises does not line up perfectly. For example, as the figure 3 depicts, the number of banks already began to decline a year before the Black Monday 1987. Nevertheless, the decline prior to the Black Monday was nothing out of ordinary as compared to relatively stable period before 1987.

The causality is not crystal clear either. The same pattern could be explained by the concentration-instability view, as it could be argued that the increased concentration is the reason behind the financial crises. Of course it is also possible that the concentration-instability and hysteresis reinforce each other. Furthermore, as mentioned in the introduction, concentration-instability view is in itself controversial. In contrast, under the concentration-stability view higher concentration should lead to more stability. Also it is worth noting that the industrial concentration is not necessarily synonymous with number of banks in the industry. However, it is certainly correlated with it and number of concentration measures, such as Herfindahl-Hirschman index (HHI) takes the number of active firms into account (Bikker and Haaf, 2002). Here we focus on the number of banks mainly because of better data availability and because this model directly predicts hysteresis in the number of banks[13].

Some of the decrease in the number of banks in the U.S. can also be explained by technological progress. Recent decades witnessed unprecedented revolution in non financial technology, such as communication and information technology, as well as financial technology, such as statistical analysis and financial engineering. Theoretically the impact of technology on bank concentration is ambiguous, but Berger (2003), Berger and DeYoung (2002) provide an evidence that technological progress during the period from mid 1980s to late 1990s favoured higher bank concentration.

Other culprits for the decline in the number of banks across the U.S. are interstate branching and banking deregulation. Prior to late 1970s and early 1980s, U.S. states had a strict banking regulation that protected local banks from outside competition. Once these restrictions were lifted the costs and prices of banking services fell, pushing some banks out of market (Jayaratne and Strahan, 1998, Stiroh and Strahan, 2003).

Looking at the state level data it is possible to see that there is quite a lot of heterogeneity between the U.S. states and territories (see the figure 4). U.S. states like Connecticut, Idaho, Indiana, Kentucky, Maine, Maryland, Massachusetts, Michigan, Mississippi, New Jersey, New York, North Carolina, Ohio, Pennsylvania, South Dakota, Vermont and Washington experienced steady decline in the number of banks during the sample period. However, in most of the U.S. states the number of banks gradually increase up until some time in between early 1970s and 2000s and then declines rapidly. Yet other states such as Alaska, Arizona, Delaware and Utah follow a more complex cyclical pattern. This heterogeneity in the timing and overall

---

[13] Although in the Monti-Klein framework there is an inverse relationship between number of banks and market concentration (Freixas and Rochet, 2008).

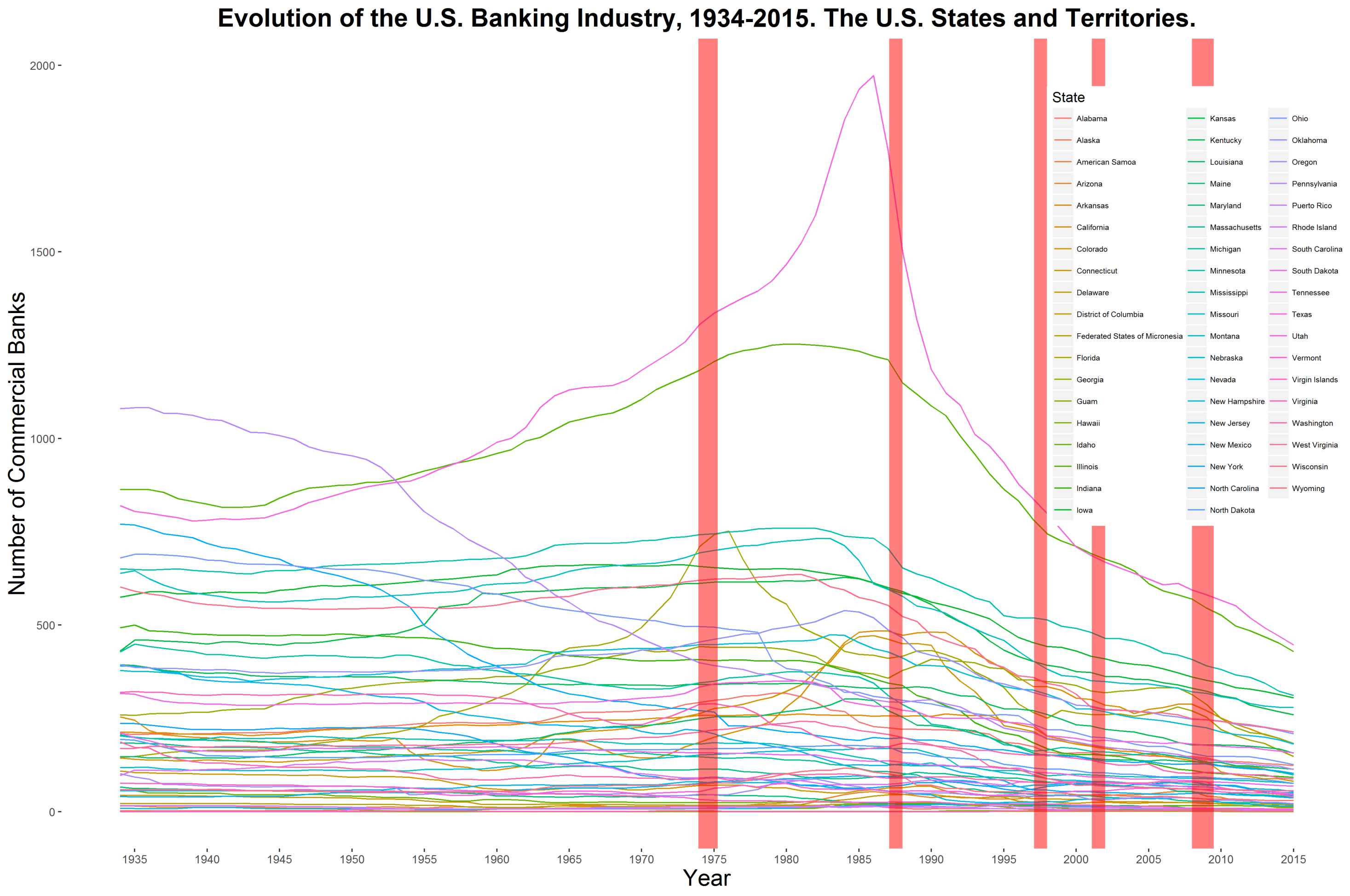


Figure 4: The number of commercial banks by the U.S. states and territories (Source: Federal Deposit Insurance Corporation (2017)).

pattern of the decline in the number of banks indicates that there are most likely other forces in play. This is so because both deregulation and technological change was affecting the U.S. states at more or less same time.

4.1 Unit Root Tests With Structural Break(s)

The starting point for applying the unit root approach to test for hysteresis in the banking sector is formal description of hysteresis as a unit root process. Following Franz (1990), hysteresis in a linear system and discrete time can be represented as a simple law of motion:

$$N_{t+1} = aN_t + u_t. \tag{36}$$

where u_t are exogenous shocks to the system. If $a \neq 1$ then the steady state of the number of active banks (denoted by bar) would be given by

$$\bar{N} = \frac{\bar{u}}{1-a} \tag{37}$$

and thus the steady state number of banks would be independent of the path followed by u_t, and would only depend on the steady state level of $\bar{u}$.

However, if $a = 1$, as it is in this case, there will be no unique steady state as

$$\bar{N}_t = N_0 + \sum_{t=1}^{T} u_t. \tag{38}$$

Here the steady state of the number of active banks is fully dependent on the shocks. In this case the series can be described by a simple unit root process:

$$N_{t+1} = N_t + u_t \tag{39}$$

The equation 39 is also a null of ADF unit root test (Dickey and Fuller, 1979) without trend or drift term and these could be added to it without any loss of generality. Adding a drift term μ, to allow for non-zero mean, and a trend term t, to control for possible time trend, gives:

$$N_{t+1} = \mu + \psi t + N_t + u_t \tag{40}$$

Thus, in principle, these tests can be applied to test for hysteresis in the number of active banks. Unfortunately, in this case, the testing for unit root is more complex. The concurrence of profound technological change, deregulation and heterogeneity in the patterns across the U.S. states, shown in the previous section, makes testing for the presence of unit root more challenging. In practice it is quite difficult to distinguish between genuine unit root and stationary series containing structural breaks and as a result, standard unit root tests can lead to misleading inferences (Perron et al., 2006).

