An Analysis of the Federal Reserve's Policies Over the Last Century

by

David Pavón-Prado

A dissertation submitted by in partial fulfillment of the requirements for the degree of Doctor of Philosophy in

Economic History

Universidad Carlos III de Madrid

Advisor:

Carlos Santiago-Caballero

Tutor:

Carlos Santiago-Caballero

July, 2019

-- Acknowledgments --

As the strong defender of meritocracy that I am, this dissertation could not see the light without first naming those who deserve credit for the work here developed. The tremendous amount of effort and will deployed on this project was only possible thanks to the intellectual, economic and emotional support that I have received over the last four years.

I must begin with my supervisor Carlos Santiago Caballero, who has been involved in this project since day one, despite the initial uncertainty and changes in the topic, providing guidance and support for the first doubts and fears, smoothing every problem that needed his help and what deserves my greatest admiration, its integrity and dedication even in his worst days. During my stay at the London School of Economics, I was also lucky to receive Albrecht Ritschl's guidance for the beginning of this thesis, who introduced me into the world of the methodology used in this dissertation (source of despair in several occasions), which later led me to what I consider fruitful developments. Learning those tools would have not been possible without the time and leeway that he and Carlos allowed me. I am also thankful to Hernan D. Seoane, who joined this project after my stay in London, and whose suggestions at that time, supposed a turning point in the way I approached this dissertation.

The Marie Curie grant that I received from the Union Europe under the MACROHIST project was responsible for my time at the London School of Economics and it supposed a substantial change in my life. It was a significant intellectual boost, product of the opportunities it offered me and the people I met. Among those people are those belonging to the MACROHIST project, with special mention to my office mates, Alexandra De Pleijt and João Sousa Da Cunha, the latter also being my flatmate and unofficial supervisor at the beginning, to whom I really appreciate its guidance, recommendations, talks and debates. For the rest of the people of the LSE family, I really appreciate the environment they created, as they made me felt comfortable there despite being far from home, specially Leonard Kukic, Enrique Jorge Sotelo, Mattia Bertazinni and Andrea Papadia with whom I spent most of the few times that I dedicated to leisure. I received a great surprise during my last months there, because few times life gives you the opportunity of knowing people like Daniel Gallardo, from whom intellectual and spiritual enrichment was and is unavoidable. Besides, we shared many funny moments in a short period of time, what again, made easier my stay in London. Not much later, I had also the fortune of coming across Maria Jose Fuentes and Germán Forero, with whom I have also laughed a lot, including the time they invited me to Barcelona, where they treated me extraordinarily well.

All these experiences could have not been possible without the other soldier that has been with me in this battle before it even started, Miguel Ángel Ortiz. We have shared worries, ambitions, ideas, conferences, travels, parties, small talks, deep talks… and I could continue with an endless list. Besides, I will be eternally in debt to him, as he was the one sending me the offer for the MACROHIST project.

Leaving the professional spectrum, I am also in debt to Maria Rodriguez for helping me in the complex process of understanding myself when it was most needed, as it led me to a wonderful period of creativity. I also appreciate the time that I have spent with my friends in my hometown, Madrid or wherever I have been able to meet them during my very needed breaks, but a special mention is needed for Manuel Ruiz and Borja Rodriguez because of their comprehension about my lack of time during these years, and for their support and shared ambitions in parallels projects.

Sometimes, words are not enough to express our feelings or ideas, and now I enter a territory where words will not make justice to the gratitude I want to express. The help and support I have received from those who personally I consider my chosen family, María Pacios and Aníbal Vera, cannot be even accounted. With the former I have shared most of my deepest thoughts, something that I hardly can do with other people. The latter, I consider him as my second brother because we have been together now for almost ten years, sharing basically our lives, and fighting the world together in search of our ambitions. Their loyalty and integrity are rarely found in this world. Last and most important, special thanks to those who most credit deserves, my family. To my brother Miguel Ángel Pavón Prado, for being there always that I have needed something, for being an example of kindness and good faith, and for bringing to these world my two precious nephews, who have taught me that to actually learn, I must see life through a child eyes again to question my most rooted ideas and thoughts. To my parents, Soledad Prado Hidalgo and Eleuterio Pavón González, who have been, are and always will be an example of hard work, resilience, strength and sacrifice. As normally I find difficulties in expressing my gratitude, I will take advantage of this opportunity to do it, and as I want them to remember this forever, the next words will be in Spanish.

Papá y Mamá, este trabajo, todo lo que soy y llegaré a ser, será gracias a vosotros, porque no solo me habéis dado mi vida, si no que también me habéis dado la vuestra. Vuestro tiempo, vuestra energía, los recursos que tanto os ha costado conseguir, todo lo habéis sacrificado para que yo haya podido realizar cosas como esta. El apoyo y la fe casi ciega en todo lo que he decidido emprender han construido puentes para los abismos que no podría haber cruzado

solo. Gracias no es suficiente para compensar todo esto, así es que mi nombre irá siempre acompañado del vuestro allá donde mi trabajo de fruto.

--Published and submitted content--

Chapters 1 and 2 of this thesis were submitted as a sole article to the Explorations in Economic History journal. The article is currently under review. It can also be found as a working paper in https://e-archivo.uc3m.es/handle/10016/28342.

-- Content --

--List of Figures --

-- List of Tables --

-- Introduction --

The Federal Reserve' creation dates from the year 1914, after the Congress approved the Federal Reserve Act on December 23, 1913. It was designed with the main purpose of solving the two main problems of the time. The first one relates to the significant fluctuations in market interest rates caused by the seasonal demand for currency, which mostly depended on the financing of crop harvest. The second reason was the desire to reduce the severity and frequency of banking crisis, as the previous thirty years had witnessed five of them. Thus, the Federal Reserve was created as an entity working as lender of last resort, for pooling reserves and lending during panics, and discounting bills to finance crop movements. Those targets are far from the maximum employment, stable prices, and moderate long-term interest rates that have been determining monetary policy for the last decades. That change is synonym of having overcome those past obstacles and of the evolution in the field of monetary policy. However, if the economic science is already relatively young in comparison to other sciences, monetary policy is even a younger field and much must be still learned. While it has evolved to prevent banking panics and smooth seasonal demand for money, new challenges have emerged. Banking, financial and debt crises, bubbles, market crashes, recessions, high inflation levels or even deflation, have been the recurrent episodes during the last decades. The lack of ability to prevent these phenomena proves that our knowledge about monetary policy is still scarce, and even, some of the already established and accepted knowledge may be erroneous. Thereby, this study was born with the intention of re-examining the Federal Reserve's history, what implies its policies and their consequences on the American economy, relearning from the lessons provided, and subsequently, to discover missing or misunderstood mechanisms operating between monetary policy and the real economy. The goal is to place us in the best possible position to foresee and prevent the harmful episodes named above by understanding those mechanisms whereby monetary policy operates.

It has to be recognized that despite the incapacity to avoid such episodes, the study of monetary policy is an ongoing process in a very early state, that nevertheless, has been able to improve policy management, leading the economy through steadier paths as supposedly exemplified in the denominated Great Moderation (McConnell and Perez-Quiros (2000) and Stock and Watson (2003)) starting in the mid-1980s and consolidated from 1990 to 2007, unlike the previous period of the Great Inflation (from the mid-1960s to the mid-1980s). The strong contrast between these periods triggered a debate in which researchers wondered whether that regime change was due to an alteration in the policies undertaken by the Fed, or factors external to monetary policy. This was framed into the “good luck, good policy” debate (Clarida, Gali and Gertler (2000) or Canova and Gambetti (2009) among many others). However, this debate was far from being a novelty because researchers had already experienced different regimes in the past. The Great Depression and the previous and posterior periods were an example (Wicker (1965), Brunner and Meltzer (1968), Hetzel (1985), Bordo and Schwartz (1999)). The particular difference between past and contemporaneous analysis is found on how those regime changes were evaluated. While more advanced methods have been used for contemporaneous studies, those examining past periods, resorted mostly, to narrative and qualitative techniques. This contrast exemplifies how monetary policy has evolved and improved the accuracy and credibility of the conclusions reached in those analysis. Surprisingly, both approaches share the lack of consensus regarding when the Federal Reserve was responsible for those periods of success or failure. Thus, the fact that the economy is still undergoing the episodes commented above, along with the researchers’ inability to unveil and agree on the factors and mechanisms operating behind the good or bad American economy’s performance, is a clear signal that our knowledge about monetary policy is defective at some stage.

With the purpose of discovering missing knowledge or correcting any possible misunderstanding, the starting point of this defective knowledge hypothesis was to assume that although the low volatilities in inflation and output, as well as the stable and low inflation of the Great Moderation is the clearest and largest example of a period when the economy has been in state relatively desired by the central bank, obviously it was not unique. Given that the good or bad performance of the economy, in higher or lower degree, even for shorter periods, had already been experienced in the past, those periods should present common patterns, despite the possible existence of factors distorting them. To concentrate efforts on unveiling those hidden patterns, I considered that a long and wide vision was necessary for that purpose and hence, as many periods as possible should be compared together, regardless of its ordinary or extraordinary characteristics. Actually, the outstanding characteristic of a specific period should not be appealing for the researcher, but the opposite. The ordinary periods, the good and the bad ones, must be examined first, so that we are able to explain the particularities later. Analyzing separately the Great Depression, the period post-Great Depression, the 1950s, the 1960s, the Great Inflation, the Great Moderation or the

Great Recession will prevent us from observing the entire picture and extract those common patterns that may be the key explanation for the mechanisms operating between the monetary policy and the economy.

Thus, Chapter 1 reviews the narrative of the Federal Reserve's history from the end of the World War I to 2007, just before the Great Recession, but skipping the 1940s, mainly because of the World War II and to make it consistent with the periods analysed in Chapter 2. The narrative covers the main events happening within the Fed regarding the learning process of its members about monetary policy and their reactions to economic and political episodes. After reviewing how the Fed used its instruments in the process of achieving their targets, not only some insight is acquired regarding the possible factors conditioning the American economy's performance, but also that review leads to questioning whether researchers have been measuring monetary policy stance correctly. While it can be a recklessness to direct the attention to what should be considered as settled bases, one must be aware that if the measuring of monetary policy were erroneous, the error would be transmitted to the results. Consequently, the comprehension of the mechanisms whereby monetary policy is able to reach the economy would be poorly recognized and would disable the possibility and necessity of unveiling the common patterns that can help us in foreseeing and preventing undesired episodes as those previously commented, as policy advice will be based on false premises. Subsequently, the lessons learned will be translated into bad policies.

This questioning appeared once I observed the common procedure to measure monetary policy stance used in the literature, in contrast with the knowledge acquired in the review of the Federal Reserve's history. Thus, Chapter 2 explains how that common procedure is erroneous and a new approach is proposed. The standard procedure used intermediate targets such as short-term rates or reserves measures, although mainly the federal funds rate (Christiano, Eichenbaum and Evans (1998), Boivin and Giannone (2006), Primicery (2005)…) as Fed's instruments. However, this approach entails a conceptual and a measuring problem. Conceptually, short-term rates or reserves measures are not Fed's instruments but intermediate target. The actual Fed's instruments are, mainly, the discount rate and open market operations. This leads us to the second problem. The use of intermediate targets to measure monetary policy stance introduce bias into the models. To understand this statement, the money market must be conceptualized as two submarkets. In the first submarket the Fed is the supply side and the banking sector the demand side. In the second submarket, the banking sector switches to the supply side, the other agents of the economy being the demand side. The Federal Reserve through the discount rate and open market operations sets the price and the amount of money in this first submarket, depending on its intermediate target. Taking the federal funds rate as example (although the same argument applies to reserves measures or other intermediate targets), when the Federal Reserve uses its instruments, it has an impact on the federal funds rate according to the level targeted. However, the federal funds rate also depends on how banks are pricing their reserves. That pricing is conditioned on the demand for loans, forecasts regarding growth, inflation, risks, Fed's policies and so on. When the Fed is targeting a specific federal funds rate, to maintain it equal at two different periods, it will have to use its instruments differently to achieve such target, because, as already said, it depends also on banks pricing decisions. Thus, the different amount and price at which reserves are provided for the different periods, despite keeping the federal funds rate at the same rate, will have a different impact on the real economy. The reason is that banks set their loans rate, mainly conditioned on the cost of money to them, namely, the cost of the reserves. As the Fed provided them at different prices, but maintaining equal the federal funds rate, the impact on the loans rate will be different, and that different impact will be converted into a higher or lower demand for money, output growth and inflation.

Thus, using the federal funds rate or other intermediate target to measure monetary policy stance, what should suppose the measurement of only the supply side of the first submarket, also captures the demand side of the first submarket. Therefore, the results will not only measure monetary policy stance, but also the banking sector's decisions. This argument was born from the fact that during the last century, when the spread between the short-term rate of reference and the discount rate was positive, borrowing at the discount window increased, signaling that banks were taking advantage of the arbitrage opportunities by borrowing cheaper reserves at the discount widow. The relatively cheaper cost of those reserves could have enabled banks to increase the loans rates less than the Fed raised short-term rates. That is, when, wittingly or unwittingly, the Fed raised short-term rates, its impact on the loans rate was below the ratio 1:1 (as banks obtained cheaper reserves at the discount window), triggering an insufficient restrain of credit, and consequently, higher inflation, as the results confirm. In relation to this, Chapter 2 also unveils that the price puzzle, whereby increases in interest rates are accompanied by higher inflation (Barth and Ramey (2001), Boivin and Giannoni (2003), Bernanke, Boivin and Eliasz (2005), Uhlig (2005) and Hanson (2006), among many others), is non-existent. That puzzle is a consequence of bad policies that allowed positive spreads between the federal funds rate and the discount rate. Thus, the lesson learned in this chapter is that monetary policy stance has been measured erroneously. To solve that, a new procedure is developed using actual Fed's instruments and the spread between short-term rates and the discount rate. According to the results, this chapter shows that monetary policy was almost useless during the interwar period, that the Fed used its instruments differently after the mid-1960s and that the banking sector was likely to have changed its behavior around 1990. It also hints that monetary policy seems to be transmitted through prices and not through quantities.

Encouraged by the hypothesis that banks set their loans rate depending on the cost at which they obtain reserves and the possible banking sector's behavior change around 1990, Chapter 3 focuses on the study of that potential channel whereby monetary policy can be transmitted to the real economy, and where the banking sector as transmitter of monetary policy, can modify Fed's policies. While some authors debated whether monetary policies were transmitted through the "money channel" (Bernanke and Blinder (1992), Kashyap, Stein and Wilcox (1993)) or the "lending channel" (Ramey (1993), Oliner and Rudebusch (1995, 1996), Romer and Romer (1990)), the premise for both channels is based on the fact that banks need deposits to lend. Thus, when the Fed removes those reserves or avoids their rise, banks either increase the loans rate (money channel) or reduce the amount of lending (lending channel). However, authors such as Moore (1983), Bindseil (2004) or Jakab and Kumhof (2015) presented strong arguments against that premise and claimed that banks are not in need of deposits to lend as the deposit multiplier theory suggests. The reality is that banks lend as long as it is profitable and their solvency is not in danger, and just later, the required reserves are obtained. Therefore, causality runs from loans to reserves and not the other way around. Also, they argue that the demand for loans (money) is endogenous. "The ability of central banks to control the rate of growth of monetary aggregates therefore hinges on their ability to control the rate of growth of bank lending, rather than the monetary base" (Moore (1983), p. 544). Further: "The assumption…is that banks set the prime rate and then attempt to meet the loan demand that results" (p.545). Thus, once this misunderstanding is overcome and taking into account Chapter 3, I elaborate what I have named the "reserves-cost" theory. It claims that given the Fed sets short-term rates, wittingly or unwittingly, by providing reserves and using the discount rate, the only way whereby it can have an impact on the real economy is by influencing the reserves cost directly, and indirectly, with the impact of this cost on the loans rate. Therefore, while the spread between the federal funds rate and the discount rate is closed, it is more likely that the Fed has total control of monetary policy. Otherwise, the banking sector will modify those policies and the impact on the real economy will be different. The reason is that those positive spreads will cause that the federal funds rate has an impact ratio below 1:1 on the loans rate, which will depend on banks decisions, conditioned on the cost of their reserves. As banks can obtain nonborrowed and borrowed reserves from different sources, the cheapest one will determine the evolution of the loans rate, and consequently, the demand for credit.

Once this theory is completely understood, I decide to drop the Fed's instruments from the model, unlike in Chapter 2, for the reason implicit in the new theory. That is, to measure monetary policy stance, as policies have its impact on the economy through their influence on the reserves cost, it is sufficient to capture the price at which reserves are available in the money market. Thus, the spread used in Chapter 2, namely, the difference between the federal funds rate and the discount rate, is also used in this case as measure of monetary policy stance. In addition, a new variable is created to capture how the banking sector responds to those policies. This variable measures the spread between the prime loans rate and the federal funds rate. When this spread is smaller, it signals that banks are obtaining cheaper reserves at the discount window than those available at the short-term rates in the money market. Consequently, the loans rate evolves differently in relation to the federal funds rate, what implies the ratification of the assumption taken in Chapter 2, namely, the impact of the federal funds rate on the loans rate is below the ratio 1:1. This phenomenon produces an insufficient restrain in lending, causing higher inflation. The results confirm the new "reserves-cost" theory and therefore, a new mechanism discovered. It implies that the banking sector is able to modify Fed's policies and that when positive spreads are allowed, the Federal Reserve lose, partially or totally, its power to drive the path of the economy. The understanding of this mechanism teaches us that to avoid episodes as the Great Inflation, besides setting a short-term rate target high enough to restrain the demand for credit, the spread between the federal funds rate and the discount rate must be closed. The "reserves-cost" theory may be also the explanation for Great Moderation (and the Great Inflation). The reason is that this episode coincides with the period when banks set a constant spread between the loans rate and the federal funds rate, being the largest of the period under analysis, unlike during the Great Inflation when the spread was smaller (or even negative) and more volatile. Thus, the stability or instability of this spread seems to be transmitted to the economy as the volatilities seen for the spread, coincide with those seen for output and inflation for the Great Inflation and the Great Moderation. At the same time, the amplitude of the spread seems responsible for the inflation levels.

However, the "reserves-cost" theory, as exposed in Chapter 3, leaves one case without explanation. While for most of the period under analysis in that chapter excess reserves were scarce and banks decisions on the loans rate were based on the interest rate at each period, after the mid-1980s banks began to accumulate excess reserves. For this scenario, from the "reserves-cost" theory can be deduced that the cost of past reserves may also have an impact on the current loans rate. If it were the case, it would mean that the theory was incomplete, because it was able to explain how banks react to Fed's policies only when current interest rates are taken into account. Consequently, Chapter 4 undertakes the completion of the "reserves-cost" theory. Further, the renewed theory is proposed as the explanation for the accumulation of reserves. While the literature has suggested low interest rates (Frost (1991), Bindseil, Camba-Mendez, Hirsch and Weller (2006), Dwyer (2010)), risk, uncertainty, (Goodhart (2010), Ashcraft, McAndrews and Skeie (2011), Chang, Contessi and Francis (2014)) or low demand for loans (Bindseil (2004) and Todd (2013)) as the most important factors determining the accumulation of excess reserves, I argue that the accumulation or use of excess reserves is explained also by the cost at which reserves are obtained.

When banks accumulate reserves, if the cost at which reserves can be obtained at a particular period is below the cost of the reserves held, as profit maximization agents, banks will prefer to obtain reserves at current rates to maximize profits by obtaining the largest margin between the loans rate and the cost paid for the reserves required for loans. On the contrary, if reserves costs are above the price at which they can be found at that moment, banks will prefer to use their reserves. For the former scenario, banks will decide to accumulate excess reserves, while for the later, they will use those already held. Thus, the "reserves-cost" theory is revamped as follow: The only way whereby the Federal Reserve can have an impact on the real economy is through the manipulation of the cost at which banks obtain reserves. This cost will steer the loans rate, which is the actual rate that the real economy endures. If the level of excess reserves is scarce, just modifying the federal funds rate (or other the short-term rate of reference) and the discount rate will have a direct impact on the cost of reserves held by banks, as long as the spread between these rates is closed. In that case, the movements of those rates are likely to be reflected one to one into the loans rate. If the excess reserves levels are significantly above the levels of required reserves and precautionary factors, the impact of interest rates aforementioned on the loans rate will be proportionally diluted to the quantity hoarded, and more aggressive policies and longer time will be necessary to drive the real economy through the desired path. The reason is that banks' decisions about the loans rate will make it to evolve differently in relation to short-term rates given the reserves cost. Consequently, there will be three hands behind the steering wheel of monetary policy. That is, past Fed's policies, banks' decisions and present Fed's policies.

To evaluate this new theory a new variable is created. This variable measures the average cost of the reserves held by the banking sector at every period. It is used along with the short-term rate of the money market, so that the spread between both variables indicates whether banks hold reserves with a cost above or below short-term rates. When the spread is positive, it will indicate that banks are holding reserves more expensive that the ones available in the market. Hence, they will obtain new reserves at current rates, accumulating even more, as they can maximize profits using the new reserves and lending at the loans rate of that period. It will also imply that banks will increase the loans rate in relation to short-term rates, in order to decrease the probabilities of incurring in loses in case they had to use their expensive reserves. On the contrary, when the spread is negative, it signals that banks hold cheaper reserves than current rates. Therefore, accumulated reserves will be used to maximize profits and the loans rates will be raised less in relation short-term rate increases, as there is no risk of using reserves with a cost above the loans rate. In addition, a relatively lower loans rate will increase loans demand and therefore, profits. The results confirms that when the cost of excess reserves is above the current short-term rate, banks accumulate more reserves and increase the loans rate in relation to the short-term rate. Thereby, this chapter is the final response to the ambitions that originated this study, as by finding an error in how to measuring monetary policy stance, the mechanism whereby Federal Reserve's policies have an impact on the real economy was unveiled, adding to the needed knowledge to manage monetary policy. The chapter ends by describing the different scenarios that the Fed has to face depending on the amount and price of reserves holds by the banking sector, the spreads between the reserves costs and short-term rate, and the impact ratio of that short-term rate on the loans rate. The possession of that knowledge supposes the first step to prevent the harmful episodes commented at the beginning.

-- Chapter 1 --

A review of the Federal Reserve's history

The first step to discover or correct the potential mechanisms whereby monetary policy operates is to review the Federal Reserve's history. This review is necessary to gain insight regarding the learning process of those responsible for monetary policy, Fed's responses to economic and political events, and the general beliefs or conclusions regarding how those responses have determined the path of the American economy. Several authors have invested their time and effort in providing that insight. Fforde (1954), Friedman and Schwartz (1963), Wheelock (1991), Toma (1997), Moore (1990) and Mayer (1999) are some of them. However, according to the level of completeness (from a personal opinion), I have decided to expose mostly those facts narrated in Meltzer (2003, 2009) and Hetzel (2008), but without positioning on whether the facts narrated are correct or not, from an ideological perspective.

1.1 1919-1939

By 1919, governor Benjamin Strong, chairman of the Fed of New York, realised that the spread between the short-term rates and the discount rate would avoid a decline in inflation, as it was profitable for banks to borrow at lower rates and lend. In 1921 began what today is known as the federal funds market. Banks with surplus reserves sold reserves to banks with deficient reserves. However, its relevance was far from what that market is nowadays. By 1922, some members of the Fed noted that reserve banks could increase "momentum" purchasing in the open market and at the same time reducing the discount rate. Strong claimed that buying in the open market was equivalent to a member bank borrowing. Thus, open market operations experiments began in 1922. The discount rate during those years was at discretion of each Reserve Bank and its use was intended to be a penalty rate. Thus, it used to be above short-term rates, at least until 1921-1922, because later, market rates on commercial paper started to be above the discount rate. Apart from that, it was supposed to be the tool to follow the gold standard rules, namely, increasing it in periods of surplus and decreasing it when there was deficit. Also in 1922, Strong noted and commented what would determine monetary policy for the next years. His observation was that when banks were in debt, they used their surplus reserves to reduce borrowing. On the contrary, when they were out of debt, they reduced rates and put their surplus to work. "The reduction in our rate had no influence in the market. It was the competition to lend money that did it" (Meltzer 2003, p. 126). Thus, changes in the discount rate were supposed to be ineffective, but by selling in the open market, the Reserve Banks could reduce bank reserves and force them to borrow, thereby restoring the effectiveness of the discount policy. In 1964 Burgess reported:

> "First, as fast as the Reserve Banks bought government securities in the market, member banks paid off more of their borrowing; and, as a result, earning asset and earnings of the Reserve Bank remained unchanged…" (Meltzer 2003 p.153).

Given Fed's concern about speculative credit during the 1920s, interest rates were raised to avoid the growth of stock exchange lending, thereby attracting more gold. Consequently, to maintain price stability, the Fed sterilized the gold inflows and the monetary expansions triggered by those gold inflows were reduced or totally cancelled. One characteristic of these years that led Federal Reserve Banks to follow the wrong policy is explained in what was the Riefler-Burgess doctrine. It was believed that banks were reluctant to borrow and they only did it when their reserves were deficient. This triggered that during many years the monetary base and borrowing moved procyclically. That is, the Fed believed that increased aggregate borrowing signalled a restrictive policy even if the monetary base and money stock accelerated. Following this reasoning, it increased the purchases in the open market and decreased the discount rate for those periods, and did just the opposite when borrowing decreased.

At the April 1925 meeting for the Governors Conference, the concern that would continue for the rest of the decade was expressed; credit to securities brokers and dealers. They feared speculative borrowing. Consequently, some Reserve Banks increased the discount rate later, as the open market account was thought to be too small to have a significant effect on reducing bank reserves. The Fed carried out open market purchases in May 1927 because of the recession, and continued in July. However, they were offset by a decline in borrowing and in the reported gold stock. Despite the increase in inflation in 1926, Strong's view of supporting the pound with lower rates dominated. By the end of 1927 Strong had complete authorization to offset gold flows without limit. From 1927 to 1929 the Fed sterilized gold inflows, preventing monetary expansions and triggering, or at least adding to the deflation witnessed for those years. In addition, the members of the Fed were misled by the lower levels of discounts and borrowing, as commented above, believing that policy was already expansive. Apart from that, during this period the discount rate was higher in real terms given the levels of deflation, but they did not distinguish between real and nominal terms. Moreover, the spread between the stock exchange call loans rate and the discount rate became significantly large after the beginning of 1928. Given the increase in credit to brokers, the Fed thought that its policies were expansive and therefore, rates were not reduced and open market purchases decreased. During 1928, the Federal Reserve offset part of the net gold outflow, but it was insufficient and the monetary base declined. Thus, when New York stopped sterilizing gold losses, discounting increased significantly but the monetary base continued falling. Discounting and bank credit were in the highest levels of the three last years.

"…[O]n December 31, 1928, the Board adopted a resolution that blamed the spread between discount rates and rates for stock exchange loans for the temptation to borrow from the Fed and lend to help buy or carry securities" (Meltzer 2003, p. 237). The Fed spent 1929 trying to reduce bank lending to brokers, as it was thought to be speculative credit, but most of the lending came from corporations and other nonbanks. Hence, the Fed was not successful in reducing call loan rates. Previous to the crash on October 23, call loan rates had already started to decrease and by that date, they were around 6%. However, the Fed refused to reduce the discount rate. At the end of November 1929 the Fed noted that there was being a liquidation of credit against securities, what could suppose a serious threat to business stability, having already in mind that there were indications of a business recession. As the short-term rates had fallen and were expected to fall further, and discounting increased, the Fed approved limited purchases in the open market. Later, industrial production, stock of money, monetary base and borrowing fell. Fed governors thought that the open market purchases had failed to revive the economy. Actually, the Fed had failed to offset the decline in borrowing by purchasing insufficiently. However, the members of the Fed thought that the purchases had already permitted banks to repay borrowing, as borrowing levels had declined. Bank failures along with the increasing demand of the public for currency, contributed to contracting the money supply.

By the end of 1930, Fed's members were deliberating how much to sell in the open market as they considered that policy was loose, since banks were keeping twice the level of excess reserves of the previous year. This offset the inflow of gold at the beginning of the year. Later, as those inflows did not decrease interest rates, the governors decided to purchase in the open market from April to June of 1931. Nonetheless, the rise in currency holding and excess reserve counteracted the effect. In 1932, gold outflow started again, along with a higher demand for currency. The Glass-Steagall Act of April 1932 and threats of additional legislation led the Fed to purchase in the open market. Signs of improvements were soon recognized but as the purchases stopped, the improvements quickly disappeared and gold outflow continued. Accordingly, it was thought that the program has failed. As already occurred between 1927-1929, since the short-term rates were historically low, the members of the Fed thought that policy was easy and no further purchases were done. However, because of the deflation, real rates were higher. The System kept an inactive role during the next months. From time to time, as in January 1933, the System sold in the open market to keep excess reserves close to \$500 million. In the beginning of 1933 short-term rates increased, but the Fed failed by discounting at higher rates than the market, when it should have set the discount rate below the short-term rates.

The last part of the interwar period, from 1933 to 1941 is characterized by the inaction of the Fed, as the Treasury took most of the responsibility for monetary policy. The Fed's open market account and the discount rate hardly changed during this period, and the variations in the monetary base were due to changes in gold stock and the devaluation of the dollar in 1934. Marriner S. Eccles, who became governor of the Board in 1934 believed that the Fed should keep market rates low, in order to facilitate private spending and government finance. Also, he thought that the growing volume of reserves at member banks could mean a threat of future inflation. Thus, reserve requirement ratios became the main instrument of the Fed during these years. In October and November of 1933 the Fed made the last purchases in the open market and it would not purchase again until April 1937.

In 1934, Roosevelt bought gold and silver to raise prices. The base and money stock increased. Also that year, he carried out devaluation up to 60% of the gold. As the president had acquired the gold held by the Federal Reserve banks previously, this devaluation supposed a \$2 billion profit, which was used to set up the Exchange Stabilization Fund (ESF). It aimed to retire national banknotes and finance industrial loans. This devaluation supposed an increase in prices and flow of gold into the country. The ESF was also used to buy bonds, in order to keep rates low and finance the deficit. In August 1935, as excess reserves rose and there was fear of future inflation, the Fed decided to increase the reserve requirements ratio. As reserves had increased, discounting decreased and hence, the discount rate could exert little influence, or at least the members of the Fed thought so. This first increase in the reserve requirement ratio had little effect because of the gold inflow. Given the fear of speculative gold inflows and increase in the monetary base, the Treasury sterilized gold inflows between December 1936 and July 1937. Just in this period, deflation appeared again. There were two further increases in the reserves requirement ratio in January and May of 1936. Later, and together with the sterilization, there was an increase in bond yields, what supposed the restart of the Fed's purchases in the open market to lower short-term rates and indirectly, long-term rates. The increases in the reserves requirement ratio along with contractive fiscal policies supposedly made the money stock to fall, causing a recession by May of 1937. Thus, in September 1937 more open market purchases were undertaken and the Treasury desterilized part of the gold inflows. Consequently, inflation levels increased for the end of the year. Around that time, the Fed proposed to manage open market operations (OMO, henceforth) in response to the level of excess reserves instead of the amount of borrowing. Thus, by 1938 the Fed's purchases in the open market were in small quantities, also because rates were low and it was believed that the monetary policy was easy. However, deflation came back again for the end of the year and the Fed had to reduce reserve requirements ratios, while the Treasury continued desterilizing the gold sterilized for the previous years.

1.2 1950-1957

The main characteristics of monetary policy from 1950 to 1960 were that the Fed still considered low interest rates as loose policy, even though M1 be decreasing; the gain in relevance in the use of OMO in relation to discounting, and that the main target was free reserves, using bank borrowing as an indicator. Money growth did not receive attention despite it was thought to cause inflation in the long run. Still, they did not differentiate between real and nominal rates and continued applying procyclical policies. In 1952 banks borrowed relatively large amounts from the discount window, taking advantage of the spread between the open market rates and the discount rate. In 1954 the federal funds market emerged again, as it had been inactive since the late 1920s.

Regarding discount rate policies the following is found. "At the August 23 1955 FOMC meeting Martin raised two issues: whether the discount rate should be a penalty rate, and whether it should lead or follow market rates… Following the meeting… in mid-September; the discount rate was a penalty rate" (Meltzer 2009a, pp. 127-128).

Later, "The Board reconsidered the role of discounting in its 1957 Annual Report… The Board, at last, recognized that when one bank repaid its borrowing, another might be forced to borrow, so that aggregate reserves did not decline. And it recognized that increased borrowing offset open market sales and that the attitude of member banks toward operation with borrowed resources varies from bank to bank… The Board found no conflict between discounting and open market operations. Market and discount rates were interdependent. By raising the discount rate above the market rate, the System encouraged banks to adjust by selling securities instead of discounting. Short-term rates rose, reinforcing an open market policy of sales" (Meltzer 2009a, p. 78).

1.3 1958-1970

Already for the period under analysis in the next chapter, there was a recession in 1957-1958. In 1958, the president asked to make price stability an explicit goal of economic policy. The Great Inflation was underway during the 1960s, sustained by rapid money growth to finance the government budget and government spending for the Vietnam War. One of the Fed's problems during this period was that it acted to reduce inflation only until unemployment rose. Given the unpopularity of inflation since 1965, emphasis shifted between those two goals.

Free reserves were increasing in 1960 and the Fed interpreted it as easy policy. As the economy slowed, discount rates decreased and free reserves and federal funds rate rose. During 1961-1962, the free reserves target was questioned and for the first time there was a target for the T-bill rate. On January 1962, an increase in deposit ceiling rates was approved. When the Fed began to control interest rates during those years, the problem was that they contained less information about the market position. "Instead of the market being a window through which we can observe indications of private actions that might call for policy changes, we have made it—in part at least—a mirror of our own intentions with respect to rates" (Meltzer 2009a, p.429). In 1963 the free reserve target was abandoned and more attention was dedicated to the federal funds rate and less to the T-bill rate. In early 1965, the first of several errors to control inflation was made, when the president's Economic Report announced the need for further expansion, even though signs of strength had already appeared. The same year, the federal funds rate became again higher than the discount rate.

Since the Fed targeted a short-term rate, to prevent a change on it when the deficit increased, the Fed had to allow the monetary base to increase. Fiscal policy contributed to inflation with the president Johnson's large deficits in 1967-1968. Before Martin left the Fed in January 1970, the Fed had adopted growth money as a policy indicator and instructed the manager to change money market conditions if money growth deviated from a 2% annual rate. When Burns became chairman of the Fed, the manager in charge of OMO lost much of his autonomy and the "tone and feel policy" ended forever. Money growth became the target and the FOMC would take decisions based on a total reserves target. Money and bank credit growth were used as target rather than as projection, meaning, the manager would change the federal funds rate when he missed the target. As the procedure began raising federal funds rate, the System ended up supplying more reserves, in part to prevent the failure of the Treasury financing. Thus, the FOMC soon gave up on monetary control. Higher rates supposed higher unemployment, and Burns disliked the result. By August 2, wage and price controls were imposed. Also the gold window was closed, and the currency was allowed to float.

"In short, the simple Keynesian model as applied in the late 1960s had three major flaws. It did not generally distinguish between nominal and real interest rate changes. It presumed that the government could permanently reduce the unemployment rate by permitting the inflation rate to rise. And it did not distinguish between one-time price level changes and maintained rates of price change. Each of these errors continued throughout the 1970s" (Meltzer 2009a, p.490).

1.4 1971-1980

This decade started with a freeze in prices and later, in interest rates, rates charged on mortgages, and consumer credit. Burns believed that in order to achieve full recovery without inflation, it was necessary to increase profits and lower wages growth, as it was understood that cost-push by unions was causing inflation. Another three phases of price controls were extended until 1973. Although in the beginning they decreased inflation, once finished, inflation increased even more. In the end, the public lost credibility on these controls.

During these years, it was usual to target growth in monetary and credit aggregates, setting at the same time lower and upper bounds on the federal funds rate. Before the elections in 1972, it is said that Burns was pressured by the government to increase growth and decrease unemployment. Once the elections passed and the administration loosed price and wage control, there was an inflationary outbreak due to those expansionary policies. In 1973, member bank borrowing rose to levels not reached since 1921 and the discount rate rose only in August 1973. Again, due to positive the spread between the discount rate and federal funds rate from 1972 to 1975. The Fed was also targeting the growth of reserves against private deposits during this period. The procedure contributed to increasing inflation as the staff estimated the growth of reserves and the level of federal funds rate consistent with the desired growth of money. Several times, they did not match and the band on federal funds rate had to be changed. In the end, the manager maintained the federal funds rate and exceeded the reserve target. As inflation increased, the federal funds rate and the discount rate reached levels never seen before. However, such was the spread between both rates that member bank borrowing increased significantly, contributing to the growth of the monetary base. It seems that along with the excessive monetary and fiscal expansion of 1972, the removal of price controls in 1973, the devaluation of the dollar after 1971, poor harvest abroad and the increases in oil prices, added to the rise in inflation. Beginning in 1975, borrowing declined and the federal funds rate came down rapidly. The FOMC continued using the federal funds rate as it principal target during this decade. "This period is unique in that the Fed controlled the funds rate so closely that market participants could identify most changes in the funds rate target on the day they were first implemented by the Fed, and these changes were reported by the market participants in the financial press the following day" (Meltzer 2009b, p. 892).

By October 1978, inflation had become a political issue as public opinion saw it as a major problem. In December 1978, the oil-producing countries decided to raise prices again. Inside the FOMC, more arguments appeared regarding how to control inflation and how this could reduce employment at the same time. Other concern was that the lack of credibility was damaging the effect of raising the federal funds rate, as the public expected that the Fed would not continue its restrictive policies and prices would increase again. Mark Willes claimed:

> "We can in fact have less inflation without more unemployment in 1980 if we have policies in 1979 that are…firmly held to so that people really believe we are going to follow through on them (FOMC Minutes, February 6, 1979, 19)" (Meltzer 2009b, p.940).

1.5 1980-1990

In August 1979, Volcker became chairman of the Fed. At his confirmation he already distinguished between real and nominal interest rates and expressed inflation as his main concern. He, following Milton Friedman, accepted that inflation could not be reduced unless money growth declined relative to growth of real output. The FOMC used a federal funds rate target and announced objectives for growth of M1 and M2 to reduce money growth. For that task, Volcker targeted nonborrowed reserves. However, they also paid attention to total reserves to move the target for nonborrowed reserves. "Monetarists criticized the procedures at the time, arguing that they made both interest rates and money growth more volatile. Growth of the money stock depends on reserve growth (or the monetary base). By holding to a fixed value (or growth) of nonborrowed reserves, banks had to borrow any deficiency to meet required reserves on deposits outstanding two weeks earlier, thereby increasing total reserves. …Further, keeping the discount rate as a penalty rate slightly above the average federal funds rate, would reduce borrowing. Most often, the Fed subsidized borrowing in 1979-82 by allowing a wide spread, 4 or 5%, between the average federal funds rate and the discount rate. This encouraged borrowing and weakened control of money" (Meltzer 2009b, p. 1028).

The attitude change on the FOMC was apparent once despite the recession of 1980, it favoured slower money growth. However, although by March the discount rate was raised, the action was insufficient and late. Heavy borrowing continued because of the subsidy of around 4 percentage points. Twelve years after Friedman's insistence on the effect of expectations, the Fed accepted that it could not permanently reduce unemployment by increasing inflation. Now, it was claimed that low inflation increased employment. The recession of 1980 and the posterior decrease in the discount rate in July avoided that the credibility on the Fed augmented. Policy tightened sharply in the spring of 1981, when the FOMC continued increasing the federal funds rate despite the recession and the unemployment rate near 8% in the fall. Market participants recognized that the Fed was fighting inflation. Thus, credibility increased. This was a turning point. By October, CPI inflation decreased quickly. The speed of the fall surprised the Fed. In part, it was ameliorated by the dollar appreciation undergone from 1980 to 1985. In 1982, the FOMC finished targeting nonborrowed reserves, and Volcker clearly began to shift to an interest rate target. He did not trust on M1 anymore:

> "On these money growth targets, in substance, I don't care. I think either of these two sets of numbers [5.5 and 6.5%] will make no difference, virtually, in what we actually

do… [W]e are within the limits of the growth targets anyway". (Meltzer 2009b, p. 1114).

"I, frankly, cannot live in these circumstances, given what is going on in the money markets, with violent moves in short-term rates in either direction. It would just be so disturbing in terms of expectations, market psychology, and fragility that it's just the wrong policy, period, during this particular period." (Meltzer 2009b, p.1115)

Shortly later, Volcker targeted borrowing to around \$500-\$600 million to prevent a raise in the interest rate. Thus, the System returned to the target used in 1920s and the basis for the free reserves target in the 1950s and 1960s. Policy became discretionary based on Volcker's judgment. Despite inflation had decreased significantly by the end of 1982, it was still high and again, created skepticism about the Fed's purpose of reducing it. In 1983-1984, long-term rates increased again. Consequently, inflationary expectations proxied by bond rates replaced money as intermediate target. Later, Greenspan concerns and focus on shaping the expectational environment would turn Volcker's experiment into a new monetary standard. A renewed rise in bond rates in the spring of 1984 tested Fed's compromise to fight inflationary expectations. Again, in 1984, Morris pointed out that the differential between the federal funds rate and the discount rate was of 2 percentage points and borrowing had reached \$1 billion. He proposed to increase the discount rate 1 percentage point but Volcker replied that it would mean "an explosion in Washington". Thus, during those years, many increases in the discount rate were not undertaken because of the pressure from the administration. When Greenspan replaced Volcker in August 1987, he set a narrow band around a federal funds rate target, which was adjustable depending on inflation and stable growth. In 1987, there was another inflation scare due to the depreciation of the dollar.