Fortunately, Perron (1989), Vogelsang and Perron (1998), Zivot and Andrews (1992) developed a framework that allows to account for structural break in unit root test in univariate setting. Allowing for the structural break in the data makes the test more consistent. However, the downside of Zivot and Andrews is that it allows for only one structural break, and the null hypothesis does not allow for breaks. This can be quite restrictive and as a consequence test can suffer from low power (Lee and Strazicich, 2003). In this case, visual inspection of data and also historical narrative is consistent with both one or two structural breaks depending on how one looks at the data and past events. Figure 3 shows a possible one time break in the number of banks right after 1975 and another potential break close to 1985. However, the shift in the series around 1975 is very small, compared to the noticeable break in 1985, and hence it may not be a true breakpoint. The historical perspective does not shed more light on this problem either. It is reasonable to assume that both the wave of banking deregulation and the profound technological change in banking industry should result in a structural break. However, it is unclear whether these breaks overlapped or not as the timing of these effects is not precise and consistent with both one or two breakpoints. Furthermore, the panel extensions of the Zivot and Andrews model are too impractical to be worthwhile as this would require expected values and variances of the ADF t-statistics for all possible break locations in the sample.

An alternative to Zivot and Andrews was developed by Lee and Strazicich (2003) and Lee et al. (2004) who propose minimum LM unit root test that allows for more than one structural break. The Lee and Strazicich test also has an additional advantage that under the null hypothesis, the expected critical values are invariant to the exact specification[14] and have relatively better power compared to the Zivot and Andrews test, as they allow for also breaks under the null hypothesis. Another advantage of the minimum LM test is that it can be applied to panel setting as well thanks to the invariance of the expected critical values under the null. This is an advantage as utilizing panel data helps to improve the power of unit root tests.

Unfortunately panel LM test tests the null hypothesis that all series contain a unit root against an alternative that at least one series is stationary. Taylor and Sarno

---

[14] Although using bootstrapping, Chou (2007) finds that in a finite samples the expected critical values under the null may slightly deviate in cases where break in both intercept and trend is assumed.

Table 1: Descriptive Statistics on the Number of Active Commercial Banks in the U.S. States (1934-2015)

U.S. State or territory	Mean	Median	Maximum	Minimum	Standard deviation
Alabama	220·024	221	317	124	51·048
Alaska	7·512	8	2	15	3·782
American Samoa	0·354	0	1	0	0·481
Arizona	24·537	17	55	6	15·124
Arkansas	217·5	229·5	262	103	43·103
California	250·842	206	484	110	106·205
Colorado	214·244	169	472	82	103·112
Connecticut	61·902	63	109	14	28·113
Delaware	29·890	28·5	48	17	9·918
District of Columbia	14·805	16	26	4	6·388
Fed. States of Micronesia	0·366	0	1	0	0·485
Florida	329·354	271	752	142	156·286
Georgia	346·427	352·5	443	183	64·956
Guam	0·842	1	2	0	0·867
Hawaii	7·646	7	22	0	6·036
Idaho	28·439	25	60	11	11·966
Illinois	915·293	895·5	1253	429	225·019
Indiana	346·659	405	500	92	137·339
Iowa	558·342	596·5	661	305	108·515
Kansas	491·671	475·5	628	260	107·759
Kentucky	313·073	341	394	157	65·484
Louisiana	188·878	173	302	114	51·349
Maine	35·768	40·5	66	6	17·831
Maryland	114·317	112	184	39	40·803
Massachusetts	122·512	146·5	206	21	60·337
Michigan	310·537	355·5	449	99	110·367
Minnesota	619·976	652	760	312	122·499
Mississippi	157·037	182	203	76	44·115
Missouri	547·524	580·5	733	279	135·002
Montana	117·854	114	169	54	31·022
Nebraska	364·268	374·5	474	181	76·934
Nevada	14·988	9	39	6	9·536
New Hampshire	49·232	57	80	4	24·819
New Jersey	198·927	209	391	51	107·957
New Mexico	60·695	53·5	96	37	17·857
New York	342·11	269	771	103	218·295
North Carolina	129·085	90	237	44	68·438
North Dakota	144·366	150	194	76	29·509
Ohio	437	495·5	690	126	193·652
Oklahoma	379·098	380·5	539	209	80·655
Oregon	53·561	49	102	22	16·658
Pennsylvania	520·402	395·5	1082	113	331·486
Puerto Rico	8	8	16	0	4·779
Rhode Island	10·976	11	16	5	3·485
South Carolina	98·073	90·5	139	48	26·618
South Dakota	141·817	158	212	69	35·455
Tennessee	270·037	289	351	158	53·254
Texas	994·512	892·5	1972	447	346·632
Utah	55·524	55	76	44	6·576
Vermont	39·915	32	76	7	22·765
Virgin Islands	1·183	1	3	0	0·818
Virginia	229·427	247·5	322	88	80·013
Washington	97	92	187	40	25·385
West Virginia	159·537	177	243	56	56·649
Wisconsin	491·22	547·5	636	214	134·511
Wyoming	61·768	56	117	30	20·62

(1998) argue that, as a consequence of this, the panel tests may often reject the null of joint non-stationarity even if only one series is stationary. This is problematic as data shows that there is a large heterogeneity between the banking sectors of individual states. The descriptive statistics in table 1 show large differences in the average, range and variation in the number of commercial banks across the U.S. These most likely reflect the underlining heterogeneity in market conditions, regulations, shocks and many other factors. Due to heterogeneity amongst the U.S. states it is possible that there are states where hysteresis does not occurs. Thus the alternative hypothesis of only one series being stationary can be misleading. This is one of the reasons why it is important to take the unit root panel results in context when they are used as tests for hysteresis.

Due to these reasons, following sections also apply the univariate unit root tests to the data aggregated across all states as a robustness check, starting with an application of Zivot and Andrews unit root test to the total number of commercial banks across the United States. Next the Lee and Strazicich (2003) minimum LM unit root test is applied to the same series. Furthermore the minimum LM unit root test will be also applied to the panel data.

4.1.1 Zivot and Andrews univariate breakpoint unit root test

This sub-section tests the random walk result by applying the Zivot and Andrews (1992) unit root test to the total number of commercial banks in the United States (see figure 3). Zivot and Andrews unit root test is an extended version of the ADF test that allows for breakpoints in both level and trend. The breakpoint date is chosen endogenously by the test in order to avoid data mining.

This application of the Zivot and Andrews unit root test to the total number of commercial banks in the United States includes both break in the levels and in the trend. The breaks in a level are included to control for the aforementioned wave of branching and banking deregulation which occurred in late 1970s and early 1980s. It is reasonable to assume that this wave of deregulation should mainly affect the level of number of active commercial banks rather than the trend as it was relatively swift. The break in trend is included to capture the technological progress that started transforming the banking industry since early 1980s, and continues to do so till today. Because the technological progress in information technology and statistical modelling is more or less gradual, it is reasonable to assume that it would lead to break in a trend too rather than only in levels.

One complication that arises here is that the Zivot and Andrews breakpoint unit root test presupposes that both level and trend breakpoint happen at the same time. In our case, the regulatory and technological change indeed overlap to some extent as they occurred around 1980s. Nonetheless, the dating of these events is not completely clear. This could possibly lead to misspecification. However, the Lee and Strazicich two break point test is used as a robustness check in the next section and, as was already mentioned, both changes happen more or less concurrently and thus use of one breakpoint can be justified. With these caveats in mind we can turn our attention to details of the Zivot and Andrews model we will be using to test for the presence of unit root in the total number of active commercial banks in the United States. The model takes the following form:

$$N_t = \mu + \theta DM_t(\lambda) + \psi t + \phi DT_t(\lambda) + aN_{t-1} + \sum_{j=1}^{k} b_j \Delta N_{t-j} + \epsilon_t. \tag{41}$$

Here μ represents a 'drift' term, that allows for non-zero mean in the series. The $\theta DM_t(\lambda)$ is the break in the intercept. This dummy takes 0 before the break date (T_b) and 1 after. The break dummy is a function of λ because the Zivot and Andrews endogenizes the break point date. The λ represents a point which minimizes the test t statistics and is calculated by recurrent estimation of the t statistics at each potential break point. The point with the lowest (i.e. the most negative) t value has the highest probability of rejecting the unit root hypothesis, and thus can be interpreted as a point which gives the least favourable result for the null hypothesis of non-stationarity. Time trend is captured by ψt and $\phi DT_t(\lambda)$ represents the break in the trend. As in previous case, the break is function of λ since the breakpoint is endogenously determined at a point where the test statistics is most negative. The last part of equation 41, that is $aN_{t-1} + \sum_{j=1}^{k} b_j \Delta N_{t-j} + \epsilon_t$, consists of just ordinary ADF terms. N_{t-1} is the lag of levels of the dependent variable and ΔN_{t-j} are lags of the first difference of the dependent variable (see Verbeek (2008) for more detailed treatment), where the number of lags is chosen to eliminate the autocorrelation.

The null hypothesis for the Zivot and Andrews (1992) test is given by

$$N_t = \mu + N_{t-1} + \epsilon_t. \tag{42}$$

Thus the null hypothesis implies that the series is integrated without structural break. Moreover, it is worth noting that the equation 42 is virtually the same as the formal description of hysteresis provided earlier.

The natural starting point for application of the Zivot and Andrews unit root test is the determination of the optimal number of lags to include in the Zivot and Andrews test. This is done using the standard lag selection criteria including Akaike, Hannan-Quin, Schwartz and the Final predictor error[15]. In this case all four criteria select 3 lags (see table 2) as the most optimal.

Table 2: Lag selection

Criterion	1	2	3	4	5	6
AIC	10.175	8.573	8.532*	8.553	8.576	8.604
HQ	10.213	8.623	8.595*	8.629	8.664	8.705
SC	10.271	8.700	8.691*	8.744	8.799	8.859
FPE	26,237.920	5,285.254	5,074.964*	5,183.896	5,304.169	5,457.432
	7	8	9	10	11	
AIC	8.623	8.647	8.674	8.690	8.710	
HQ	8.737	8.774	8.813	8.842	8.874	
SC	8.910	8.966	9.024	9.073	9.124	
FPE	5,565.367	5,704.357	5,861.978	5,965.054	6,086.438	

FPE: Final prediction error.
AIC: Akaike information criterion.
SC: Schwarz information criterion.
HQ: Hannan-Quinn information criterion.
* denotes the selected number of lags.

The results of estimating the Zivot and Andrews breakpoint unit root test with three lags can be seen in the table 3. These show that the test cannot reject the null of stationary series at any standard confidence level. Therefore the test cannot reject the null hypothesis of a unit root process. This supports the hysteresis hypothesis outlined previously.

The figure 5 plots the test statistics of the Zivot and Andrews unit root test. The breakpoint is selected endogenously at the point where the test statistic (shown in figure 5) is the most negative. In our case the breakpoint occurs at the beginning of the 1970s. The exact point selected by the model is 1969. This is indeed a bit early, but still consistent with the historical narrative discussed earlier. Moreover, as the figure 5 shows the test statistic is very low during the whole period between 1970 and 1980. For example, the year 1979 has only slightly higher test statistics than 1969. However, in any case at no potential breakpoint date the test statistics crosses significance threshold, as the breakpoint is already chosen at a point where the test statistic is minimal.

Moreover, because the table 3 shows that break in a trend is not significant we also estimate Zivot and Andrews with breakpoint only in level (intercept). The results

---

[15] The maximum number of lags was determined using the Schwert (1989) principle. According to this principle the optimal maximum lag k to consider is given by $l_{max} = 12\left(\frac{t}{100}\right)^{(1/4)}$, where t is the length of the time dimension. With 82 observations the maximum number of lags rounded to closest whole number is 11.

Table 3: Zivot-Andrews Breakpoint Unit Root Test

t-stat	Critical values		
-4.259	1% lvl. t-stat.	-5.57	
	5% lvl. t-stat.	-5.08	
	10% lvl. t-stat.	-4.82	
	Break Date:	1969	
Dep. variable :	Total Number of Commercial Banks in the U.S. (N_t)		
Coefficients:	Estimate	Std. Error	t value
N_{t-1}	0.960	0.010	100.834***
ΔN_{t-1}	0.979	0.113	8.644***
ΔN_{t-2}	-0.290	0.157	-1.847*
ΔN_{t-3}	0.129	0.107	1.210
μ	549.898	131.041	4.196***
t	-0.311	1.208	-0.257
DM	115.237499	39.660055	2.906***
DT	-10.167	2.645	-3.844
Adjusted R-squared: 0.9995			
F-statistic (7,70): 2.296E+04			
Residual standard error: 61.68 on 70 degrees of freedom			

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

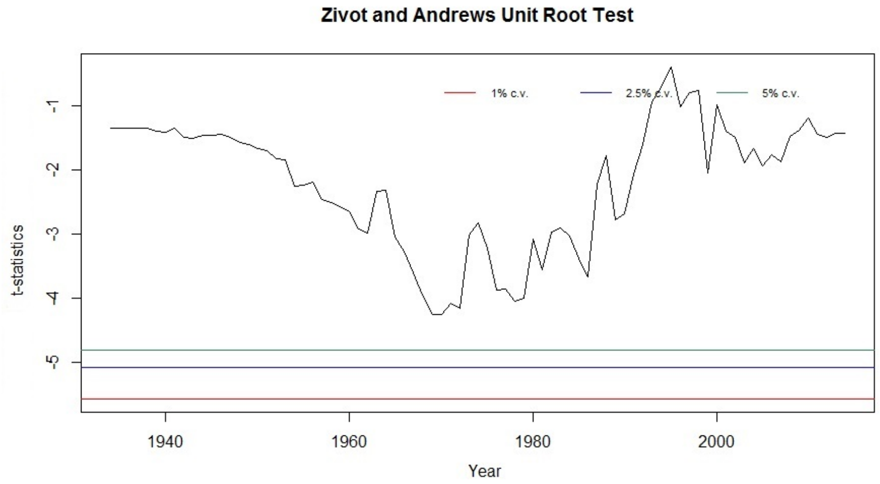


Figure 5: The Zivot and Andrews test statistics over every potential breakpoint. The breakpoint selected by the test is the point where the test statistics is minimized.

of this estimation are shown in the table 4 and on figure 6. Again the test cannot reject the null of unit root at any conventional significance level and at any potential breakpoint. However, in this case the potential breakpoint is identified much later in the series. In this case, minimization of the t-statistics puts the breakpoint to 1986 which is almost at the middle of the period where the deregulation and technological change was reshaping the U.S. banking industry.

Table 4: Zivot-Andrews Breakpoint Unit Root Test

t-stat	Critical values		
-3.5644	1% lvl. t-stat.	-5.34	
	5% lvl. t-stat.	-4.80	
	10% lvl. t-stat.	-4.58	
	Break Date:	1986	
Dep. variable :	Total Number of Commercial Banks in the U.S. (N_t)		
Coefficients:	Estimate	Std. Error	t value
N_{t-1}	0.982	0.005	192.549***
ΔN_{t-1}	0.889	0.120	7.383***
ΔN_{t-2}	-0.317	0.155	-2.054 **
ΔN_{t-3}	0.027	0.106	0.256
μ	235.2	73.93	3.181***
t	0.537	0.652	0.824
DM	-229.5	54.89	-4.183***
Adjusted R-squared: 0.9995			
F-statistic (6,71): 2.745E+04			
Residual standard error: 60.93 on 71 degrees of freedom			

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

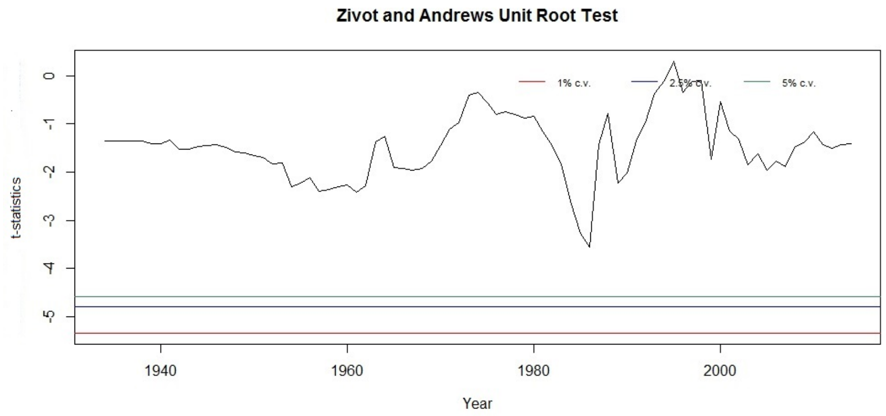


Figure 6: The Zivot and Andrews test statistics over every potential breakpoint. The breakpoint selected by the test is the point where the test statistics is minimized.

It is worth noting that in both cases the estimates for N_{t-1} are close to unity.

In the first and second specification the estimates of N_{t-1} coefficient are 0.960 and 0.982 respectively. Even if the tests could reject the null hypothesis, such high coefficients would indicate a high near unit root degree of persistence, and thus would still provide some evidence for hysteresis. Many authors, including Blanchard and Summers (1986), consider even near unit roots as an evidence for hysteresis. This is because unit roots can be considered only as a special case of hysteresis where all past shocks determine the present outcomes. Nonetheless, hysteresis is better described by systems where only certain shocks determine the present outcomes and these are often associated with near unit roots. However, beyond measuring the degree of persistence and testing for the presence of unit roots, these tests do not have deeper economic meaning as models following the structural approach.

4.1.2 Lee and Strazicich minimum LM unit root tests

This section checks the robustness of results from the Zivot and Andrews unit root test using the Lee and Strazicich minimum LM unit root test. Moreover, this section also takes advantage of the fact that Lee and Strazicich allows for multiple structural breaks and uses test specification which includes two structural breaks. This is done to allow for the possibility that breakpoints caused by the technological change and deregulation occurred at separate points in time.