1.6 1990-2007

After the recession in 1990, the FOMC followed a "soft recovery" strategy and Greenspan focused on reducing expected inflation by reducing bond rates rather than just focusing on them during inflation scares. However, inflation concerns appeared in mid-1990 when Iraq invaded Kuwait and oil prices raised. Another inflation scare was faced in the beginning of 1993, but the Fed kept raising the federal funds rate during the next months until February 1995. "By the end of the decade, financial market had stopped associating high real growth with a resurgence of inflation. The Fed had defeated the "bond market vigilantes"" (Hetzel

2008, p. 205). By 1998, falling unemployment rate and low inflation created expectations about an increase in the federal funds rate. Nonetheless, low world growth perspectives led to lower the federal funds rate. Thus, the FOMC exacerbated an unsustainable rise in equity prices. From mid-1997 through mid-1999, the FOMC changed its procedures and raised rates when resource utilization rates were high. Greenspan believed that monetary policy should counter irrational expectations. He did not increase the federal funds rate again until February 2000 because he did not consider inflation as a threat. He believed that productivity growth was restraining inflation. As inflation and unemployment fell together after 1995, Greenspan explained:

> "The lack of pricing leverage has once again concentrated the minds of business people on the need to increase productivity… [E]conomic experience appears to be running full circle, back to the early 1960s: a period of low-inflation and strong productivity growth …[L]ower inflation historically has been associated…with faster growth of productivity… Lower inflation and inflation expectations reduce uncertainty in economic planning and diminish risk premiums for capital investment" (Hetzel 2008, p.231).

After the Asian financial crisis, the FOMC began to raise the federal funds rate. The equity market began its rise in 1995 until its peak in early 2000. After that peak, the NASDAQ began a prolonged fall after September 2000 and investors lost a significant amount of wealth. Consumption growth rates fell and the economy weakened. In January 2001, the Fed decreased rates, slowly and late. Thus, policy was contractionary by then. In 2001, policy followed the lean-against-the-wind pattern, whereby the FOMC raised (lowered) the funds rate in a persistent, measured way if the economy grew above (below) trend. The characterization for the last years of the Greenspan era is that the FOMC pursued its basic expected inflation/growth gap procedures but raised its implicit inflation target from price stability to low inflation.

-- Chapter 2 --

Have we been measuring monetary policy correctly? Analysing the Federal Reserve's policies over the last century

Unlike the standard and erroneous practice of using the federal funds rate or another intermediate target to measure the monetary policy stance, a new procedure is developed using the actual Federal Reserve's instruments and the spread between short-term rates and the discount rate. Accordingly, I estimate a time-varying coefficient Bayesian SVAR for the interwar period and 1958-2007. The new technique unveils a new mechanism operating between Fed's policies and the real economy. The results show that monetary policy was mostly irrelevant for the interwar period, but the situation changed after 1958. For this last case, however, the new mechanism, which focuses on the cost at which banks obtain reserves, explains that positive spreads between the federal funds rate and the discount rate contributed to increasing inflation, revealing that the "price puzzle" is non-existent.

- *JEL classification: E43, E51, E52, E58*
- *Keywords: monetary policy, Federal Reserve, Bayesians, SVARs, price puzzle, federal funds rate*

2.1 Introduction

The Great Recession and, inherently, the inability to avoid it supposed, or should have supposed, an inflection point. It showed that monetary policy, as a tool to achieve central banks' goals, was not managed correctly and therefore not properly understood. That lack of understanding means that our knowledge about the channels whereby monetary policy operates must be incorrect at some stage. To address that issue, this chapter takes not only a step forward, but also to the side, uncoupling from the standard and widespread approach to measuring the impact of monetary policy on the real economy. The step forward is materialized in the review of the Federal Reserve's history from its early days to the years prior to the Great Recession, regarding how this institution and its members reacted to political and economic events, how those actions supposedly influenced the American economy's performance and how monetary policy evolved until the years prior to the Great Recession. Thus, while the literature has focused only on certain periods of interests, the purpose of this review is to draw common patterns from the long-term picture and learn how monetary policy instruments were used and interacted with inflation, output and the money supply. Once a better understanding of the Fed's instruments is acquired, the step to the side is inevitable and supposes the major contribution of this chapter. The federal funds rate, another short-term money market rate or reserve measures have been used extensively as Fed's instruments in the literature to analyze the monetary policy stance and its impact on the real economy; however, by definition, they are intermediate targets. The actual instruments are open market operations (henceforth OMO), the discount rate and the reserve requirements ratio[1]. The conceptual mistake and the subsequent erroneous use of intermediate targets to measure the monetary policy stance introduce bias into the model (explained in section 2.2) and provide the wrong conclusions in relation to the impact of Fed's policies on the economy. Moreover, the correct use and understanding of the actual instruments, and the incentives they produce for the banking sector, inevitably lead to the creation of a new variable that measures arbitrage opportunities for bank reserves and provides an explanation for the known price puzzle, whereby when interest rate are raised, inflation increases. Therefore, the novelty of this chapter is the analysis of longer periods with the use of the actual instruments and the new variable. The new set-up facilitates the study of and shed new light on the relationship between instruments and the real economy, and potential regime

[1] After the 2008 crisis, new instruments were incorporated. See https://www.federalreserve.gov/monetarypolicy/policytools.htm.

changes in Fed's policies. As a result of measuring monetary policy stance correctly, a new hypothetical mechanism whereby monetary policy operates is discovered. This mechanism focuses on the source from where banks obtain reserves, and how the different cost of each source has a different impact on the economy. It also provides an explanation for the price puzzle and claims that there is not puzzle, but bad policies applied by the Fed, when it allowed positive spreads between short-term rates and the discount rate.

While a vast literature exists covering how the Federal Reserve's monetary policy stance has influenced the American economy's performance, references to the instruments have been used only and occasionally for those analyzing the interwar period, although from a narrative perspective; whereas to the best of my knowledge, there is no reference to the instruments for the second half of the twentieth century, despite the use of advanced methodologies. For the interwar period, Miron (1988) focused on monetary aggregates, and referencing how the Fed used the discount rate, he claimed that Fed's policies might have created more volatility in inflation and output during the 1920s and part of the 1930s. Bordo (1993) analyzed and compared how different monetary regimes determined the evolution of real variables. By estimating a bivariate VAR on the price level and output, he stated, "the gold standard and interwar period emerge as a relatively unstable period stressed by widely dispersed supply shocks" (Bordo 1993, p. 16). For those attributing the responsibility for the Great Depression, the debate has focused on whether the death of Governor Strong produced a change in policy implementation, mostly when using OMO, although also commenting changes in the discount rate. For example, Friedman and Schwartz (1963), Bordo and Schwartz (1999) and Hetzel (1985) considered Strong's years as governor of the New York Fed as a period of mostly successful monetary policy, but, once he died, those who opposed his ideas took charge, which could have created or worsened the Great Depression. On the other hand, Wicker (1965), Brunner and Meltzer (1968) and Wheelock (1989, 1990) argued that, had Strong lived during those years, the outcome would have been the same, as the policies were already ineptly administrated during his lifetime. By looking at the evolution of some real variables, market rates, the discount rate and OMO, Hamilton (1987) concluded that despite the slight change in policies during the 1920s, it was insufficient to explain the Great Depression and that some other factors were involved. Along this line and as exception in methodology, Ritschl and Woitek (2000) estimated a BVAR, using non-borrowed reserves, the discount rate or the short-term money market rates to measure the impact of monetary policy on real variables. They found that positive shocks to the discount rate and intermediate targets had, in general, a positive impact on inflation (price puzzle) and negative on output. They concluded that the monetary policy before the stock market crash did not cause the recession. At most, it could have produced a mild recession, which is also in line with Temin (1973). For the period after 1933, Friedman and Schwartz (1963) and Orphanides (2004) claimed that, despite the Fed's inactivity during those years, the inflow of gold was the factor guiding the well functioning of the economy.

For the second part of the twentieth century, since the Fed recovered its independence from the Treasury in 1951, the literature has been approaching Fed's monetary policy stance rather methodologically and mainly immersed in the "good luck, good policy" debate[2], which is also related to monetary policy switching regimes. At the same time, this literature gathers common characteristics, such as the use of VARs and the controversy about the price puzzle. As commented previously, the use of the federal funds rate as Fed's instrument has been the common procedure to evaluate monetary policy actions and I will specify when some variations are applied. On the one hand, some authors have focused on Fed's responses to movements in output and inflation. Such is the case of Clarida, Gali and Gertler (2000) using a GMM for a monetary policy reaction function. They observed that a policy change occurred during the Volcker–Greenspan era. This change is assumed to have brought stability to the economy, avoiding the indeterminacy equilibria existing before the Volcker era, by responding more aggressively to inflation. With a similar model, Favero and Rovelli (2003) obtained analogous results. Cogley and Sargent (2005), applying a similar model to the one in Canova and Gambetti (2009) indicated below, but using the 3-month Treasury bills rate, observed changes in the Federal Reserve's stance toward inflation, but their conclusion was not decisive in disentangling the good luck from the good policy hypothesis. Orphanides (2004b), comparing Taylor rules with real data and the data available for the Fed, claimed that bad policies played a relevant role during the Great Inflation, as the Fed wrongly understood how the economy worked, mistakenly predicting larger output gaps and intervening in the economy more than necessary, thus creating instability. Once it focused on inflation rather than the output gap, and the interventions became fewer and more accurate, the situation improved.

On the other hand, the object of study is rather the effect of positive shocks to the federal funds rate or a similar variable, on the real economy and money aggregates. Firstly, however, the use of VARs is evaluated. Giannone, Lenza and Reichlin (2008) examined VAR models of different sizes and trusted more those that included more variables, arguing

[2] This debate evaluates whether the American economy's performance has been the result of Fed's policies or external shocks. The main focus is on the pre- and post-Volcker era.

that the VARs supporting the good luck hypothesis, which contain fewer variables, are naïve models; Benati and Surico (2009), who, using a New Keynesian model via Bayesian methods, whereby they moved from determinacy to indeterminacy states, were able to demonstrate that those works based on VARs and supporting the good luck hypothesis may have misinterpreted good policy as good luck. In this case, inflation responded negatively to a positive nominal interest shock. Beyond VARs evaluation, Boivin and Giannone (2006), who used an SVAR, founded that a positive shock to the federal funds rate originated a price puzzle when analyzing the inflation responses, whereas output responded negatively. With a VAR, Christiano, Eichenbaum and Evans (1998) found that the Federal Reserve's activism declined in the late 1960s and was neutral in the early 1970s. Then, the Fed became passive for the remainder of the 1970s, not increasing the federal funds rate enough to counteract inflation. In 1981, monetary policy became activist again until Greenspan's term, when it decreased slightly but regained strength since 1993. By using alternatively the federal funds rate or nonborrowed reserves, a positive shock to these variables showed mostly a negative impact on the money supply and output, and initially no effect on inflation, which turned negative after roughly a year. However, when commodity prices were excluded from the estimation, a price puzzle was found. Primicery (2005), using a time-varying coefficient Bayesian structural vector autoregression (TVC-BSVAR), argued that, despite observing a change in monetary policy, it was not significantly different between the pre- and post-Volcker periods, and Canova and Gambetti (2009), with a similar method, found that the policy was the same for both periods, showing that the Taylor principle was not satisfied in any of the periods and that the transmission of monetary policy shocks to output and inflation remained stable over the periods analyzed, but inflation's persistence changed over time. For the former, a small price puzzle is found depending on the period, which disappears quickly, while for the latter, both, inflation and output respond negatively to a positive federal funds rate shock. Using a semi-structural VAR, including the federal funds rate, nonborrowed and total reserves, Hanson (2006) showed that a change in the policy is noticeable after Volcker's era but that it seems more probable that shocks coming from variables such as output or prices were important in determining the economic performance. Again, a price puzzle was found.

For those analyzing the evolution and volatility of inflation and output, but still using the same variables to measure monetary policy actions, Gali and Gambetti (2009), using a TVC-BSVAR and analyzing the variations in non-technology and technology shocks, concluded that monetary policy could have been among the factors explaining the decrease in volatility in output after the Great Inflation. Stock and Watson (2003), examined time-varying standard deviations and VARs with break tests and attributed the reduction of volatility to the decrease in shocks but also ascribed 10%–25% of the importance to improved monetary policy. Moreno (2004) developed a rational expectations model and showed that CPI inflation volatility declined in the 1980s and 1990s because of the propagation mechanism, but, considering the GDPD volatility, the decline is explained by smaller shocks. Federal funds rate and 3-month Treasury bill rate were used as Fed's instruments.

This chapter uses the same model as Primiceri (2005), namely a TVC-BSVAR with Del Negro and Primiceri's (2013) corrigendum, but including the actual Fed's instrument along with a variable measuring the difference between the short-term rate of reference for the period under analysis and the discount rate. The results suggest that monetary policy was almost irrelevant for the interwar period, as none of the instruments or the new variable is significant when analyzing its impact on output and inflation. For the second period, monetary policy gains relevance and the discount rate and the spread are able to influence the evolution of the variables under analysis. Furthermore, two regimes changes are observed around 1965 and 1990. While the first one may correspond to a change in Fed's policies, the second one is probably due to a change in the banking sector's behavior. Last and most important, the use of the spread unveils a new mechanism that explains the prize puzzle found in papers such as Barth and Ramey (2001), Boivin and Giannoni (2003), Bernanke, Boivin and Eliasz (2005), Uhlig (2005) and Hanson (2006), among many others. The reason for the creation of this variable is that for those periods when this spread was positive, banks had the possibility of obtaining higher profits by borrowing more cheaply at the discount window and lending at higher rates. The obtaining of cheaper reserves increased the possibility that banks set relatively lower loans rates in relation to the raises in the federal funds rate. Consequently, the insufficient restrain of credit and the money supply, triggered higher inflation levels. Therefore, I claim that there is no puzzle in prices behavior but bad Fed's policies by allowing those positive spreads.

The chapter is structured as follows. Section 2.2 develops the theoretical framework necessary to understand the model, which is described in Section 2.3. Section 2.4 covers the data sources, the identification structure of the VAR and the priors used for the model. Section 2.5 analyzes the results obtained. Section 2.6 gathers all the lessons and patterns obtained from section 2.5. Finally, section 2.7 summarizes the main conclusions.

2.2 Theoretical framework

As commented previously, the common way of approaching and evaluating the monetary policy stance, in methodological terms, has been the use of the federal funds rate or some measure of reserves (typically total reserves or nonborrowed reserves) as Federal Reserve's instruments. By definition, this assumption is erroneous. The federal funds rate or any type of reserve measure has always been and will be an intermediate target to achieve a final target such as price stability, stable growth or low unemployment. Actually, the Federal Reserve has available three instruments to achieve its intermediate and final targets. Such instruments are OMO, the discount rate and the reserve requirements ratio. Accordingly, to analyze the Federal Reserve's role and the impact of its policies on the American economy's performance, it is appropriate to use those instruments. In my model, I count with two of them, the discount rate and OMO, two intermediate targets, the spread between the short-term rate of reference (call loans rate[3] or federal funds rate) and the discount rate,[4] what is the new variable commented above, and M1, and two final targets, the industrial production index (IPI) and CPI inflation. The reserve requirements ratio is not incorporated given that it hardly varies over time.

To understand why it is erroneous to use of the federal funds rate (or another short-term rate)[5] to measure monetary policy, the money market must be conceptualized as two submarkets. The first submarket includes the central bank, in this case the Fed on the supply side, and the banking sector on the demand side. In the second submarket, the banking sector switches to the supply side, the other agents of the economy being the demand side. The Fed

---

[3] The call loans rate, for the interwar period, was the most similar rate to the present federal funds rate. "The market for brokers' loans, as it is generally conceived, is centered around the New York Stock Exchange. Although some of these loans grow out of a customer relationship between banks as lenders and brokers and dealers as borrowers, the majority are made in the open market on a strictly impersonal basis. The market in which these loans are made was until recent years the most active and the most sensitive of the money markets of the country. It was the market where surplus funds of banks, and sometimes of other lenders, could generally be readily placed or from which funds could be quickly withdrawn when needed. Because of the dominance of call loans, the branch of the money market dealing in brokers' loans has been designated as the call money market" (Board of Governors of the Federal Reserve System (U.S.), 1935–. *Banking and Monetary Statistics, 1914-1941*, 1943, https://fraser.stlouisfed.org/title/38, p. 434).

[4] In this case, it is not really an intermediate target but a hybrid between an intermediate target and an instrument, because the short-term rate is an intermediate target and the discount rate an instrument.

[5] The same applies to any measure of reserves, as a reserve target will determine the federal funds rate or the other way around.

controls the federal funds rate by purchasing and selling securities in the open market and increasing or decreasing the discount rate. Thus, the Fed controls the amount (with OMO) and price (with the discount rate and federal funds rate) of money in the first submarket, subject to the banking sector's demand for money and subsequent decisions about the federal funds rate and to considerations regarding whether that level of demand could harm the stability of the economy. Given the price and amount of money set by the Fed, banks will also decide, the amount (loans), although just in part, and price (loans rate) of money in the second submarket. Unlike Fed's intentions, banks will set prices and amounts conditioned on their profitability. Their decisions about prices and quantities will depend, mainly, on the cost of money to them, namely, reserves cost, and the prospect of profits from lending.

Thus, the federal funds rate is affected by supply and demand forces in the first submarket. The banking sector (as the demand side) will set a federal funds rate depending on several factors, such as the demand for reserves necessary to back the amount of loans, banks' surplus reserves, expectations about the adequate level of reserves to hold for the future or their own desire regarding the optimum level of reserves held under certain circumstances. Meanwhile, the Federal Reserve (supply side) will target a federal funds rate to control the evolution of the demand side by using OMO and the discount rate. However, the simultaneous use of both instruments to set the same level of the federal funds rate at two different points in time will be different. The reason is that as the federal funds rate depends also on the banking sector, the Fed will have to purchase or sell different amounts of securities and the discount rate will have to be set at different levels, to provide banks with the required amount of reserves determined by the demand forces, under the federal funds rate targeted. In the end, the same level of the federal funds rate can produce different equilibria. A greater demand for money faced with a greater supply of money can have a price of equilibrium (the federal funds rate in this case), which can be the same with a lower demand and supply of money. Nonetheless, this lower or greater supply of money will have different effects on the economy, because under the same federal funds rate, the cost and amount of reserves provided will have different impacts on the loans rate, which in the end, is the rate determining a greater or lower demand for loans. Hence, including the federal funds rate in the econometric model to represent the monetary policy stance is erroneous, because, by capturing demand and supply forces, the supposed policies carried out by the Fed will produce misleading results, as the banking sector can, to some extent, modify them and the results will capture those demand side modifications. By using the real instruments, Fed's policies, namely the supply forces, are isolated from the demand forces. However, despite being isolated from the Fed's instruments, the federal funds rate entails another problem if included in the model alone. Its impact on the second submarket is relative, as it depends on the discount rate levels. As both rates measure the cost of reserves from two different sources, banks will opt for the cheaper cost once arbitrage opportunities appear. This is where the new variable (the spread between the short-term rate of reference and the discount rate) plays its role.

Sometimes, even though the Fed increased the federal funds rate, if the spread between the federal funds rate (or another short-term rate) and the discount rate was positive, it offered profitable opportunities for banks. Thus, they continued borrowing cheaper reserves at the discount window and lending at higher rates than the discount rate. Lending increased or at least did not diminish as much as the Fed intended, because with a cheaper source for reserves, the federal funds rate was likely to not have a one to one impact on the loans rate, since the most important factor determining the loans rate is the cost of money to banks. As the impact ratio would be below one, the loans rate could not exert enough restrain on the demand for credit. Accordingly, the Fed partially lost the control of its targets.

Thus, in the model presented here, while Fed's instruments capture the supply side of the money market, this new variable will be a representation of those periods when banks could obtain cheaper reserves and therefore, the loans rate could not be moving one to one with the federal funds rate. Accordingly, positive spreads are likely to exert inflationary pressures, while negative spreads will do the opposite once the interest rate is raised, as long as that interest rate is high enough to restrain the demand forces.

2.2.1 Discarding some concerns

Given the novel exposure of this mechanism, it is necessary to clarify that this new procedure can be applied regardless of the period, the monetary regime or the financial environment.

Regarding the monetary regime, the advantage of using real instruments is that the analysis of the policies undertaken by the Fed will not be altered by the different active monetary regimes during the period of study, because a monetary regime implies an intermediate target, such as gold, the money supply, interest rate, reserves, exchange rate and so on. Consequently, if an intermediate target is used as a measure of the monetary policy stance, apart from the problem explained above about the federal funds rate (which also applies to the other intermediate targets), that variable must be removed from the model (a VAR in this case) or ordered differently as the monetary regime changes. Conversely, using

Fed's instruments, regardless of the monetary regime, the same instruments have to be used to achieve both intermediate and final targets. The only way whereby the Fed can achieve its goals under any monetary regime goes through the use of the discount rate and OMO.

Concerns may be raised as well about how the financial environment has evolved and how it could have modified the relationship between the Federal Reserve and the banking sector. The most important source of change could be Regulation Q, as it imposed interest rate ceilings on deposits rates. This regulation was active from 1933 to 1986 but was binding only when market rates reached ceiling levels around the 1960s. That meant that banks could not offer enough yields to attract depositors. Consequently, saving and loans associations would catch those clients. According to Koch (2015), interest rate ceilings contracted banks' credit growth and affected the lending channel, because without more deposits, banks could not increase lending. This extra tool included in the Fed's armory would be, therefore, a significant omitted variable once ceiling rates were below market rates. There is, however, a significant flaw in the premise that banks need deposits to lend, as banks use those deposits as reserves. The reality is quite different. Banks first lend, and just after, look for the necessary reserves to back the demand for loans. There is an extensive literature explaining this issue, such as Moore (1988), Bindseil (2004) or Jakab and Kumhof (2015) among many others. This fact is also supported by some data. In Koch (2015, Figure 1), he displayed when and for how much markets rates were above ceiling rates. The periods of higher rates, except for 1960, coincides with those periods when the spread between the federal funds rate and the discount rate was positive (Figure 1.1). If those periods are now contrasted with Figure 1.2, where the amount of borrowed reserves is displayed, the evidence is clear, as borrowing increased when the spread was positive. Thus, even if banks had not been borrowing for the opportunity of obtaining cheaper reserves at the discount window, which is the case, the lower amount of reserves from deposits could have been counteracted with borrowing and credit would have not been restrained. Also, it is not surprising Koch's conclusion that when the interest rate ceiling was binding, credit decreased. As expected, for periods of increased market rates, at some point, lending levels will decrease. On top of that, there are also references in Meltzer (2009a, p.470, 608 and 648) to how banks evaded regulation by offering different kinds of deposits or services. However, the reader may be confused when reading that banks complained about ceiling rates (Meltzer 2009a, p.383). If they did not need deposits to lend, why did they complain? When market rates were above those ceilings, banks could obtain cheaper reserves from deposits, but the larger the spread, the fewer quantity of deposits was demanded. Therefore, even if ceiling rates were raised and banks

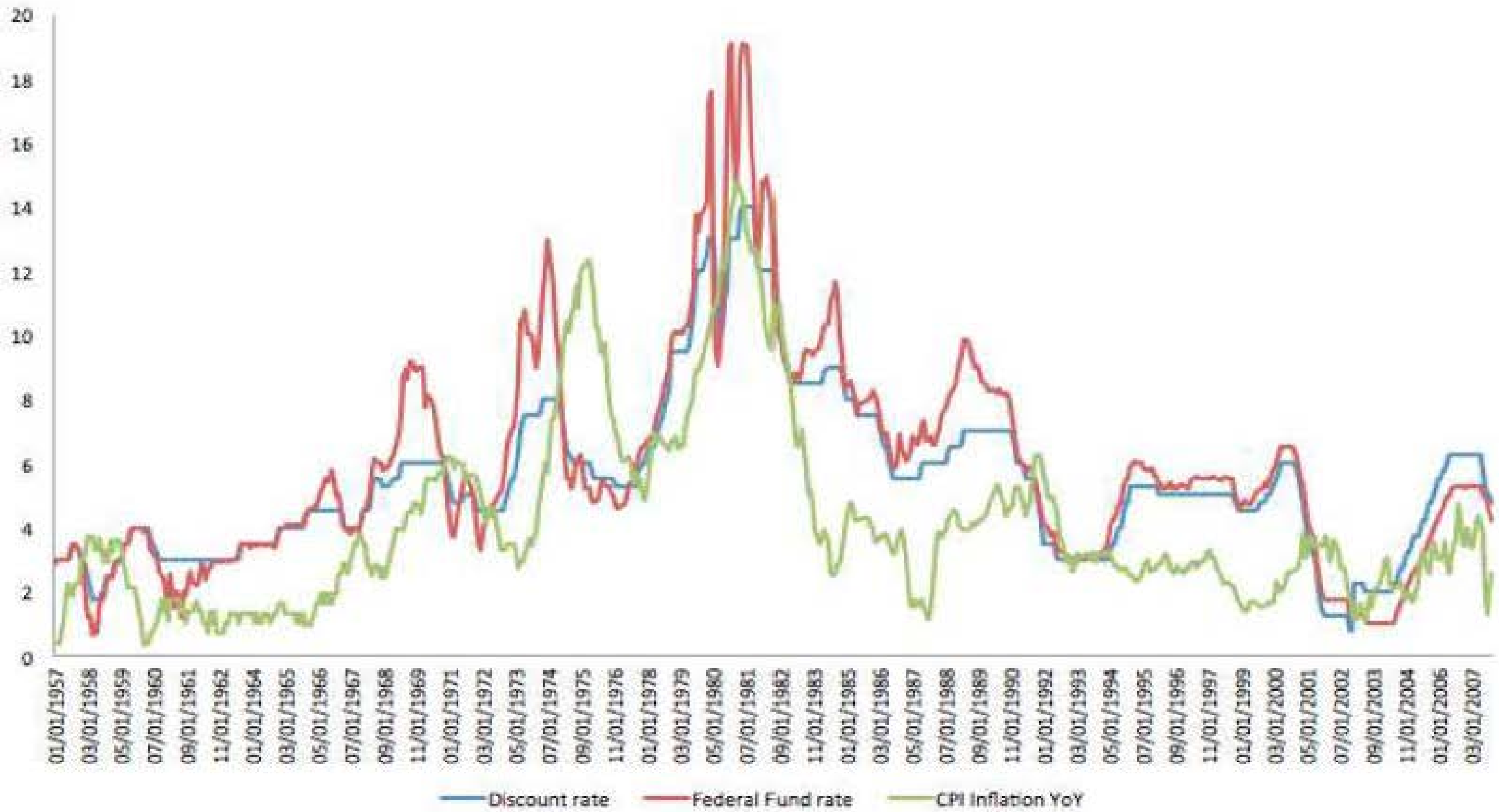

Figure 1.1 – Data Source: FRED

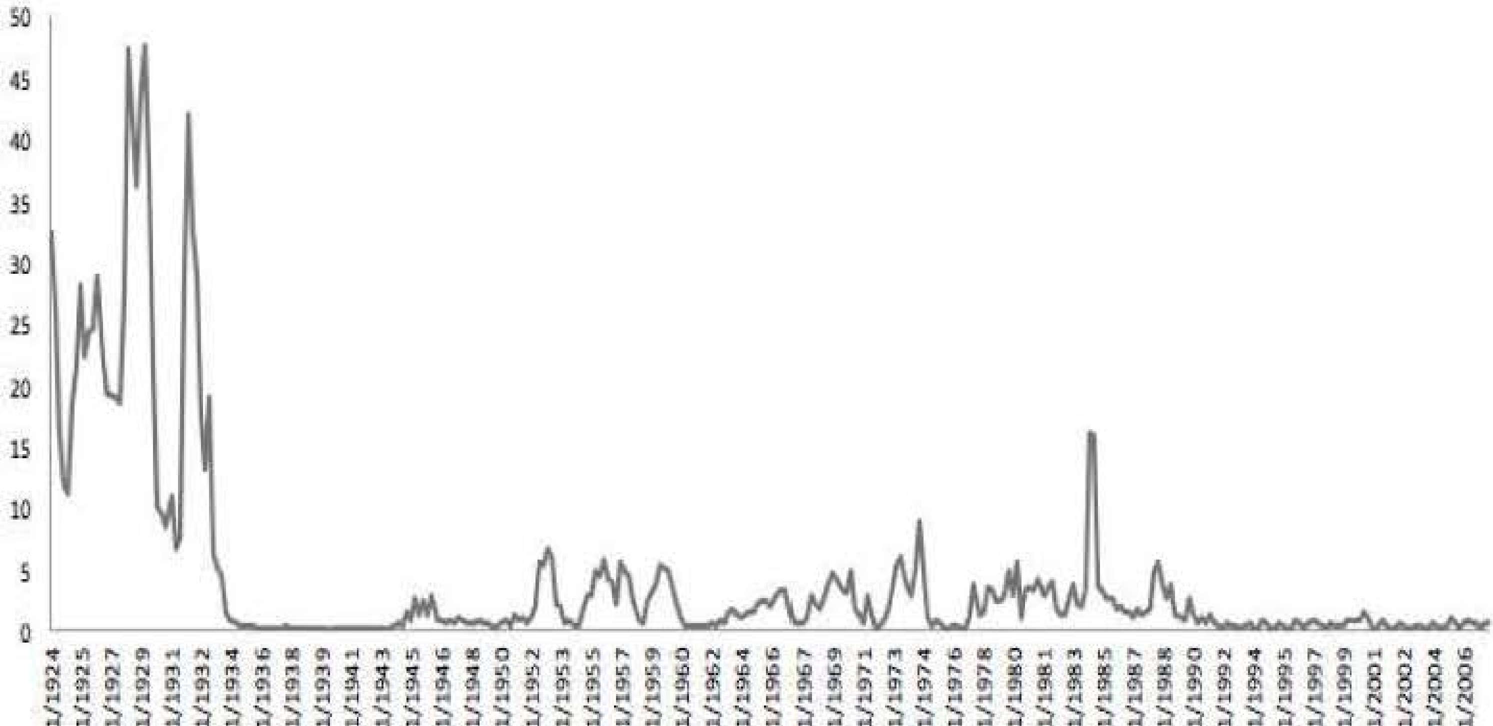

Figure 1.2 – Data Source: FRED and FRASER

had to pay more to their clients, they could still obtain cheaper reserves than from the discount window or the federal funds market.

Other source of concern is the evolution of the discount window. There are some misconceptions regarding its use and even though the model used in this chapter will capture any related variation or regime change, a couple of things need to be clarified. First, in 1955, the Board issued regulation A, where System orthodoxy was that banks did not borrow for profit but only reluctantly for need. Later, the discount rate stopped being a ceiling on the federal funds rate. It triggered more borrowing as seen in Figure 1.2, but the System needed a long time until it changed its mind and accepted that banks also borrowed for profits. Therefore, the supposed stigma for borrowing for that period is false. Second, the Depository Institution Deregulation and Monetary Act of 1980 allowed more institutions access to the discount window. Despite this fact, the analysis undertaken here is in aggregate level. That is, before some institutions could have access to the discount window, they were likely to borrow from the banking sector, which in turn, would borrow at the discount window if more reserves were necessary. Hence, no significant regime change is expected under such deregulation act.

2.3 Methodology

The model used in this chapter is the same as that in Primicery (2005), a TVC-BSVAR, in which, unlike other similar models, not only the coefficients vary but also the variance covariance matrix. The code used to estimate the model was downloaded from Gary Koop's website.[6] The advantage of this model is that the drifting coefficients are able to capture nonlinearities or time variation in the lag structure of the model, while the multivariate stochastic volatility is able to capture possible heteroscedasticity of the shocks and nonlinearities in the simultaneous relations among the variables of the model. Thereby, it allows the data to determine whether the possible variations observed in the relation among variables emanates from the shocks (impulse) or changes in the propagation mechanism (response). The adequacy of this model for the purpose of the chapter is founded on its capacity to capture continuous and smoothed switching regimes, unlike those works that modeled time variation with discrete breaks. For the topic addressed in this case, it is

---

[6] https://sites.google.com/site/garykoop/home/computer-code-2.

The only modifications applied to the code are made to adapt it to the data used here as well as for those tools necessary for the representation of the results. The process of estimation is entirely as found in the file.

expected that the Federal Reserve, banking sector and other agents of the economy learn from the evolution of the economy and each other. The learning process is considered to be slow and not to happen overnight. Hence, changes in the behavior of those agents, as a consequence of their learning process, will evolve smoothly.

Using the same notation as Primiceri, the model is the following:

$$y_t = c_t + B_{1,t}y_{t-1} + \ldots + B_{k,t}\,y_{t-k} + u_t \qquad t = 1, \ldots, T \qquad (1)$$

where y_t and c_t are $n \ x \ 1$ vectors of observed endogenous variables and a vector of time-varying coefficients multiplying constant terms, respectively. $B_{i,t}$, $i = 1, \ldots, k$, represents $n \ x \ n$ matrices of time-varying coefficients. Last, u_t are heteroscedastic unobservable shocks. The variance covariance matrix Ω_t is triangularly reduced and defined by

$$A_t \, \Omega_t A'_t = \textstyle\sum_t \sum'_t \qquad (2)$$

where A_t is a lower triangular matrix with ones in the main diagonal, $\alpha_{ij,t}$ being the non-zero and non-one elements of the matrix. $\sum_t$ is a diagonal matrix with $\sigma_{n,t}$ elements in the diagonal. Hence,

$$y_t = B_{0,t} + B_{1,t}y_{t-1} + \ldots + B_{k,t}y_{t-k} + A_t^{-1}\textstyle\sum_t \varepsilon_t \qquad (3)$$

$$V(\varepsilon_t) = I_n$$

Stacking all the $B_{k,t}$s in a vector,

$$B_t = vec(B'_t) = [B_{0,t}, B_{1,t}, B_{2,t}, \ldots B_{k,t}]'$$

and with

$$X_t = I_n \otimes [1, y_{t-1}, y_{t-2}, \ldots y_{t-k}]',$$

the VAR can be represented and modeled as:

$$y_t = X'_t B_t + A_t^{-1}\textstyle\sum_t \varepsilon_t \qquad (4)$$

Stacking by rows the elements $\alpha_{ij,t}$ of the matrix A_t and the elements $\sigma_{n,t}$ of the matrix $\sum_t$, the state vectors or transition equations representing the dynamics of the model are:

$$B_t = B_{t-1} + v_t \tag{5}$$

$$\square_t = \square_{t-1} + \zeta_t \tag{6}$$

$$\log \sigma_t = \log \sigma_{t-1} + \eta_t \tag{7}$$

where both the B_ts and the non-zero and non-one elements of the matrix A_t, α_t, follow random walks, while the standard deviations of equation (7) follow a geometric random walk, accordingly belonging to the stochastic volatility models. The innovations of the model are assumed to be jointly normally distributed, supposing the following variance covariance matrix:

$$V = \mathrm{Var}\left(\begin{bmatrix} \varepsilon t \\ vt \\ \zeta t \\ \eta t \end{bmatrix}\right) = \begin{bmatrix} In & 0 & 0 & 0 \\ 0 & Q & 0 & 0 \\ 0 & 0 & S & 0 \\ 0 & 0 & 0 & W \end{bmatrix} \tag{8}$$

where I_n is an n-dimensional identity matrix and Q, S and W are positive definite matrices. As Primiceri pointed out, the zero blocks could be replaced by non-zero blocks, but there are two reasons for the assumptions taken. First, as Primiceri (2005) already considered the number of parameters to be high and adding non-zero blocks would require a sensible prior to prevent ill-determined parameters, I include the double of the variables in the model. Second, I do not have any structural interpretation to impose on the different sources of uncertainty. S is assumed to be block diagonal, with blocks corresponding to parameters belonging to a separate equation; that is, the coefficients of the contemporaneous relations evolve independently in each equation. For the estimation of the model, I refer the reader to Appendix A of Primiceri (2005), taking into account Del Negro and Primiceri's (2013) corrigendum, whereby the algorithm used for the Gibbs sampling undergoes a modification regarding the blocks from which the draws are taken.

In this model the Bs are restricted to being non-explosive to impose stability. As Koop and Potter (2011) (K-P henceforth) explained, "in the absence of such inequality restrictions (or a very tight prior), Bayesian TVP-VARs will place a large amount of a priori weight on nonsensical paths for the states." Primiceri used Carter and Kohn's (1994) algorithm, which draws an entire vector of states and rejects any that violate the constraint imposed. The problem of applying this algorithm is that, when the number of parameters is relatively high, it is easy for the algorithm to become stuck drawing explosive Bs. Thus, all the draws are discarded and computation is not feasible, as is the case in this chapter for some cases. To solve this problem, K-P developed a single-move algorithm.[7] While the MCMC algorithm in Primiceri (2005) draws from state space models without the inequality restriction, K-P's single-move algorithm draws from the state space model subject to the inequality restriction, drawing the states one at a time. This single-move algorithm does not become stuck rejecting every candidate draw, like the multi-move algorithm. It draws B_t from $p(B_t|y^T, Q, B_{t-1})$,[8] accepting the single draw B_t with a certain acceptance probability if it has satisfied the restriction imposed. Now, although the probability of becoming stuck diminishes significantly, the algorithm mixes more slowly.

2.3.1 Data, identification strategy and priors

The sample under analysis is split into two periods. The first one covers the interwar period with monthly data from 1925:I to 1939:XII, and the second period encompasses the interval between 1958:I and 2007:IV with quarterly data. The reason for using different periodicity is that for the interwar period the sample size is excessively small if using quarterly data, taking into account that a longer sample is necessary for the priors. For the first period, the variables used are the Industrial Production Index (IPI), the Consumer Price Index (CPI), M1, the difference between the stock exchange call loans rate and the discount rate (C-D), open market operations (OMO)[9] and the discount rate [10]. The first three variables and OMO are growth rates, while the spread and the discount rate are in levels. For the second period, the variables are the same except for C-D, as the stock exchange call loans rate is substituted for the federal funds rate (F-D). While the call loans rate and the federal funds rate represent different (although similar) money markets, they perform the same role. In both cases, they represent the price at which bank could obtain reserves, when they were not borrowed at the discount window. Another difference is that for the interwar period, the Fed did not target short-term rates directly as in the second period, when OMO were used for that purpose. However, those short-term rates were involuntarily conditioned by the movements in the discount rate anyways. That is, the short-term rate responded to demand forces for the interwar period rather than to Fed's desires as in the second period, but still, it was influenced

[7] Koop and Potter (2011). Section 2.3, pp. 13–15.

[8] The other blocks are not included in the notation (although they are in the algorithm), as the modification only affects the draws of B_t and Q.

[9] U.S. Government securities (bought outright and repurchases) and acceptances held by the Federal Reserve.

[10] Until 1948, the discount rate belongs to the New York Federal Reserve Bank as representative of all other Reserve Banks. Later, all of them offered a homogeneous rate.

by Fed's instruments. To have all the variables on the same scale, they are standardized $(y_t - E(y_t^*))/std(y_t^*)$. The order of the variables (contrary to that indicated above) takes the IPI as the last variable in the VAR and the discount rate as the first one. In this way, the relation among the variables has a structural interpretation: while the IPI reacts only after one lag to all the other variables' movements, the discount rate reacts contemporaneously to all of them. The order assumed is based on the mechanism described in section 2. The first price that the Fed sets in the first submarket is the discount rate. From there, it controls the federal funds rate or whatever intermediate target it has through OMO. Even though sometimes OMO and the discount rate will move at the same time, the discount rate remain at the same levels for longer periods. During those periods, the intermediate target is adjusted through OMO. Both instruments, voluntarily or not, will determine the spread. The data were collected from the National Bureau of Economic Research (NBER), the Federal Reserve Economic Data (FRED), the Bureau of Labor Statistics and the Reserve Archival System for Economic Research (FRASER).

Regarding the lag structure, I find limitations to this model. Given the high dimensionality of the parameters estimated, adding more than one lag in any of the periods triggers the draws from the B's distribution to be non-stationary. Having previously imposed the stationarity restriction, the multi-move algorithm used by Primiceri becomes stuck in the zone of the distribution where the draws are non-stationary; therefore, no draw is taken. Using K-P's single-move algorithm, I am able to introduce one lag more (including more than two lags makes the algorithm collapse). Thus, for the interwar period with monthly data, I use K-P's algorithm directly to have at least two lags, which are already few. For the second period, I present the results for one lag (multi-move algorithm) and comment the few relevant variations obtained with two lags (K-P's single-move algorithm).

2.3.2 Priors and computational details

For the first period, an invariant VAR from 1920:I to 1924:XII (60 observations) is estimated to calibrate the priors' distributions, while, for the second period, the priors are obtained from the period from 1948:I to 1957:IV (39 observations)[11]. The set-up for the priors (as written in Gary Koop's code) is the following:

$$B_0 \sim N(\hat{B}_{OLS}, 4 \cdot V(\hat{B}_{OLS})),$$

[11] For the federal funds rate series, data is only available from 1954. As a proxy, I have used the 3-months T-bill rate from 1948 to 1954 to estimate the priors.

$$A_0 \sim N(\hat{A}_{OLS}, 4 \cdot V(\hat{A}_{OLS})),$$
$$\log \sigma_0 \sim N(\log \hat{\sigma}_{OLS}, 4 \cdot I_n),$$
$$Q \sim IW(k^2_Q \cdot 60(\text{or } 39) \cdot V(\hat{B}_{OLS}), 60(\text{or } 39)),$$
$$W \sim IW(k^2_w \cdot (1+\dim(W)) \cdot I_n, 1+\dim(W)),$$
$$S_1 \sim IW(k^2_s \cdot (1+\dim(S_1)) \cdot V(\hat{A}_{1, OLS}), 1+\dim(S_1)),$$
$$S_2 \sim IW(k^2_s \cdot (1+\dim(S_2)) \cdot V(\hat{A}_{2,OLS}), 1+\dim(S_2)),$$

where S_1 and S_2 are the two blocks of S, $A_{1,OLS}$ and $A_{2, OLS}$ are the corresponding blocks of A_{OLS} and $k_Q = 0.01$, $k_S = 0.1$ and $k_W = 1$. Thus, the priors are not flat but diffuse.

For the first and second periods, when the single-move algorithm is used, 400,000 draws are generated, discarding the first 200,000 and using 1 in every 100 to avoid correlation between them. For the second period and the multi-move algorithm, 450,000 draws are generated, discarding the first 200,000 and using 1 in every 125. The difference in the number of draws is explained by the computational time necessary and the percentage of acceptance of draws for each algorithm. Regarding the computational time to estimate the model, the multi-move algorithm required around 60 hours, and the single-move algorithm about 23 days. Convergence tests are displayed in Appendix C.