Let us start with a short description of the test. Following Im et al. (2005), Nunes et al. (2004), N_t is considered to be determined by a data generating process such that

$$N_t = \delta_1 + Z_t\delta + \eta_t, \tag{43}$$

where

$$\eta_t = \rho N_{t-1} - \epsilon_t \iff \eta_t - \rho N_{t-1} = \epsilon_t. \tag{44}$$

As previously, N_t is the number of active banks. The Z_t is a vector which includes exogenous variables such as the level and trend breaks. Thus,

$$Z_t = \left(DM_{ti}(\lambda), t_i, DT_{ti}(\lambda)\right), \tag{45}$$

here the i subscript stands for the number of structural breaks. This work considers maximum of 2 breaks and so here $i = \{1, 2\}$. The level breakpoint dummy is defined as $DM_t = 1$ when $t > T_b$ and 0 otherwise. The breakpoint trend dummy will be defined as $DT_t = 1$ if $(t > T_b)(t - T_b)$ and 0 otherwise. Again breakpoint dummies are endogenously chosen, based on the test statistic λ to avoid data mining. This is once again done at the point where the test statistic is minimized (Lee et al., 2004). Finally, ϵ_t is assumed to be *iid* $N(0, \sigma^2)$.

Under the null hypothesis $\rho = 1$, that is the series contains an unit root. Following, Nunes et al. (2004) and Chou (2007) restricted minimum likelihood estimates of the coefficients are obtained from

$$\Delta N_t = \Delta Z_t \delta + \upsilon_t, \tag{46}$$

where Δ denotes the first differences of a series. Moreover, the series is also detrended using

$$\tilde{S}_t = N_t - Z_t \tilde{\delta} - (N_1 - Z_1 \tilde{\delta}). \tag{47}$$

Allowing for autocorrelation by including lags, the LM unit root test statistics for testing null of $\varphi = 0$ is estimated from the following regression

$$\Delta \tilde{S}_t = \Delta Z_t \delta + \varphi \tilde{S}_{t-1} + \sum_{j=1}^{k} g_j \tilde{S}_{t-j} + \upsilon_t. \tag{48}$$

The null and alternative hypothesis of this test is given by equation 49 and 50:

$$H_0 : N_t = \mu_0 + d_1 DM_{t1} + d_2 DM_{t2} + d_3 DT_{t1} + d_4 DT_{t2} + N_{t-1} + \nu_{t1} \tag{49}$$

$$H_A : \mu_1 + d_1 DM_{t1} + d_2 DM_{t2} + d_3 DT_{t1} + d_4 DT_{t2} + d_5 t + \nu_{t2}. \tag{50}$$

After this brief introduction of the Lee and Strazicich minimum LM unit root test, let us turn to its aplication to the series on total number of active banks in the United States. Starting with a one break minimum LM test which provides direct robustness check to the Zivot and Andrews unit root test.

The table 5 shows the results from one-breakpoint Lee and Strazicich minimum LM unit root test with three lags. As in the Zivot and Andrews case we cannot reject the null of unit root in the series at any confidence level. In this case the test identifies breakpoint much later in the sample than the Zivot and Andrews unit root test which allowed for breakpoint in trend. Whereas previously the breakpoint was determined to occur in 1969 the minimum LM test puts it further to 1988. This is more in line with the historical perspective as 1980s seen large share of the U.S. banking deregulation and also numerous technological advances. As in the Zivot and Andrews unit root test, the trend breakpoint is found to be insignificant. This may seem puzzling as the series clearly shows a downward trend. This may be due to the sharpness of the decline, as the number of banks decrease by a half in a span of only two decades. Because of this, a version of the test that includes only breakpoint in levels is again estimated as well.

Table 5: Lee and Strazicich Breakpoint Unit Root Test Allowing for Both Intercept and Trend Breaks

LM t-stat	Critical values:	1% lvl. t-stat	5% lvl. t-stat.	10% lvl. t-stat.
	Location of T_b (T_b/T)			
-3.827	0.1	-5.11	-4.50	-4.21
	0.2	-5.07	-4.47	-4.20
	0.3	-5.15	-4.45	-4.18
	0.4	-5.05	-4.50	-4.18
	0.5	-5.11	-4.51	-4.17
	Break Date:	1988 (0.7)		
Dep. variable :	Detrended Total Number of Commercial Banks in the U.S. ($\tilde{S}_t$)			
Coefficients:	Estimate	Std. Error	t value	
$\tilde{S}_{t-1}$	-0.054	0.014	-3.827***	
$\Delta\tilde{S}_{t-1}$	0.902	0.09245	9.759***	
$\Delta\tilde{S}_{t-2}$	0.103	0.118	0.869	
$\Delta\tilde{S}_{t-3}$	-0.004	0.102	-0.037	
t	-111.3	13.25	-8.399 ***	
DM	194.5	77.65	2.505**	
DT	-32.24	28.98	-1.113	
Adjusted R-squared: 0.8961				
F-statistic (7,71): 97.13				
Residual standard error: 69.79 on 71 degrees of freedom				

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The result of this estimation is shown on table 6. Since in this case only a level breakpoint is used the critical values are no longer dependent on the breakpoint position. In this case, the test can reject the null hypothesis at 10% confidence interval but only barely. The test statistics is -3.242 while the critical value is -3.211. Therefore, the inference changes at the 10% confidence interval but remains unchanged at more stringent 5% and 1% confidence intervals.

Because the plots of test statistics from the Zivot and Andrews indicate that there could be potentially two breakpoints, the two breakpoint minimum LM test is preformed as well. In this case the test can again reject null of unit root at the 10% confidence level but not at 5% or 1%. Therefore, the evidence for unit root hysteresis cannot be rejected at the standard 5% confidence.

Moreover, as in the Zivot and Andrews case the coefficients of the Lee and Strazicich with one breakpoint are high enough to be considered as an evidence for near unit roots even if the test would significantly reject the unit root hypothesis. In this case the coefficient values $(1 - \varphi)$ are 0.946 for model which allows for break in both intercept and trend and 0.984 in the case where only break in intercept is allowed. In the two breakpoint case the estimated coefficient is only 0.879. Nonetheless, this could still be considered as a modest near-unit root case. Thus these results would point to the hysteresis even at the 10% confidence level where the null of unit root could be rejected.

Table 6: Lee and Strazicich Breakpoint Unit Root Test Allowing for Intercept Break

LM t-stat	Critical values		
-3.242*	1% lvl. t-stat.	-4.239	
	5% lvl. t-stat.	-3.566	
	10% lvl. t-stat.	-3.211	
	Break Date:	1988	
Dep. variable :	Detrended Total Number of Commercial Banks in the U.S. ($\tilde{S}_t$)		
Coefficients:	Estimate	Std. Error	t value
$\tilde{S}_{t-1}$	-0.016	0.005	-3.242***
$\Delta\tilde{S}_{t-1}$	0.923	0.093	9.888***
$\Delta\tilde{S}_{t-2}$	0.097	0.120	0.806
$\Delta\tilde{S}_{t-3}$	-0.036	0.095	-0.381
t	-77.382	13.437	-5.759 ***
DM	205.465	80.066	2.566**
Adjusted R-squared: 0.8908			
F-statistic (6,72): 107			
Residual standard error: 71.56 on 72 degrees of freedom			

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 7: Lee and Strazicich Breakpoint Unit Root Test With Two Breaks

LM t-stat	Critical values:	1% lvl. t-stat	5% lvl. t-stat.	10% lvl. t-stat.
	Loc. of $T_{b1}\backslash T_{b2}$	0.4	0.4	0.4
-5.446*	0.2	-6.16	-5.59	-5.27
	0.4	NA	NA	NA
	0.6	NA	NA	NA
	Loc. of $T_{b1}\backslash T_{b2}$	0.6	0.6	0.6
	0.2	-6.41	-5.74	-5.32
	0.4	-6.45	-5.67	-5.31
	0.6	NA	NA	NA
	Loc. of $T_{b1}\backslash T_{b2}$	0.8	0.8	0.8
	0.2	-6.33	-5.71	-5.33
	0.4	-6.42	-5.65	-5.32
	0.6	-6.32	-5.73	-5.32
	Break Date:	T_{b1} 1968 (0.4)	T_{b2} 1989 (0.7)	
Dep. variable :	Detrended Total Number of Commercial Banks in the U.S. ($\tilde{S}_t$)			
Coefficients:	Estimate	Std. Error	t value	
$\tilde{S}_{t-1}$	-0.121	0.022	-5.446***	
$\Delta\tilde{S}_{t-1}$	0.850	0.095	8.960***	
$\Delta\tilde{S}_{t-2}$	-0.001	0.133	-0.008	
$\Delta\tilde{S}_{t-3}$	0.151	0.104	1.451	
t	-153.0	18.42	-8.305***	
DM_1	-87.29	0.735	-1.188	
DT_1	143.5	361.5	3.970***	
DM_2	140.8	76.83	1.833*	
DT_2	-174.5	40.51	-4.307***	
Adjusted R-squared: 0.8994				
F-statistic (9,69): 78.44				
Residual standard error: 68.7 on 69 degrees of freedom				

Note: the break location is expressed as T_{bi}/T

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

4.1.3 Panel Minimum LM Unit Root Test With Breakpoints

This subsection extends the minimum LM test described in the previous section to panel setting. Following Im et al. (2005) panel minimum LM test is estimated by applying individual minimum LM tests to each member of a panel allowing for one or two structural breaks. Afterwards a joint panel test statistics is applied to avoid the problems of false negatives and positives that may easily occur when a large number of series is involved.