2.4 Results

The TVC-BSVAR provides two different tools to evaluate the impact of the instruments. First, the impulse response functions will show the posterior mean of the response of certain variables to another variable shock. The use of time-varying coefficients allows the discovery of whether the interaction between two variables has undergone any significant change over the period under analysis. This may be the first hint regarding whether the Fed, at some point, modified its policies or used its instruments differently. Second, the posterior mean of the standard deviation of the residuals for each equation of the VAR will shed light on possible external shocks affecting the results, namely variables that are not included in the model, which could distort the relations observed between the instruments and the variables under analysis.

Before analyzing the results, I advise the reader to pay keen attention to the relations described in Figure 1.3 for the interwar period and Figure 1.1 for the second period between the discount rate, the short-term rate of reference and the levels of inflation, because they are important to understand what the impulse responses display. It needs to be clarified that the analysis undertaken intends to comprehend the impact of the instruments on inflation, output and the money supply, regardless of the reasons behind Fed's decisions to use the instruments. That is, once the Fed increases the interest rate, because of correct or incorrect forecasts, anticipation to political events, international factors or whatever reason, that increase has an impact, and that impact is the only thing of interest in this study. This is the only way whereby a better understanding of the interaction between instruments and real variables or money aggregates can be acquired.

Given the detailed analysis of the impulse response functions and the number of figures in three dimensions, in this section, I only analyze the response of the final targets to the instruments and the new variable. The evaluation of the rest of the figures is available in Appendix A. Ordinary impulse response functions to analyze the responses' significance is evaluated only for the relationships between instruments and final or intermediate targets, and for selected years.

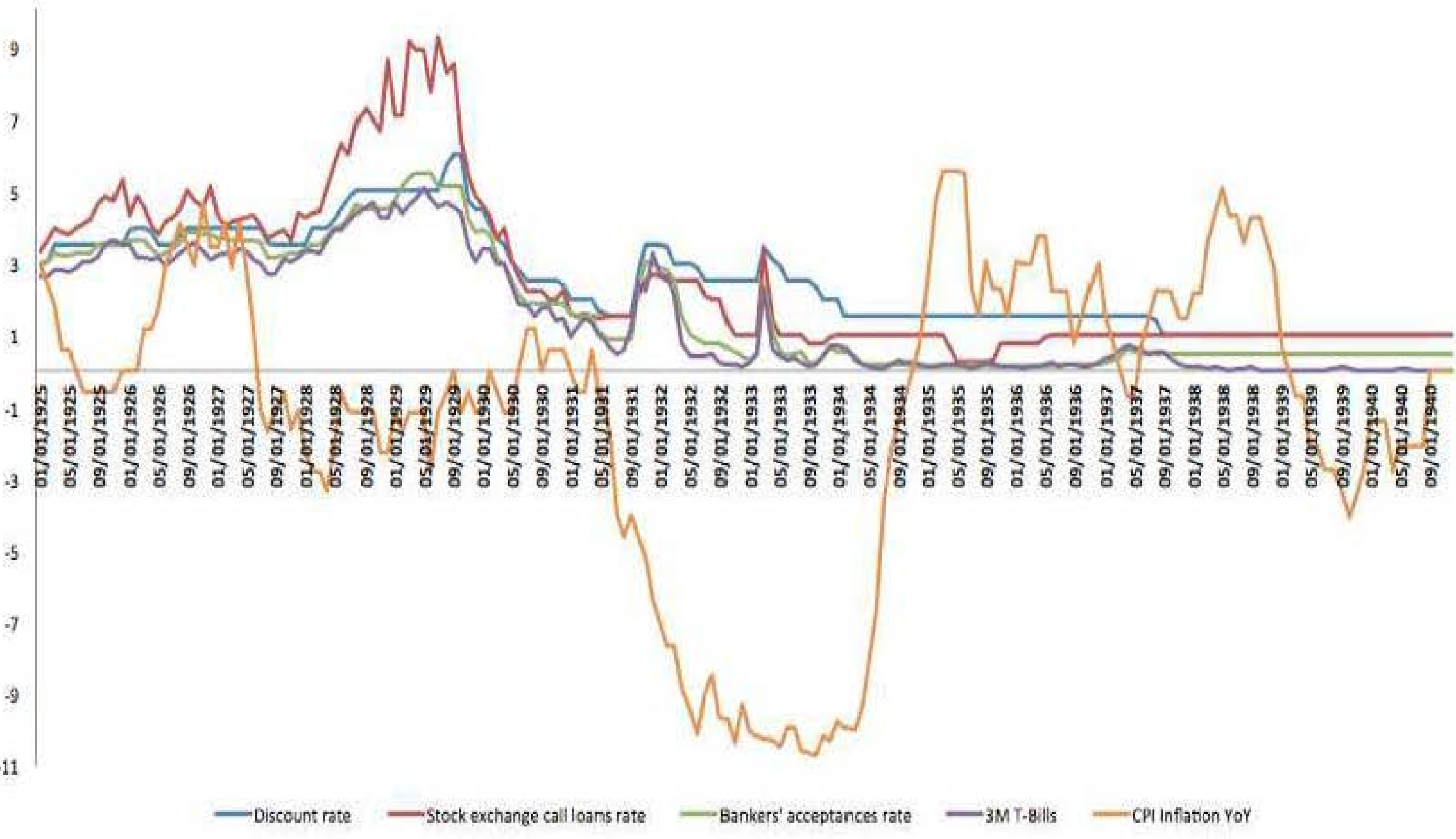

Figure 1.3 Source: Banking and Monetary Statistics, 1914-1941 (FRASER) and Bureau of Labor Statistics

2.4.1 Interwar period (1925:I–1939:XII)[12]

The figures analyzed below display each period of the sample on the X-axis, the response to the shock from one to twenty months/quarters on the Y-axis and the scale of the response on the Z-axis.

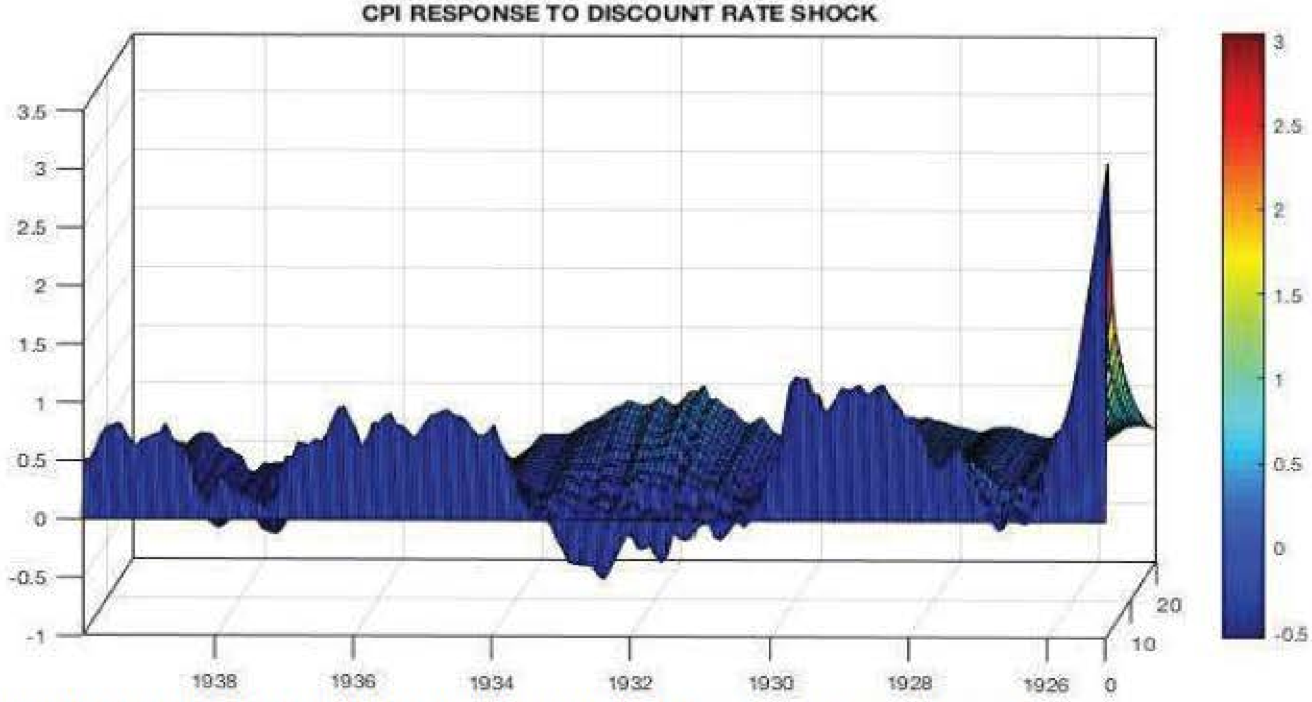

Figure 1.4.1 CPI inflation impulse response to a discount rate shock. Note: Posterior means.

Looking at the discount rate shock (Figure 1.4.1), the inflation response is positive from 1925 to the end of 1929, with the exception of the almost-zero response around mid-1926. The positive response coincides with the discount rate being below or near the call loans rate (Figure 1.3). Those positive responses could be also capturing the inflows of gold once the discount rate was raised, triggering increases in the money supply and inflation, as long as those inflows were not offset. For 1926, the zero response corresponds to the fact that the discount rate was near the call loans rate. Between 1930 and 1934, the negative or zero response coincides with the facts that the discount rate in real terms was higher than represented because of the deflation. Also, it was above the call loans rate. After six months, the response becomes slightly positive. From 1934 to 1939, the response is positive but

[12] The results shown for this first period belong to the model with monthly data and two lags using K-P's algorithm. Even though the analysis starts in 1925, Chapter 1 presents the knowledge of the Fed since 1919, which is relevant to a better understanding of the policies undertaken later.

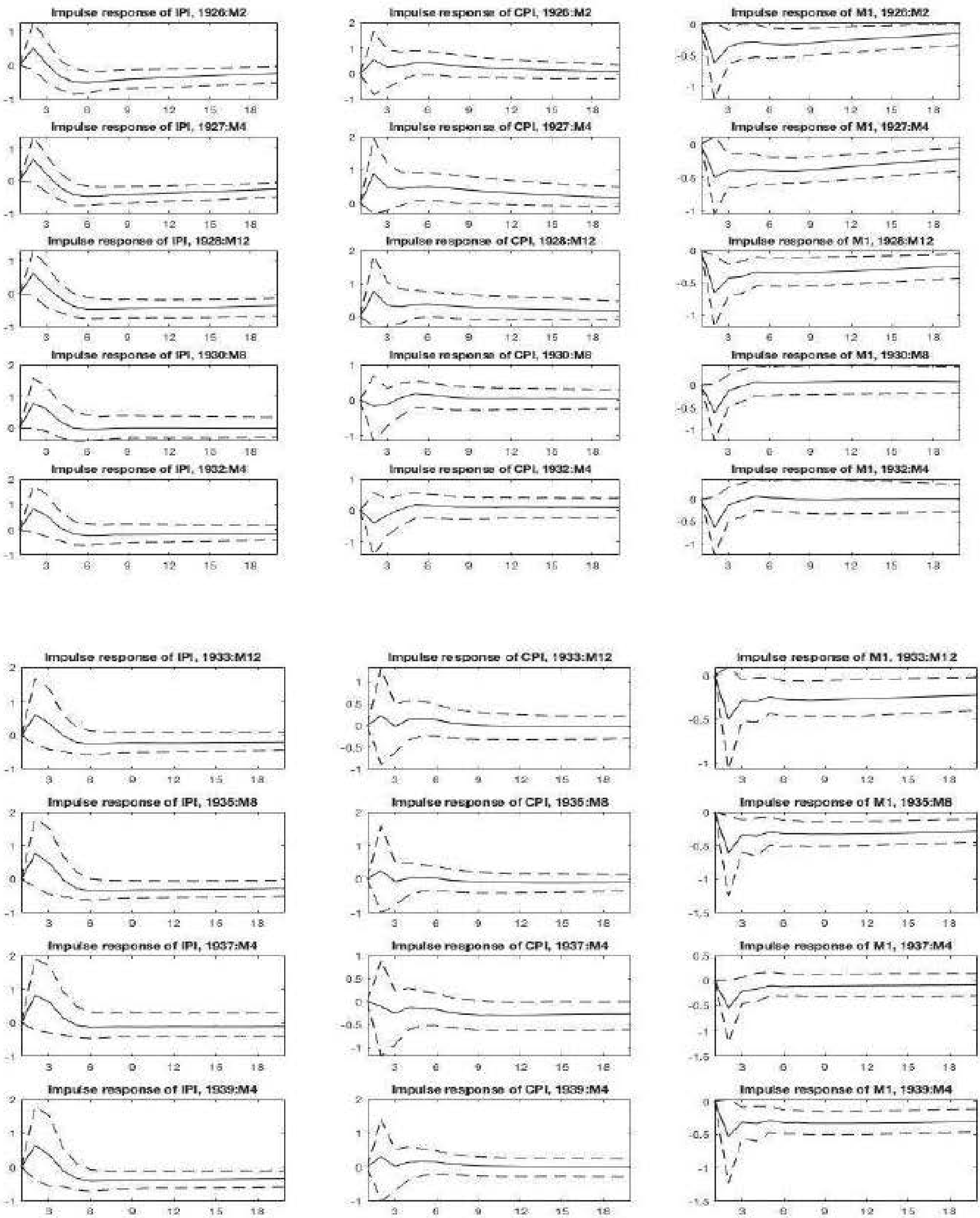

Figure 1.5.1 – Impulse responses to a discount rate shock. IPI, CPI and M1 in columns 1, 2 and 3 respectively. Note: The solid lines depict the 50-th percentile with the 16-th and 84-th percentiles for the dashed lines.

becomes negative after two to three months. This time, although the discount rate was above or at the same level as the call loans rate, it was mostly below the inflation levels, avoiding an initial negative response. Despite the variations observed, none of the responses are significant (Figure 1.5.1). The response of CPI inflation to an OMO shock (Figure 1.4.2) is the same for almost the whole period: initially positive but negative after two months. The exception is from 1927 to the end of 1929 (a deflation period), when it is negative. It corresponds to the Fed's gold sterilization. In general, this figure shows that, for the entire interwar period, open market purchases had an ephemeral effect.

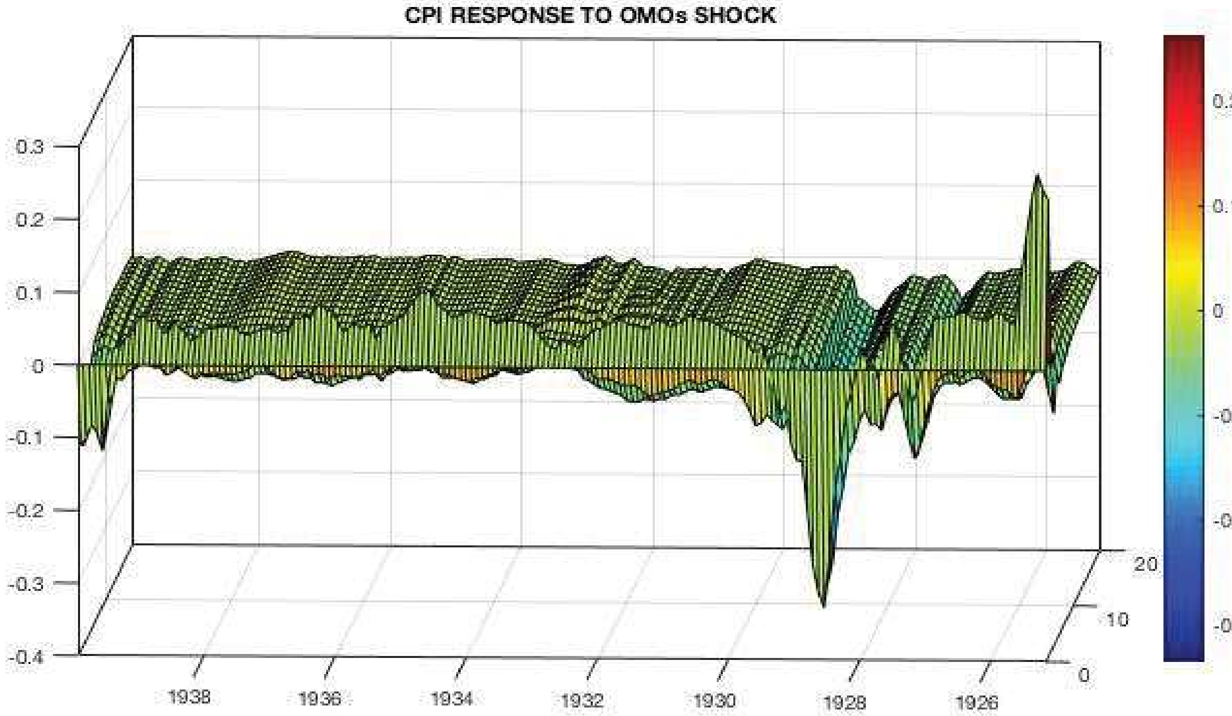

Fig. 1.4.2 – CPI inflation impulse response to an OMO shock. Note: Posterior means.

For the years 1930, 1931 and 1932, the Fed purchased more intensively than before but the consequences of the Great Depression regarding bank failures, a higher demand for excess reserves and currency, and gold outflows offset those purchases. After 1933, the Fed was relegated to the backseat, the Treasury being mostly in charge of the monetary policy. OMO and the discount rate were hardly used since then. These results are supported by Figure 1.5.2, in which again, a shock to this variable has no significant effect on CPI at any period.

Considering the shock to the new variable C-D (Figure 1.4.3), the response of CPI inflation varies depending on whether the spread is negative or positive, whether the economy is experiencing inflation or deflation and the position of both rates in relation to the inflation levels.

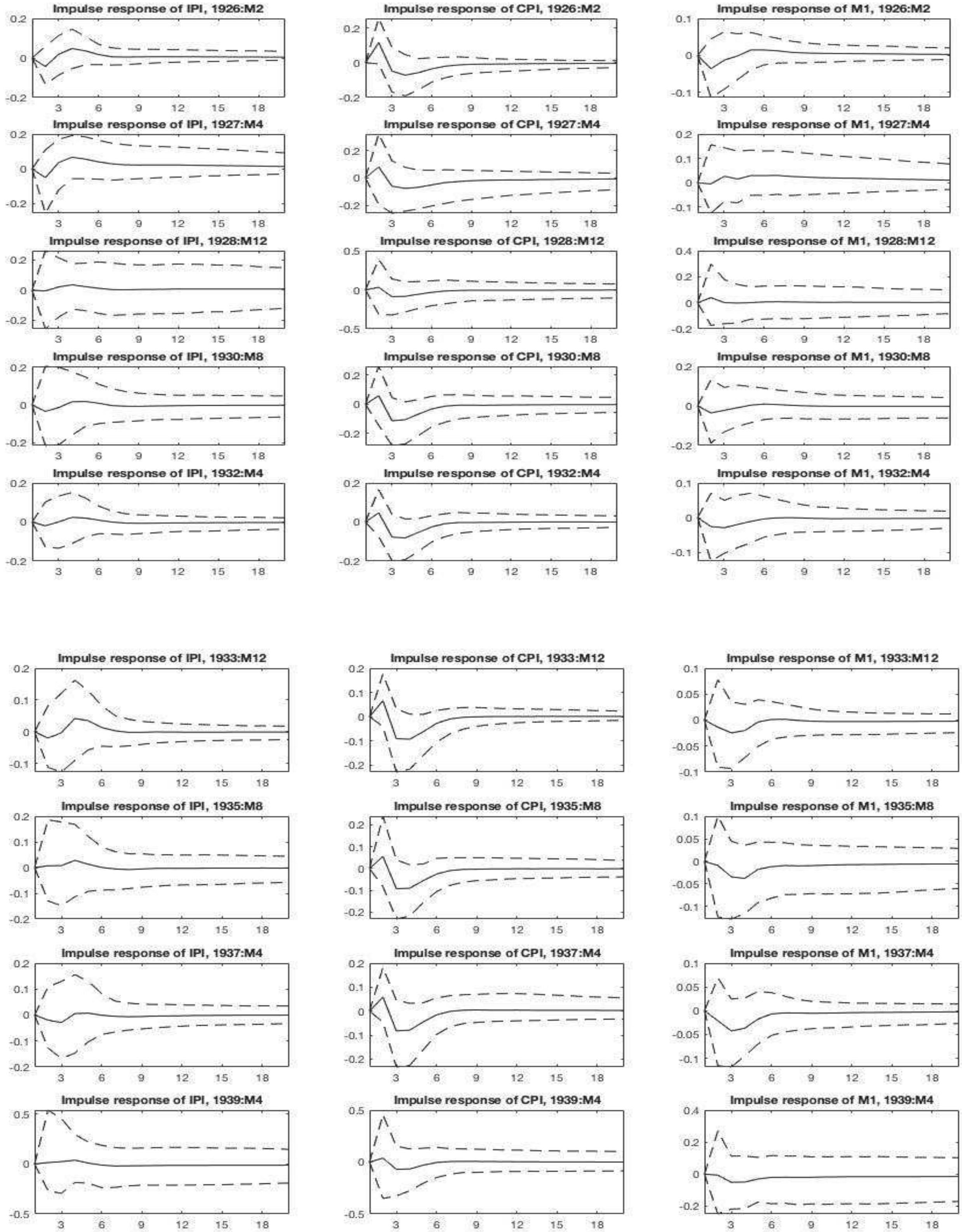

Fig. 1.5.2– Impulse responses to an OMO shock. IPI, CPI and M1 in columns 1, 2 and 3 respectively. Note: The solid lines depict the 50-th percentile with the 16-th and 84-th percentiles for the dashed lines.

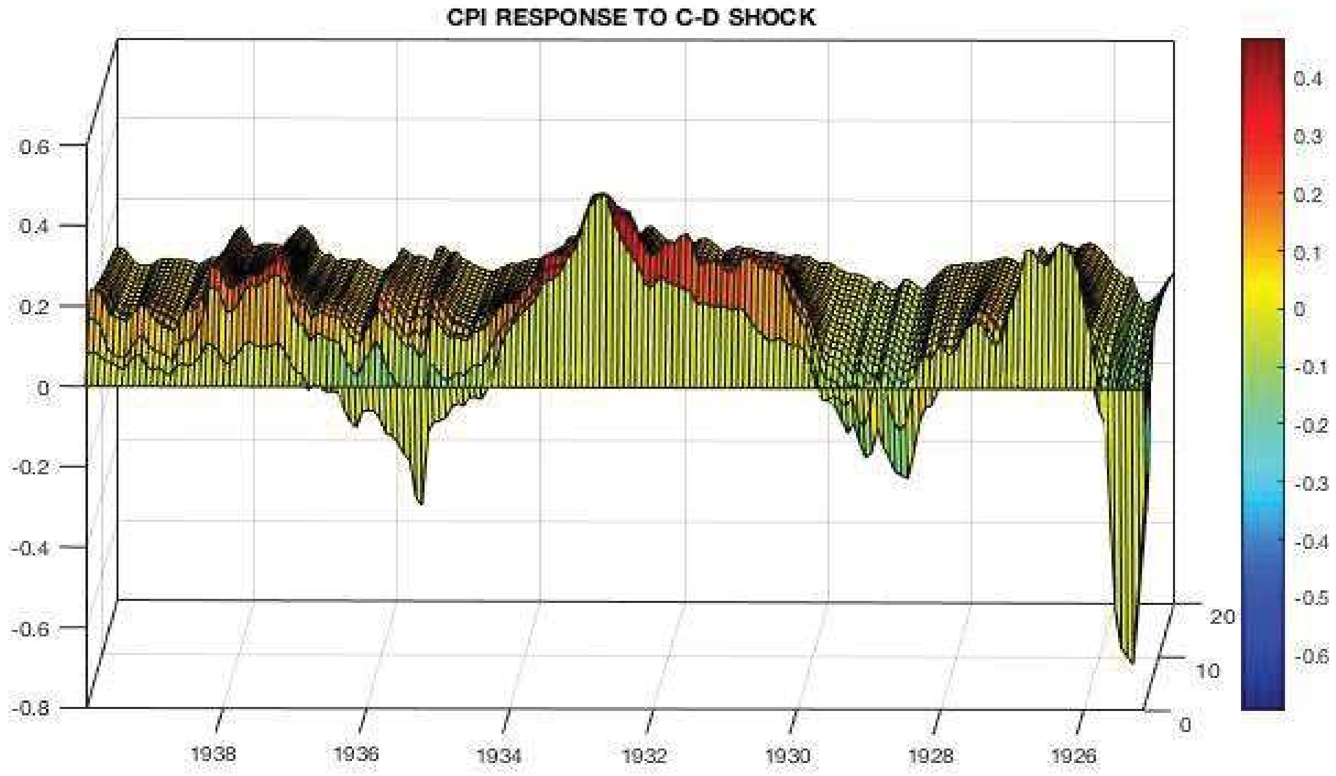

Figure 1.4.3 – CPI inflation impulse response to a C-D shock. Note: Posterior means.

When the spread is positive but there is deflation, the response is negative, as for the second half of 1925 and between 1928 and 1929. This could mean that the rates were too high. Consequently, a larger spread, meaning a higher call loans rate in relation to the discount rate would have tightened the economy even more. The response is positive from the beginning of 1926 to the beginning of 1928. This was a period with a positive spread (sometimes quite narrow or zero) until mid-1927, with both rates around the positive levels of inflation. For the last half of 1927, a larger spread would have exerted inflationary pressures, counteracting the deflation triggered by the sterilization of gold. From 1930 to 1933, CPI inflation responds positively to an increase in the spread, as for the last half of 1927, but under a deflationary scenario. For this period, the spread was almost zero or negative. It seems that a positive or less negative spread would have supposed an increase in inflation. Related to the last statement, the highest positive peak during this period occurs around the beginning of 1933, when the short-term rates increased, as seen in Figure 1.3. However, the Fed did not allow them to be above the discount rate. For the rest of the subperiod, the response is mostly positive.

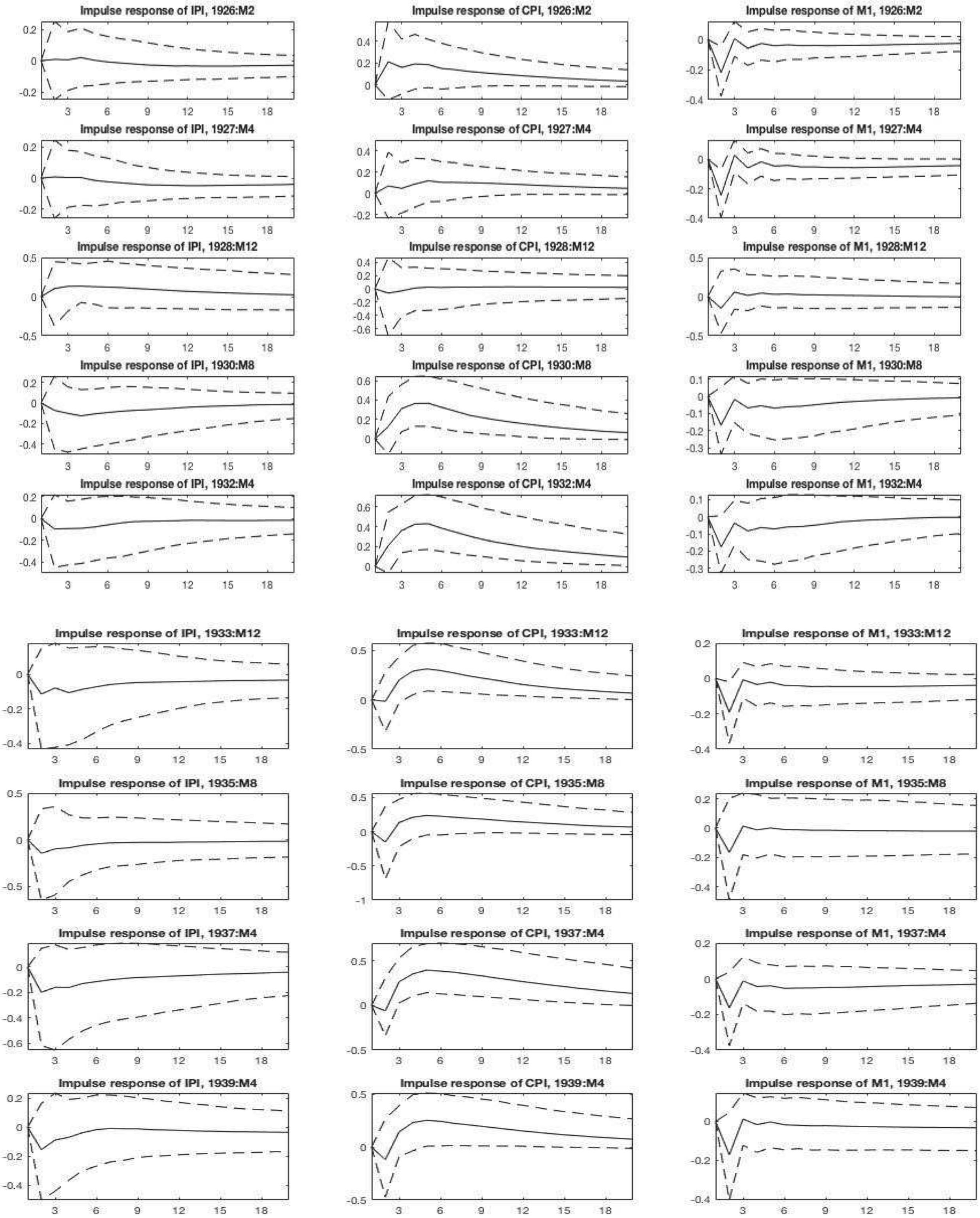

Figure 1.5.3– Impulse responses to an C-D shock. IPI, CPI and M1 in columns 1, 2 and 3 respectively. Note: The solid lines depict the 50-th percentile with the 16-th and 84-th percentiles for the dashed lines.

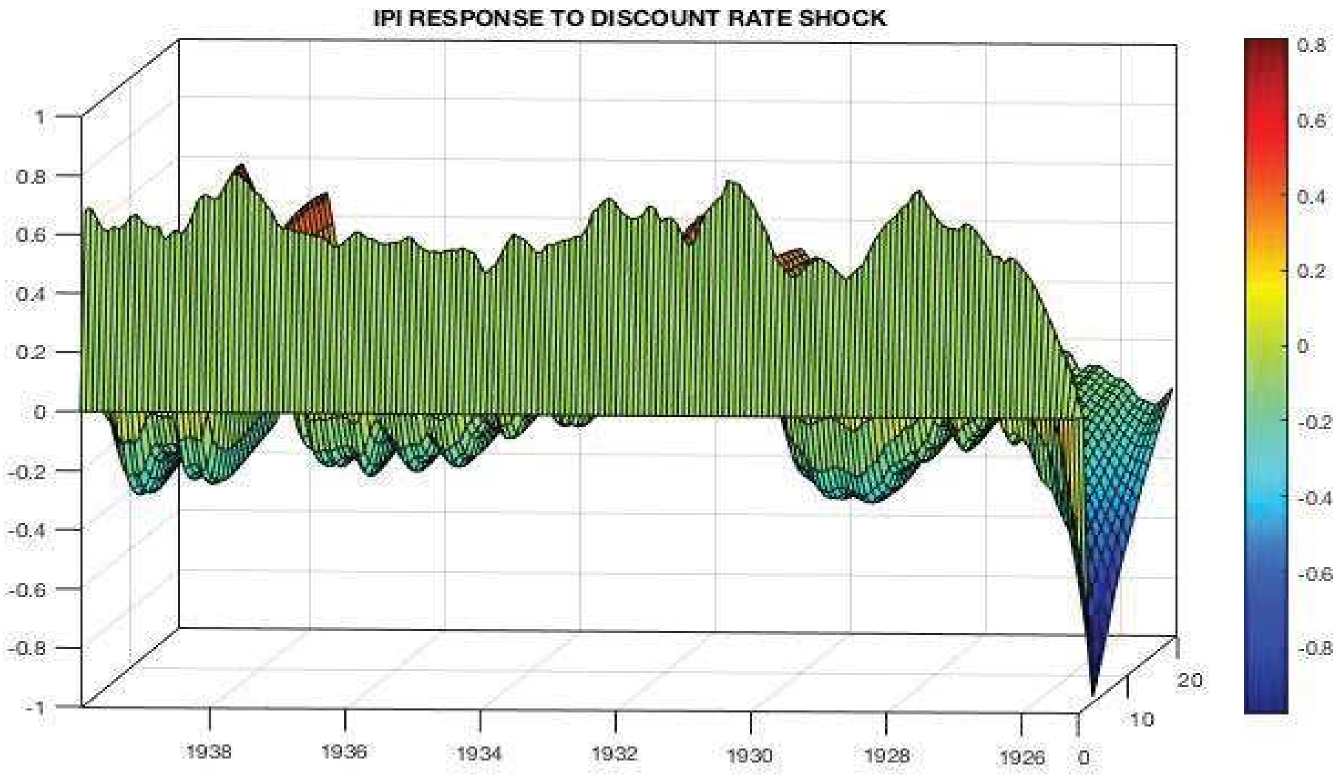

Figure 1.4.4– IPI impulse response to a discount rate shock. Note: Posterior means.

During those years, the spread was zero or negative, but both interest rates were below inflation most of the time, which did not contribute to reducing it. For this shock, the response is significant, approximately, for the period 1930-1934 and again around 1937 (Figure 1.5.3). In both cases, the spread was negative and there was deflation. Therefore, a lower discount rate in relation the short-term rate would have increased inflation. The responses are significant from the second or forth month and last beyond twenty months.

Focusing now on the responses of the IPI, a shock to the discount rate (Figure 1.4.4) produces quite a homogeneous response for the entire period. The response is initially positive but becomes negative after two months. The only difference occurs from 1930 to 1932, when the response never becomes negative. Gold inflows and open market purchases could be the reason, along with the great decrease in the discount rate. In general, either the transmission mechanism of the discount rate towards the output needs more time to materialize or the use of two lags with monthly data may not be enough to capture the real effect. Furthermore, its impact is not significant for the entire period (Figure 1.5.1). A shock to OMO (Figure 1.4.5) has, mostly, a positive impact. There is an exception from 1926 to 1930, when the initial response is negative for two months before turning positive. This figure has a similar pattern to the response of inflation to an OMO shock (Figure 1.4.2).

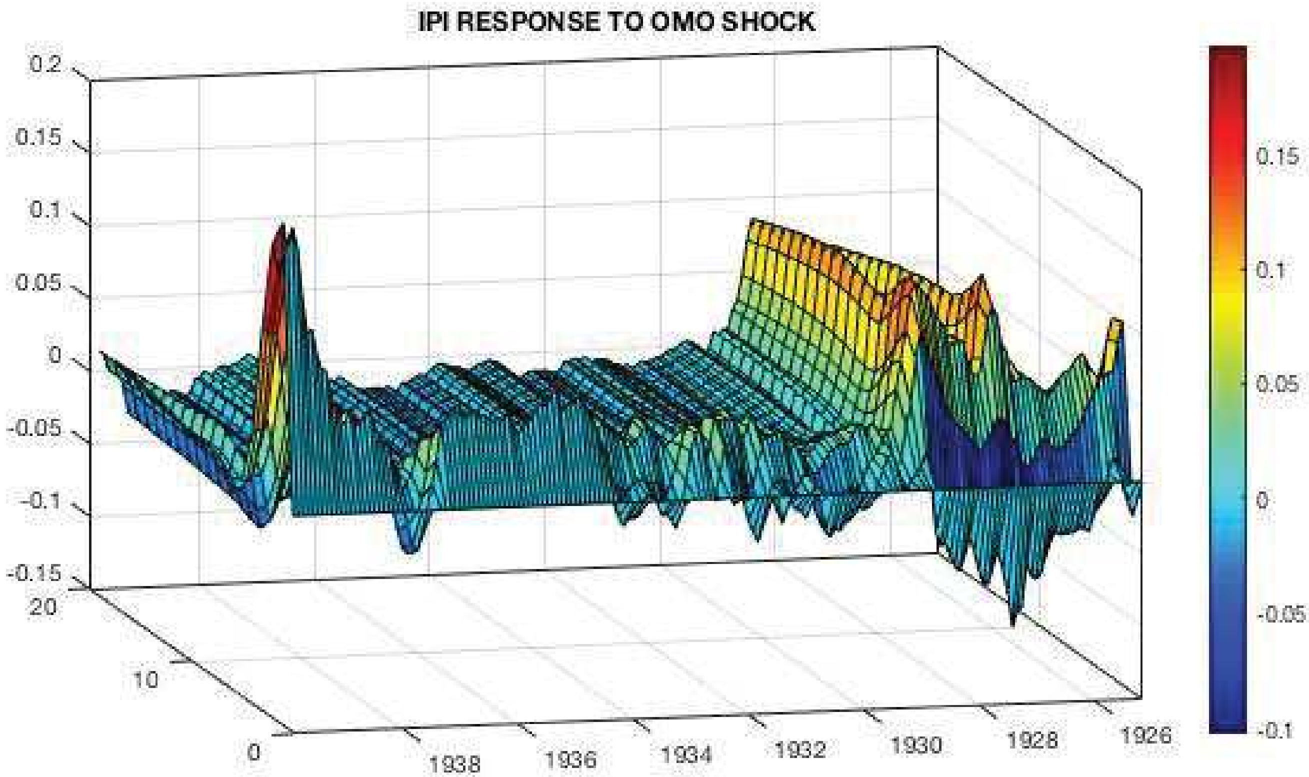

Figure 1.4.5– IPI impulse response to an OMO shock. Note: Posterior means.

Regarding the positive peak seen at the end of 1928, it seems that the Fed purchased more in the open market after the crash, but only for some months, because the response becomes less positive for the following periods. This is in line with Chapter 1. Once more, its effect is not significant (Figure 1.5.2). Regarding a shock to C-D (Figure 1.4.6), the response is mostly positive until 1929. For parts of 1926 and 1927, when the spread was almost zero, the response is slightly negative, although only initially. Afterwards, when the largest spreads are observed, the response is positive again. This positive response also occurs during periods of deflation, meaning that the spread could have contributed to increasing lending and growth, despite the sterilization of gold. After 1930, the response is negative, when the discount rate started to be at the same level or above the call loans rate (Figure 1.3). Again, the responses are not significant at any period (Figure 1.5.3). The residuals presented in Figures 1.6.1 and 1.6.2 are in line with the narrative described in Chapter 1. For the equations of the final targets, namely the IPI and CPI, the residuals are higher during the period in which the Fed was relegated to the backseat, after 1933, suggesting that other variables, which are not the Fed's instruments included in the model (as they were hardly used), could be driving the results of that period. Such factors could be fiscal policies, gold flows or the devaluation of the dollar in 1934. For the CPI equation, the residuals are also higher from 1925 to 1929 (Figure 1.6.1), likely corresponding to gold flows. For the residuals of M1, there is a peak at the end of 1929, probably related to the crash, bank failures and holdings of currency, and between the end of 1932 and the beginning of 1933, when more bank failures occurred (Figure 1.6.1). Apart from those peaks, the residuals are constant for the entire period. For the C-D equation (Figure 1.6.2), the increase in the residuals appears between 1928 and 1930, when the difference between rates was the largest and the Fed was unable to reduce the call loans rate because other institutions were giving credit. Regarding the OMO and discount rate equations, the residuals behavior is similar to the C-D equation, but they start to decrease in 1932 (Figure 1.6.2). That means that the Fed was targeting other variables beyond those included in the model while in charge of monetary policy. According to the narrative, the Fed may have been responding to gold flows or bank reserves. Afterwards, the residuals tend to zero, as the instruments were hardly used. These results are in line with the lack of significance seen in most of the impulses response.

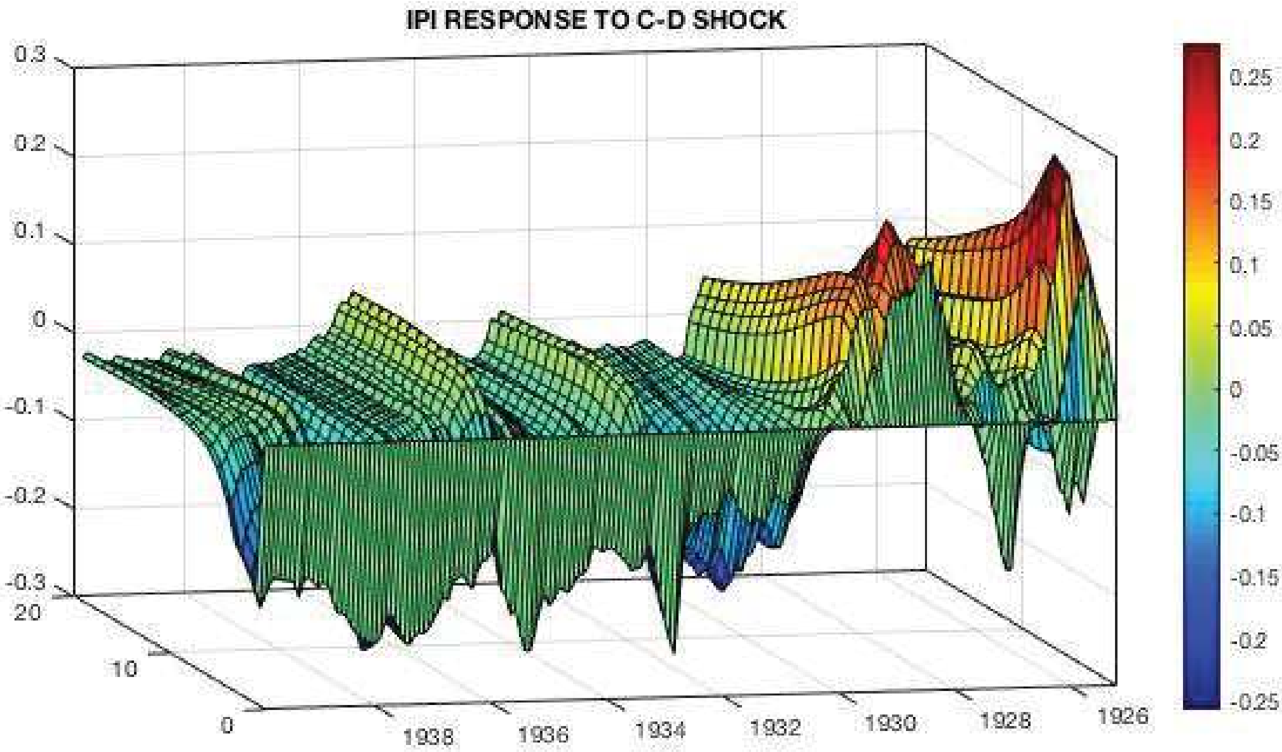

Figure 1.4.6– IPI impulse response to C-D shock. Note: Posterior means.

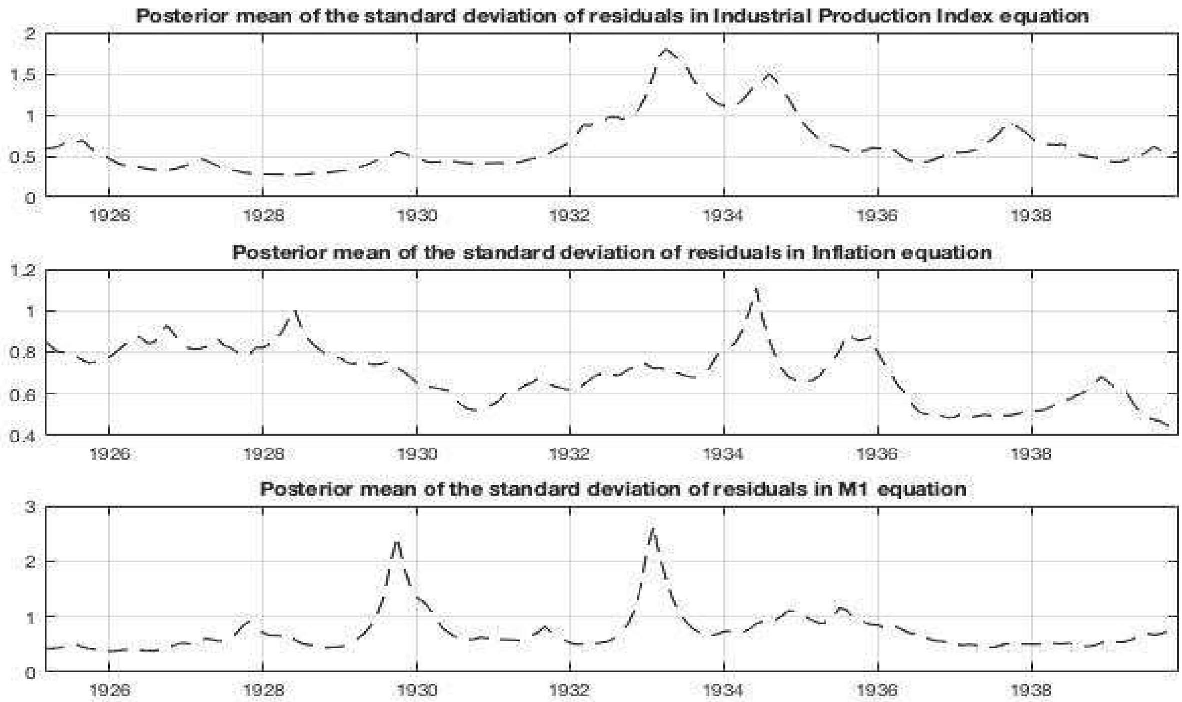

Figure 1.6.1– Posterior mean of the standard deviation of the residuals in IPI, CPI inflation and M1 equations respectively.

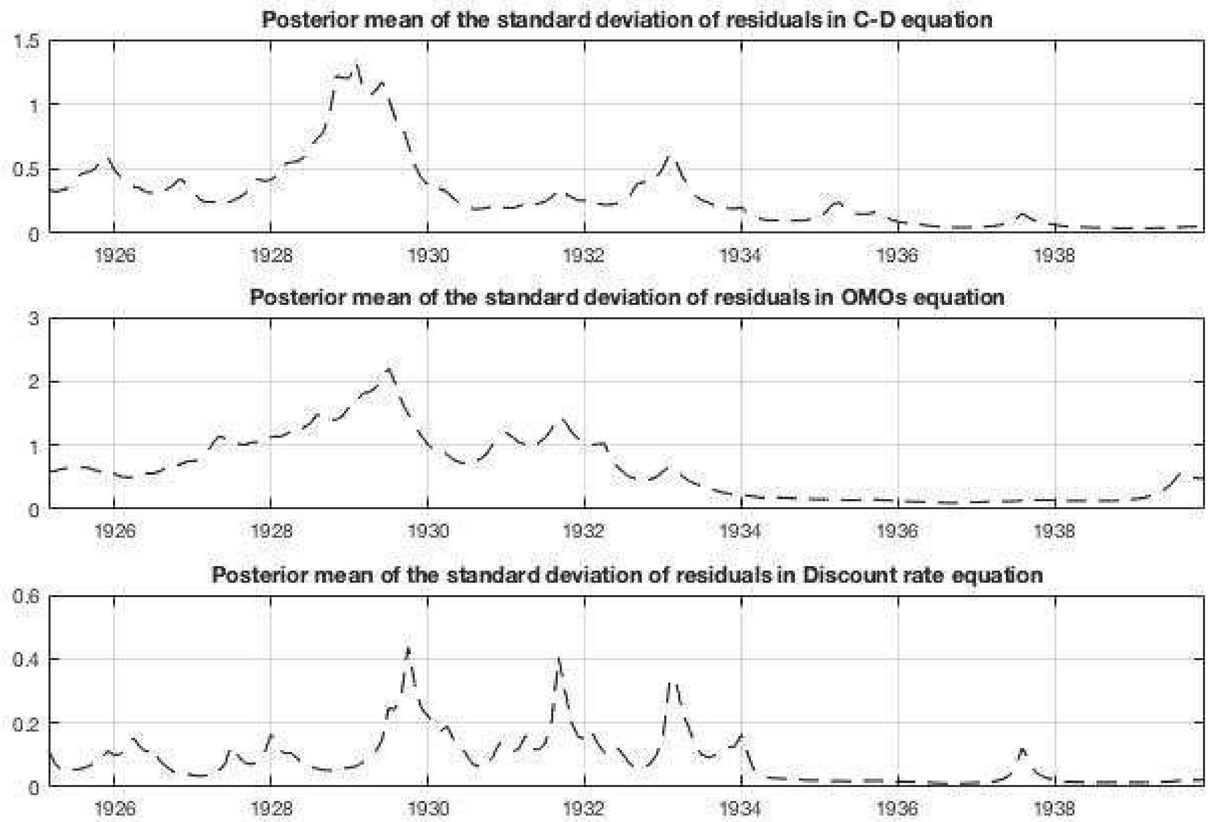

Figure 1.6.2– Posterior mean of the standard deviation of the residuals in C-D, OMO and discount rate equations respectively.

2.4.2 1958:I–2007:IV[13]

For the impulse response function analysis of this second period, I will only analyze the results obtained with multi-move algorithm with one lag. The results obtained with K-P's algorithm with two lags are very similar. While the figures are not presented, I will comment the relevant variations.

This time, I start by analyzing the response of inflation to an F-D shock (Figure 1.7.1), as it provides the perfect beginning for the explanation of the next impulse responses. Until 1968, the response of inflation is almost zero, coinciding with the period when the spread between the rates was zero or negative. However, that response is not significant (Figure 1.8.1). From 1968 to 1982, the response of inflation practically mimics the evolution of the spread (Figure 1.1). Thus, when it becomes larger, inflation is higher. Those positive peaks are reversed once the spread is zero or negative, and after the federal funds rate has visited maximum levels.

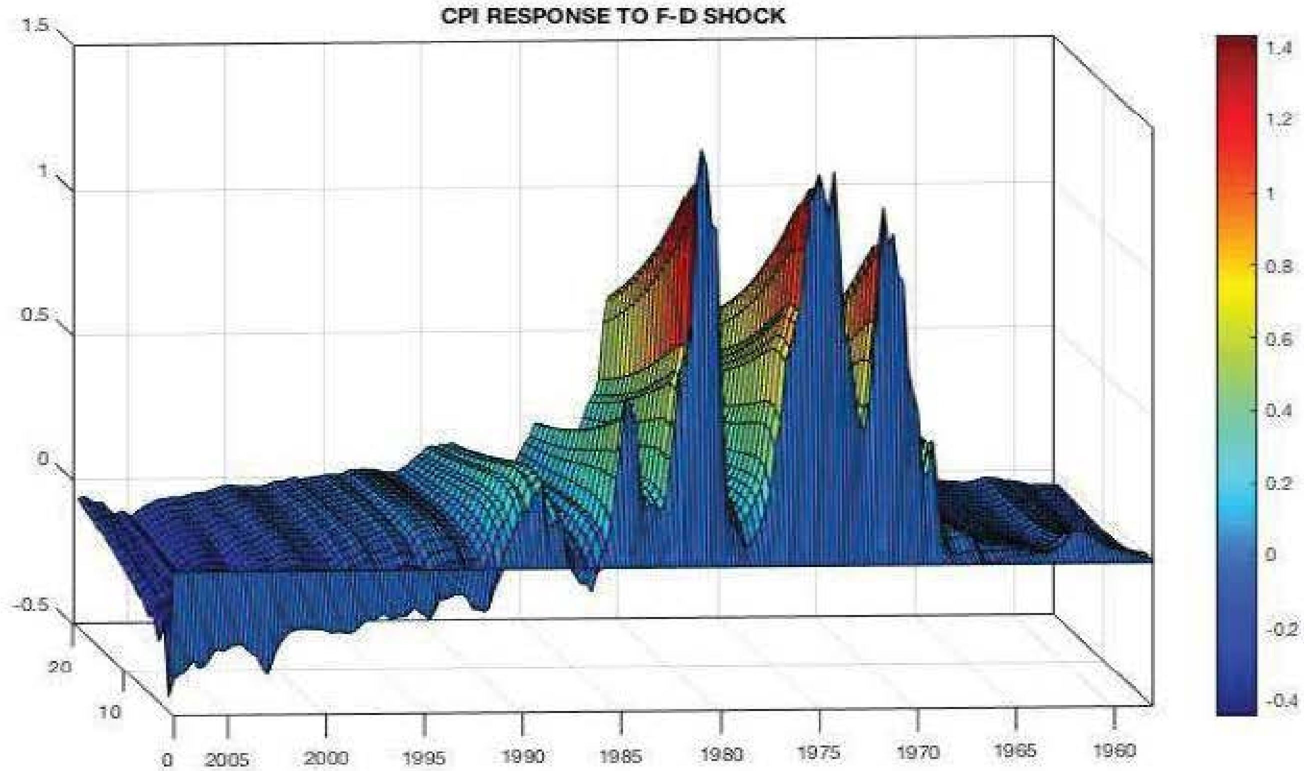

Figure 1.7.1– CPI Inflation Impulse Response to F-D Shock. Note: Posterior means.

[13] The historical context is in Chapter 1 and the remaining figures in Appendix A.2.

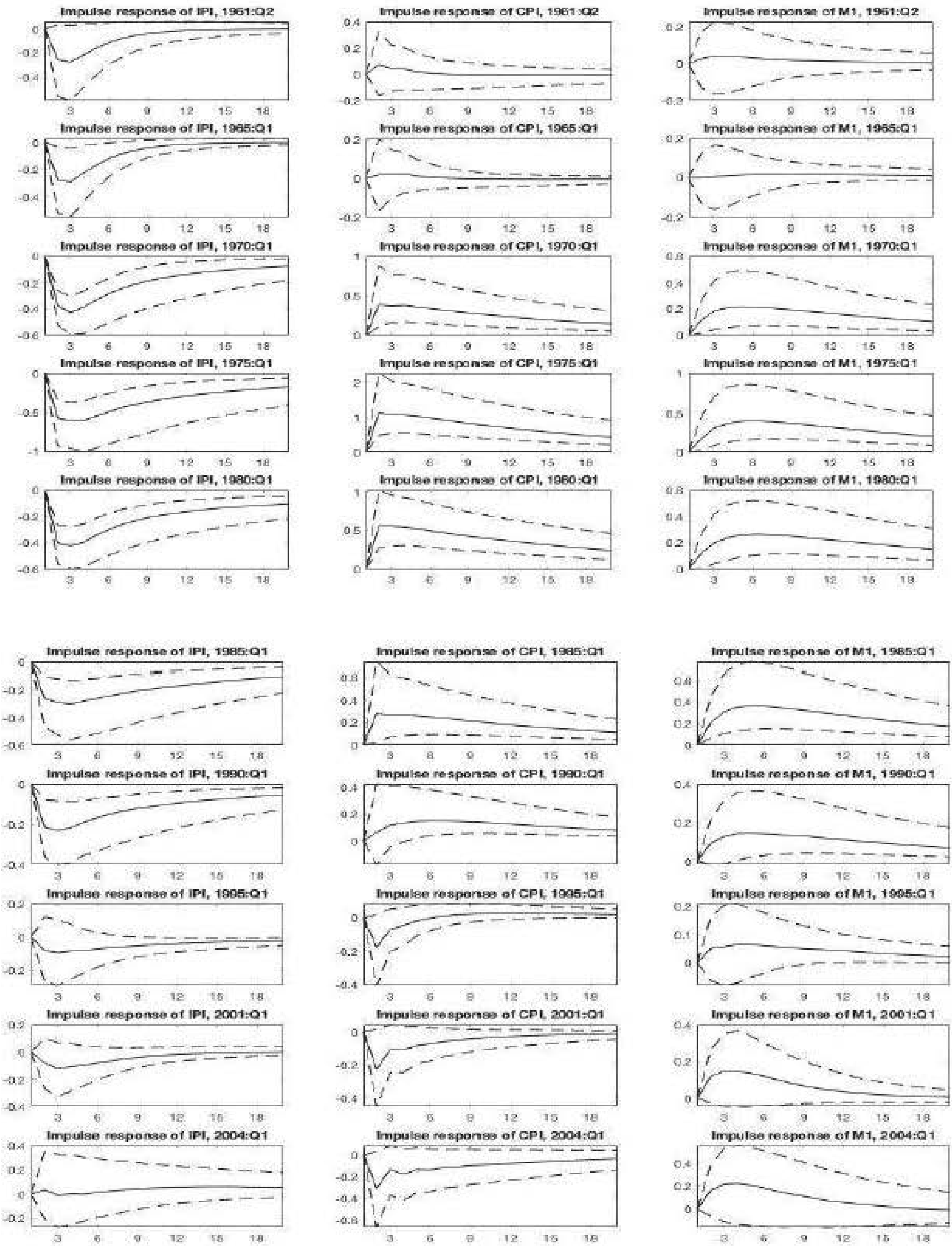

Figure 1.8.1– Impulse responses to an F-D shock. IPI, CPI and M1 in columns 1, 2 and 3 respectively. Note: The solid lines depict the 50-th percentile with the 16-th and 84-th percentiles for the dashed lines.

However, the responses are still positive, as inflation is above both rates. From 1982 to 1990, the pattern is the same. However, despite observing even larger spreads, inflation responses are not as positive as in the 1970s or early 1980s. For these periods of positive spreads, Figure 1.8.1 shows that the spread has a significant and long-lasting impact on inflation. Shortly after 1990, around mid-1992, the response of inflation turns negative until the end of the sample. It could be expected that, at least in 2001 and 2002, or even 2003, the response of inflation would be positive given that inflation is above the discount and federal funds rate, or from 1994 to 2000, when, despite not being large, the spread is positive. However, this does not happen and coincides with the decrease in borrowing observed in Figure 1.2. Besides, the responses are not significant for this last period (Figure 1.8.1).

Encouraged by this regime change, I discovered some literature that sheds light on it. In the Federal Reserve Bulletin of November 1994, Clouse (p. 965), apart from supporting the fact that a larger spread led to higher borrowing and that the relationship was quite stable until 1980, given the failing bank situation during the 1980s and 1990s, stated:

> … changes became evident during the 1980s in the willingness of healthy institutions to turn to the discount window. Many banks apparently became more reluctant to turn to the window for fear of provoking market concerns about their financial condition. The greater reluctance to borrow weakened the historical relationship between the discount borrowing and the spread of the federal funds rate over the discount rate.

Furthermore, "This reluctance became acute during the economic downturn in the 1990-1991 …" (Clouse 1990, p. 969). Kasriel and Merris (1982) added that the Fed, before 1979, did not pay attention to borrowed reserves and the relation between the discount window and the federal funds rate. Pearce (1993) showed that the relationship between borrowing and the spread changed under different target regimes. Thus, from January 1975 to October 1979, under a federal funds rate target, there was a strong nonlinear relationship between the spread and borrowing, whereby a larger spread led to higher borrowing, although to a certain extent.[14] From October

[14] Peristani (1991) found an S-shaped pattern when analyzing the period 1959–1988. Thus, although the spread led to higher borrowing, when the difference was around 4%, borrowing hardly, if at all, increased.

1979 to October 1982, under a nonborrowed reserve targeting procedure and lagged reserve accounting, the relationship weakened and the amount of borrowing decreased. Last, after October 1982, under a borrowed reserve targeting procedure and contemporaneous reserve accounting, the relationship was even weaker and the borrowing decline more pronounced. Therefore, it seems that this negative no significant response and regime switch indicate that some factor related to the banking sector triggered that positive spreads did not increase inflation. Figure 1.7.1 also shows that, as commented in the introduction, the "price puzzle" is non-existent, because inflation increased with rises in the federal funds rate, mostly when the spread was positive and borrowing increased. Therefore, there is no puzzle but an inadequate Fed's policy by allowing those positive spreads. The response of inflation to an OMO shock (Figure 1.7.2) is negative until 1967, with the spread being generally negative or zero. Thus, it seems that while the Fed did not target short-term rates, purchases in the open markets were scarce to boost inflation. Since then, under an interest target and in line with the figure analyzed previously, positive peaks occur for those periods when the spread was positive. This means that, despite banks already were borrowing, taking advantage of that spread, the Fed purchased (although perhaps in a relatively smaller proportion than when the spread was zero or negative) in the open market,

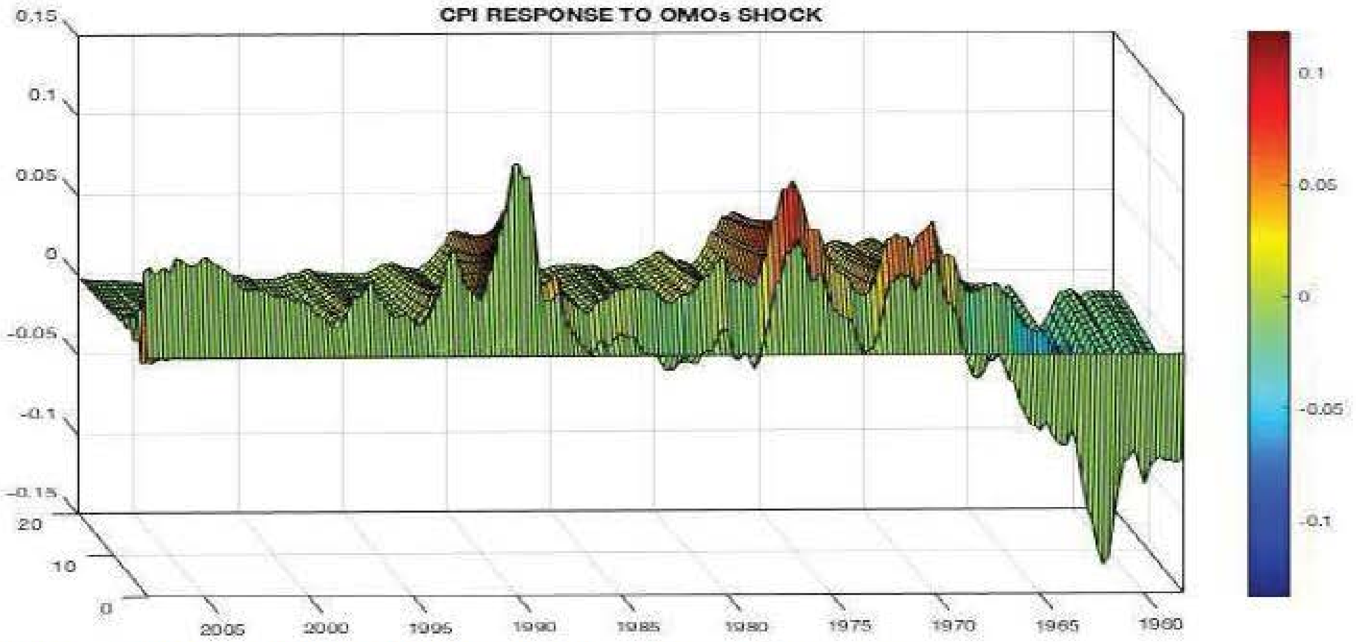

Figure 1.7.2- CPI Inflation Impulse Response to an OMO Shock. Note: Posterior means.

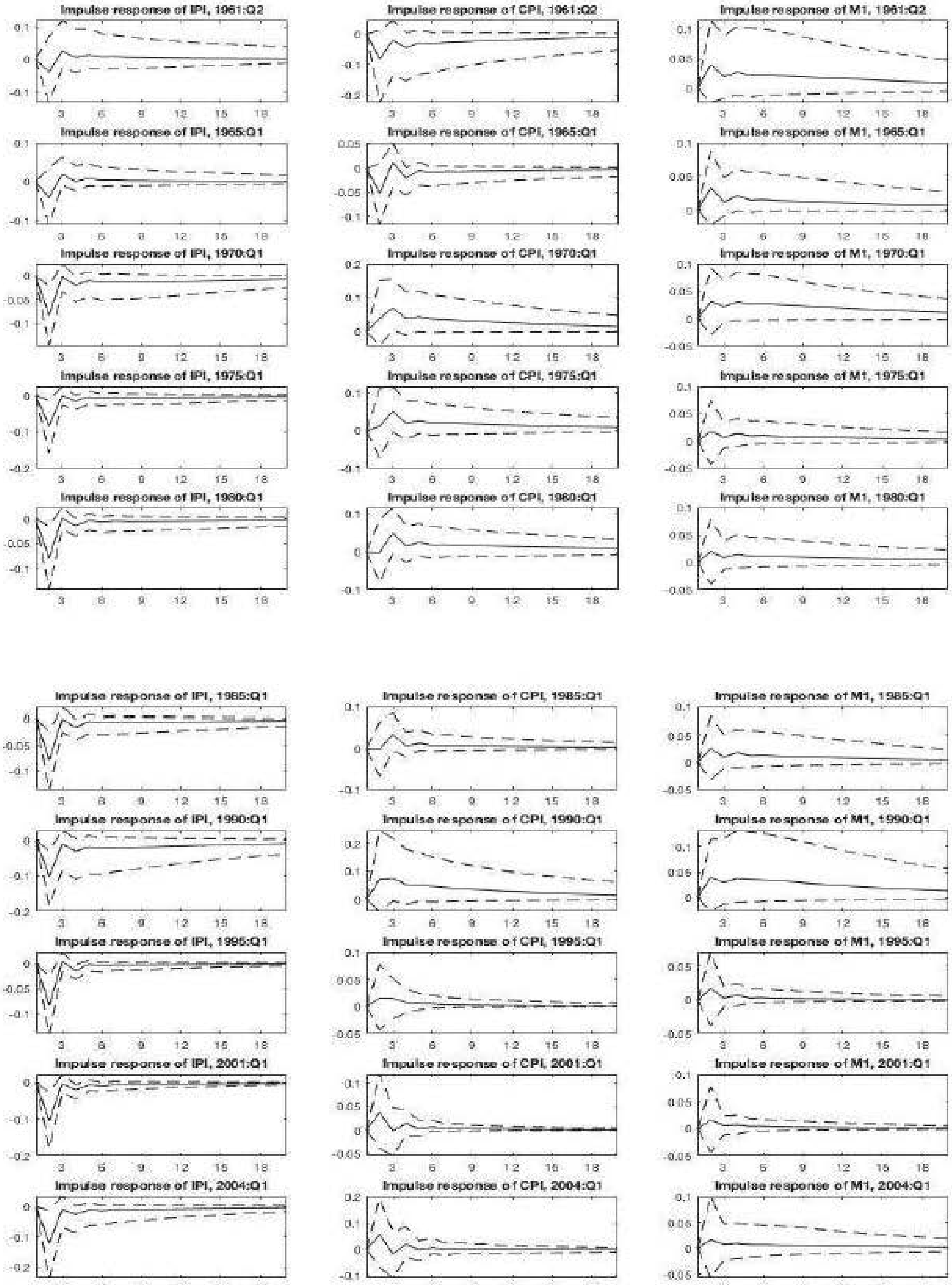

Figure 1.8.2 – Impulse responses to an OMO shock. IPI, CPI and M1 in columns 1, 2 and 3 respectively. Note: The solid lines depict the 50-th percentile with the 16-th and 84-th percentiles for the dashed lines.

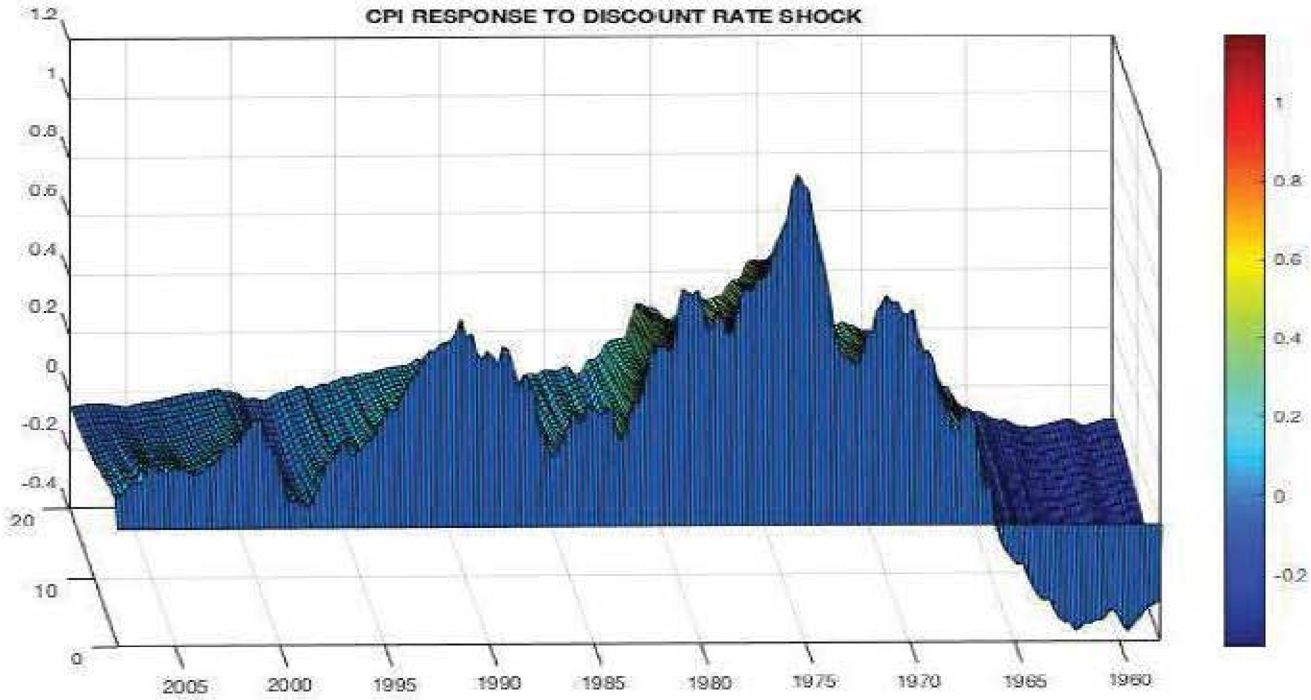

Figure 1.7.3– CPI inflation impulse response to a discount rate shock. Note: Posterior means.

contributing even more to increasing inflation. Nonetheless, the responses are not significant for the entire period (Figure 1.8.2). The results above support the response of inflation to a discount rate shock (Figure 1.7.3), which is not expected. Its response is negative until 1966, a period when the discount rate is either above or equal to the federal funds rate. For the first row in Figure 1.8.3 (1961:Q2), the response is slightly significant. Afterwards, the response becomes positive either because, when the spread was narrow or negative, the Fed offset the rises in the discount rate by increasing its purchases in the open market (although in those cases the positive response was reduced) or because, when the spread was positive, the Fed exacerbated the amount of borrowing at the discount window with more purchases. Since positive spreads emerged, the discount rate has a positive and significant impact on inflation, except for the last years of the sample, when the response is not significant (Figure 1.8.3). It is important to highlight that since 2003, the discount rate was set above the federal funds rate as a penalty rate, but it allowed borrowing with no question asked.

The response of the IPI to an F-D shock (Figure 1.7.4) is, for almost the entire period, negative and almost proportional to the levels of the federal funds rate. The lowest peaks coincide with the highest federal funds rates. Almost similar to inflation, the response is

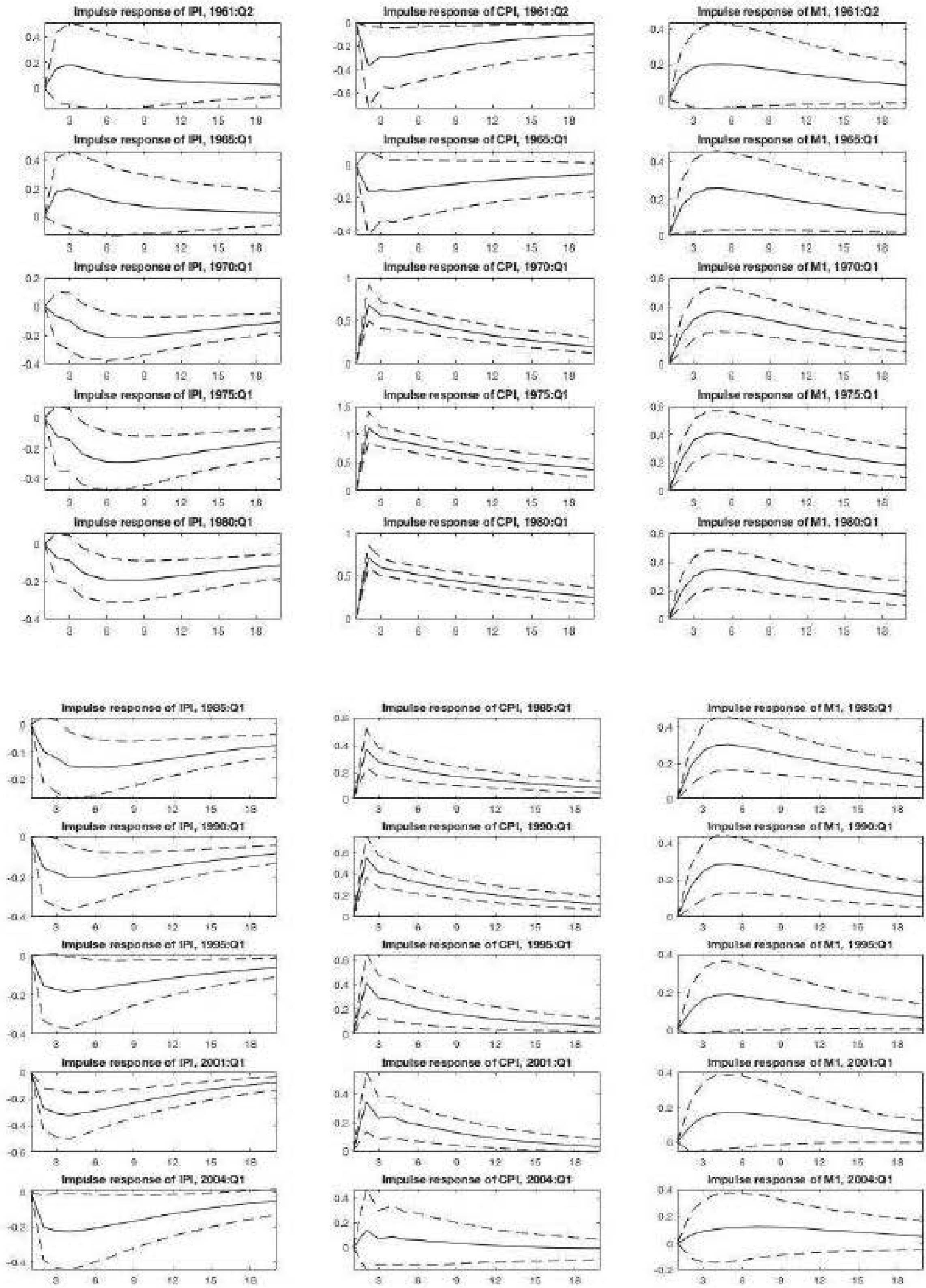

Figure 1.8.3 – Impulse responses to a discount rate shock. IPI, CPI and M1 in columns 1, 2 and 3 respectively. Note: The solid lines depict the 50-th percentile with the 16-th and 84-th percentiles for the dashed lines.

significant from the mid-1960s to the mid-1990s, with an horizon of two years or more (Figure 1.8.1). For the discount rate shock (Figure 1.7.5), the response of the IPI is positive until 1965 and then becomes negative. Since then, as in the last figure, the negative peaks occur around the highest peaks of the federal funds rate. For this case, however, the impact is not proportional since the beginning of Volcker's era. After approximately 1970, the responses are significant, although most of them just after some quarters (Figure 1.8.3).

Regarding the response of the IPI to an OMO shock (Figure 1.7.6), it is negative with the exception at the beginning of the sample, when it becomes positive after two quarters. Similar to the last figure, the responses are mostly significant after 1970, although this time they are weak and only for one quarter (Figure 1.8.2). This response is not expected. As explained for Figure A.7.11 in Appendix A.2, the transmission between the two variables may need more lags, because for those periods when lending was decreasing as a consequence of the high rates, what also led to a decline in output, the Fed would have been purchasing more securities to decrease interest rates.

To sum up, a regime change is observed around 1965, when the Fed began to pay attention and target short-term rates. The other regime change, although represented not in the instruments' impulse responses but in the spread, is around 1990. Although the figures of the instruments mostly maintain their sign and shape since 1965, this happens under different

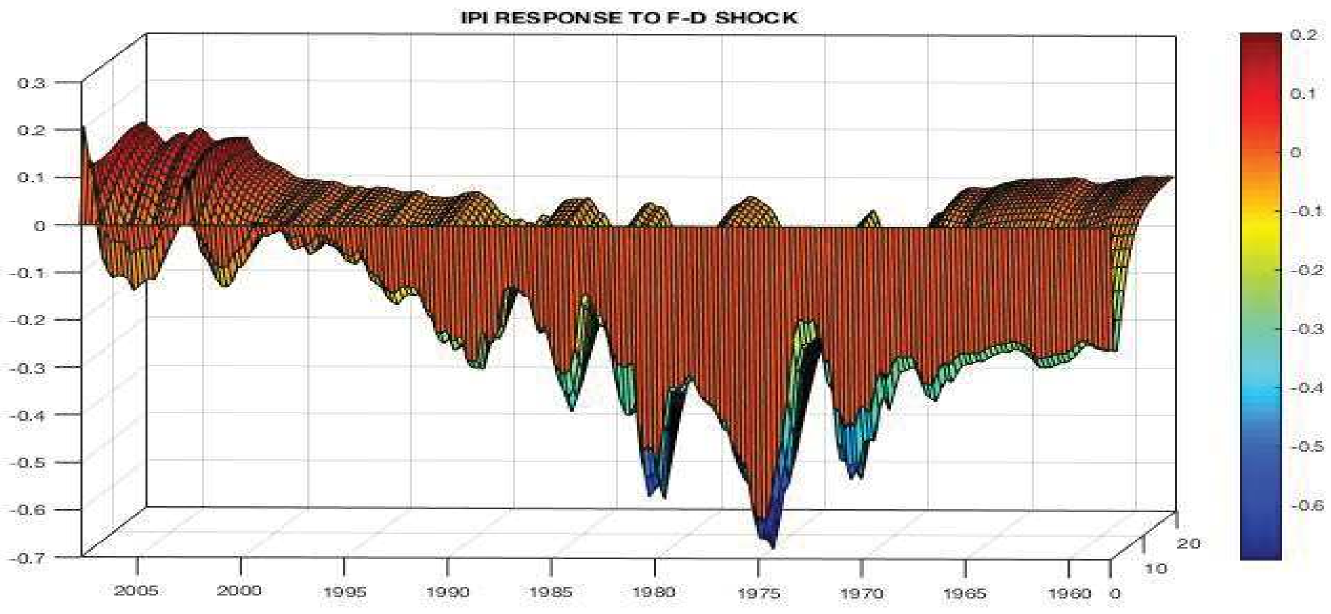

Figure 1.7.4– IPI impulse response to a F-D shock. Note: Posterior means.

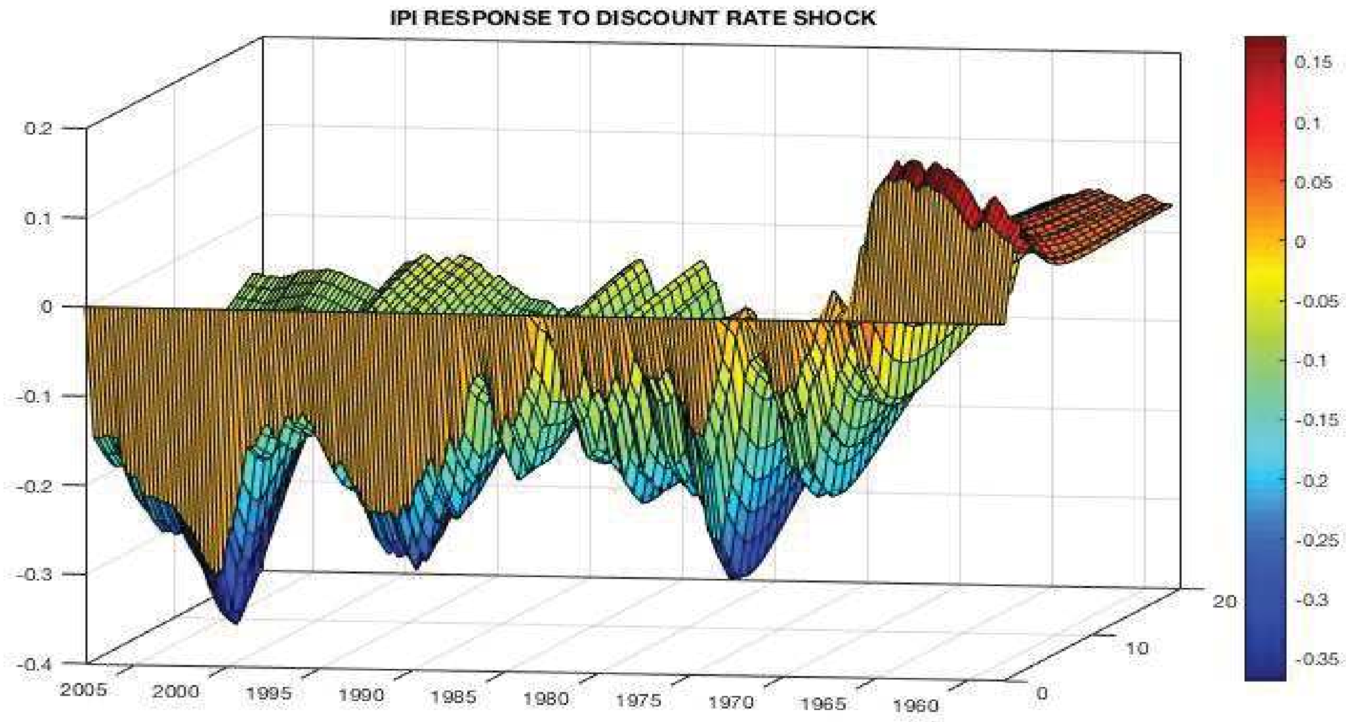

Figure 1.7.5– IPI impulse response to a discount rate shock. Note: Posterior means.

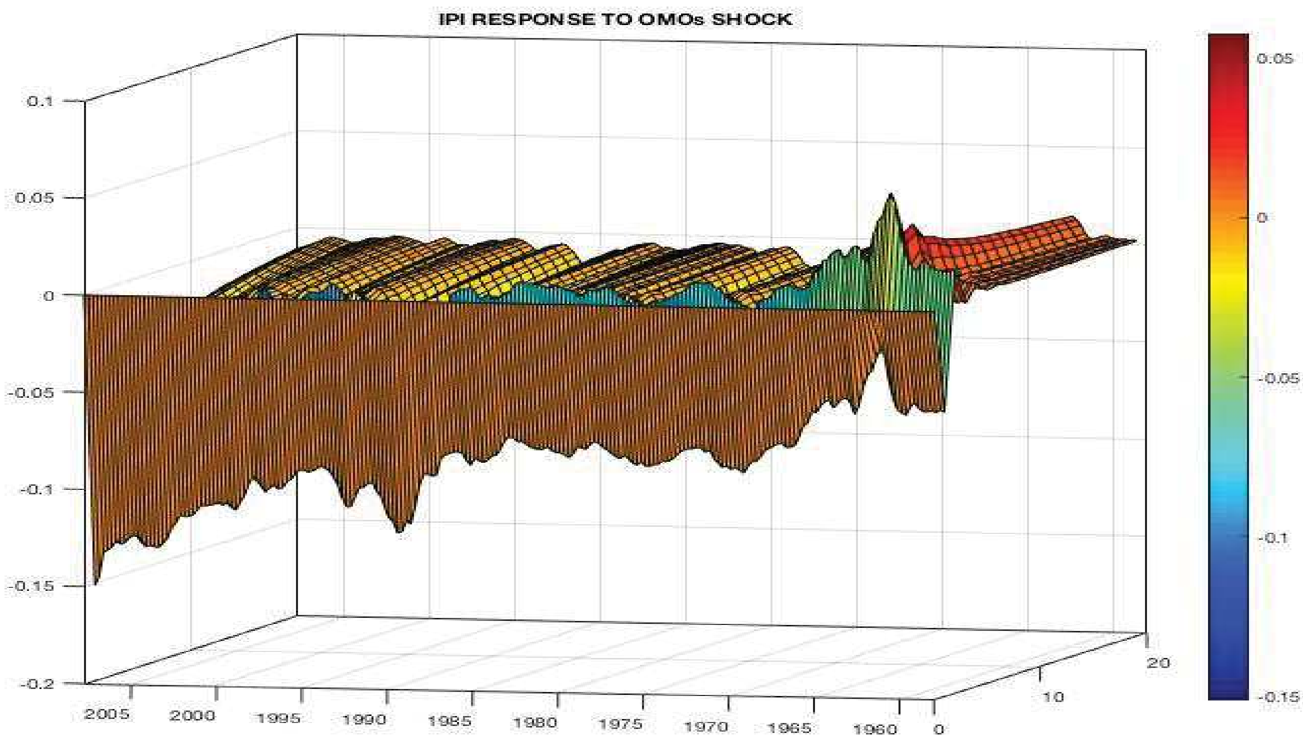

Figure 1.7.6– IPI impulse response to an OMO shock. Note: Posterior means.