The panel test statistics tests the null of joint unit root against the alternative that at least one series is stationary. The first step in calculating the panel test statistic is averaging of LM t statistics across the panel:

$$LM_{\overline{NT}} = \frac{1}{\mathcal{N}} \sum_{q=1}^{\mathcal{N}} \text{LM } t_q\text{-stat}, \tag{51}$$

where $\mathcal{N}$ is the total number of states and LM t_q-stat are the LM t statistics of each individual member of the panel.

Finally the panel test statistics is given as the standardized difference between the average LM t statistics and the expected LM t-statistics under the null hypothesis (Im et al., 2005).

$$LM_{\text{Panel}} = \frac{\sqrt{\mathcal{N}}\,[LM_{\overline{NT}} - E(L_T)]}{\sqrt{V(L_T)}} \tag{52}$$

The expected values and variances of the test statistics under the null ($E(L_T)$ and $V(L_T)$) are provided by Im et al. (2005) using stochastic simulations with 500,000 replications.

The results of panel LM test can be seen in table 8. The exact test specification in each series was determined using the general to specific approach. As in the univariate case the critical values depend on the location of break(s) (T_{b1}, T_{b2}), but due to space consideration only the most relevant critical values are reported. Table 8 also shows the intermediate results for each state or territory separately. At 5% confidence level LM unit root test is able to reject the null of unit root in 4 out of 56 cases. The states for which LM unit root test rejects the null are Hawaii, Texas, West Virginia and Wyoming. Consequently the panel LM test rejects the null of joint unit root test in favour of alternative hypothesis stating that at least one series being stationary.

This seems to provide an evidence against the hysteresis in individual states but the results from the panel unit root tests should be interpreted with great caution.

Table 8: Lee and StrazicichPanel Minimum LM Unit Root Test With Breakpoints

U.S. State or territory	Break dates (T_{b1}, T_{b2})	Break type	# of lags (k)	LM t-stat	Critical values (LM t-stat.)		
					1% lvl.	5% lvl.	10% lvl.
Alabama	(1971, 1989)	Both	6	-5.515*	-6.42	-5.65	-5.32
Alaska	(1950, 1991)	Both	2	-5.371*	-6.33	-5.71	-5.33
American Samoa	(1981)	Levels	0	-1.732	-4.24	-3.57	-3.21
Arizona	(1980)	Both	5	-4.191*	-5.05	-4.50	-4.18
Arkansas	(1953, 1991)	Both	3	-5.444*	-6.33	-5.71	-5.33
California	(1978)	Both	7	-3.629	-5.11	-4.51	-4.17
Colorado	(1979, 1996)	Both	1	-4.478	-6.32	-5.73	-5.32
Connecticut	(1955, 1974)	Both	3	-4.789	-6.45	-5.67	-5.31
Delaware	(1962, 1993)	Both	4	-4.380	-6.42	-5.65	-5.32
District of Columbia	(1981, 1999)	Both	7	-4.319	-6.32	-5.73	-5.32
Fed. States of Micronesia	(1985)	Levels	0	-1.733	-4.24	-3.57	-3.21
Florida	(1970)	Both	6	-4.408*	-5.11	-4.51	-4.17
Georgia	(1971)	Both	3	-3.185	-5.11	-4.51	-4.17
Guam	(1971, 1990)	Levels	1	-3.473	-6.32	-5.73	-5.32
Hawaii	(1981, 1997)	Both	5	-7.103***	-6.32	-5.73	-5.32
Idaho	(1958, 1975)	Both	7	-4.532	-6.45	-5.67	-5.31
Illinois	(1966, 1987)	Both	1	-3.799	-6.42	-5.65	-5.32
Indiana	(1969, 2002)	Both	5	-5.394*	-6.42	-5.65	-5.32
Iowa	(1974, 1991)	Both	5	-4.704	-6.32	-5.73	-5.32
Kansas	(1954, 1980)	Both	3	-4.868	-6.41	-5.74	-5.32
Kentucky	(1987, 1997)	Both	1	-4.813	-6.32	-5.73	-5.32
Louisiana	(1980, 1992)	Both	1	-5.431*	-6.32	-5.73	-5.32
Maine	(1970, 1986)	Both	3	-5.210	-6.45	-5.67	-5.31
Maryland	(1985)	Both	2	-3.468	-5.05	-4.50	-4.18
Massachusetts	(1975, 1989)	Both	6	-5.146	-6.45	-5.67	-5.31
Michigan	(1971, 1992)	Both	3	-5.016	-6.32	-5.73	-5.32
Minnesota	(1978)	Both	4	-3.801	-5.11	-4.51	-4.17
Mississippi	(1991)	Both	4	-3.067	-5.05	-4.50	-4.18
Missouri	(1970, 1989)	Both	4	-4.237	-6.32	-5.73	-5.32
Montana	(1976, 1994)	Both	2	-3.977	-6.32	-5.73	-5.32
Nebraska	(1958, 1981)	Both	7	-4.389	-6.41	-5.74	-5.32
Nevada	(1992)	Both	4	3.766	-5.05	-4.50	-4.18
New Hampshire	(1990)	Levels	6	-2.430	-4.23	-3.57	-3.21
New Jersey	(1989, 2006)	Both	5	-4.758	-6.32	-5.73	-5.32
New Mexico	(1968, 2000)	Both	4	-4.545	-6.42	-5.65	-5.32
New York	(1960, 1975)	Both	7	-4.979	-6.16	-5.59	-5.27
North Carolina	(1964, 1994)	Both	3	-4.621	-6.45	-5.67	-5.31
North Dakota	(1948, 1983)	Both	2	-3.607	-6.45	-5.67	-5.31
Ohio	(1983)	Levels	3	-1.332	-4.24	-3.57	-3.21
Oklahoma	(1972, 1990)	Both	4	-5.399	-6.32	-5.73	-5.32
Oregon	(1975, 1988)	Both	2	-4.577	-6.32	-5.73	-5.32
Pennsylvania	(1952)	Levels	7	-3.172	-4.24	-3.57	-3.21
Puerto Rico	(1952, 1974)	Levels	1	-2.412	-6.16	-5.59	-5.27
Rhode Island	(1966, 1971)	Levels	6	-2.865	-6.45	-5.67	-5.31
South Carolina	(1992)	Levels	4	-1.574	-4.24	-3.57	-3.21
South Dakota	(1996)	Both	5	-4.374*	-5.15	-4.45	-4.18
Tennessee	(1971, 1997)	Both	6	-5.210	-6.32	-5.73	-5.32
Texas	(1988)	Both	4	-4.794**	-5.05	-4.50	-4.18
Utah	(1973, 1994)	Both	4	-4.465	-6.32	-5.73	-5.32
Vermont	(1966)	Both	1	-2.487	-5.05	-4.50	-4.18
Virgin Islands	(1958, 1975)	Levels	1	-3.006	-6.45	-5.67	-5.31
Virginia	(1977, 1979)	Levels	3	-3.901	-6.45	-5.67	-5.31
Washington	(1954, 1975)	Both	3	-3.731	-6.16	-5.59	-5.27
West Virginia	(1971, 2006)	Both	6	-6.616***	-6.32	-5.73	-5.32
Wisconsin	(1958, 2006)	Both	6	-5.019	-6.33	-5.71	-5.33
Wyoming	(1976, 1996)	Both	4	-5.793**	-6.32	-5.73	-5.32

Panel LM test statistics: -28.315***. The 1, 5 and 10% critical values for the panel LM unit root tests (with or without breaks) are -2.326, -1.645 and -1.282, respectively.

Notes: break type is an indicator of whether break was assumed to occur only in levels (Levels) or also in trend (Both).
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

There are several reasons for this. First, although it was argued that the hysteresis effects in banking sector should be relatively more likely due its characteristics, it is still possible that these did not occur in few states. This could be due to pure luck or because of the heterogeneity between the individual banking markets across the states. It is possible that in some states the fixed per-period costs are actually higher than the fixed entry costs and in this case hysteresis would not occur. Other prerequisites for hysteresis may be lacking as well.

Second, as equation 52 shows, the test statistics of the panel LM unit root test is increasing (although at a decreasing rate) in the size of cross-section of the panel. This is because the panel LM unit root test tests the null of joint unit root against an alternative that at least one series is stationary. Naturally, as the cross-section increases, it becomes more likely that some of the negative results may just be false negatives. Thus, it is often observed that the panel LM unit root test often rejects the null in a presence of even one series that can reject the null at a given significance level, as observed by Taylor and Sarno (1998), or even in cases where there are many findings near the significance threshold.

Third, although the breakpoint unit root tests are more consistent when testing for a unit root, there is a chance that hysteresis effect may be confused for a structural break. Of course, it is still important to control for structural breaks caused by other factors such as the technological change or change in regulation. Because of this, it is possible that some of the endogenously selected breaks may actually capture the hysteresis effects themselves.