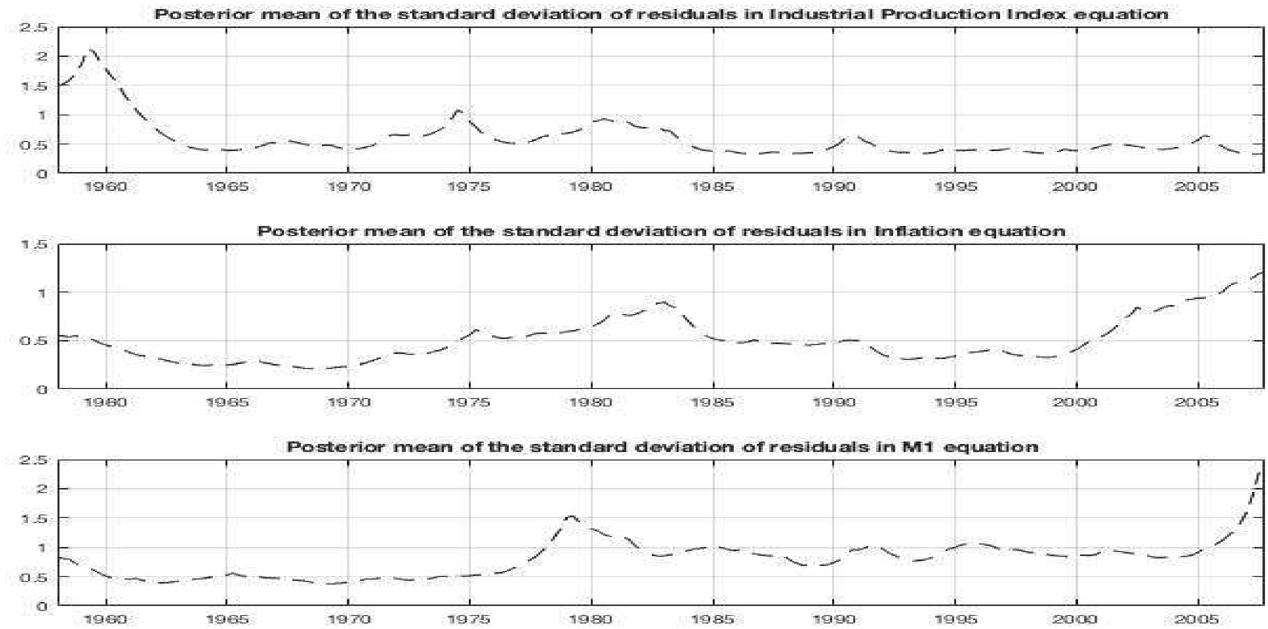

Figure 1.9.1– Posterior mean of the standard deviation of the residuals in IPI, CPI inflation and M1 equations respectively.

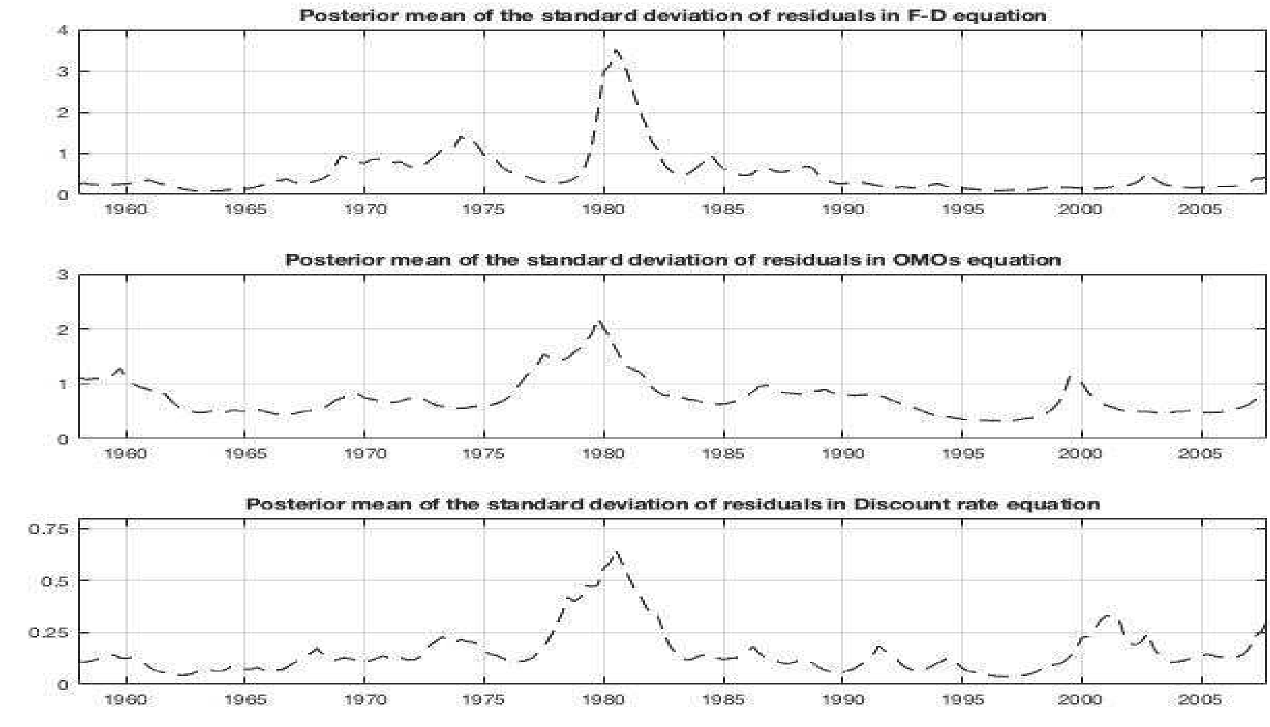

Figure 1.9.2– Posterior mean of the standard deviation of the residuals in F-D, OMO and discount rate equations respectively.

circumstances. Therefore, a change in the banking sector's behavior seems to be the most likely explanation. This statement is supported by the counterfactual exercise developed in Appendix B, where I find no variation in how the Federal Reserve applied its policies for the period under analysis, and where inflation expectations are also taken into account.

For the residuals of the IPI equation (Figure 1.9.1), an increase is observed at the beginning of the sample, which quickly disappears. Thereafter, the residuals maintain roughly the same levels, being slightly higher in the 1970s and the first half of the 1980s. For inflation (Figure 1.9.1), the residuals increase in the 1970s and at the beginning of the 1980s. This is a period when money growth or some type of reserve measure was supposedly targeted. The residuals decrease until 2000, to increase again until the end of the sample. For M1 (Figure 1.9.1), the residuals increase in the second half of the 1970s, remain at roughly the same levels until 2005 and then rise again. While, for the IPI behavior, external shocks could have determined the results at the beginning of the sample and slightly during the 1970s and part of the 1980s, the residuals of the inflation equation show that external shocks could affect the results in the 1970s and since 2000. The same is applicable to M1. Bank reserves and banking sector's behavior, expectations, fiscal policies, exchange rate or political pressures are factors that are not included the model and could have influenced the results. For the equations of the Fed's instruments (Figure 1.9.2), the 1970s also seem to be a period when external shocks could have influenced their behavior, as the Fed may have been aiming other targets. The same applies to the small peak around 2000, and the rise since 2006.

2.4.3 Alternative identifications

For both periods, alternative orders were tested by locating OMO the first and third in the VAR. While most of the responses to an OMO shock change and remain no significant, the relationships between the other variables stay almost identical. Another identification scheme was tried allowing the instruments and the spread to react contemporaneously to each other and imposing zeros in some of the relationships between OMO and the spread with M1 and the final targets (so that the decomposition of the variance covariance matrix was no longer lower triangular), using the algorithm developed in Canova and Forero (2014). However, none of the draws overcame the stationary restriction and the exercise could not be carried out.

2.5 What lessons can we learn?

The picture of monetary policy in the U.S. for the last century shows some recognizable patterns in the instruments of monetary policy and intermediate and final targets. Before analyzing the instruments individually, from a general perspective according to the results, the Fed's instruments were not the main drivers of the American economy's performance for the interwar period, while for good or for bad, they gained relevance for the second period. The lack of significance obtained for the interwar period is supported with some of the facts described in Chapter 1, such as the Fed's inability to differentiate between nominal and real interest rates, the lack of experience operating in the open market or the period when the Treasury took the responsibility for monetary policy. Also, as commented in the literature review, it seems that gold flows had an important role in determining the path of the economy.

Focusing now on the instruments, I start by analyzing the discount rate. Its impact on inflation varies according to two scenarios. First, when it is above the short-term rate and above the inflation levels, its impact on inflation is negative. There is an exception for the last years of the sample, when its impact is positive but not significant, coinciding with the increase in the residuals of the inflation equation. Therefore, other factors could be distorting that relationship for those years. The second scenario occurs when the discount rate is below the short-term rate, the inflation levels or both. In this case, inflation responds positively. While these patterns were also found for the interwar period, the results displayed for the residuals and the narrative analysis, seem to explain why those impulses responses are not significant for that period. For the IPI, when the discount rate impact has a significant response (what happens only after 1965), its sign is negative. From 1958 to 1965 the response is positive as in the interwar period, although in this last case, the sign is negative after some months. In any case, those responses are not significant. The conclusion here is that while no other factor is influencing the relationship discount rate-output, increases in the discount rate should decrease output. Last, the M1 response is different for each period and could be related to the different Fed's procedure about OMO and targets. One the one hand, for the interwar period (negative and significant response) when purchases in the open market were scarce, the discount rate had a more important impact on short-term rates and therefore, on M1. The reason is that short-term rates, conditioned on the discount rate but not manipulated by the supply of reserves with OMO as under an interest target, responded directly to demand forces, influencing the path of M1. On the other hand, around 1965 (positive significant response) the Fed began to target short-term rates, what required a more intensive use of

OMO. As the raises in the interest rates intended to reach the Fed's target were insufficient to restrain the demand for credit and inflation continued increasing, to keep the federal funds rate under its target, more purchases were needed and the money supply increased along with the discount rate. Despite the more "natural" relationship between interest rates and the demand forces for the interwar period, unlike under an interest rate target, the use of the discount rate, implicitly and unwittingly, conditioned short-term rates anyways. This mechanism makes both periods comparable despite the different targets and short-term rates.

Evaluating the spread, it is the only one having a significant response for both periods in relation to the inflation levels. A shock to this variable shows different responses depending on the sign of the spread, whether both rates are below or above inflation levels and whether those levels are positive or negative. When the spread is positive or both rates are below the inflation levels, the response is positive. For the interwar period, those two scenarios had a positive but no significant impact on inflation. However, the response was significant and positive as well, when the spread was negative for periods of deflation. This suggests that decreases in the discount rate in relation to short-term rates would have contributed to increasing inflation. The reason why the positive spread had a no significant inflation response for the interwar period is probably due to the fact that real interest rates were too high given the levels of deflation, which were caused by other factors such as gold sterilization. For the second period, the analysis shows that positive spreads above inflation levels contributed to increasing inflation. Moreover, it is likely that the banking sector changed its behavior around 1990, because the positive and significant responses turned negative and no significant even for periods with positive spreads. The different Fed's procedures explained previously, together with the positive and significant responses of inflation when the spread was positive, are in line with the fact that M1 responded positively and significantly to positive spreads, despite it implied higher rates. Regarding the IPI, increases in the spreads are negative and significant only for the period 1965-1990, unlike the discount rate case, in which that significant and negative impact was extended to the end of the sample. Therefore, the spread itself seems to not decrease output, and the high inflation levels related to that spread could be the cause of that negative impact.

Last, although OMO has a significant impact on the IPI for some periods after 1965, they are ephemeral and weak. Therefore, the results suggest that monetary policy can be transmitted through prices but not quantities. That is, even though the amount of money supplied will drive short-term rates in the first submarket, only interest rates will determine the demand for money in the second submarket.

These lessons also provide an explanation for the price puzzle, a problem that has occupied the literature for years. Apparently, a shock to the federal funds rate produces, at least initially, a positive inflation response. It has been argued that missing variables and the consequent lack of information in VARs produce the so-called “price puzzle.” According to the results obtained for the second period (Figure 1.7.1), most of the increases in the federal funds rate were accompanied by large spreads. This led to an increase in money growth and inflation because banks could borrow cheaper reserves at the discount window and set a relatively lower loans rate. This would trigger a lower restrain in credit than the intended by the Fed. Thus, there is not a price puzzle but a real positive relation between the increase in the federal funds rate and inflation as a consequence of the bad policies that allowed positive spreads.

Last, Primicery (2005) used unemployment, inflation and the federal funds rate in his model and concluded that the change in policies did not differ between the pre- and post-Volcker periods, and no regime switch was observed. According to the counterfactual, I reach the same conclusion regarding the change in policies for those periods; however, I observe two regime switches. The first one is around 1965 in OMO and the discount rate. The second regime switch appears in the F-D spread around 1990. In the latter case, it seems probable that the responsibility could belong to the banking sector. Apart from that, while Primicery obtained standard deviations of the residuals of the CPI inflation equation three times higher in the 1970s than at the beginning of the sample, here they are smoother, increasing in 1975 by only half of the levels seen in the 1960s. Therefore, it seems that the inclusion of the spread along with the instruments explains a great part of the inflation behavior.

2.6 Conclusions

A re-evaluation of the mechanisms operating between the Federal Reserve’s policies and its intermediate and final targets is undertaken in this chapter, from nearly the birth of the Federal Reserve to the period before the Great Recession. For that purpose, however, the standard procedure to evaluate monetary policy is questioned and declared erroneous, given the measuring problems associated to the use of intermediate targets as Fed’s instrument, and a new procedure is proposed. This new procedure uses the actual instruments and the spread between short-term rates and the discount rate. Thus, a TVC-BSVAR was performed, applying the algorithms already used by Primiceri (2005) and Koop and Potter (2011). To gain a better understanding of the results obtained, they were contrasted with a narrative review of the Federal Reserve's history. Summing up, the Fed's lack of knowledge and inactivity for the interwar period, supported by the results, indicated that monetary policy was unable to influence output and inflation for that period. After 1958, monetary policy gained power to influence the path of the American economy. However, the management was inadequate and the results showed that increasing the federal funds rate is not enough to decrease inflation. The increase in the federal funds rate should be accompanied by discount rate increases, avoiding positive spreads. Otherwise, they will provide profitable opportunities for banks, triggering an increase in borrowing, and preventing enough restrain in the demand for loans. The mechanism behind these relationships is that as banks obtained cheaper reserves at the discount window, they were likely to raise the loans rate less in relation to the increases in the federal funds rate. Consequently, the demand for credit was not restrained as much as the Fed intended. These facts deny the existence of the "price puzzle," as the increase in inflation when the federal funds rate was raised was due to bad policies by allowing those positive spreads. In the case of aiming to reduce output, it seems sufficient to increase the discount rate. Last, the results suggest that the monetary policy transmission channel is effective only through prices, even though those prices are driven by quantities and other prices. That is, OMO can modify short-term rates, but in the end, the demand side will respond to the price at which money is supplied, namely, the loans rate influenced by the short-term rate and the discount rate. These results leave open questions for future research as a consequence of the regime changes observed around 1965 and 1990. What is the role of the banking sector in transmitting to the real economy the monetary policies undertaken by the Fed? Could bank determine the impact of monetary policy regardless of the Fed's intentions and to what extent?

-- Chapter 3 --

The transformer of monetary policy: The banking sector

A new theory is proposed, called "reserves-cost", which explains the mechanism whereby the Federal Reserve's policies reach the economy through the banking sector. The theory states that the Fed can have an impact on the real economy by directly influencing the cost of reserves, and indirectly, with the impact of that cost on the loans rate. To evaluate it, a time varying coefficients Bayesians SVAR is estimated for the period 1958-2007. The results support the new theory and show that when the spread between the federal funds rate and the discount rate was positive, banks could obtain cheaper reserves at the discount window and increase the loans rate less than the Federal Reserve raised the federal funds rate. The consequence was an insufficient restrain in the demand for loans, triggering more lending and higher inflation. This theory is also proposed as the explanation for the different inflation and output volatilities and levels witnessed during the Great Inflation and the Great Moderation.

- *JEL classification: E43, E51, E52, E58*
- *Keywords: monetary policy, Federal Reserve, Bayesians, SVARs, reserves cost*

3.1 Introduction

During the last decades, monetary policy has gained in relevance and attention as the instrument capable of achieving stability in, mostly, prices and output. Accordingly, central banks have evolved as institutions with an increasing responsibility for that task. To accomplish such goals, the successful use of central banks' instrument depends on the understanding of the channels through which monetary policy works. When certain policies are undertaken or evaluated, if the channels are forgotten or misunderstood, the effectiveness of monetary policy will be conditioned on the degree of modification carried out by the agents responsible for those channels. While this situation will lessen the central banks' power to drive the path of the economy, it will not exempt them from responsibility. The main channel through which central banks' policies operate is the banking sector. Central banks decide the amount of reserves to be provided and their price. Influenced by those policies, the banking sector will determine the conditions for lending. Accordingly, this chapter aims to analyse whether the banking sector, guided by its profitable prospects, is capable of modifying the policies undertaken by the Federal Reserve for the period 1958-2007 and how these possible modifications impacted on the real economy in relation to the Fed's intentions. While some authors have already examined this transmission channel, I address it from a totally different perspective, enlightened by other two literature blocks, which are fundamental to understand the right functioning of that channel. Therefore, three different literature blocks are linked here for the purpose of the hypothesis under analysis. The first one exposes some of the most accepted theories regarding how the banking sector channel operates. The second literature block, more contemporaneous, overthrows those theories by explaining that they were developed under wrong premises and subsequently, sets the correct ones. The last block introduces a fundamental detail regarding the interaction between the banking sector and the Fed, which is essential for the posterior elaboration of the theory that explains the channel whereby monetary policy is able to have an impact on the real economy.

The first literature block introduces the debate between those who defend that monetary policy has its effect on the real economy through the "money channel" (or "money view") and those who defend the "lending channel" (or "credit view"). Bernanke (1993) explained the "money view" as that whereby to slow down aggregate demand, the Fed sells in the open market to lessen bank reserves, reducing the money supply and putting upward pressure in interest rates, what will decrease the aggregate demand. For this story to hold, it is assumed that all nonmonetary assets are perfect substitutes, and that money and bank deposits have no perfect substitutes. He describes the "credit view" as the channel whereby monetary policy, apart from affecting short-term rates, determines aggregate demand by altering the availability of bank loans. This channel requires that loans and other forms of credit are not substitutes for borrowers. In a nutshell, the "money view" implies that the Federal Reserve removes reserves from the banking sector to reduce their deposits. Having less money on their balances, banks will increase interest rates. The "credit view" involves also the removal of reserves, but in this case, with the aim of forcing banks to reduce the amount of lending. As Bernanke exposed some examples of nonmonetary assets which are not perfect substitutes, and assets which are substitutes for money and bank deposits, he focused on the "credit view" and factors that weakened and will weaken the effect of monetary policy over this channel. Supporting the "lending view" were also Bernanke and Blinder (1992). They used a VAR and claimed that the federal funds rate was a better forecaster of real variables, as it was less contaminated by endogenous responses than the money growth rate. The results showed that an increase in the federal funds rate triggered, firstly, a decrease in securities in banks' balance sheets, which recovered after eight months, and a fall in loans after approximately six months, which did not recover in the horizon of the impulse response displayed. Kashyap, Stein and Wilcox (1993) supported also the "lending view" by showing a decline in loans supply once there was monetary tightening. Kashyap and Stein (2000), using a VAR, added that the "lending channel" was stronger for banks with less liquid balance sheets, namely, small banks.

On the "money view" side, Ramey (1993) using a VECM showed more predictive power in M2 velocity for output than in bank loan velocity. Oliner and Rudebusch (1995, 1996) claimed that the lending channel did not operate, as a monetary shock did not affect bank debt very differently from nonbank debt. Therefore, a negative monetary shock to bank reserves had no effect on lending, because, as argued in Romer and Romer (1990), banks have available other means to obtain funds with little cost in terms of reserve holdings. Thus, monetary policy influences the economy through the stock of transaction balances. The Romers supported the "money view" and the previous conclusion by analysing the behaviour of money and bank lending around episodes when the Fed carried out shifts in monetary policy regardless of the real economy performance.

However, both "views" are born from a flawed premise. Actually, banks do not need reserves for either opening deposits or lending. Causality runs the other way around. Deposits and loans lead to more demand for reserves, which under an interest rate target, will be provided to keep the interest rate targeted[15]. Thus, the second literature block discusses the concept of the money multiplier (intrinsic part of the channels described above) and its existence, as it is essential to comprehend that causality runs from loans to reserves. Against the monetarists' argument, Moore (1983) claimed that money growth is endogenous and that "the ability of central banks to control the rate of growth of monetary aggregates therefore hinges on their ability to control the rate of growth of bank lending, rather than the monetary base" (p. 544). Further: "The assumption…is that banks set the prime rate and then attempt to meet the loan demand that results" (p.545). Lombra (1992) and Goodhart (2007) also mentioned the endogeneity of the monetary aggregates once interest rates are set. Holmes (1969) stated: "…commercial banks are … creators of money and credit…" (p.70); "In the real world, banks extend credit, creating deposits in the process, and look for the reserves later" (p.73). Following the same argument, Jakab and Kumhof (2015) explained that banks do not need savings to build deposits and use them for lending, as the deposit multiplier theory suggests. Actually, banks lend while it is profitable and does not endanger their solvency. They lend depending on the demand for loans and afterwards, borrow the necessary reserves. Under the current interest rate target as the authors claimed: "modern central banks… are committed to supplying as many reserves (and cash) as banks demand at that rate, in order to safeguard financial stability. The quantity of reserves is therefore a consequence, not a cause, of lending and money creation" (p.5). As they explained, when a customer enters the bank asking for a credit, the bank opens a liability (a deposit) and an asset (a loan). Thereby, the bank is able to create money without the necessity of savers' deposits as many textbooks suggest. Regarding deposits, when a customer carries his savings to a bank, the funds in bank A increases. However, at the same time, a deposit from bank B is removed. Therefore, in aggregate levels, those deposits are unable to create any money.

[15] If the reader is remembering now that during the Volcker's experiment monetary aggregates were targeted instead of interest rates, the opposite can be deduced from the following quotes. In Bindseil (2004) the following Goodhart's statement is found: "if properly analyse [the episode], reveal that the Fed continued to use interest rates as its fundamental modus operandi, even if it dressed up its activities under the mask of monetary base control…there was a degree of play-acting even deception…" (Goodhart (2001), p.30). In Bindseil words "The "smokescreen" created by Volcker would thus have been simply a necessary condition for bringing inflation to an end under conditions of imperfect central bank independence"(Bindseil (2004),p.30). Even Volcker in 1982 stated: "On these money growth targets, in substance, I don't care. I think either of these two sets of numbers [5.5 and 6.5%] will make no difference, virtually, in what we actually do… [W]e are within the limits of the growth targets anyway" (Meltzer 2009b, p. 1114). Further: "I, frankly, cannot live in these circumstances, given what is going on in the money markets, with violent moves in short-term rates in either direction. It would just be so disturbing in terms of expectations, market psychology, and fragility that it's just the wrong policy, period, during this particular period" (Meltzer 2009b, p.1115).

Likewise, the reserves required to bank A for the new deposit will translate in a reduction in required reserves in bank B. Hence, again, reserves levels remain the same in aggregate levels. These authors, using a DSGE model and the financing through money creation (FMC) model view, based on the previous argumentation, found larger and faster changes in lending and money and a larger effect on the real economy than when identical shocks were applied to the same model with the deposit multiplier theory, which they called the intermediation of loanable funds (ILF) model view. As conclusion, the FMC model was more consistent with procyclical bank leverage and rationing of credit during downturns. Other authors explaining the fallacy of the deposits multiplier were Bindseil (2004), Borio and Disyatat (2009) and McLeay, Radia and Thomas (2014a and 2014b). Also, Carpenter and Demiralp (2010) using a SVAR explained that the causality works from loans to banks liabilities and there is not money multiplier. They based their argument on the current situation, in which despite the Fed has provided banks with an enormous amount of reserves, they are kept in banks' balance sheet without reaching the real economy through loans.

The last block of literature refers to the spread between the federal funds rate and the discount rate. Why is this important? It has been argued that the demand for money is endogenous and will depend on the loans rate. But what does determine that rate? To the same question sent by email to the Federal Reserve "Contact Us" option, the following was answered: "The prime rate is a rate established by commercial banks as a lending rate or base off which their commercial loans are priced. In other words, the banks set their own rates based on the demand for various kinds of loans, on the cost of money to the banks, and on the administrative costs of making loans…" Here, the important piece for the third block is "on the cost of money to the banks". That is, the spread is important because it determines the source from where banks obtain their reserves and at which cost. Subsequently, that cost influences the prime loans rate, what in turn, determines the endogenous demand for money. Therefore, this spread is the starting point for explaining the channel whereby Fed's policies reach the real economy (which is developed in the next section).

The literature exposes that a larger spread between the discount rate and the federal funds rate triggered more borrowing at the discount window. Pierce (1993) claimed that the changes in the Fed's operating procedures from 1975 to 1991 transformed the relationship between the spread and the borrowing function. According to Pierce, the period from 1975 to 1979 was the period of federal funds rate target, from 1979 to 1982, of nonborrowed reserves target and a lagged reserve accounting, and from 1982 to 1991, of borrowed reserves with lagged accounting, which changed to contemporaneous reserve accounting after 1984. He showed that for the first period there was a strong non-linear relationship between positive spreads and borrowing. Thus, the larger the spread, the more borrowed reserves were demanded. For the second period, when the spread was larger and for longer time, the relationship was looser. For the last period, the relationship became weak. Also, he pointed out that during the contemporaneous reserve accounting period, the excess reserves ratio increased and borrowing fell to its lowest level despite the large spreads. Hence, while the lagged reserve accounting was in place, the relationship between spreads and borrowing was more or less strong depending on the Fed's operating procedure, but it weakened under the contemporaneous reserves accounting. Peristiani (1991) identified a nonlinear relationship between the spread and borrowing (for the period 1959-1988) with an inverted S-shape. The number of banks going to the discount window increased when the spread was larger, but borrowing decreased at the highest levels of the spread due to restrictions, further costs and the collateral required to back the amount of borrowing. In the same line, Hamdani and Peristani (1991) with a disaggregated approach, differentiating between small and large banks, observed the same non-linear relationship as the authors above. Also, they found that borrowing was positively autocorrelated for small banks but not for large banks. Kasriel and Merris (1982), following the preceding conclusions, claimed that borrowing also depended on expectations about the spread. That is, if banks expected a larger spread in the future, they would borrow less at that moment. They also added that before 1979, under the federal funds target, the Fed was careless about the relation between the spread and borrowing. Later, under the nonborrowing reserves target and lagged reserves accounting, the relationship weakened because of the greater uncertainty and volatility of the spread, as also mentioned in Goodfriend (1981). Clouse (1994) mentioned the banks' reluctance to borrow after the active involvement of the discount window during the 1980s and 1990s to avoid bank failures. As explained by the New York Fed (Fedpoint 2015), institutions that borrowed at the discount window expressed their concern about the signal of weakness that it represented. More facilities were provided in 1999 to avoid this situation. Although after 1990 the discount rate was fixed 0.25-05% below the fed funds rate, what incentivised institutions to borrow, they must have exhausted before other available funds sources. It could be also important to highlight that after the Monetary Control Act in 1980, from 1980 to 1982, when the spread was positive, the Fed imposed a surcharge in addition to the discount window rate, varying between two and three percentage points, for institutions with deposits of $500 million or more or those who borrowed frequently. Last, since 2003 the discount rate was set above the federal funds rate as a penalty rate, but borrowing was allowed with no question asked.

In order to evaluate the hypothesis regarding whether the banking sector has potential to transform monetary policy, and learn about the functioning of that channel, as well as about its impact on the real economy, a time-varying coefficient Bayesians vector autoregressive (henceforth, TVC-BSVAR) is estimated as Primiceri (2005), with the Del Negro and Primiceri's (2013) corrigendum. To assess that hypothesis, firstly, monetary policies must be measured correctly. Unlike the standard procedure that uses the federal funds rate (or another short-term rate and some reserves measure), the approach undertaken here is based on Chapter 2 where it was explained that the federal funds rate is not an instrument but an intermediate target and its inclusion in the model would lead to misleading results. The reason is that, beyond the conceptual error, to measure monetary policy stance, only the supply side of the money market, namely Fed's policies, must be captured. The federal funds rate, however, captures demand and supply forces of that market. Taking a step forward, in this chapter I drop the Fed's instrument form the model used in Chapter 2 and focus on two variables to measure monetary policy stance and banking sector behavior. The first variable, already used in Chapter 2, is the spread between the federal funds rate and the discount rate, which represents the Fed's policies when using its instruments, as it captures the different prices at which reserves can be obtained. The second variable is the spread between the prime loans rate and the federal funds rate, which measures whether the banking sector is modifying Fed's policies, according to the cost at which reserves were acquired. The results confirm the new proposed "reserves-cost" theory and show that when the Federal Reserve allowed positive spreads between the federal funds rate and the discount rate, the banking sector set a lower loans rate in relation to the federal funds rate, inasmuch as the reserves needed to back the demand for loans could be obtained at the discount window at a lower rate than the federal funds rate. Thus, as the federal funds rate had an impact ratio below one on the loans rate, the demand for credit was not restrained sufficiently and triggered higher inflation. Thereby, the Fed's lack of attention or knowledge about this channel, led to a reduction in its control over monetary policy.

This chapter is structured as follow. Section 3.2 explains the theoretical framework. Section 3.3 describes the model and data, as well as the identification strategy, priors for the estimation of the model and the computational details. In section 3.4, the results are presented and all the lessons and patterns from those results are connected and interpreted in section 3.5. Finally, section 3.6 summarizes the main conclusions.

3.2 Theoretical framework

To evaluate whether the banking sector is able to modify the Fed's policies, as already mentioned, it is necessary to capture correctly the monetary policy stance. In the literature is common to use the federal funds rate (or other intermediate target) as Federal Reserve's instrument. As explained in Chapter 2 this approach is erroneous because the federal funds rate is an intermediate target, not an instrument. The real instruments available for the Fed before the Great Recession were the reserve requirement ratio, open market operations (OMO henceforth) and the discount rate[16]. Beyond this conceptual error, the consequence of introducing the federal funds rate in the model for measuring monetary policy stance is that the results will be misleading, as this interest rate captures the demand and supply forces of the money market, when only the supply side should be captured, namely, Fed's policies. To understand this statement, the money market must be visualized as two submarkets. The first one includes the central bank, in this case the Fed, on the supply side, and the banking sector on the demand side. In the second submarket, the banking sector switches to the supply side, the other agents of the economy being the demand side. In both cases, the supply side receives that denomination because of its ability to create money. The Fed controls the federal funds rate by purchasing and selling securities in the open market and increasing or decreasing the discount rate. Thus, the Fed controls the amount (with OMO) and price (with the discount rate and the federal funds rate) of money in the first submarket, according to the impact it intends to exert on the real economy and subject to the banking sector's demand for money and its consequent decisions about the federal funds rate. Given the price and amount of money set by the Fed in the first submarket, banks will also decide the amount (loans), partially, and the price (loans rate) of money in the second submarket. However, on the demand side of the first submarket, banks also set the federal funds rate according to the credit demand and other factors determining the amount of reserves required or desired. Forecasts about the evolution of those factors and Federal Reserve's policies will also influence the federal funds rate at which they will offer their reserves surplus. Thus, in order to achieve the same targeted federal funds rate at two consecutive periods, the Fed will have

[16] https://www.federalreserve.gov/monetarypolicy/policytools.htm

to use its instruments differently, by injecting different amounts of money into the market and setting the discount rate at different levels, depending on the aggregate demand for money and banks decisions on the amount of reserves they desire to hold, as well as how they price them. Consequently, the different amount of reserves and their price in the first submarket will suppose that the banking sector establishes a different loans rate in the second submarket, resulting in a different impact on the real economy under the same federal funds rate. That different impact is the final and definitive reason against the use of the federal funds rate (or other intermediate target) as a measure of monetary policy, because it also captures the modifications carried out by the banking sector, namely, the demand side of the first submarket.

While this argument was already developed in Chapter 2, the different impact of the federal funds rate on the loans rate was only hypothesized. Starting from there, now I elaborate the mechanism (which is evaluated and confirmed later) operating between both rates, which explains the functioning of the banking sector channel. To begin with, it must be acknowledged that banks are profit maximization agents. As in every profit maximization problem, profits depend on costs. When banks lend, they need to back those loans with reserves to fulfil their level of required reserves or the levels they desire to hold. Notwithstanding, as commented previously, banks do not need reserves to lend. Rather, they lend and later, obtain the reserves. Accordingly, banks observe previously the associated cost to reserves, and set the loans rate[17]. Regarding the cost associated to reserves, banks can obtain two types of reserves: borrowed and nonborrowed reserves. Borrowed reserves can be obtained at the discount window, where the cost is the discount rate. Regarding nonborrowed reserves, the source from where banks obtain them is either the federal funds market, where the cost is the federal funds rate, or the open market[18]. For this last case, when banks sell

[17] Risk, liquidity, required capital ratios…are also factor influencing the loans rate, as well as the amount of credit already extended. However, I am using the prime loans rate, which is the rate that commercial banks charge to their most credit-worthy customers. This rate drives the other loans rates. Therefore, the aforementioned potential factors influencing the loans rate either do not apply to the case under study, namely, the prime loans rate, or they are significantly diminished.

[18] Another source from where banks can obtain reserves is deposits. In any case, in Figure 2.3, taking as reference the rate paid on the 3-months Eurodollar deposits, it is seen that the cost is either the same as the federal funds rate or slightly higher. Until the late 1980s, a cheaper source was time and saving deposits, as under regulation Q banks could not offer an interest rate above the ceiling rate established, and several times short-term rates were above the ceiling. Therefore, those reserves were cheaper than the costs displayed in Figure 2. 3. Nonetheless, given the alternatives yields, depositors looked for better interest rates in Eurodollar deposits or in loans and saving associations. Additionally, commercial banks evaded those regulations by

securities in the open market to obtain reserves, they are renouncing to the interest rate paid on those securities. The cost (or opportunity cost) of getting those reserves is the interest rate not received. Minimum, the cost of those reserves will be the 3-months T-bill rate, which in general, is around the federal funds rate levels[19]. If the term of those securities is longer, the cost will be higher. Hence, it can be considered that the cost of borrowed reserves is the discount rate and the cost of nonborrowed reserves is the federal funds rate. Obviously, these measures are not accurate, but they are a good approximation for the reserves cost. Thereby, when the spread between the federal funds rate and the discount rate (FR-DR) is positive, the cost of reserves (at least, a high percentage) will be the discount rate, as banks will prefer to borrow more cheaply at the discount window. When both rates are similar, the source will be indifferent, as the cost will be approximately the same. When the spread is negative, banks will avoid the discount window, unless it is strictly necessary. Once banks are aware of the cost of those reserves, they will set the prime loans rate to maximize their profits and avoid any potential solvency problem. This is the key for explaining the functioning of the banking sector as the channel for monetary policy. When banks obtain "cheaper" reserves at the discount rate because the spread FR-DR is positive, even though the federal funds rate is raised, banks have the capacity of increasing the loans rate relatively less than the federal funds rate. Consequently, the demand for credit is not restrained as much as the Federal Reserve intends. This is the main reason why the use of the federal funds rate is not adequate. Thus, what I have decided to call the "reserves-cost" theory, explains the channel whereby monetary policy operates and is described as follows: Given the Fed sets short-term rates, wittingly or unwittingly, by providing reserves and using the discount rate, the only way whereby it can have an impact on the real economy is by influencing the reserves cost directly, and indirectly, with the impact of this cost on the loans rate. Therefore, while the spread between the federal funds rate and the discount rate is closed, it is more likely that the Fed has total control of monetary policy. Otherwise, the banking sector will modify those policies and the impact on the real economy will be different.

Although to the best of my knowledge this theory is a novelty in the literature, Thornton (1982), when analysing the effect of the discount rate on market interest rates,

---

offering different services and deposits (Meltzer 2009a, p.470, 608 and 648). Thus, it is not expected that a significant amount of reserves came from that source.

[19] If looking at Figure 2.1, it is seen that the 3-months T-Bill rate normally tracks the federal funds rate, except for those periods when the spread FR-DR is larger. Yet, the 3-months T-Bill rate remained above the discount rate in those cases and banks would prefer to borrow at the discount window.

already visualized part of this transmission channel: "Changes in the discount rate affect market interest rates through their impact on borrowing from the Federal Reserve" (p.1). Later, he claimed "It is not simply the level of the discount rate that influences a depositary institution's decision to borrow, but the level of the discount rate relative to rates on

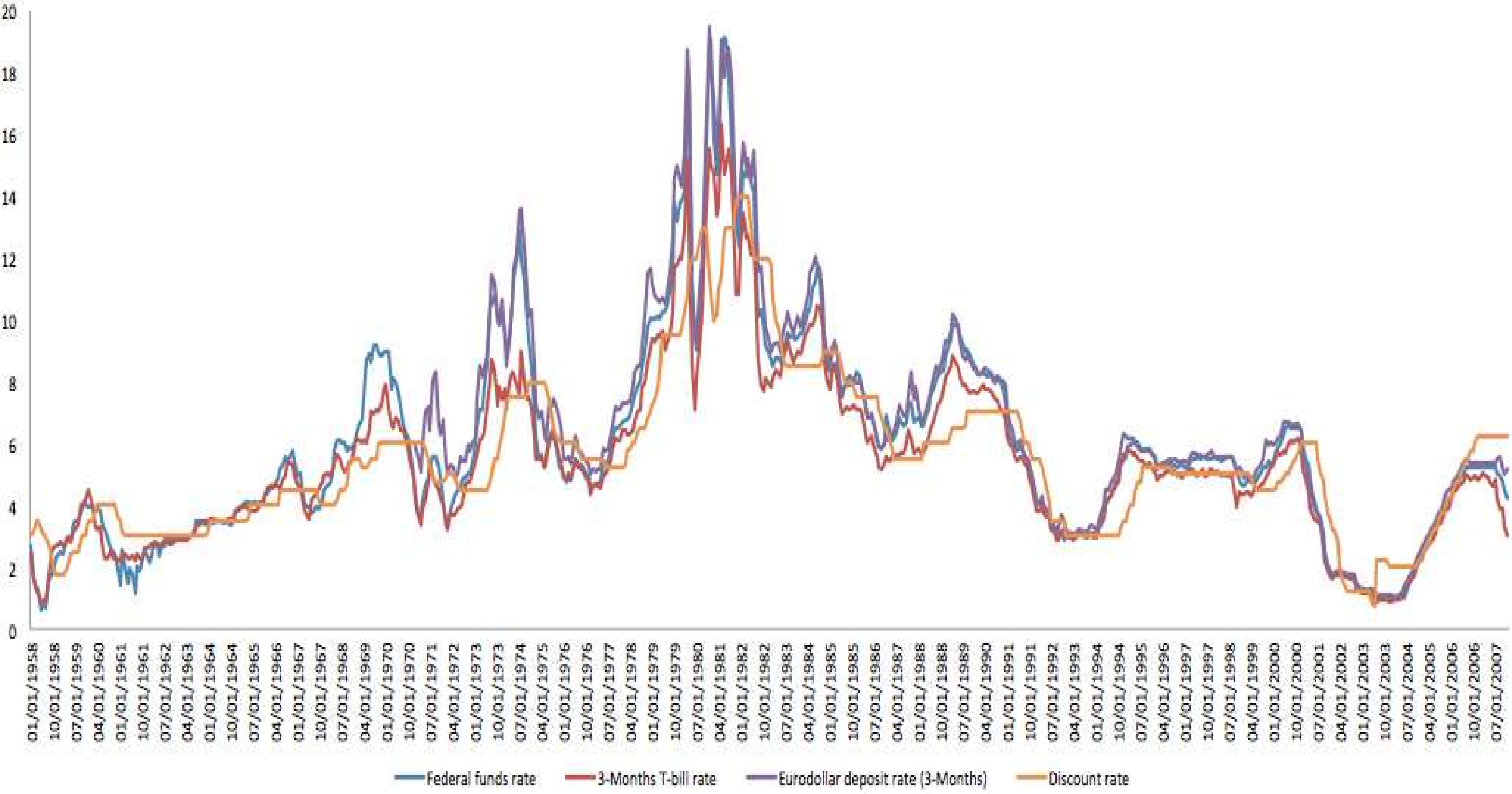

Figure 2.1 – Data source: FRED

alternative adjustment asset. […] Thus, the important variable in the decision to borrow is the so-called least-cost spread between the rate on the next best reserve adjustment asset and the discount rate" (p.2).

Once the theory has been explained, there is empirical evidence that seems to support it. First, in Figure 2.2 three facts are to be highlighted. The first one is that when the spread FR-DR was larger, it coincided with higher inflation levels. However, this effect smoothly disappeared after the second half of 1980s, as showed in Chapter 2[20]. Apart from that, when this spread was larger, the spread between the prime loans rate and federal funds rate (LR-FR) became smaller, what was more noticeable in the late 1960s and early 1970s, when the spread even turned negative. Likely, the loans rate did not restrain the demand for money as

[20] See Figure 1.7.1

much as the Fed intended with the raises in the federal funds rate, because banks could obtain more cheaply reserves at discount rate cost and increase less the loans rate. The smaller the spread LR-FR, the lower pressure was exerted on the money market, causing a higher demand for loans and inflation. The second fact is that after the late 1960s, the spread between the loans rate and discount rate (LR-DR) was quite volatile, but between mid-1985 and 1987 was constant, occurring again from 1989 to 1994. For those years, the loans rate seems to have been pegged to the discount rate. The spread between the loans rate and the federal funds rate was also more volatility after the late 1960s and became constant from 1990 to the end of the sample. Hence, the loan rate seems to have been pegged to the federal funds rate for that period. Surprisingly, in Chapter 2 a regime change was also observed around 1990 for the response of inflation to a FR-DR shock (Figure 1.7.1). The third fact is that not only the spread LR-FR was pegged since 1990, but also it was almost the largest of the whole sample. Therefore, it was the period when the loans rate was tightening most the money market, in relation to the federal funds rate.

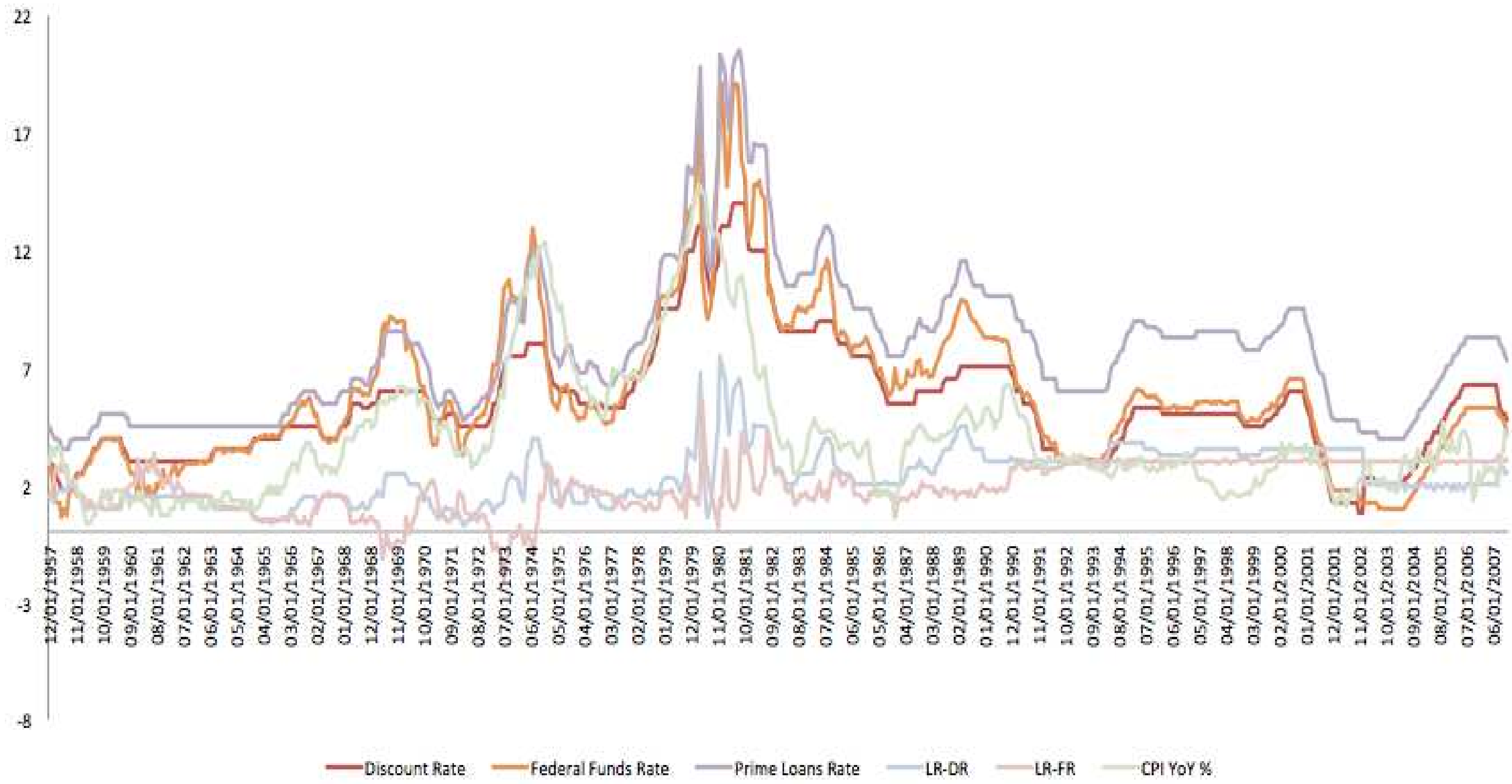

Figure 2.2 – Data source: FRED

Figure 2.3 displays free, excess and borrowed reserves as percentage of total reserves. The variable of free reserves offers a good picture of the source from where most reserves are obtained. When the Fed provides banks with enough nonborrowed reserves through open market purchases to cover the level of required reserves, free reserves are positive. Therefore, banks have most of their reserves at the cost, minimum, of the 3-months Treasury bill rate or the federal funds rate. When the amount of nonborrowed reserves is not enough to back the

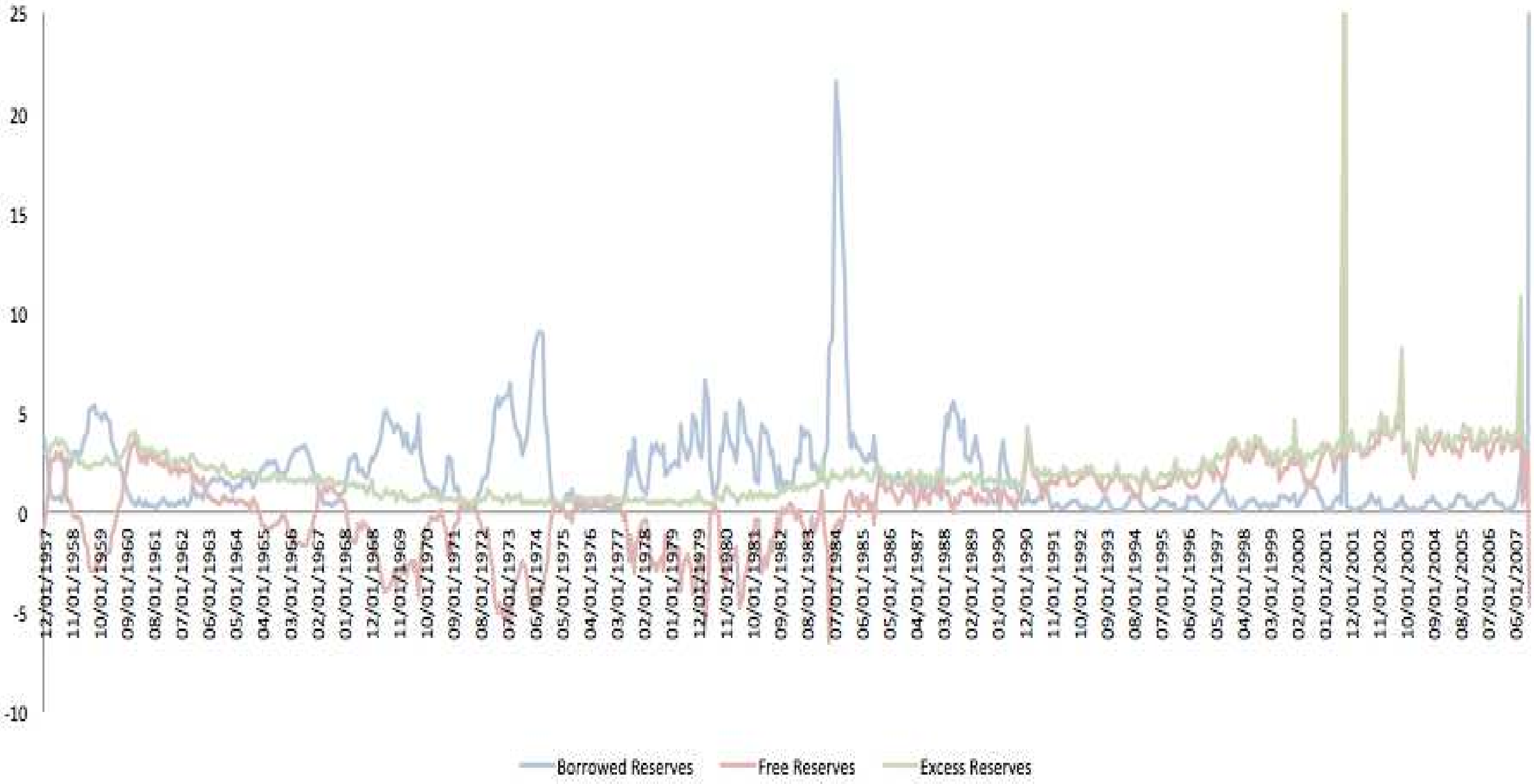

Figure 2.3 – Data source: FRED

demand for credit, banks have to borrow at the discount window to reach the level of required reserves. Thus, the level of free reserves is negative. Those reserves have, mostly, the cost of the discount rate. Another three interesting facts are found in this figure in relation to those observed in Figure 2.2. First, the higher percentages of borrowed reserves coincided with the periods of positive spreads FR-DR, and also with those periods when the spread LR-FR was more volatile. Second, after approximately the end of the supposed Volcker's experiment, free reserves levels became positive and increased until the end of the sample. Third, since free reserves became that up trending path, excess reserves have kept track of free reserves. Linking these three facts and those form Figure 2.2 the following can be extracted: borrowed reserves increased for the periods of positive spreads FR-DR and decreased significantly after 1990. For the latter case, the vast amount of reserves held by the banking sector came from nonborrowed reserves. This means that the use of the discount window decreased and banks had their reserves mostly at the nonborrowed reserves cost, namely, minimum at the 3-month T-bill rate (or at the federal funds rate) [21]. This occurred after 1990, just when the loans rate was pegged to the federal funds rate. Thus, since banks received most of the required reserves from the Fed through the open market, banks fixed the loan rate to the federal funds rate. On the other hand, when banks borrowed reserves at the discount window at a cheaper cost (given the positive spread FR-DR), they had more leeway to set a lower loans rate in relation to the federal funds rate. In consequence, the spread LR-FR was more volatile and smaller. This caused that the Fed lost, at least partially, control over monetary policy, as the loans rate produced different effects on the real economy in relation to monetary policy intentions.

3.3 Methodology

The model used in this chapter is the same as that in Primicery (2005), a TVC-BSVAR. The code used to estimate the model was downloaded from Gary Koop's website[22]. In this model, not only the coefficients vary, but also the variance covariance matrix. It uses drifting coefficients, in order to capture nonlinearities or time variation in the lag structure of the model, while the multivariate stochastic volatility is meant to capture potential heteroscedasticity of the shocks and nonlinearities in the simultaneous relations among the variables of the model. The combination of drifting coefficients and the variance covariance matrix allows the data to resolve whether the possible variations observed in the relation among variables emanate from the size of the shocks (impulse) or changes in the propagation mechanism (response). The advantage of this model is founded on its capacity to capture continuous and smoothed switching regimes, unlike those works that modeled time variation with discrete breaks. For the topic addressed in this case, it is expected that the Federal Reserve, banking sector and other agents of the economy learn from the evolution of the economy and each other. The learning process is considered to be slow and not to happen overnight. Hence, changes in the behavior of those agents, as a consequence of their learning process, will evolve smoothly.

[21] A factor that could have contributed to the accumulation of free and excess reserves in the mid-1980s is the fact that the contemporaneous reserves accounting started in 1984. The uncertainty created by not knowing the demand for credit and consequently, the reserves necessary for loans, could have led banks to be more cautious and hold more excess reserves. That regime, however, ended in 1998.

[22] https://sites.google.com/site/garykoop/home/computer-code-2.
The only modifications applied to the code have been to adapt it to the data used here, as well as for those tools necessary for the representation of the results. The process of estimation is entirely as found in the file.

Using the same notation as Primiceri, the model is the following:

$$y_t = c_t + B_{1,t}y_{t-1} + \ldots + B_{k,t}\,y_{t-k} + u_t \qquad t = 1,\ldots, T. \qquad (1)$$

where y_t, and c_t are n x 1 vectors of observed endogenous variables, and a vector of time-varying coefficients multiplying constant terms respectively. $B_{i,t},\ i = 1,\ldots, k,$ represents n x n matrices of time varying coefficients. Last, u_t are heteroscedastic unobservable shocks. The variance covariance matrix Ω_t is triangularly reduced and defined by

$$A_t\,\Omega_t A'_t = \sum_t \sum'_t \qquad (2)$$

where A_t is a lower triangular matrix with ones in the main diagonal, $\alpha_{ij,t}$ being the non-zero and non-one elements of the matrix. $\sum_t$ is a diagonal matrix with $\sigma_{n,t}$ elements in the diagonal.

Hence,

$$y_t = B_{0,t} + B_{1,t}y_{t-1} + \ldots + B_{k,t}y_{t-k} + A_t^{-1}\sum_t \varepsilon_t \qquad (3)$$

$$V(\varepsilon_t) = I_n$$

Stacking all the $B_{k,t}$'s in a vector,

$$B_t = vec(B'_t) = [B_{0,t},\ B_{1,t}, B_{2,t}, \ldots B_{k,t}]'$$

and with

$$X_t = I_n \otimes [1, y_{t-1}, y_{t-2}, \ldots y_{t-k}]',$$

the VAR can be represented and modelled as:

$$y_t = X'_t B_t + A_t^{-1}\sum_t \varepsilon_t \qquad (4)$$

Stacking by rows the elements $\alpha_{ij,t}$ of the matrix A_t and the elements $\sigma_{n,t}$ of matrix $\sum_t$, the state vectors or transition equations representing the dynamics of the model are:

$$B_t = B_{t-1} + v_t \qquad (5)$$

$$\alpha_t = \alpha_{t-1} + \zeta_t \qquad (6)$$

$$\log \sigma_t = \log \sigma_{t-1} + \eta_t \qquad (7)$$

where both, the B_ts and the non-zero and non-one elements of the matrix A_t, α_t, follow random walks, while the standard deviations (7) follow a geometric random walk, accordingly belonging to the stochastic volatility models.

The innovations of the model are assumed to be jointly normally distributed, supposing the following variance covariance matrix:

$$V = \mathrm{Var}\left(\begin{bmatrix} \varepsilon t \\ vt \\ \zeta t \\ \eta t \end{bmatrix}\right) = \begin{bmatrix} In & 0 & 0 & 0 \\ 0 & Q & 0 & 0 \\ 0 & 0 & S & 0 \\ 0 & 0 & 0 & W \end{bmatrix} \qquad (8)$$

where I_n is an n-dimensional identity matrix and Q, S and W are positive definite matrices. As Primiceri pointed out, the zero blocks could be replaced by non-zero blocks, but there are two reasons for the assumptions taken. First, for the case under study, I include in the model two more variables than Primiceri (2005), who already considered the number of parameters to be high. Therefore, adding non-zero blocks would require a sensible prior to prevent ill-determined parameters. Thereby, given the even higher number of parameters used, the algorithm gets stuck in non-stationary draws if more parameters need to be estimated. As a restriction of stationarity has been imposed, it would mean that no draw is taken. This limitation prevents the inclusion of more variables or lags apart from those already included, and of course, the substitution of these zero blocks for non-zero blocks. The second reason is that I do not have any structural interpretation to impose on the different sources of uncertainty. S is assumed to be block diagonal, with blocks corresponding to parameters belonging to separate equation. That is, the coefficients of the contemporaneous relations evolve independently in each equation. For the estimation of the model, I refer the reader to the Appendix A of Primiceri (2005), taking into account the Del Negro and Primiceri‘s (2013) corrigendum, where the algorithm used for the Gibbs sampling undergoes a modification regarding the blocks from which the draws are taken.

3.3.1 Data and identification strategy

The period under analysis is 1958:II - 2007:IV and the source for the entire database is FRED (St. Louis Fed). The model includes the following variables:

- Real GDP (RGDP): Growth rate
- CPI Inflation (CPI): Growth rate
- Loans: It measures the growth rate of the sum of commercial, industrial, real state, consumer loans and other loans and leases.
- Federal funds rate-Discount rate spread (FR-DR): It measures the difference between the rate at which depositary institutions trade federal funds with each other overnight, and the rate charged to commercial banks and other depository institutions when they borrow at the Federal Reserve's discount window.
- Prime loans rate-Federal funds rate spread (LR-FR): This variable is the difference between the base rate used by banks to price short-term business loans, posted by a majority of 25 insured U.S.-chartered commercial banks, and the aforementioned federal funds rate.

This model is intended to capture the banking sector's response to Federal Reserve's policies, and in turn, how it affects the demand for credit, inflation and output. The Fed's policies are captured by the spread FR-DR. The spread LR-FR measures the banking sector's behavior in relation to Fed's policies. This spread is the first variable in the VAR because it is expected that banks, as already explained, before setting the loans rate, take into account the price at which reserves can be obtained (FR-DR). Then, loans, inflation and real GDP will respond to monetary policies and the loans rate after one period, in that order, to have a structural interpretation. That is, the order of the variables in the VAR is LR-FR first followed by FR-DR, loans, CPI inflation and real GDP.

There are two reasons why the Fed's instruments included in Chapter 2, namely, OMO and the discount rate, are now removed. First, the results obtained in Chapter 2 showed that OMO had no significant impact on output, inflation or money supply and therefore, monetary policy was transmitted through prices. Second, the "reserves-cost" theory implies that to measure monetary policy stance, as policies have its impact on the economy through their influence on the reserves cost, it is sufficient to capture the price at which reserves are available in the money market.

In order to have all the variables on the same scale they have been standardized $(y_t-E(y_t^*))/std(y_t^*)$. The periodicity chosen for the model is quarterly and it has been estimated with 1 lag. Given the periodicity of the data, I consider one quarter enough for the spreads to affect the rest of variables. Further, the use of more lags would be incoherent for the relationship between the spreads, given the mechanism described in section 3.2.

3.3.2 Priors and computational details

An invariant VAR from 1949:II to 1957:IV (35 observations)[23] is estimated to calibrate the priors' distributions. The set-up for the priors (as written in Gary Koop's code) is the following:

$$B_0 \sim N(\hat{B}_{OLS}, 4 \cdot V(\hat{B}_{OLS})),$$
$$A_0 \sim N(\hat{A}_{OLS}, 4 \cdot V(\hat{A}_{OLS})),$$
$$\log \sigma_0 \sim N(\log \hat{\sigma}_{OLS}, 4 \cdot I_n),$$
$$Q \sim IW(k^2_Q \cdot 35 \cdot V(\hat{B}_{OLS}), 35),$$
$$W \sim IW(k^2_w \cdot (1+\dim(W)) \cdot I_n, 1+\dim(W)),$$
$$S_1 \sim IW(k^2_s \cdot (1+\dim(S_1)) \cdot V(\hat{A}_{1,OLS}), 1+\dim(S_1)),$$
$$S_2 \sim IW(k^2_s \cdot (1+\dim(S_2)) \cdot V(\hat{A}_{2,OLS}), 1+\dim(S_2)),$$

where S_1 and S_2 are the two blocks of S, $A_{1,OLS}$ and $A_{2,OLS}$ are the corresponding blocks of A_{OLS}, and $k_Q = 0.01$, $k_S = 0.1$, and $k_W = 1$. Thus, the priors are not flat but diffuse. 500,000 draws were generated, discarding the first 200,000 and using 1 in every 150 to avoid correlation between them. Convergence tests are displayed in Appendix D. Regarding the computational time, the model needed around 20 hours.

3.4 Results

The TVC-BSVAR provides two tools to evaluate the "reserves-cost" theory. First, the impulse response functions show how a certain variable responds to another variable shock. Time varying coefficients allow disentangling or at least, providing some insight, regarding when there is a change either in the Federal Reserve's policies or in the banking sector's behavior. Later, the posterior mean of the standard deviation of the residuals for each equation of the VAR will shed light on the possibility of external variables, namely, variables not included in the model, with potential to distort the results obtained. The figures analyzed

[23] For the federal funds rate series, data is only available from 1954. As a proxy, I have used the 3-months T-bill rate from 1948 to 1954 to estimate the priors.

below display each period of the sample on the X-axis, the response to a positive shock from one to twenty months/quarters on the Y-axis and the scale of the response on the Z-axis.

The first set of figures belongs to the banking sector and the real economy's responses to Fed's policies, represented by FR-DR. Figure 2.4 displays the banking sector's response (LR-FR) to an FR-DR shock. The response is completely negative, although its values are quite heterogeneous. The same happens regarding the quarters at which the responses tend to zero. In Figure 2.15, ordinary impulse responses are represented for isolated periods but the same impulse response function. Except for the very last years of the sample, the response is significant for every period. Thus, the main point of the "reserves-cost" theory is confirmed. That is, when banks can obtain cheaper reserves because of positive spreads between the federal funds rate and the discount rate, banks raises the loans rate relatively less than the federal funds rate is increased. Figure 2.5 represents the response of loans to an FR-DR shock. The response in this case is subtler and needs a deeper analysis. The first relevant fact is that the response is positive or becomes more positive for periods of positive spread FR-DR. Those periods are 1965-1967, 1968-1970, 1972-1975, 1979-1981 and 1983-1984, which connects with the period 1985-1989. Between 1993 and 2000 the response is positive, but the values are lower in comparison with previous periods. However, positive responses occur also between 1960-1963, 1990-1993 and 2000-2004. A common characteristic of these last periods is that the federal funds rate was around two to three percentage points, being the lowest values of the entire sample. Also, an interesting detail is found for the period 1960-1963, because while the discount rate and the federal funds rate were increasing being approximately equal, the loans rate remained at the same level. That is, even though the Fed increased interest rates, the loans rate endured by the economy did not vary. Last, the negative responses or those positive responses after a decrease from the highest peaks correspond to periods of negative or zero spread FR-DR. For the positive spread in 1989, however, the positive peak starts to decrease before the highest level is reached. The responses are, nevertheless, rarely and weakly significant. Thus, in Figure 2.16, the loans response is significant only for the periods of positive response of the impulses response belonging to 1961 and 1980. Figure 2.6 shows that the response of CPI inflation to an FR-DR shock is positive for the entire period except after 2004, when the discount rate was already set as a penalty rate. In Figure 2.16 is seen that the positive responses are significant for around 17 quarters for every period, except after 2000 when the responses are no longer significant. Last, Figure 2.7 displays the response of real GDP to an FR-DR shock.

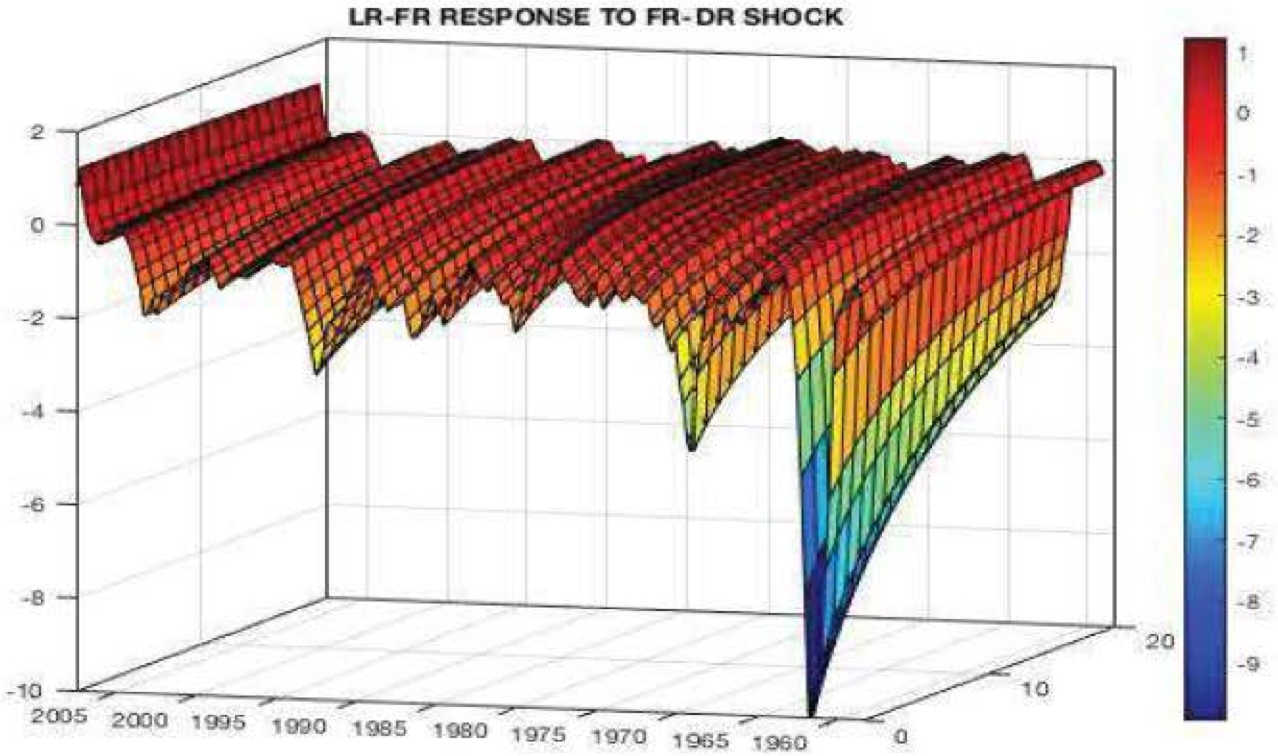

Figure 2.4 LR-FR impulse response to an FR-DR shock. Note: posterior means.

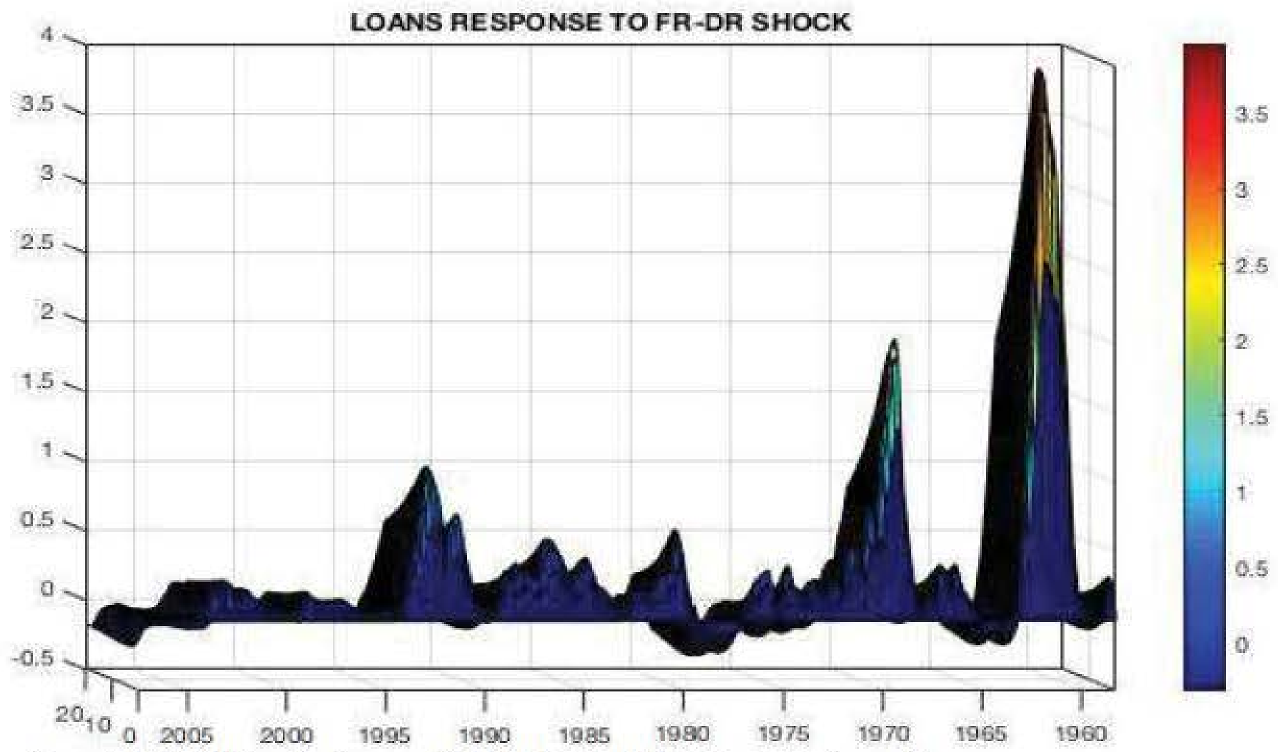

Figure 2.5 LOANS impulse response to an FR-DR shock. Note: posterior means.

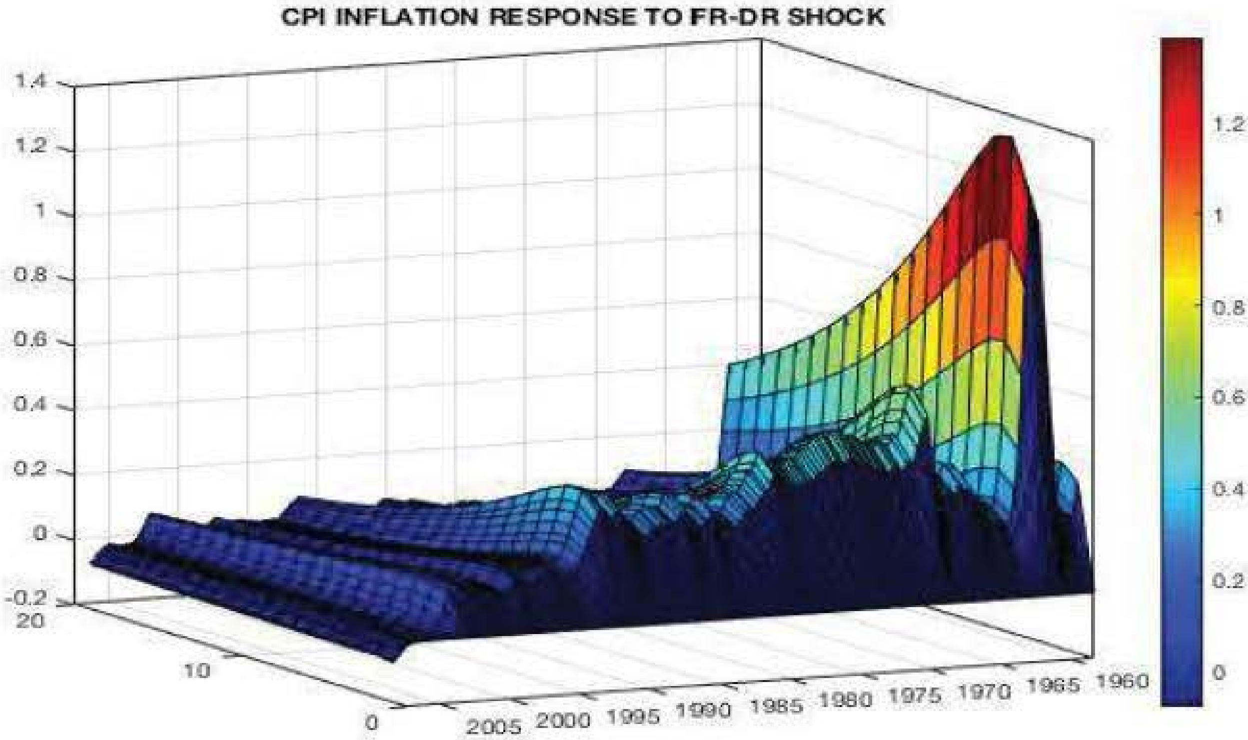

Figure 2.6 - CPI Inflation impulse response to an FR-DR shock. Note: posterior means.

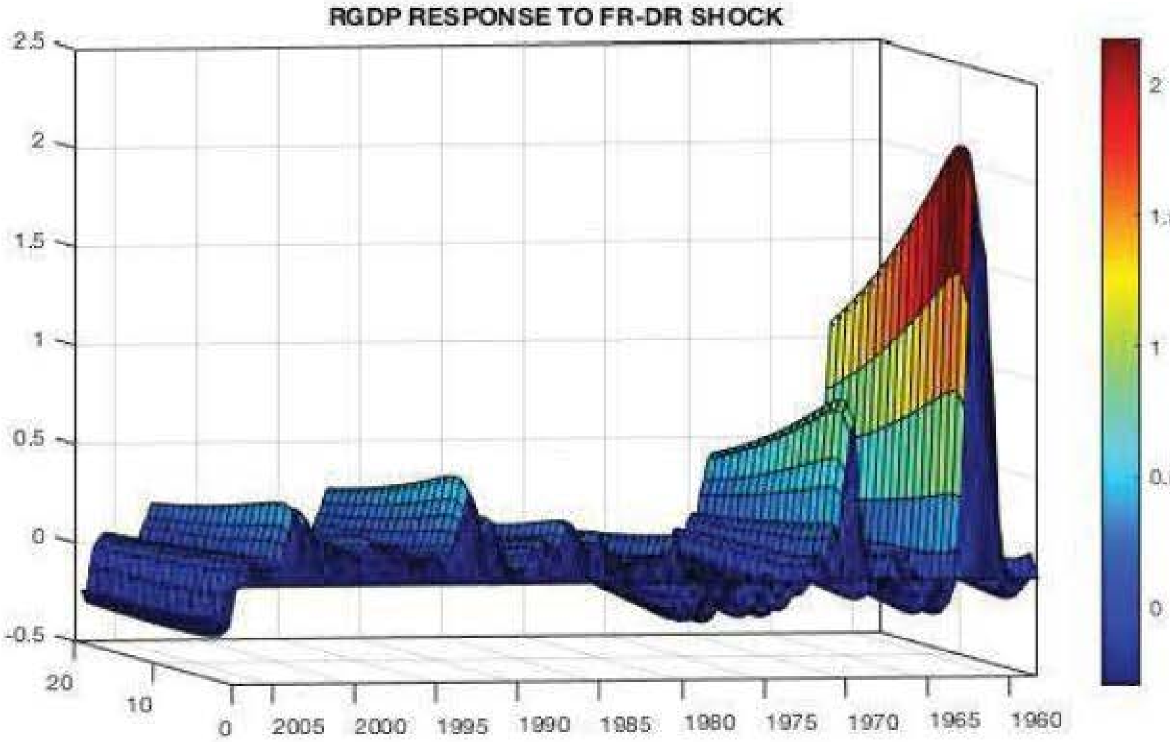

Figure 2.7 - Real GDP impulse response to an FR-DR shock. Note: posterior means.

The shape of this figure is very similar to Figure 2.5, with the exception that from 1970 to 1975 the response is negative. The interesting thing about this is that whereas the response is not significant for the periods with similar response to Figure 2.5, the negative response for 1975 is significant after 3 quarters (Figure 2.16). This may indicate that despite the positive spread, the federal funds rate was already high enough to restraint the demand for money, because it was the longer period that the loans rate remained below the federal funds rate. That is, banks could not increase the loans rate more, given the demand for credit.

The next set of figures corresponds to the analysis of the banking sector's behavior, conditioned on Fed's policies. Figure 2.8 shows that the spread FR-DR responds positively to an LR-DR shock except after 2002. This relationship is however, unimportant for the analysis. Figure 2.9 depicts the response of loans to an LR-FR shock. The response is negative for the entire period. The response is significant and tends to zero only after 20 quarters, except again for the last years of the sample (Figure 2.17). Figure 2.10 shows a negative inflation response to a LR-FR shock. Nonetheless, it is not significant at any period (Figure 2.17). The similarity between the real GDP and loans response to an FR-DR shock, occurs again between the real GDP and loans response to a LR-DR shock (Figure 2.11). In this case, the negative response is significant for the entire period and tends to zero only after 20 quarters (Figure 2.17). Last, Figures 2.12 and 2.13 represent the inflation and real GDP response to a loans shock, respectively. Both of them respond positively, although after 1995 the former has a negative response after two quarters and the latter, from the beginning. While inflation responses are not significant, for real GDP, the response corresponding to the periods 1965, 1970, 1975, 1980 and 1990 are significant for the three first quarters (Figure 2.18). When the residuals are evaluated, Figure 2.14 shows that real GDP and loans present higher residuals until the early 1980s. However, RGDP residuals remains low until the end of the sample, while loans residuals increase again around 1995, at the same time that inflation residuals. Inflation residuals were also from the early 1970s to the late 1980s. For the variables capturing Fed and banking sector's behavior, in both cases residuals are slightly higher in the first half of the 1970s. Later, for the early 1980s, an outstanding peak is found. Therefore, the early 1980s seems to hide some factor or factors that could distort the results obtained. The same can be said for the loans and inflation results after 1995 and for real GDP and loans until 1980.

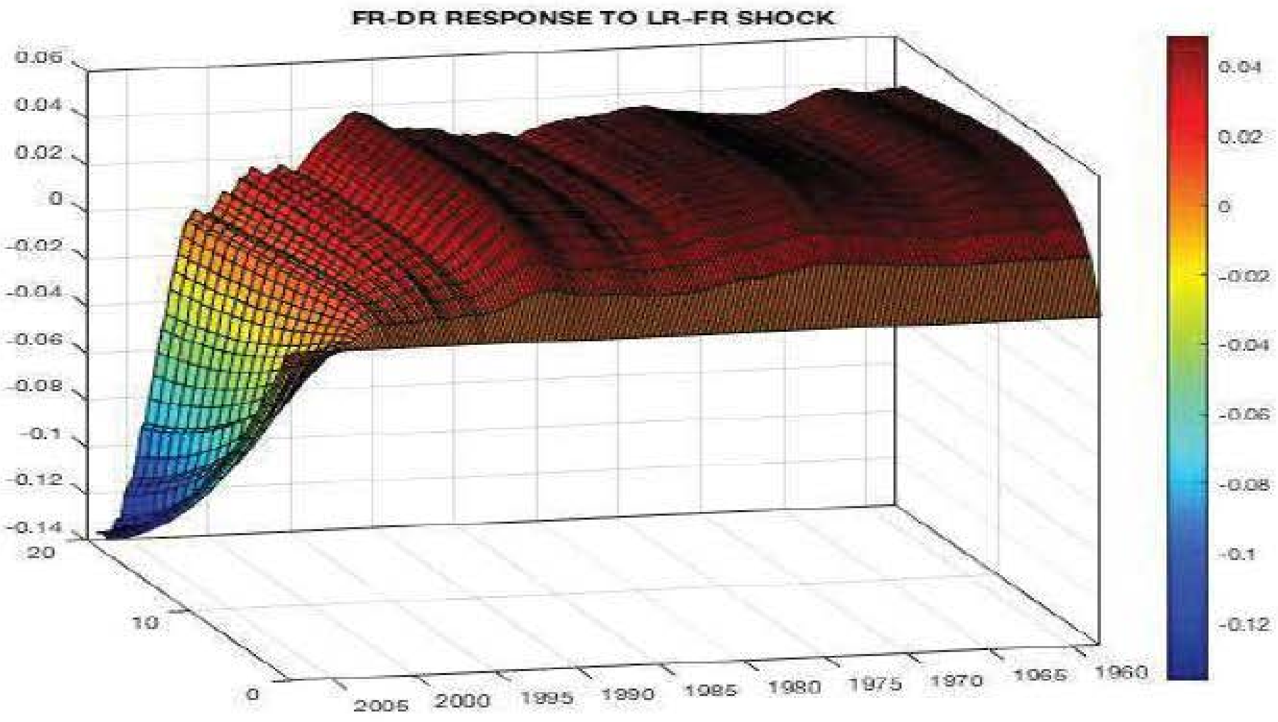

Figure 2.8 – FR-DR impulse response to an LR-FR shock. Note: posterior means.

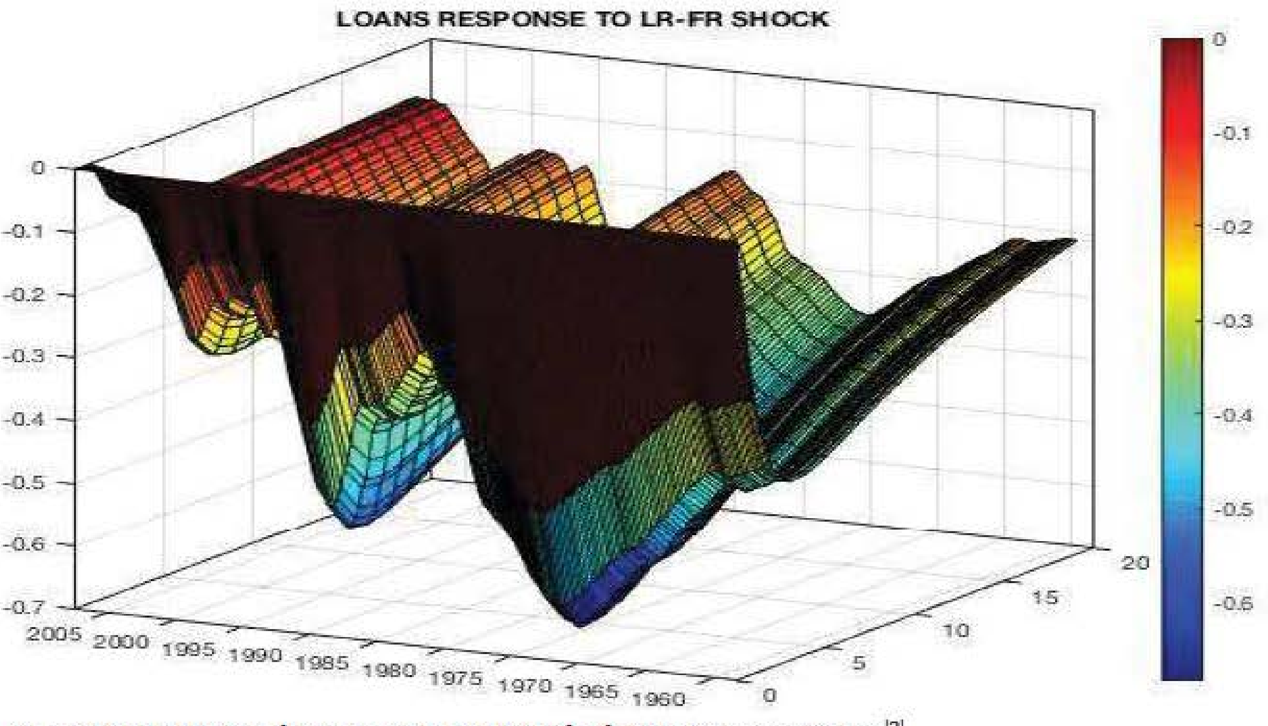

Figure 2.9 – LOANS impulse response to an LR-FR shock. Note: posterior means.