As table 8 shows, out of the four states, including Hawaii, Texas, West Virginia and Wyoming where test rejects the null of a unit root, in Hawaii, West Virginia and Wyoming test identifies structural break quite late in the series. For example, in both West Virginia and Wyoming the second break is selected to be in 2006, and in the case of Hawaii to be in 1997. This dating seems to be too late to capture the effect of deregulation and the start of implementation of new technology that affected the industry. Therefore, as a robustness check the LM unit root test was applied again to these states allowing only for one structural break.

The table 9 shows the results of this robustness check. As it is possible to see the results from Hawaii, West Virginia and Wyoming completely reversed themselves. The results for Texas did not change as the original estimation included only one break. Moreover, if these four series are considered as a separate panel, the panel LM unit root test statistics cannot reject the null of joint unit root. However, if the test statistics for the full panel is calculated using these new results the test can still reject the hypothesis of joint unit root in favour of at least one series being stationary at all standard confidence levels. These results highlight both the problems of properly controlling for the structural shift and that the alternative hypothesis of one or more series being stationary can be too strict. Due to these reasons and the fact that in most series, the test was not able to reject the null of unit root, these results should be interpreted as showing that the hysteresis effects may not be present in all states rather than rejecting the hysteresis hypothesis altogether. This being said clearly more research is needed to resolve the aforementioned issues and to provide more evidence either in favour or against the hysteresis in the banking sector.

Table 9: Lee and Strazicich Panel Minimum LM Unit Root Test With Breakpoints

U.S. State or territory	Break dates (T_{b1}, T_{b2})	Break type	# of lags (k)	LM t-stat	Critical values (LM t-stat.) 1% lvl.	5% lvl.	10% lvl.
Hawaii	(1992)	Both	1	3.190	-5.05	-4.50	-4.18
Texas	(1988)	Both	4	-4.794**	-5.05	-4.50	-4.18
West Virginia	(1988)	Both	3	-3.672	-5.05	-4.50	-4.18
Wyoming	(1991)	Both	3	-3.511	-5.05	-4.50	-4.18

Panel LM test statistics: -0.074. The 1, 5 and 10% critical values for the panel LM unit root tests (with or without breaks) are -2.326, -1.645 and -1.282, respectively.

Notes: break type is an indicator of whether break was assumed to occur only in levels (Levels) or also in trend (Both).
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

5 Policy Implications of Hysteresis in the Banking Industry

This section looks briefly at some of the policy implications of the generalised Monti-Klein model introduced in the third section. It also highlights the potential of this model to provide further insights into the controversy between the concentration stability and concentration instability views. The implications of hysteresis itself for financial policy are discussed as well.

An interesting feature of the hysteresis model, developed herein, is that the model is broad enough to encompass both the financial stability and instability view of banking concentration. According to Allen and Gale (2000), the higher concentration enhances financial stability because it provides higher market power and consequently higher profits that can serve as a 'cushion' during the times of financial distress. Similar result can be seen in the model presented in the third section. This is because the width of the 'band of inaction' is inversely related to the number of banks in the market. As a result of this, as the number of banks decreases larger shocks are required to push the equilibrium rate of profit out of the 'band of inaction'. In this model bank entries and failures can happen only outside the 'band of inaction', and thus wider band implies, ceteris paribus, that bank failures are less likely. However, because the volume of loans and deposits that each bank manages increases as the market gets smaller, the bank failures may be more fatal even if less likely, thus leading to the 'megabank' problem (See Yeyati and Micco, 2007). On the other hand, as the number of banks approaches infinity, the 'band of inaction' gets narrower, making the bank failures more likely but less costly since the fraction of loans and deposits managed by each bank decreases. Thus there is a trade-off between the probability of failure and the size of potential loses that the failure entails.

The model also entails a welfare trade-off as at higher levels of market concentration banks' profits and surplus increase while the consumer surplus of lenders and producer surplus of depositors decrease. In a standard Monti-Klain framework the welfare losses of lenders and depositors from higher banking concentration would outweigh the gains of banks (Freixas and Rochet, 2008). But if lenders and depositors value the stability of banking sector, restricting banking completion may be potentially welfare improving.

The exact nature of these trade-offs depends both on the parameters of the model and social preferences. Thus the controversy between the concentration-stability and concentration-instability camps can be ultimately resolved only empirically. Unfortunately, the empirical evidence is so far inconclusive (Beck, 2007, Berger et al., 2004, Boyd and De Nicolo, 2005, Nicoló et al., 2004). Nevertheless, this model provides a potential framework for explicitly modelling these trade-offs.

Moreover, hysteresis in itself can represent a policy challenge. The reason for this is that hysteresis ultimately implies that banking concentration and competition are intrinsically unstable. The results proven in the third section show that the equilibrium number of banks active in a market critically depends on both the initial conditions and the past shocks. Moreover, it was proven that the number of banks in the industry follows a random walk. In the Monti-Klein model, both bank concentration and competition are directly related to the number of banks in the market (Freixas and Rochet, 2008). Therefore, by extension, the concentration and competition can also be expected to evolve according to the random walk. An important caveat worth mentioning is that while this holds for the Monti-Klein model, there are other models where competition does not directly depend only on the number of competing banks. Nevertheless, it is still considered an important factor and many measures of banking competition at least implicitly also include the number of banks in the market (Bikker and Haaf, 2002).

Such a result has serious implications for a policy as the proper regulation of financial activities depends on the concentration and competitiveness of the industry (Barth et al., 2008, Brunnermeier et al., 2009, Matutes and Vives, 2000). Hysteresis implies that the concentration and competitiveness of banking industry can change as a result of random shocks without any deeper structural change. As a consequence of this, proper regulatory approach should pay attention even to temporary shocks, taking more macroprudential and dynamic approach to the bank regulation. This is far from being the first paper to call for such an approach to regulation. Since the Great Recession, number of authors called for both more macroprudential and dynamic approach to the bank regulation (see Blanchard et al., 2010, Gauthier et al., 2010, Gersbach and Rochet, 2012, Hanson et al., 2011, and the sources cited therein). Nonetheless, this work provides additional arguments in favor of this approach.

Furthermore, the model developed herein also provides several ways of potentially avoiding hysteresis. This is because, as argued in the third section, banking regulation can have an impact on both entry and per-period fixed costs. In turn the width of the 'band of inaction' depends on these costs. Therefore, by increasing fixed entry costs and decreasing fixed per-period costs, regulators can decrease the chance that random shocks will lead to the hysteresis effect. However, while hysteresis should not be ignored, avoiding it at any cost should not be the sole objective of a public policy. In some cases hysteresis can be potentially beneficial. For example, if future evidence provides more support for the concentration instability view then policy makers may strive to encourage the hysteresis effects that lead to increasing competition. Thus hysteresis should not only be viewed as a problem but also as an opportunity since it can potentially be used by regulators to adjust the level of concentration and competition of a market.

There is also another channel through which the policy can intervene in the hysteresis effects. This is the monetary policy channel. The expression for the equilibrium profits, described by equation 16, shows that the profits depend on the central bank's rate r and the reserve coefficient α. Hysteresis occurs when exogenous shocks force the equilibrium profit outside the 'band of inaction' and the changes in the central bank's rate and the reserve requirements can counterbalance the effects of such shocks. Thus central banks can potentially intervene to prevent or to engineer hysteresis in the financial markets. However, just because a central bank can intervene it does not necessarily mean that it should. Monetary policy has serious implications for the whole economy and consequently the monetary policy has to take a macroeconomic 'birds eye' view of the economy (Romer, 2012). This being said, the banking sector and financial markets, more broadly, are very important for wider economy. Empirical studies show that the financial sector has first order impact on economic growth and that the benefits of healthy financial sector are shared disproportionately by poor and small firms (see Beck, 2007, Beck and Demirguc-Kunt, 2006, Honohan, 2004, Levine, 1997, 2005, and sources cited therein). Therefore, while there may be more important considerations when crafting the monetary policy, the possibility of hysteresis should not be completely absent from considerations as well. Additionally, this paper assumes the monetary authority to be completely passive but in practice, changes to the central bank's rate itself could lead to the hysteresis effect. Thus, monetary authorities should be aware of this and take their potential impact on competition and concentration into consideration as well.

This is far from being an exhaustive list of policy implications, as the Monti-Klein framework itself offers a rich set of insights. Introducing hysteresis in such a rich framework is bound to change many of these and provide a lot of new ones. However, these are too numerous to be crammed in one paper. Moreover, the policy discussion here was only brief as in many cases policy implications of the hysteresis model developed herein are contingent on the results from wider empirical literature, but these are often mixed. Therefore, more research is needed to provide better analysis of the implications that hysteresis has for policy.

6 Conclusion

This paper introduced hysteresis into the banking competition using the generalised Monti-Klein model of banking industry. This was done by adding fixed per-period and bank specific non-recoverable entry costs to the canonic version of the model, and by extending it into quasi-dynamic setting. The latter was done by a comparative static exercise, considering small deviations from the static solution to the model.