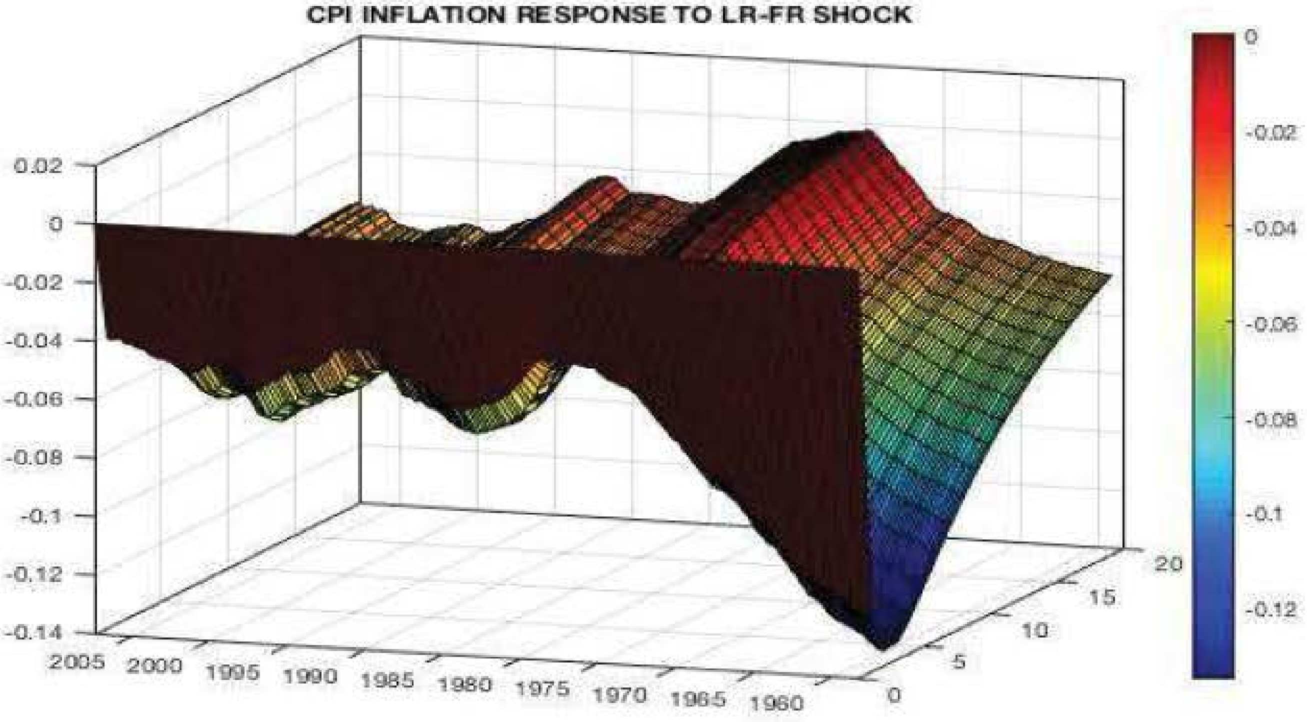

Figure 2.10 - CPI Inflation impulse response to an LR-FR shock. Note: posterior means.

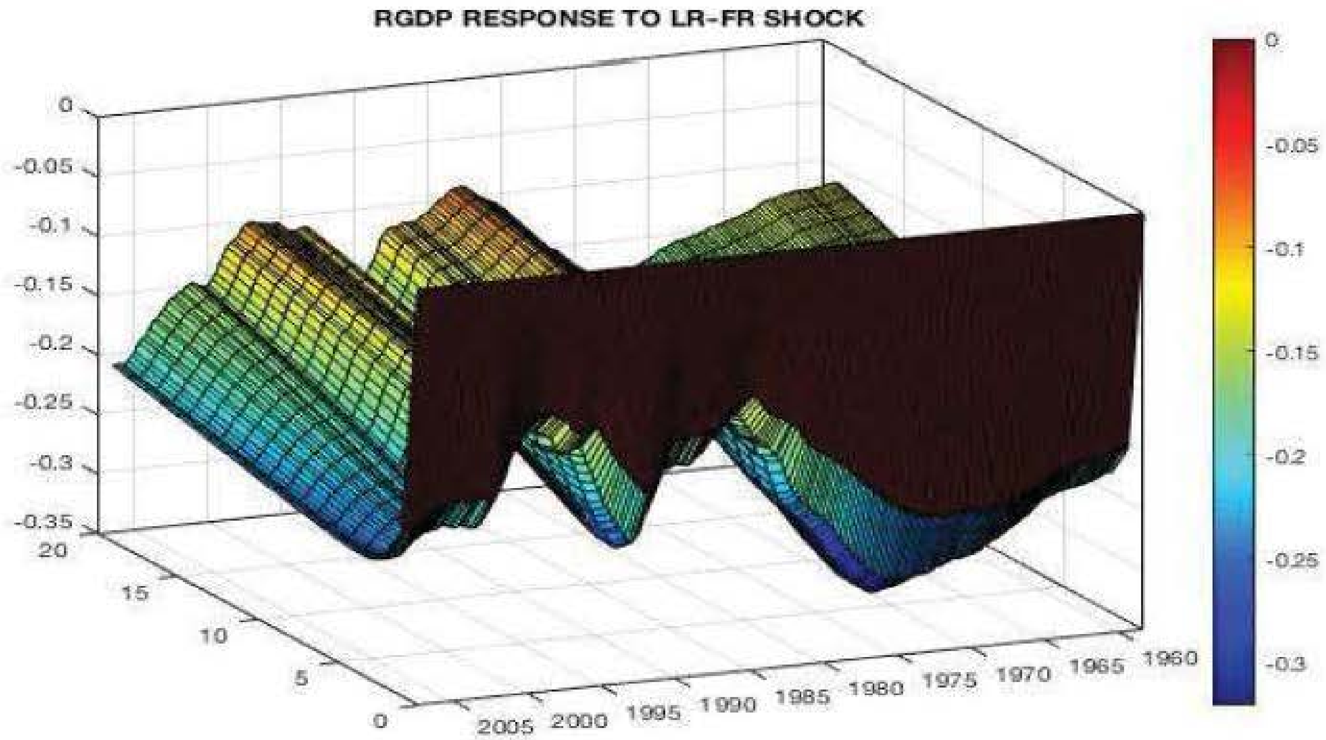

Figure 2.11 - Real GDP impulse response to an LR-FR shock. Note: posterior means.

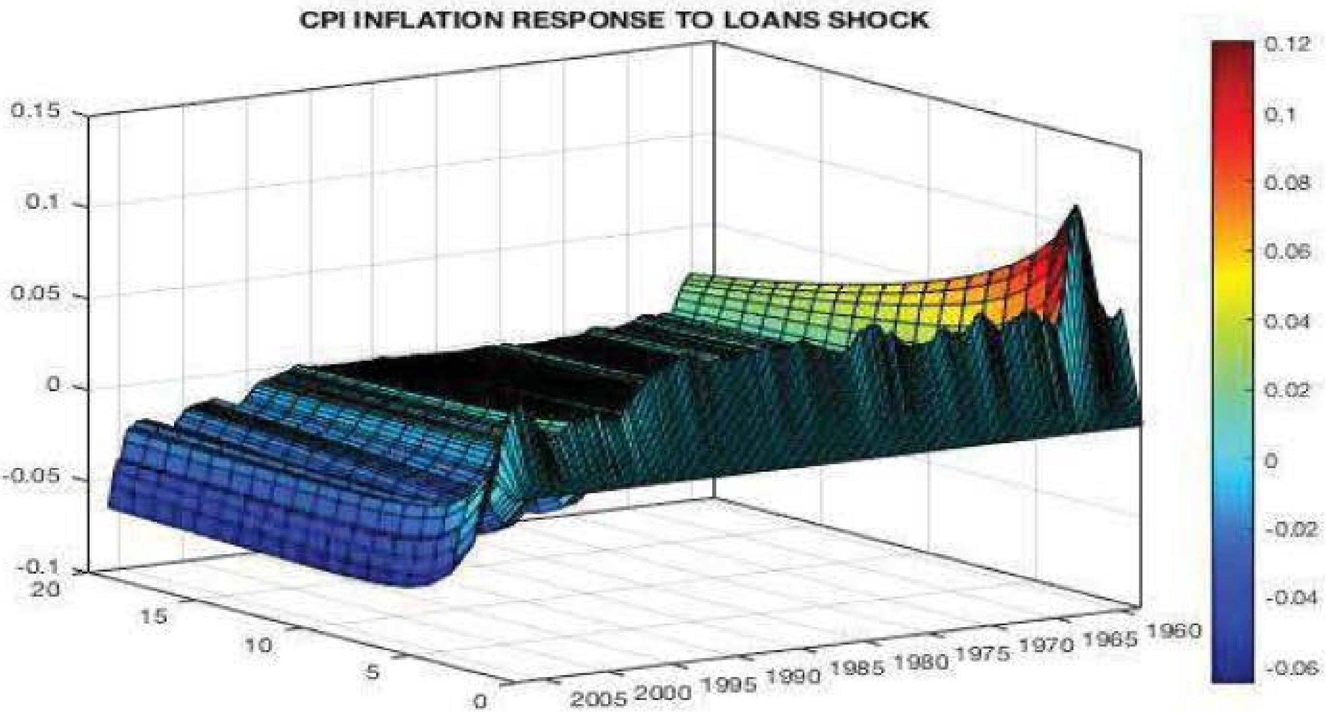

Figure 2.12 - CPI Inflation impulse response to a LOANS shock. Note: posterior means.

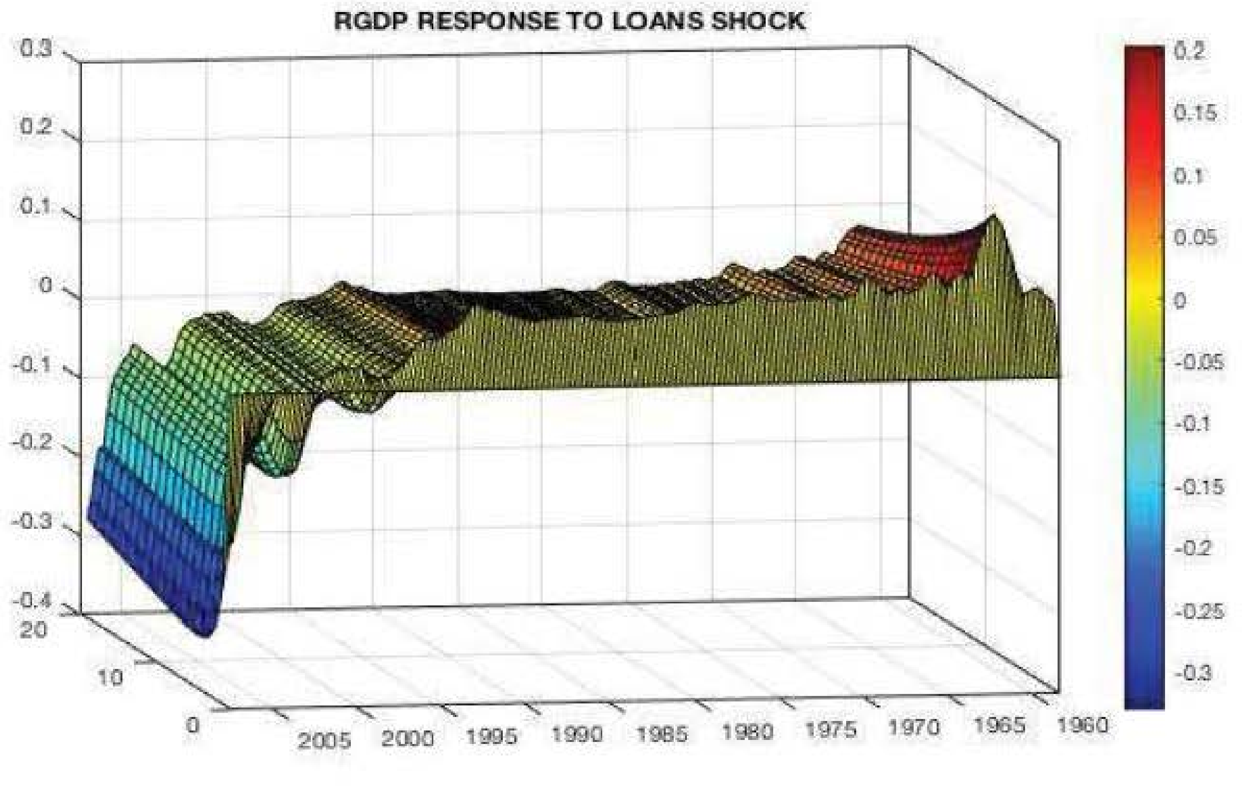

Figure 2.13 - Real GDP impulse response to a LOANS shock. Note: posterior means.

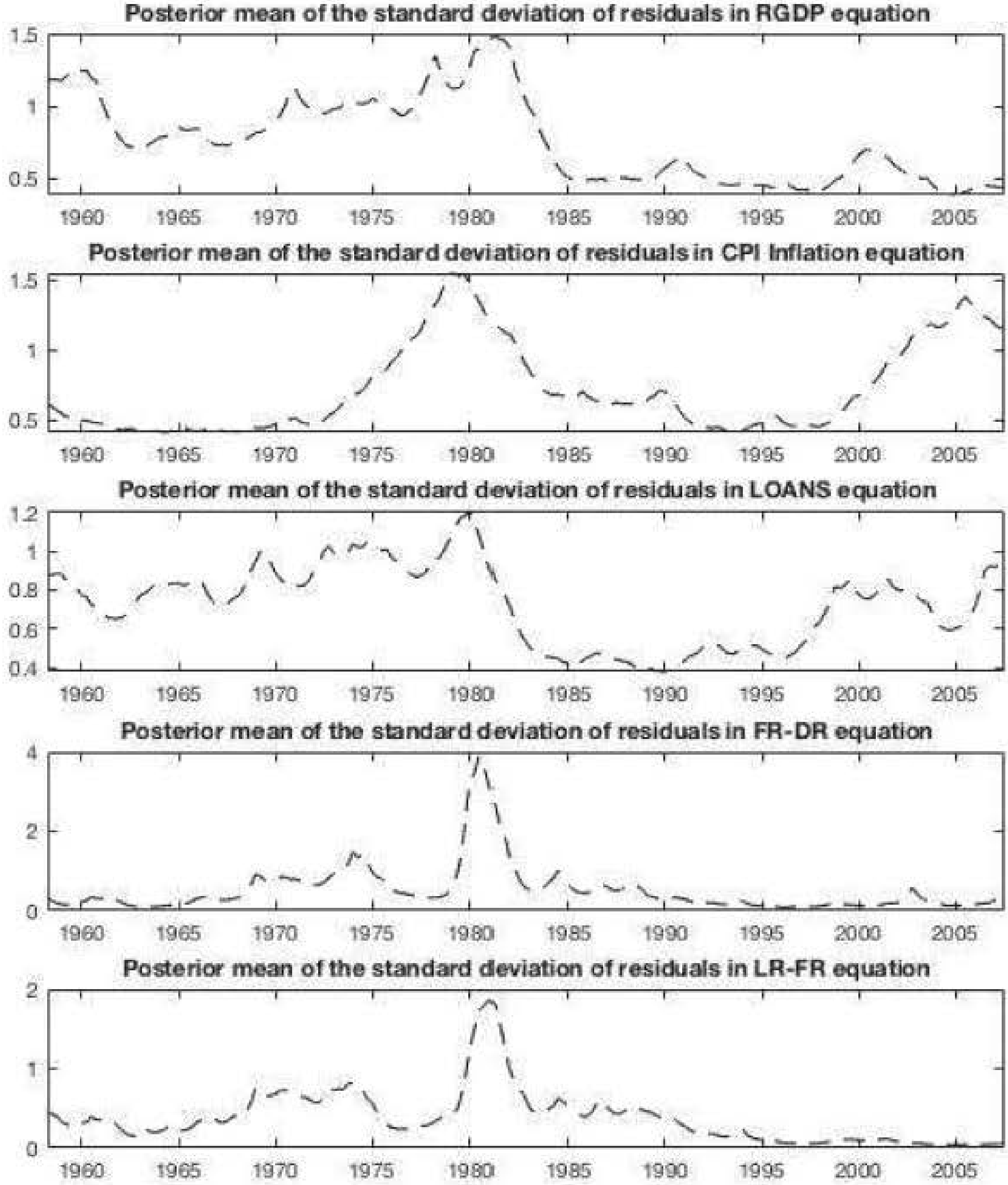

Figure 2.14 - Posterior mean of the standard deviation of the residuals in RGDP, CPI Inflation, LOANS, FR-DR and LR-FR equations respectively.

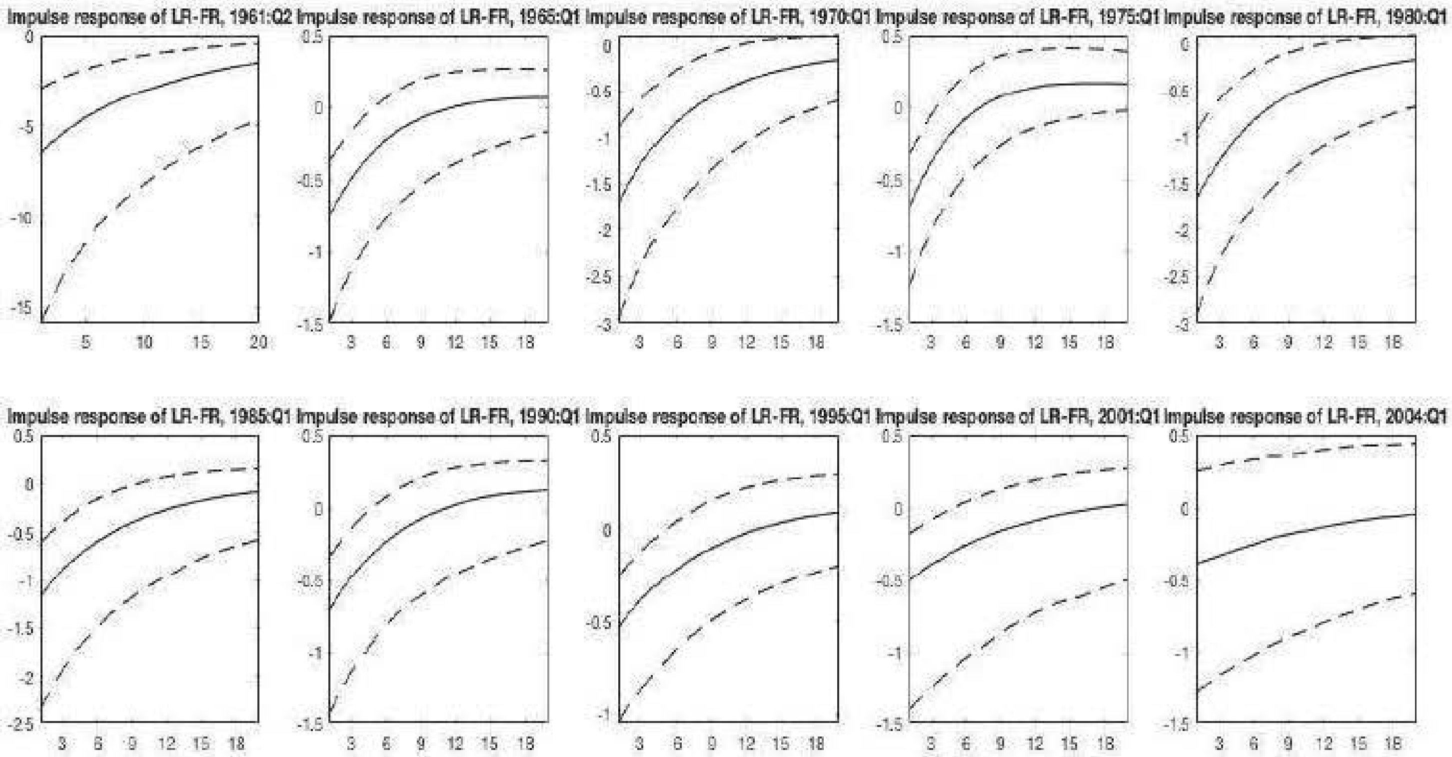

Figure 2.15 LR-FR impulse response to an FR-DR shock. Note: The solid lines depict the 50-th percentile with the 16-th and 84-th percentiles for the dashed lines.

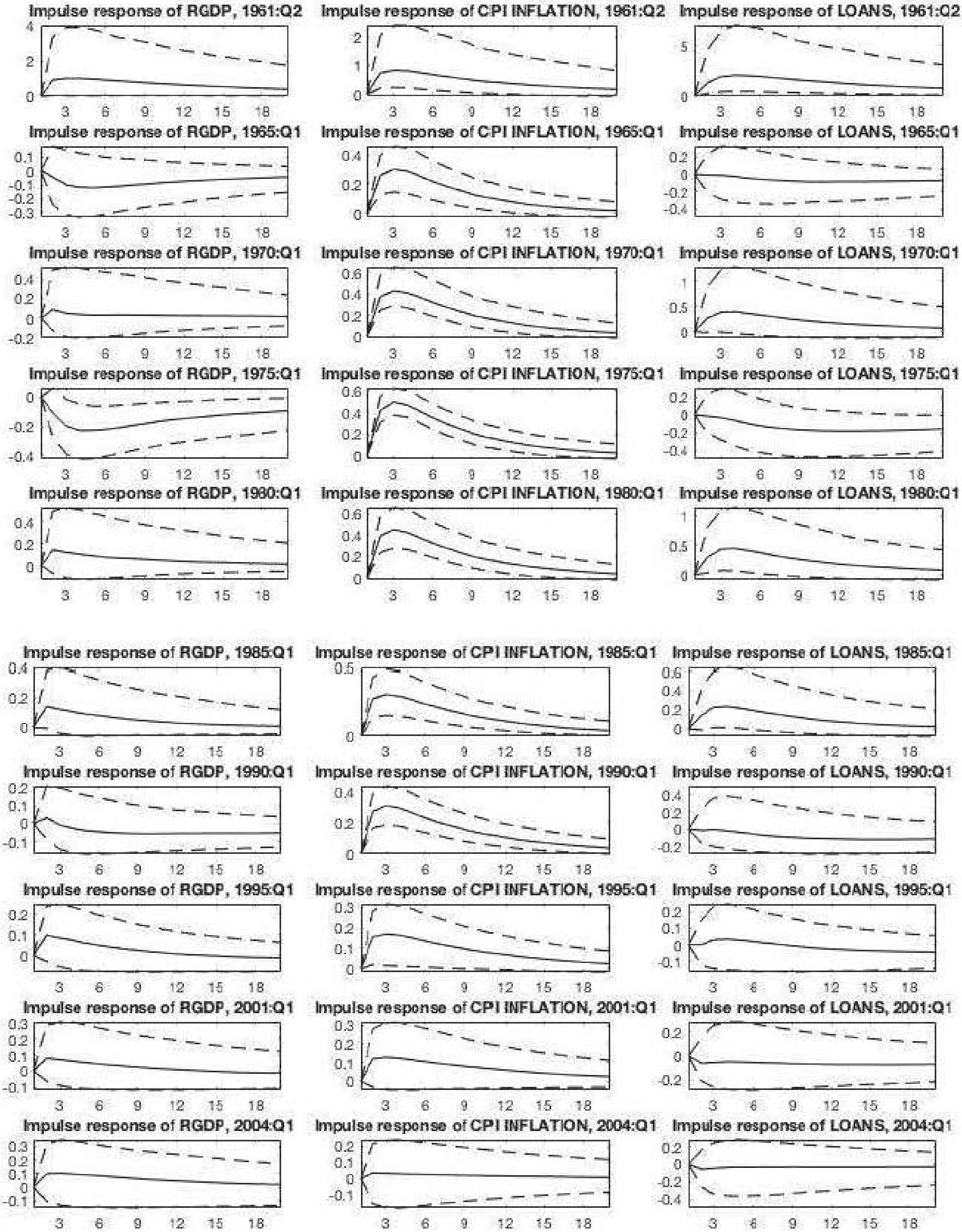

Figure 2.16 Impulse responses to an FR-DR shock. RGDP, CPI Inflation and LOANS in columns 1, 2 and 3 respectively.
Note: The solid lines depict the 50-th percentile with the 16-th and 84-th percentiles for the dashed lines.

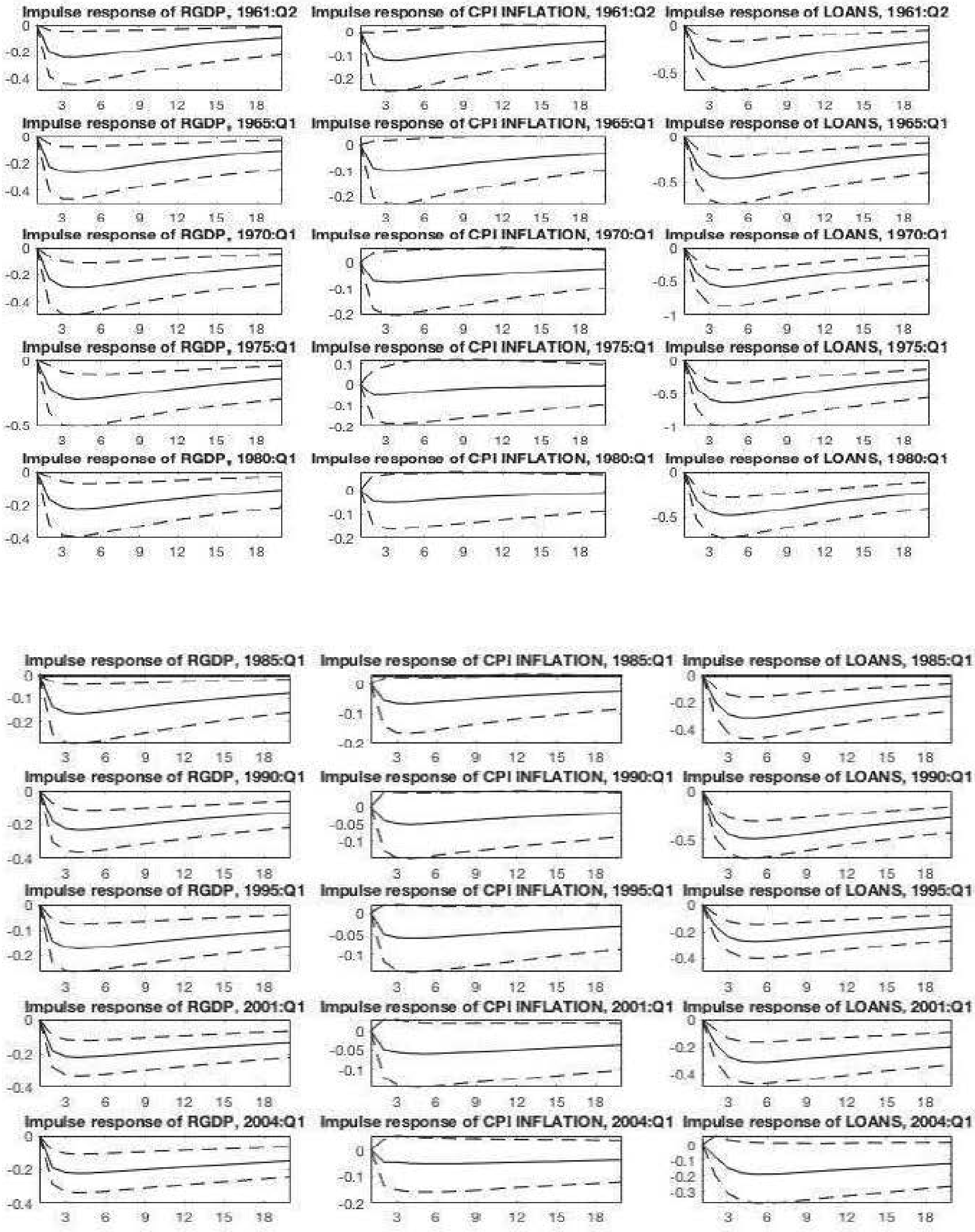

Figure 2.17 Impulse responses to an LR-FR shock. RGDP, CPI Inflation and LOANS in columns 1, 2 and 3 respectively. Note: The solid lines depict the 50-th percentile with the 16-th and 94-th percentiles for the dashed lines.

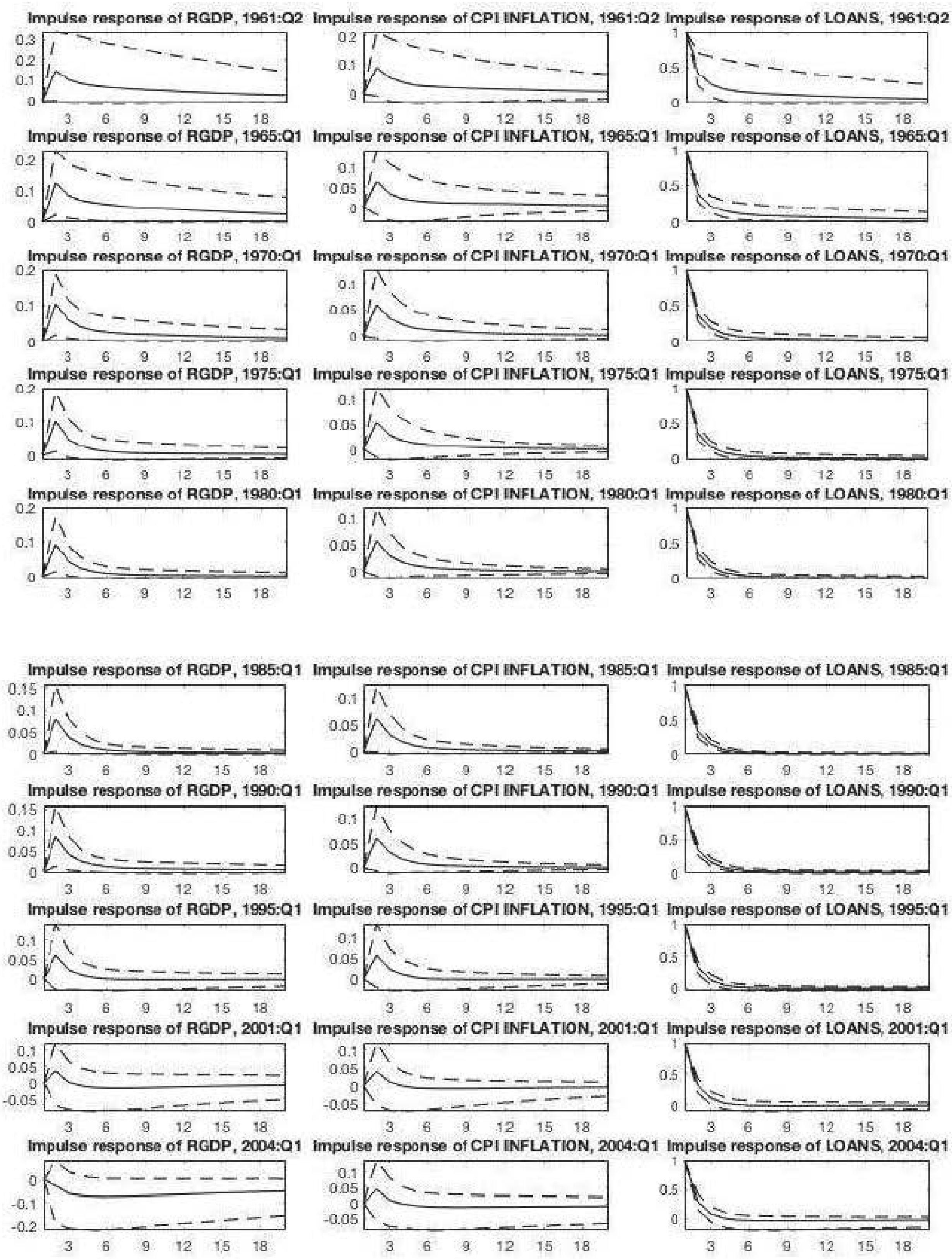

Figure 2.18: Impulse response to a LOANS shock. RGDP, CPI inflation and LOANS in columns 1, 2 and 3 respectively. Note: The solid lines depict the 50-th percentile with the 16-th and 84-th percentiles for the dashed lines.

Given that a shock to the spread LR-FR has a no significant inflation response, unlike the shock to FR-DR, but the opposite is found for the loans and real GDP, the same model was estimated dropping the variable FR-DR, in order to evaluate whether LR-DR has no power to decrease inflation, or it is simply that FR-DR is stronger in capturing the inflation behaviour. The results (not presented) show that a shock to LR-FR has a negative and significant impact on inflation. However, the real GDP response turns no significant for most of the periods. At the same time, a shock to loans compensate that lost of significance, as the real GDP responses are significant almost for every period, and the significant responses last slightly more than those in the baseline model.

3.5 What lessons can we learn?

The results confirmed the proposed mechanism operating between the Federal Reserve's policies and the banking sector, whereby when the Fed's influence on the reserves cost opens the door for arbitrage opportunities, banks obtain their reserves from the cheapest source. The implication of this phenomenon is that Fed's raises of the federal funds rate, due to the cheaper cost of the reserves obtained at the discount window, has an impact on the loans rate below the ratio 1:1. Consequently, the demand for credit is restrained less than the Fed's policies intend, triggering higher inflation. According to the model, while a smaller spread LR-FR has a positive and significant impact on loans and output, the variable having a positive and significant influence on inflation is the one capturing Fed's policies, namely, a positive spread FR-DR. Therefore, the Fed should include as a target the tracing and control of the reserves cost, in order to have full control over monetary policy. This statement confirms that monetary policy has its impact through prices and not quantities, despite the quantities influence those prices, as claimed in Chapter 2. These results answer the unknown regarding the switching regime around 1990 also commented in Chapter 2. Indeed, that regime switch belongs to a change in the banking sector's behaviour when the loans rate was pegged to the federal funds rate. Its confirmation is based on the fact that the results analyzed are mostly homogeneous, without presenting any switching regime. It is the consequence of including the adequate variable in the model, namely, the spread LR-FR that captures the banking sector's behavior. In relation to this, it is important to stress that after 1990, when banks obtained most of their reserves as nonborrowed reserves at the federal funds rate cost, not only the loans rate was pegged to the federal funds rate, but also it was the period when the spread was constant and the largest of the whole period (Figure 2.2). This period coincides perfectly with the decline in the volatility of inflation and output levels that gave name to the known Great Moderation (1990-2007), discussed by authors such a McConnell and Perez-Quiros (2000)[24], Stock and Watson (2003), Moreno (2004) or Gali and Gambetti (2009). In the same line, the Great Inflation (1965-1982), characterized by high volatility in inflation and output, coincides with the smallest and more volatile spread LR-FR. Therefore, the "reserves-cost" theory could also be the explanation for both episodes, because when the federal funds rate impact ratio is not constant, the economy faces a different loans rate, which transmits more volatility into the economy.

However, this theory leaves one case without explanation. It can be perfectly applicable while excess reserves are scarce, because banks will take their decisions based on the interest rates of the moment. Conversely, when they accumulate a significant amount of excess reserves, as for the last years of the sample (and posterior years), monetary policy can be even more powerless. With a significant amount of excess reserves, banks could set their loans rate depending on the rate at which reserves were obtained in the past. A possible consequence is that they could facilitate lending conditions only when it is profitable for them, causing higher inflation levels, as under the already evaluated positive spread FR-DR scenario. Another consequence would be that if past reserves were acquired at a relatively expensive cost, and the current interest rate is lower than that cost (as it is the case now), they will accumulate more reserve[25] to not incur in loses when using them, and consequently, the inflation levels will remain low. In relation to the present situation, after the huge liquidity injection from the quantitative easing, as banks do not need reserves to lend, and their past reserves were expensive compared to the current low interest rates (Figure 2.1), it is not surprising that liquidity injections do not reach the real economy until interest rates rise or the demand for credit increases significantly, what simultaneously would suppose an increase in interest rates.

The last thing deserving attention is that even though this theory seems to explain the high inflation during the 1970s and 1980s, and offers a lesson regarding how policies should be managed to avoid those episodes again, or even why inflation has remained in low levels after the spread LR-FR was almost the largest of the sample, it also shows the scarce

[24] These authors observed a break in the volatility of U.S. output around 1984, what just matches with initial the uptrend of free reserves (Figure 2.3).

[25] Apart from the low interest rates nowadays, the current monetary policy of paying interest rates on required reserves (IORR) and excess reserves (IERR) might be contributing as well.

knowledge possessed regarding the behaviour of the demand for money. The residuals displayed in Figure 3.14 increase around 1980 for all the variable of the model. At the same time, an increase in the residuals for the inflation and loans equations was observed after 1995. While the spread LR-FR was negative and the spread FR-DR positive for the late 1960s and mid-1970s, pointing out that banks were unable to increase the loans rate above the federal funds rate as a signal of low credit demand, that phenomena did not repeat again for the following FR-DR positive spreads. This fact may suggest that the models estimating the optimum federal funds rate for the economy according to Fed's targets were erroneous and the demand for money and inflation could be also influenced by uncontrolled factors not included in those models.

As it is widely known, credit flowed more than it should for the last year of the sample and the consequence was the Great Recession. A reason why hidden or not accounted factors may be driving the demand forces is because our knowledge and measuring tools are insufficient to evaluate and control those forces. Consequently, the required interest rate to lead the economy through a steady path cannot be set accurately. For the case of the measuring tools, the available technology and the own definition of a variable could be the cause of that inaccurate measurement. For example, although the available data for output, output gap or inflation as a sole entity for representing the whole economy may be efficient for a quick and general glance about the state of the economy, that generality is likely to generates blind spots. Then, a sector may be growing faster than others, where prices increase much faster than in other segments of the economy. Thus, for the case of inflation, the standard measure can underestimate the evolution of that sector, and the interest rates set for the entire economy will be wrongly set for that specific case. The consequence will be more credit flowing towards that sector given profit prospects, causing dangerous debt accumulation or bubbles, even though the other sectors are under control given the interest rate targeted.

3.6 Conclusions

This chapter undertook the search of the responsibility for American economy's performance from 1958 to 2007, as well as the consequences of who owns it, under the proposition of a new theory that explains the mechanism working between the banking sector and the Federal Reserve's policies. The so-called "reserves-cost" theory proposes that the Fed is able to have an impact on the real economy, directly, by its influence on the cost of reserves, and indirectly, by the impact of that cost on the loans rate set by banks. For that task, a TVC-BSVAR was estimated, given its adequacy for capturing smoothed regime changes over time. The results have confirmed the mechanism proposed by the theory, showing that the Fed, by not targeting the cost at which reserves were obtained, granted the monetary policy steering wheel to the banking sector, causing episodes such as the Great Inflation. However, while the banking sector could be considered as the direct responsible for the American economy's performance for such episodes, the first responsibility falls on the Federal Reserve, as it is the one able to control the reserves cost mechanism.

-- Chapter 4 --

Why do banks accumulate reserves and how do they influence monetary policy? An explanation from the “reserves-cost” theory

Low interest rates, risk, uncertainty, or low demand for loans have been proposed as some of the most important factors determining the accumulation of excess reserves in the American banking sector during the Great Depression and the Great Recession. Estimating an SVAR for the period 1922-2017 and using the new “reserves-cost” theory I complement some of the factors aforementioned, and propose that when banks are holding reserves obtained at a higher rate than the present short-term rate, banks will hoard reserves, in order to not use those acquired at higher cost and incur in loses. At the same time, the cost of those reserves will determine how banks set the loans rate in relation to the short-term rate. This mechanism is important because it implies that Federal Reserve’s policies can be transformed by the banking sector, and their impact on the real economy may be far from the desired.

- *JEL classification: E43, E51, E52, E58*
- *Keywords: monetary policy, Federal Reserve, SVARs, excess reserves, reserves cost*

4.1 Introduction

After the Great Recession, comparisons arose between the causes and consequences of this unfortunate episode and the Great Depression. One of them was the hoard of excess reserves in the banking sector. At the end of 1940 in the United States, banks' holding of excess reserves was around seven billion dollars (about 6.8% in relation to GDP), while in the last quarter of 2014 they were holding around 2,700 billion dollars (almost 15.4% in relation to GDP). After those dates, these balances initiated their decline. For the first case, those levels reached a minimum at the beginning of the 1970s. For the second case, it is still unknown when it will happen. Were we able to completely comprehend the reasons and mechanisms behind the accumulation of reserves, its consequences could be anticipated. The reason why the consequences must be anticipated is because of the impact that reserves have on the real economy. To understand that statement, it is necessary to know what is the key point for the theory exposed later; that banks set their loans rate conditioned on the cost of their reserves. To the question "what does determine the prime loans rate?" sent by email to the Federal Reserve "Contact Us" option, it was answered: "The prime loans rate is a rate established by commercial banks as a lending rate or base off which their commercial loans are priced. In other words, the banks set their own rates based on the demand for various kinds of loans, on the cost of money to the banks, and on the administrative costs of making loans…" That "cost of money" is the price at which reserves are obtained. As the loans rate is the rate that the real economy actually endures and it is connected to reserves, that accumulation of reserves will determine, to some extent, and through certain mechanisms, the path of the American's economy performance. Therefore, three questions must be answered. What determines the levels of excess reserves? What are the consequences of those levels for the real economy? And from these two questions arise the third one, what can central banks do about it, if total control and responsibility to lead the economy through an optimum path is expected from them?