The economic story behind the model goes as follows. Because new banks face bank specific non-recoverable fixed costs, they enter the market only if profits are sufficiently high to cover these costs. However, once the bank enters the market, the entry costs become sunk ex-post. Because of this entrant banks will not exit the market even after the profit decreases, as long as the profit is sufficient to cover the unit costs and fixed per-period costs. As a result, even temporary shocks that affect the rate of profits can have permanent effects on bank concentration and competition.

Furthermore, this paper applies the unit root approach to provide an evidence in support of the hysteresis model developed herein. Specifically, this paper uses both Zivot and Andrews and Lee and Strazicich unit root tests to test the implications of the model. These empirical tests generally find a support for hysteresis in the U.S. banking industry. To be more specific, the univariate Zivot and Andrews and Lee and Strazicich are unable to reject the null hypothesis of unit root at the 5% confidence interval. This result is also robust to various test specification. However, the panel extension of the Lee and Strazicich unit root test, applied to the data from individual U.S. states rejected the null of a joint unit root in all states. Nonetheless, this result is caused by the rejection of unit root hypothesis only in four states. In three of these states the result changes if only one breakpoint is allowed. Moreover, in all other states the test cannot reject the null of unit root. These results imply that there is some evidence for hysteresis in the banking sectors of most of the U.S. states, but not in all.

The model of hysteresis provides several implications for policy. First, this model offers a framework that can be potentially used for evaluating various trade-offs between the levels of competition and financial instability in a partial equilibrium setting. Second, the hysteresis result in itself can represent both challenges and opportunities for policy makers. On one hand, hysteresis may lead to profound changes in levels of competition and industrial concentration that can catch regulators unaware if they do not pay close attention to these changes. On the other hand, it is possible to prevent negative hysteresis shocks and create positive ones through a proper policy. However, further research is needed to provide more refined policy advice.

References

Allen, F. and Gale, D. (2000). *Comparing financial systems*. MIT press.

Amable, B., Henry, J., Lordon, F., and Topol, R. (1994). Strong hysteresis versus zero-root dynamics. *Economics Letters*, 44(1):43–47.

Amable, B., Henry, J., Lordon, F., and Topol, R. (1995). Hysteresis revisited: a methodological approach. *The natural rate of unemployment: reflections on*, 25:153–180.

Archibald, G. (1995). Hysteresis and memory in the labour market. *The Natural Rate of Unemployment: Reflections on 25 Years of the Hypothesis*, page 90.

Baldwin, R. (1988). Hysteresis in import prices: the beachhead effect.

Baldwin, R. (1989). Sunk-cost hysteresis.

Baldwin, R. (1990). Hysteresis in trade. *Empirical Economics*, 15(2):127–142.

Balke, N. S. and Fomby, T. B. (1997). Threshold cointegration. *International economic review*, pages 627–645.

Barth, J. R., Caprio, G., and Levine, R. (2006). Cambridge University Press, 1st edition edition.

Barth, J. R., Caprio, G., and Levine, R. (2008). *Rethinking bank regulation: Till angels govern*. Cambridge University Press.

Beck, T. (2007). Bank concentration and fragility. impact and mechanics. In *The risks of financial institutions*, pages 193–234. University of Chicago Press.

Beck, T., Demirgüç-Kunt, A., and Levine, R. (2007). Finance, inequality and the poor. 12(1):27–49.

Beck, T. and Demirguc-Kunt, A. (2006). Small and medium-size enterprises: Access to finance as a growth constraint. *Journal of Banking & Finance*, 30(11):2931–2943.

Belke, A. and Göcke, M. (2001). Exchange rate uncertainty and employment: an algorithm describing 'play'. *Applied Stochastic Models in Business and Industry*, 17(2):181–204.

Belke, A. and Göcke, M. (2005). Real options effects on employment: does exchange rate uncertainty matter for aggregation? *German Economic Review*, 6(2):185–203.

Belke, A. H., Göcke, M., and Werner, L. (2014). Hysteresis effects in economics–different methods for describing economic path-dependence.

Berger, A. N. (2003). The economic effects of technological progress: Evidence from the banking industry. *Journal of Money, credit, and Banking*, 35(2):141–176.

Berger, A. N., Demirguc-Kunt, A., Levine, R., and Haubrich, J. G. (2004). Bank concentration and competition: An evolution in the making. *Journal of Money, Credit, and Banking*, 36(3):433–451.

Berger, A. N. and DeYoung, R. (2002). Technological progress and the geographic expansion of the banking industry.

Bernanke, B. S. (1983). Non-monetary effects of the financial crisis in the propagation of the great depression.

Besanko, D. and Thakor, A. V. (1992). Banking deregulation: Allocational consequences of relaxing entry barriers. *Journal of Banking & Finance*, 16(5):909–932.

Bikker, J. A. and Haaf, K. (2002). Measures of competition and concentration in the banking industry: a review of the literature. pages 1–46.

Blanchard, O., Dell'Ariccia, G., and Mauro, P. (2010). Rethinking macroeconomic policy. *Journal of Money, Credit and Banking*, 42(s1):199–215.

Blanchard, O. J. and Summers, L. H. (1986). Hysteresis and the european unemployment problem. *NBER macroeconomics annual*, 1:15–78.

Bowen, H. P., Hollander, A., and Viaene, J.-M. (2012). *Applied international trade.* Palgrave Macmillan.

Boyd, J. H. and De Nicolo, G. (2005). The theory of bank risk taking and competition revisited. *The Journal of finance*, 60(3):1329–1343.

Boyd, J. H. and Runkle, D. E. (1993). Size and performance of banking firms: Testing the predictions of theory. *Journal of monetary economics*, 31(1):47–67.

Brunnermeier, M. K., Crockett, A., Goodhart, C. A., Persaud, A., and Shin, H. S. (2009). *The fundamental principles of financial regulation*, volume 11. ICMB, Internat. Center for Monetary and Banking Studies.

Buiter, W. H. (1987). The right combination of demand and supply policies: the case for a two-handed approach.

Chou, W. L. (2007). Performance of lm-type unit root tests with trend break: A bootstrap approach. *Economics Letters*, 94(1):76–82.

Cross, R. (1994). The macroeconomic consequences of discontinuous adjustment: Selective memory of non-dominated extrema. *Scottish Journal of Political Economy*, 41(2):212–221.

Cross, R. (1995). *The natural rate of unemployment: Reflections on 25 years of the hypothesis*. Cambridge University Press.

Cross, R. and Allan, A. (1988). On the history of hysteresis. *Unemployment, Hysteresis and the Natural Rate Hypothesis*, pages 26–38.

Cross, R. B. (1987). Hysteresis and instability in the natural rate of unemployment. *The Scandinavian Journal of Economics*, pages 71–89.

Dell'Ariccia, G., Friedman, E., and Marquez, R. (1999). Adverse selection as a barrier to entry in the banking industry. *The RAND Journal of Economics*, pages 515–534.

DeLong, J. B. and Summers, L. H. (2012). Fiscal policy in a depressed economy. *Brookings Papers on Economic Activity*, 2012(1):233–297.

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

Dickey, D. A. and Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American statistical association*, 74(366a):427–431.

Dixit, A. (1989). Entry and exit decisions under uncertainty. 97(3):620–638.

Dixit, A. (1992). Investment and hysteresis. 6(1):107–132.

Engle, R. F. and Granger, C. W. (1987). Co-integration and error correction: representation, estimation, and testing. *Econometrica: journal of the Econometric Society*, pages 251–276.

Franz, W. (1987). Hysteresis, persistence and the nairu. a theoretical and empirical analysis for the federal republic of germany.

Franz, W. (1990). Hysteresis in economic relationships: An overview. pages 1–17.

Freixas, X. and Rochet, J.-C. (2008). *Microeconomics of banking*. MIT press.

Frisch, R. and Waugh, F. V. (1933). Partial time regressions as compared with individual trends. *Econometrica: Journal of the Econometric Society*, pages 387–401.

Gauthier, C., Lehar, A., and Souissi, M. (2010). Macroprudential regulation and systemic capital requirements.

Georgescu-Rogen, N. (1971). *The Entropy Law and The Economic Process*. Harvard University Press.

Gersbach, H. and Rochet, J.-C. (2012). Aggregate investment externalities and macroprudential regulation. *Journal of Money, Credit and Banking*, 44(s2):73–109.

Göcke, M. (1994). An approximation of the hysteresis loop by linear partial functions. econometric modelling and estimation/eine approximation der hysteresis-schleife über lineare teilfunktionen. ökonometrische modellbildung und schätzung. *Jahrbücher für Nationalökonomie und Statistik*, 213(5):572–596.

Göcke, M. (2001). A macroeconomic model with hysteresis in foreign trade. *Metroeconomica*, 52(4):449–473.

Gomes, F. A. R. and da Silva, C. G. (2009). Hysteresis versus nairu and convergence versus divergence: The behavior of regional unemployment rates in brazil. *The Quarterly Review of Economics and Finance*, 49(2):308–322.

Hallett, A. H. and Piscitelli, L. (2002). Testing for hysteresis against nonlinear alternatives. *Journal of Economic Dynamics and Control*, 27(2):303–327.

Hamilton, J. D. (1994). *Time series analysis*, volume 2. Princeton university press Princeton.

Hanson, S. G., Kashyap, A. K., and Stein, J. C. (2011). A macroprudential approach to financial regulation. *The Journal of Economic Perspectives*, 25(1):3–28.

Honohan, P. (2004). Financial sector policy and the poor selected findings and issues. (43).

Huang, R. and Ratnovski, L. (2011). The dark side of bank wholesale funding. *Journal of Financial Intermediation*, 20(2):248–263.

Im, K.-S., Lee, J., and Tieslau, M. (2005). Panel lm unit-root tests with level shifts. *Oxford Bulletin of Economics and Statistics*, 67(3):393–419.

Im, K. S., Pesaran, M. H., and Shin, Y. (2003). Testing for unit roots in heterogeneous panels. *Journal of econometrics*, 115(1):53–74.

Jayaratne, J. and Strahan, P. E. (1998). Entry restrictions, industry evolution, and dynamic efficiency: Evidence from commercial banking 1. *The Journal of Law and Economics*, 41(1):239–274.

Jobst, A. and Lin, H. (2016). *Negative Interest Rate Policy (NIRP): Implications for Monetary Transmission and Bank Profitability in the Euro Area.* International Monetary Fund.

Johansen, S. (2002). A small sample correction for the test of cointegrating rank in the vector autoregressive model. *Econometrica*, 70(5):1929–1961.

Kemp, M. C. and Wan, H. Y. (1974). *Hysteresis of Long-Run Equilibrium from Realistic Ajustment Costs*, volume volume. publisher, edition edition.

Krasnosel'skii, M. A. and Pokrovskii, A. V. (1989). *Systems with hysteresis.* Berlin: Springer-Verlag.

Labonte, M. (2015). Systemically important or "too big to fail".

Lee, H.-Y., Wu, J.-L., and Lin, C.-H. (2010). Hysteresis in east asian unemployment. *Applied Economics*, 42(7):887–898.

Lee, J. and Strazicich, M. C. (2003). Minimum lagrange multiplier unit root test with two structural breaks. *The Review of Economics and Statistics*, 85(4):1082–1089.

Lee, J., Strazicich, M. C., et al. (2004). Minimum lm unit root test with one structural break. *Manuscript, Department of Economics, Appalachian State University*, pages 1–16.

Lee, J.-D., Lee, C.-C., and Chang, C.-P. (2009). Hysteresis in unemployment revisited: evidence from panel lm unit root tests with heterogeneous structural breaks. *Bulletin of Economic Research*, 61(4):325–334.

Levin, A., Lin, C.-F., and Chu, C.-S. J. (2002). Unit root tests in panel data: asymptotic and finite-sample properties. *Journal of econometrics*, 108(1):1–24.

Levine, R. (1997). Financial development and economic growth: Views and agenda. 35:688–726.

Levine, R. (2005). *Finance and Growth: Theory and Evidence*, volume volume. Elsevier, 1st edition.

Lovell, M. C. (2008). A simple proof of the fwl theorem. *The Journal of Economic Education*, 39(1):88–91.

Maddala, G. S. and Wu, S. (1999). A comparative study of unit root tests with panel data and a new simple test. *Oxford Bulletin of Economics and statistics*, 61(S1):631–652.

Marshall, A. (1890). 1920. principles of economics. *London: MacMillan.*

Matutes, C. and Vives, X. (2000). Imperfect competition, risk taking, and regulation in banking. *European Economic Review*, 44(1):1–34.

Mayergoyz, I. (1991). Mathematical models of hysteresis.

Merton, R. C. (1993). Operation and regulation in financial intermediation: a functional perspective.

Mishkin, F. S. (1999). Financial consolidation: Dangers and opportunities. *Journal of Banking & Finance*, 23(2):675–691.

Nicoló, G. D., Bartholomew, P., Zaman, J., and Zephirin, M. (2004). Bank consolidation, internationalization, and conglomeration: Trends and implications for financial risk. *Financial markets, institutions & instruments*, 13(4):173–217.

Nunes, L. C. et al. (2004). Lm-type tests for a unit root allowing for a break in trend. *Econometric Society.*

O'hara, M. and Shaw, W. (1990). Deposit insurance and wealth effects: the value of being "too big to fail". *The Journal of Finance*, 45(5):1587–1600.

O'Shaughnessy, T. (2000). Hysteresis in an open economy model. *Scottish Journal of Political Economy*, 47(2):156–182.

Park, J. Y. (1992). Canonical cointegrating regressions. *Econometrica: Journal of the Econometric Society*, pages 119–143.

Perron, P. (1989). The great crash, the oil price shock, and the unit root hypothesis. *Econometrica: Journal of the Econometric Society*, pages 1361–1401.

Perron, P. et al. (2006). Dealing with structural breaks. *Palgrave handbook of econometrics*, 1(2):278–352.

Pesaran, M. H., Shin, Y., and Smith, R. J. (2001). Bounds testing approaches to the analysis of level relationships. *Journal of applied econometrics*, 16(3):289–326.

Phelps, E. S. (1972). *Inflation Policy and Unemployment Theory*. Macmillan.

Phillips, P. C. and Hansen, B. E. (1990). Statistical inference in instrumental variables regression with i (1) processes. *The Review of Economic Studies*, 57(1):99–125.

Piscitelli, L., Cross, R., Grinfeld, M., and Lamba, H. (2000). A test for strong hysteresis. *Computational Economics*, 15(1):59–78.

Piscitelli, L., Grinfeld, M., Lamba, H., and Cross, R. (1999). On entry and exit in response to aggregate shocks. *Applied Economics Letters*, 6(9):569–572.

Preisach, F. (1935). Über die magnetische nachwirkung. *Zeitschrift für Physik A Hadrons and Nuclei*, 94(5):277–302.

Pulley, L. B. and Humphrey, D. B. (1993). The role of fixed costs and cost complementarities in determining scope economies and the cost of narrow banking proposals. *Journal of Business*, pages 437–462.

Romer, D. (2012). *Advanced macroeconomics*. McGraw-Hill/Irwin,.

Romero-Avila, D. and Usabiaga, C. (2007). Unit root tests, persistence, and the unemployment rate of the us states. *Southern Economic Journal*, pages 698–716.

Saikkonen, P. (1991). Asymptotically efficient estimation of cointegration regressions. *Econometric theory*, 7(1):1–21.

Samuelson, P. A. (1965). Some notions on causality and teleology in economics. 3:428–427.

Schwert, G. W. (1989). Tests for unit roots: A monte carlo investigation. *Journal of Business & Economic Statistics*, 20(1):5–17.

Shin, H. S. (2009). Reflections on northern rock: the bank run that heralded the global financial crisis. *The Journal of Economic Perspectives*, 23(1):101–119.

Shin, Y., Yu, B., and Greenwood-Nimmo, M. (2011). Modelling asymmetric cointegration and dynamic multipliers in a nonlinear ardl framework.

Stern, G. H. and Feldman, R. J. (2004). *Too big to fail: The hazards of bank bailouts.* Brookings Institution Press.

Stiglitz, J. E. . A. W. (1981). Credit rationing in markets with imperfect information. 71(3):393–410.

Stiroh, K. J. and Strahan, P. E. (2003). Competitive dynamics of deregulation: Evidence from us banking. *Journal of Money, Credit, and Banking*, 35(5):801–828.

Stock, J. H. and Watson, M. W. (1993). A simple estimator of cointegrating vectors in higher order integrated systems. *Econometrica: Journal of the Econometric Society*, pages 783–820.

Stock, J. H. and Watson, M. W. (2007). Introduction to econometrics.

Taylor, M. P. and Sarno, L. (1998). The behavior of real exchange rates during the post-bretton woods period. *Journal of international Economics*, 46(2):281–312.

Teräsvirta, T. (1994). Specification, estimation, and evaluation of smooth transition autoregressive models. *Journal of the american Statistical association*, 89(425):208–218.

Verbeek, M. (2008). *A guide to modern econometrics.* John Wiley & Sons.

Verheyen, F. (2013). Exchange rate nonlinearities in emu exports to the us. *Economic Modelling*, 32:66–76.

Vogelsang, T. J. and Perron, P. (1998). Additional tests for a unit root allowing for a break in the trend function at an unknown time. *International Economic Review*, pages 1073–1100.

Walter, J. R. (2005). Depression-era bank failures: the great contagion or the great shakeout?

Wooldridge, J. M. (2015). *Introductory econometrics: A modern approach*. Nelson Education.

Yeyati, E. L. and Micco, A. (2007). Concentration and foreign penetration in latin american banking sectors: Impact on competition and risk. *Journal of Banking & Finance*, 31(6):1633–1647.

Zivot, E. and Andrews, D. W. K. (1992). Further evidence on the great crash, the oil-price shock, and the unit-root hypothesis. *Journal of business & economic statistics*, 20(1):25–44.