For the first question, three main reasons have been highlighted. First, low interest rates prevent banks from investing in alternative assets. Second, financial shocks, what could gather also uncertainty and risk factors, lead banks to accumulate precautionary reserves against reserves withdrawals or liquidity scarcity in the market. Third, oversupply of reserves and low credit demand. Regarding the accumulation of reserves for the 1930s, Frost (1971) defended what he called the "adjustment-cost" hypothesis, whereby banks hold substantial amounts of excess reserves at low interest rates because brokerage costs, commissions, spreads…prevent them from adjusting their reserve position, as these costs are greater than the interest earned on short-term securities. Thus, the demand function for excess reserves is kinked at a certain rate, when it is lower than the costs aforementioned. Lindley, Clifford and Mounts (2001) and Dwyer (2010) also claimed a negative relationship between interest rates and excess reserves. However, the former authors added that an increase in deposits decreased excess reserves. This claim is contrary to the risk factor, as banks would increase their holding of excess reserves, as a preventing measure before the possibility of deposits outflows. This result is outstanding, as the period they analyze underwent an increase in deposits and excess reserves. After new estimations, they reached the conclusion that banks received reserves which were unable to convert into income-earning assets given the low demand for credit. Subsequently, they had to hold unintended reserves balance. Wilcox (1984) refuted Friedman and Schwartz's (1963) claim that excess reserves increased as a consequence of the increase in reserve requirements (or other shocks such as bank runs), because despite it may be able to explain the initial hoarding of excess reserves in the early 1930s, it cannot explain the greater pile up after 1934 and 1937. Instead, low interest rates seemed to be a more powerful factor in explaining those increases in excess reserves. Also, Calomiris, Mason and Wheelock (2011) by using microeconomic data on Fed member banks, stated that excess reserves did not increase between June 1936 and June 1937 because banks curtailed credit as Friedman and Schwartz's (1963) claimed. On the contrary, banks increased lending by 750 million dollars, but reduced their holding of government securities by 1,100 million dollars. They concluded, "as bank profits and loan opportunities increase, and as macroeconomic risk recedes, banks will reduce excess reserves to finance loan expansion." The increase in lending accompanied by the decline in government securities aforementioned could be explained by Cagan (1969)'s claim that when the spread between the discount rate and the short-term rate of the funds market was positive, banks preferred to borrow at the discount window. When the spread was zero or negative, as it was the case for the 1930s, banks sold securities to finance the expansion in loans. In general, he argued that the accumulation of excess reserves in the 1930s was due to the cost of investing in short-term securities, supplemented by the lack of demand for loans and the risk of investing in long-term securities. At very low yields on short-term securities, the transaction costs of buying and selling may equal or exceed the return and therefore, idle reserves will remain with banks. Todd (2013), regarding the connection between credit and excess reserves, claimed that the accumulation of excess reserves in the 1930s went away only once banks were offered government-guaranteed lending alternatives, which funded defense production programs. Thus, in comparison to the Great Recession episode, he stated that the quantitative easing (QE) aimed to encourage banks to ease the terms of credit was not accomplished and excess reserves were hoarded. The reason is that while commercial and industrial loans had increased 13.6% from August 2007 to April 2013, the monetary base, as a consequence of the monetary stimulus, had increased 367%. This belongs to the third factor commented above, namely, oversupply of reserves in comparison to credit demand. This author also claimed that the new policy tool, which is the interest rate paid on reserves by the Federal Reserve, is encouraging banks to retain excess reserves.

For those analyzing the accumulation of reserves for more recent periods, Dow (2001) estimated the demand for excess reserves for the 1990s and obtained that one percentage point increase in the federal funds rate decreased excess reserves in $120 millions, while an increase in deposits of one billion dollars increased excess reserves in three million dollars. Regarding the years around the Great Recession, the debate for the factors behind the hoarding of excess reserves hardly varies. Chang, Contessi and Francis (2014), analyzing bank-level data for commercial banks and saving institutions found that banks accumulate excess reserves when there is a deterioration between capital adequacy and loans loss provisions (what could be classified as precautionary motive due to weak balance sheet), when the opportunity cost of holding low-interest-bearing assets is low, and when the penalty for holding insufficient reserves increases. However, measuring economic uncertainty by the volatility index (VIX) and industrial production index (IPI) variance, the estimations showed that the uncertainty factor did not influence the level of excess reserves. Ennis and Wolman (2015) found that banks did not substitute reserves for liquid securities, but they complemented them contributing to increasing bank liquidity, partially, because the Fed is paying interest rates on reserves. Also, the increase in reserves did not pressure insured banks' balance sheet capacity, because the Fed's purchase programs flooded with reserves, mostly, those banks with abundant capital. Therefore, banks could have lent without any pressure on their capital ratios. Ashcraft, McAndrews and Skeie (2011) analyzed the daily behavior of large and small banks in the federal funds market form 2002 to 2008, and observed that the reluctance to lend and the desperation to borrow for 2007 and 2008 triggered more volatility (extreme spikes and crashes) in the federal funds rate and accordingly, banks held more precautionary reserves.

Japan underwent the same episode of reserves accumulation during the 1990s. Ogawa (2004) also pointed out to low interest rates and precautionary reasons as explanatory factors.

While he highlighted an important increase in excess reserves after the bankruptcies of financial institutions in 1997, that increase was overshadowed by the rise of excess reserves levels in the second half of 2000, when the Bank of Japan undertook its QE. Again, banks seemed to be forced to hold unintended reserves. That is, the supply side was the main contributor to those accumulation levels. For the case of the ECB, Bindseil, Camba-Mendez, Hirsch and Weller (2006) supported the already commented negative correlation between excess reserves and interest rates. They also argued that excess reserves are not the mechanism whereby banks can expand loans and create inflation. However, it was highlighted that if the level of excess reserves is not forecasted correctly, open market operations can induce more volatility on short-term rates and make monetary policy transmission more noisy and inefficient. Goodhart (2010) summarized all the factors behind the accumulation of excess reserves into risk of lending, capital ratios, regulation, lower demand for loans, interest rates paid on reserves, increasing public deficit and almost zero interest rate in public sector debt.

Regarding the second question about the consequences of hoarding excess reserves, the most feared has been an inflation outburst as commented in Meltzer (2009) and Plaser (2011) to name someone. However, there is strong opposition against that belief. Martin, McAndrews and Skeie (2016) developed a model for the U.S. banking system, where interest rates are paid on reserves, in order to shed light on the debate about the potential consequences on credit and inflation of the current level of excess reserves. They showed that lending was not related to the amount of reserves, because the key determinant of lending is the expected profits between the return obtained on a loan and the opportunity cost of it. "Banks expand their balance sheets so long as the marginal cost of funding is less than the marginal return on bank lending, abstracting from credit and liquidity risk" (p. 195). According to Bindseil (2004) the inflation fears have no foundation and they come from the erroneous "Reserve Position Doctrine" which has been in textbook for decades. Under an interest rate target, banks do not need reserves to lend. Banks first lend, if its profitable for them, and then, obtain the necessary reserves. Therefore, the result of an injection of reserves to make banks expand credit will result in a drop of interest rates to zero (if there is no deposit facility), and once this occurs, the money multiplier should fall with every further reserves injection, as the amount of reserves provided overcomes the demand for credit. Moore (1998) had clearly exposed that argument previously as reflected in the following statements:

> "So long as reserves with the central bank do not bear interest, and interest rates are sufficiently positive to cover marginal costs, profit maximization alone implies that individual banks will have an incentive to lend out any reserves in excess of their legally required minimum ratios" (p. 372).

> "Loans make deposits. […] Increases in bank loans are made at the initiative of bank borrowers, not the bank themselves. Bank may unilaterally increase their advertising budgets, shade their lending rates, or ease their collateral requirements. But as with any other business, the amount of good or service they can sell depends ultimately on the demand for their product" (p. 373).

> "…while the Federal Reserve can directly determine de quantity of nonborrowed reserves, it cannot directly determine de quantity of total reserves" (p. 374).

> "The Federal Reserve has no choice but to accommodate and provide all increases in required and excess reserves demanded, albeit at a supply price (interest rate) of its own choosing." (p.380). Last, "The money supply is endogenously determined by market forces" (p.381).

The third question regarding what central banks can do, has been answered by paying interest rates on excess reserves, either by building floors for the interest rate levels or corridors, as explained in Goodfriend (2005), Keister, Martin and McAndrews (2008), Bernanke (2010) and Bowmman, Gagnon and Leahy (2010). According to these authors, it can be considered as a measure to separate the control of the interest rate from the amount of reserves. In that way, the Fed can inject liquidity into the market when there are liquidity problems like in September 2001, without altering the policy about interest rates. It also helps the Fed to implement more efficiently its policies, as otherwise, the volatility in the federal funds rate and reserves in the funds market, makes difficult to achieve the targeted level for the federal funds rate. Last, the Fed can control the level of excess reserves as banks are receiving interests on those reserves and they will not look for alternative assets unless they offer a higher return.

In this chapter, the main goal is to answer the enigma about the accumulation of reserves to extract some insight about its consequences. This will shed light on which actions may be necessary to undertake, so that monetary policy has a direct impact on the economy and lead it through the desired path, without undergoing any transformation on its way through the banking sector. However, I depart from the standard factors enumerated above and develop a new explanation, which can be considered mainly based on the interest rate factor. This new explanation is an extension of the "reserves-cost" theory explained in Chapter 3. That theory claimed that the Fed can control the real economy, directly by influencing the reserves cost, and indirectly by the impact of that cost on the loans rate. While that theory focused only on the reserves cost at current market rates, it did not take into account the cost of the reserves accumulated through time. In the next section, I explain the creation of a new variable that measures the reserves cost by way of inventory. That variable is important to explain the reason behind reserves accumulation and its impact on the real economy. It measures the average cost of the reserves held at every period and represents, depending on its position regarding short-term rates, the price at which banks will decide either to use their excess reserves or obtain them form other source. When banks are holding reserves with a cost higher than the current short-term rates, they will not use the reserves held and borrow at the current short-term rate, which is cheaper. In this way, they obtain the highest profit from lending, or at least, they do not incur in losses by lending at lower rates than the cost of their reserves. On the contrary, when the short-term rate of reference is higher than the cost of the reserves held, excess reserves will diminish as banks obtain more profits by using those cheaper reserves when lending. Also, this variable is able to determine the evolution of the spread between the short-term rate of the money market and the prime loans rate. The higher the difference between the cost of reserves held and the short-term rate, the higher banks will set the loans rate in relation to the short-term rate of reference, in order to diminish the risk of incurring in loses when lending. Thus, this variable is also relevant for explaining how banks' decisions about the loans rate, conditioned on monetary policies, will influence Fed's final targets, such as inflation or output. These results have been obtained by estimating a structural vector autoregressive (SVAR) for the period 1922-2017.

The chapter is structured as follow. Section 4.2 develops the extension of the "reserves-cost" theory. Section 4.3 describes the variables and the model estimated. Section 4.4 displays the results obtained and their interpretation. Section 4.5 analyses the new policy of paying interests on reserves and potential consequences of hoarding reserves. Last, section 4.6 summarizes the main results and its implications.

4.2 The “reserves-cost” theory. Part II

The “reserves-cost” theory developed in Chapter 3 can be briefly explained as follows:

Banks can obtain borrowed and nonborrowed reserves at their corresponding cost, which may differ or not. Borrowed reserves can be obtained at the discount window at the cost of the discount rate. For the case of nonborrowed reserves, the source from where banks obtain them is the federal funds market, where the cost is the federal funds rate, or the open market. When banks sell securities to obtain reserves, they are renouncing to the interest rate paid on those securities. The cost (or opportunity cost) of getting those reserves is the interest rate not received. Minimum, the cost of those reserves will be the 3-months T-bill rate, which in general, is around the federal funds rate levels. If the term of those securities is longer, the cost will be higher. Hence, it can be considered that the cost for borrowed reserves is the discount rate and the cost for nonborrowed reserves is the federal funds rate. As profit-maximizing agents, banks will prefer to obtain as much reserves as possible (as long as they are needed or desired) from the cheapest source. When there are positive spreads between the federal funds rate (or other short-term rate, depending on the period under analysis) and the discount rate, banks will borrow at the discount window. If this spread is zero, banks will be indifferent from where obtain reserves. When the spread is negative, banks will avoid the discount window, unless it is strictly necessary.

As mentioned in the introduction, banks lend and afterwards, obtain the necessary reserves for fulfilling the necessary required reserves and their desired excess reserves levels. To lend, banks will avoid to incur in loses. Therefore, they adjust previously the loans rate according to the reserves cost in the money market at that moment. As the loans rate is the rate that the real economy endures, the theory proposes that the Federal Reserve can have an impact on the real economy only through how its policies affect the reserves cost (discount rate and short-term rate of reference), because this cost will determine the loans rate and consequently, the demand for money in the real economy. One must be aware that as the Fed’s main policy is to control the federal funds rate, considered the short-term rate of reference, if positive spreads are allowed between the short-term rate and the discount rate, the Fed will be losing power to modify the performance of the economy. The reason is that banks will arbitrage as consequence of the different costs and will be able to increase the loans rate less in relation to the raises in the federal funds rate. Consequently, policies intended to tighten the economy, will not be accomplished, or at least not to the extent desired.

However, this theory was incomplete. It claimed that the cost of reserves could determine the loans rate depending solely on current interest rates (static reserves-cost theory). This argument is reasonable for those periods with scarce excess reserves, because the different sources of reserves commented above would have to be used at the rate that they offer at a particular moment. Nonetheless, when banks accumulate excess reserves, they do not need to obtain reserves to back the demand for loans. As those reserves were obtained at past rates, the interest rate of the moment would lose its influence on the loans rate for that period. It would be the cost of past reserves along with current rates what would determine the loans rate. This reasoning implies that if banks are holding reserves with a higher cost than the loans rate that can be applied in relation to the other market rates, they will retain those reserves until interest rates are higher, so as to not incur in loses. Likewise, if the cost of reserves held is above current rates they will decide to obtain more reserves at market rates to make more profits, rather than using their relatively more expensive reserves. Therefore, it is probable that they accumulate more reserves under those circumstances.

Once the “reserves-cost” theory has been renewed (dynamic “reserves-cost” theory), it is time to prove that it is able to explain the evolution of excess reserves and influence the loans rate in relation to short-term rates. The consequences of that relationship on the real economy are not the goal of this chapter, because they were already examined in Chapter 3. The results showed that smaller spreads between the prime loans rate and the federal funds rate triggered a higher demand for credit and increases in inflation, even when the federal funds rate was raised.

In order to visualize how the cost of reserves evolves, a new variable has been created as follows. Starting from 1919, when the level of excess reserves was near zero, the difference between the quantity of borrowed reserves held in period t and t-1 is calculated at each period of the sample. The same procedure is applied to nonborrowed reserves. When the difference for borrowed reserves is positive, that quantity is registered at its cost, namely, the discount rate of that period, by way of inventory. The same process has been carried out for nonborrowed reserves, but the interest rate applied is the short-term rate of reference (weighted average of open market rates for New York, 3-months Treasury bill or federal funds rate, depending on the period)[26]. When the difference between periods for any type of reserves is negative, that quantity is removed from the lowest interest rate where reserves were registered in the inventory. In that way, banks maximize profits by using the largest

[26] Appendix E

difference between the reserves cost and the current loans rate. Once reserves are added or removed from the inventory at each period, each quantity of reserves remaining and registered until that period is multiplied by its cost. Thereafter, I calculate the average cost. As result, I obtain the average cost that banks have paid in period t for one dollar of reserves. This procedure has to be repeated for every period. The new variable is called reserves cost (RC). Yet, it is far from accurate, because apart from accounting the reserves obtained at each period with a standard and simple method, which obviously is not exact, if banks are trapped holding reserves at higher costs than the interest rates of the moment, the image of the inventory becomes distorted once cheaper reserves are accumulated, as they drag down the RC average value. In the end, there would not be any clue regarding the level at which banks would be free of using all their reserves without incurring in losses. Therefore, to complement RC, I have also created a variable measuring the maximum interest rate (either the short-term rate of reference or the discount rate) at which banks obtained reserves until period t. Thus, this variable (MAX[STR], henceforth) measures the level at which banks could be trapped with expensive reserves.

These two variables are plotted in Figure 3.1 along with other variables of interest. The green line represents the logarithm of excess reserves as percentage of required reserves (ER). The graph is cut above for the sake of the explanation; therefore the ER line disappears after 2009 given the high levels since then. ER increased after the Great Depression and initiated their decline around 1940, reaching its minimum around the mid-1970s, when they increased again. The increases corresponding to the Great Depression and the Great Recession can be mostly explained by Figure 3.2, where the level of nonborrowed reserves is displayed as percentage of loans. This variable is able to measure when the Fed over exceeded the supply of reserves, given the demand for loans. Despite the common believe that the injection of reserves was greater during the Great Recession, it is observed that in relation to lending, the amount of nonborrowed reserves provided during the Great Depression was twice the levels of the Great Recession. Returning to Figure 3.1, the violet line is the short-term rate of reference (STR), the red line is the prime loans rate (LR), the dark blue line is RC, and the light blue line is MAX[STR]. Finally, the orange line shows the difference between the loans rate and MAX[STR] (LR-MAX STR henceforth).

The graph displays three facts that seem to support how the "reserves-cost" theory can explain the accumulation of reserves:

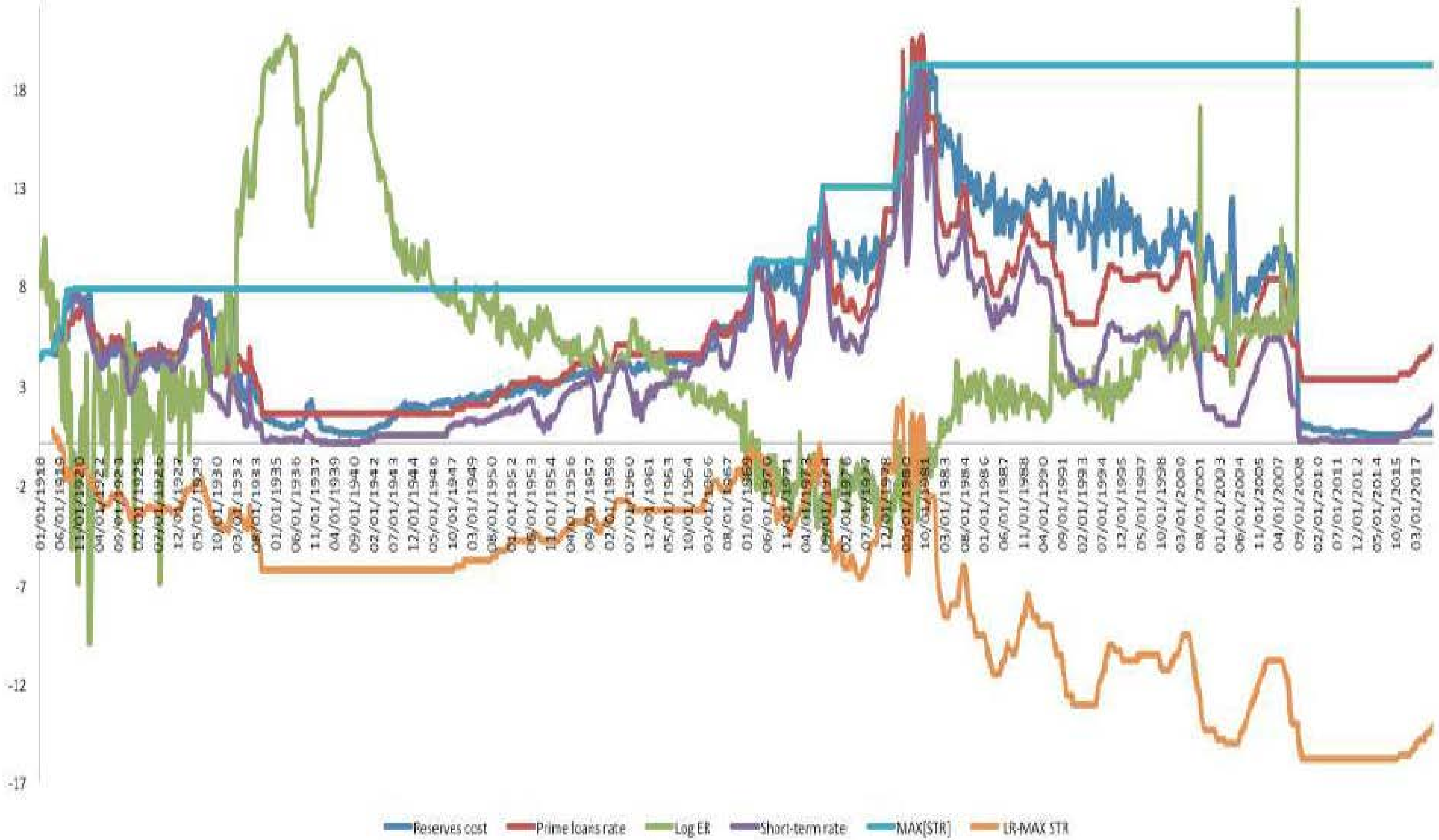

Figure 3.1 - Data Source: FRED and FRASER

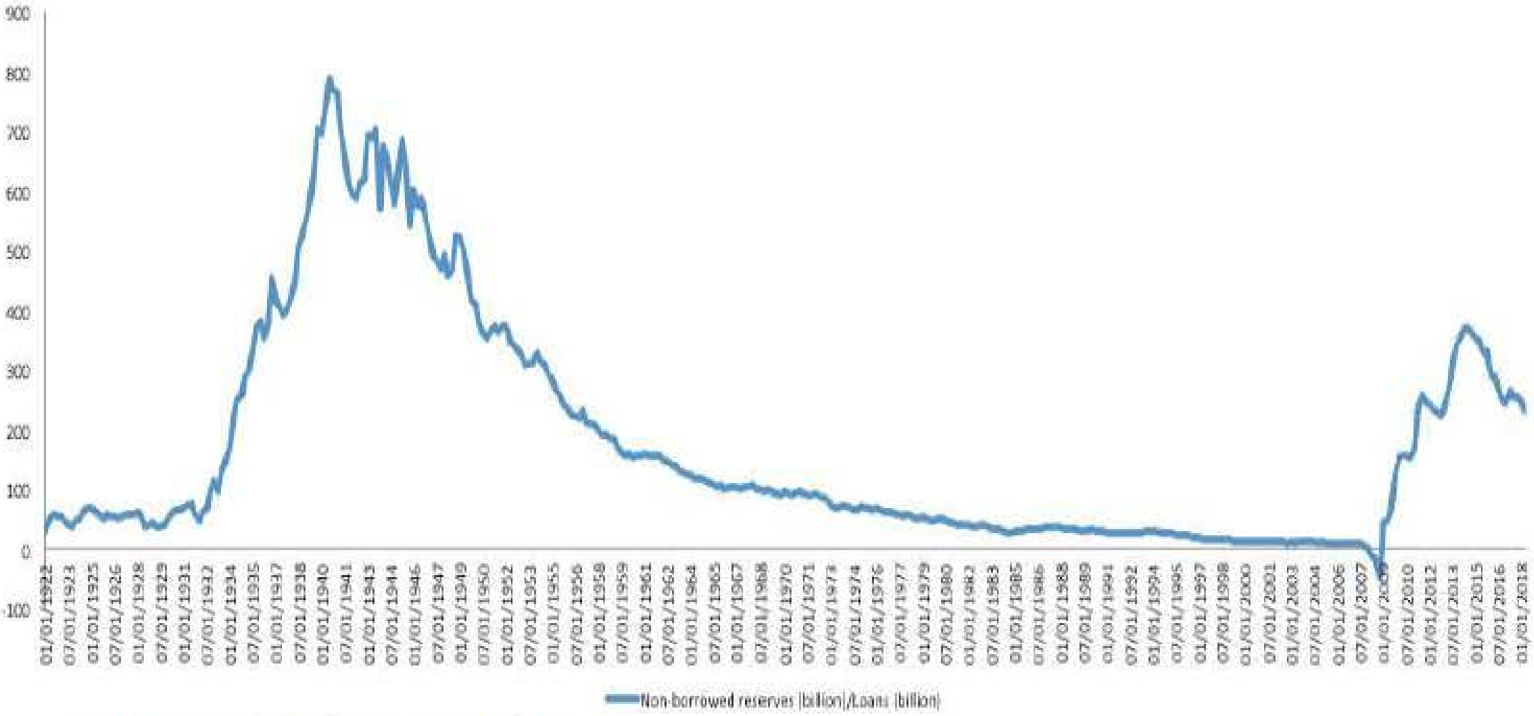

Figure 3.2 – Data Source: FRED and FRASER

- ER seems to increase every time after the short-term rate turns down from its local or global maximum (except in 1970). After reaching a minimum, when interest rates rise again, even if it does not reach the last maximum, the amount of ER diminishes. This fact is in line with some authors' arguments reviewed in the literature who highlighted the importance of interest rates in determining the evolution of excess reserves.

- After a maximum in the short-term rate, RC stays mostly above the prime loans rate (except for the periods around the Great Recession and Great Depression, as the average cost is dragged down given the large supply of reserves at low cost). That means that unless banks keep the reserves already in possession and obtain further reserves from other sources, they may have to assume losses by using those with a higher cost than the current loans rate. Also, they would obtain lower profits than using reserves at current market rates.

  Once the federal funds rate reached its maximum in the beginning of the 1980s, RC has been above the loans rate. Therefore, it seems that banks have been accumulating more and more reserves, being unable to use those obtained at higher rates.

- There is a clear negative correlation between the spread LR-MAX STR and ER (There is an exception for the period 1969-1973). Therefore, it seems that when banks are trapped with expensive reserves, the further is the loans rate from that critical level to get rid of them, the more reserves are held.

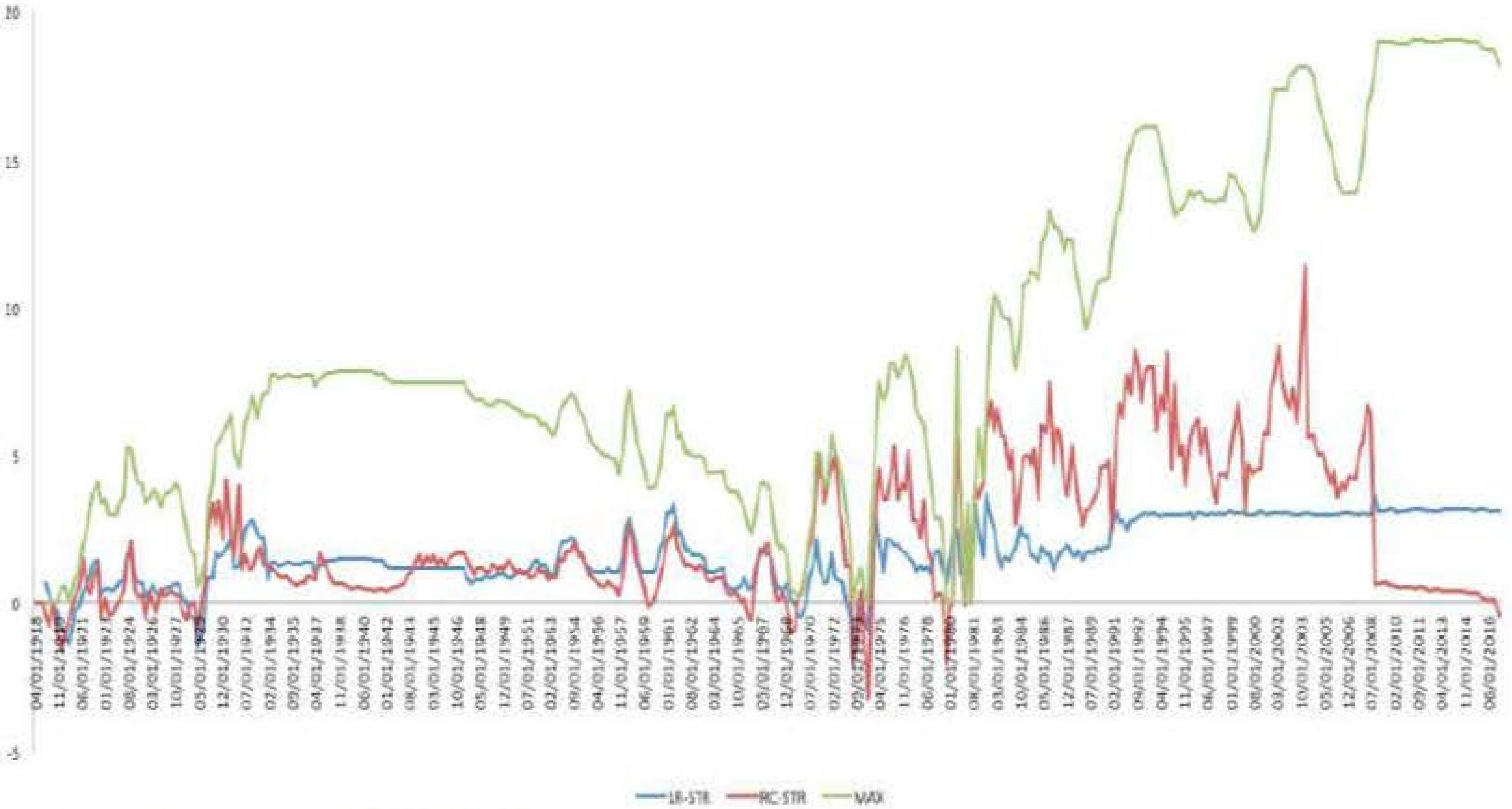

Figure 3.3 - Data Source: FRED and FRASER

Figure 3.3 shows how the "reserves-cost" theory may be also able to explain the evolution of the loans rate. The graph displays the spread MAX (green line), which is the difference between MAX[STR] and STR, the spread between the RC and STR (RC-STR, red line), and the spread between the loans rate and the short-term rate of reference (LR-STR, blue line), already used in Chapter 3. This graph displays two important facts:

- The spread MAX is positively correlated with LR-STR. This fact seems to indicate that the larger (smaller) is the spread MAX, the less (more) likely is that banks can use the reserves obtained at those maximum levels. Subsequently, they will increase (decrease) the loans rate in relation to the short-term rate of reference, in order to diminish the probability of incurring in losses, and spend those reserves as soon as possible when interest rates reach nearby levels.

- There is also a positive correlation between the spread LR-STR and RC-STR. It means that the further (nearer) is the average reserves cost from market rates, the higher (lower) banks set the loans rate in relation to the short-term rate.

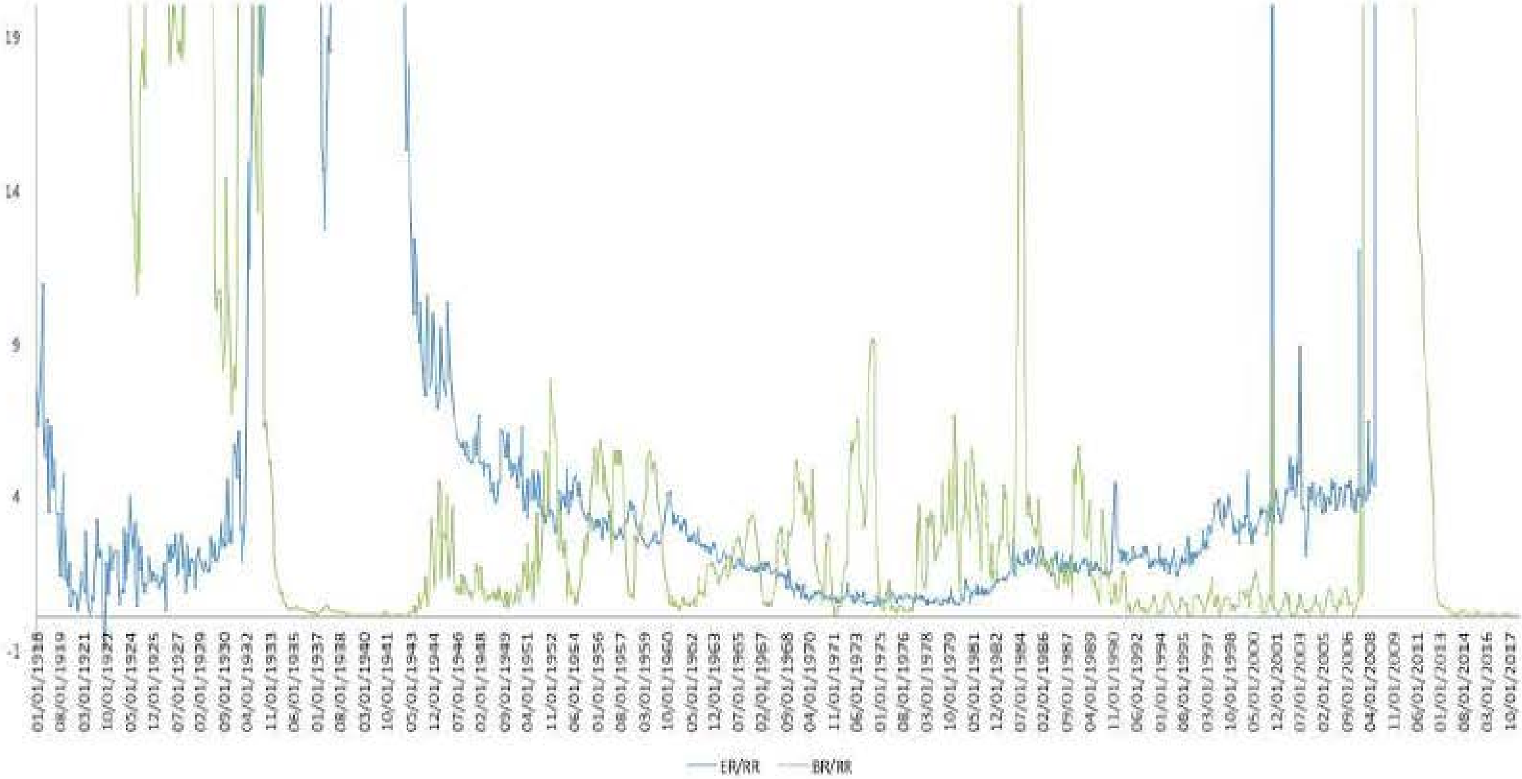

Figure 3.4 - Data Source: FRED and FRASER

The last evidence in favor of the "reserves-cost" theory is observed in Figure 3.4, where excess reserves (blue line) and borrowed reserves (green line) as percentage of required reserves are displayed. Again, the graph is cut above because the amount of reserves during the Great Depression and Great Recession. It is observed that despite the positive excess reserves levels after the 1930s, once the demand for credit increased as commented in the literature, positive borrowed reserves levels arise. Why would banks borrow reserves when they are already holding enough amounts of them? Despite the spikes found in 1984 and 2001, due to exceptional risky situations, one reason could be that they would incur in losses if they spend some of the reserves held. For the end of the 1960s and 1970s, borrowed reserved increased because excess reserves levels were low and the Fed allowed positive spreads between the federal funds rate and the discount rate. Banks, therefore, arbitraged and obtained cheaper reserves at the discount window.

Once some evidences have been showed, the “reserves-cost” theory can be revamped as follows: the only way whereby the Federal Reserve can have an impact on the real economy is through the manipulation of the cost at which banks obtain reserves. This cost will steer the loans rate, which is the actual rate that the real economy endures. If the level of excess reserves is scarce, just modifying the federal funds rate (or other the short-term rate of reference) and the discount rate will have a direct impact on the cost of reserves held by banks, as long as the spread between these rates is closed. In that case, the movements of those rates are likely to be reflected one to one into the loans rate. If the levels of excess reserves are significantly above the levels of required reserves and precautionary factors, the impact of interest rates aforementioned on the loans rate will be proportionally diluted to the quantity hoarded, and more aggressive policies and longer time will be necessary to drive the real economy through the desired path. The reason is that banks’ decisions about the loans rate will make it to evolve differently in relation to short-term rates, conditioned on the reserves cost. Consequently, there will be three hands behind the steering wheel of monetary policy. That is, past Fed’s policies, banks’ decisions and present Fed’s policies.

As clarification, although uncertainty or risk shocks hitting the economy at a specific time will contribute to accumulating reserves for a while, and although at some point variables controlling those factors will be included in the model, the theory exposed aims to unveil the underlying factors behind the accumulation of reserves for the long term at aggregate levels.

4.3 Methodology and variables[27]

The model set-up is as follows. Consider the VAR standard form

$$y_t = c + \Gamma_1\, y_{t-1} + \Gamma_2\, y_{t-2} \ldots + u_t \qquad (1)$$

where y_t is an *n x 1* vector of endogenous variables at time t, c is an *n x 1* vector of constant terms, $\Gamma_t := ß_t A^{-1}$, are *n x n* matrices of coefficients and $u_t := \varepsilon_t A^{-1} B$ is an *n x 1* vector of error terms, with u_t having variance covariance matrix $\Sigma_u = B\ A^{-1}\Sigma_\varepsilon\ A^{-1}{}'B'$. Normalizing the variances of the structural innovations to one, i.e., assuming $\varepsilon_t \sim (0, I_n)$, $\Sigma_u = B\ A^{-1}\ A^{-1}{}'B'$. Therefore, the model can be rewritten as

[27] The definitions and details regarding the construction of the variables are explained in Appendix E.

$$y_t = c + \Gamma_1\, y_{t-1} + \Gamma_2\, y_{t-2} \ldots + BA^{-1}\varepsilon_t \quad (2)$$

However, to recover the parameters of the structural form, at least $2n^2 - \frac{1}{2}n(n+1)$ restrictions are to be imposed in B and A to identify all $2n^2$ elements of these matrices. In this way, the structure of the model follows the AB-model described in Lutkepohl (2005). Before proceeding with the identification scheme, a previous knowledge of the variables included and the effects they intend to capture is necessary.

The aim of this model is twofold. First, to shed light on the influence of the reserves cost on the amount of excess reserves held. Second, to unveil the impact of the reserves cost on banks' decisions about the loans rate. Hence, the variables to consider are:

- ER (growth rate): Excess reserves as percentage of required reserves. It measures the desired percentage of reserves held by banks, regardless the variations in required reserves.

- NBR (growth rate): Nonborrowed reserves as percentage of loans. It intends to measure the supply levels of reserves in relation to the loans provided. Thus, it will capture when the Fed is oversupplying or undersupplying banks with reserves. It is expected that an oversupply of reserves have a positive effect on excess reserves. Besides, part of nonborrowed reserves belongs to deposits. Therefore, the deposits effect will be included in this variable, although inseparable from the oversupply or undersupply of reserves.

- MAX (levels): the spread between the maximum interest rate at which banks obtained reserves up to period t (MAX[STR]) and the short-term rate at period t (STR). It measures how far the short-term rate is from the maximum cost at which banks could be trapped with relatively expensive reserves. The larger (smaller) the spread, the more (less) the loans rate will be increased, so that banks are able to use as soon as possible all their reserves without losses. Furthermore, the larger the spread, the more excess reserves will be accumulated, as banks will prefer to use reserves at current rates to obtain higher profits.

- RC-STR (levels): the spread between the reserve cost variable and the short-term rate of reference. As RC is only an average of how much banks have paid for one dollar of reserves, a positive (negative) spread implies that the probability of banks using reserves incurring in losses is higher (lower), as they may have to lend at a rate below the average cost. Likewise, a positive (negative) spread entails higher (lower) probability that banks keep their reserves and obtain more at cheaper cost, so that their profits are greater.

- LR-STR (levels): the spread between the prime loans rate and the short-term rate of reference. It measures the dissonance (in case they evolve differently) between Fed's policies (short-term rate targeted) and banks' decisions about the loans rate.

Setting equation 2 as $A(I_k + ß_1L + ß_2L^2\ldots)y_t = B\varepsilon_t$ where L is the lag operator, the order of the variables and the restrictions are set as follows:

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ b21 & 1 & 0 & 0 & 0 \\ b31 & b32 & 1 & 0 & 0 \\ b41 & b42 & b43 & 1 & 0 \\ b51 & b52 & b53 & b54 & 1 \end{pmatrix} \begin{pmatrix} ER \\ NBR \\ MAX \\ RC-STR \\ LR-STR \end{pmatrix}_t (I_k+ß_1L+ß_2L^2\ldots) = \begin{pmatrix} \sigma 11 & 0 & 0 & 0 & 0 \\ 0 & \sigma 22 & 0 & 0 & 0 \\ 0 & 0 & \sigma 33 & 0 & 0 \\ 0 & 0 & 0 & \sigma 44 & 0 \\ 0 & 0 & 0 & 0 & \sigma 55 \end{pmatrix} \varepsilon_t \quad (3)$$

Initially, this identification scheme assumes that ER is affected just after one lag for all the variables. That is, only after banks observe the price at which reserves have been obtained, their loans rate has determined lending levels, and the Fed has provided them with reserves to cover the demand for loans according to their estimations and targets, banks decide how many reserves to hold. NBR responds contemporaneously to ER either because for some periods the Fed targeted the amount of excess reserves or simply because the level of excess reserves influences the short-term rate and therefore, the Fed has to provide different amounts of non-borrowed reserves to control it. However, the relationship between this variable and the three spreads is troublesome for the identification. While loans (as part of NBR), adequately, respond after one period to the loans rate in LR-STR and the supposed reserves cost that drives the loans rate (RC-STR), non-borrowed reserves determine the short-term rate instantaneously. Given the purpose of the model here and the spreads set-up, likewise that in quantum physics one particle seems to be in two different positions at the same time but the observer can only locate it at one place, here, I cannot measure at once how nonborrowed reserves affects immediately the short-term rate, at the same time that the lending levels responds to the loans rate after one period. Therefore, different orders will be tested. Regarding the two variables measuring reserves costs, the logical order is with RC-

STR responding contemporaneously to MAX, as the average RC is influenced by the reserves obtained at the maximum cost. At the same time, these variables react contemporaneously to excess and nonborrowed reserves and loans, because they determine the amount of reserves and therefore, their average cost. The spread LR-STR responds contemporaneously to all variables, because banks set the loans rate taking into account the reserves costs to maximize profits, and partially, also, the demand for credit.

Last, the period under analysis is 1922:I-2017:III (382 observations). According to BIC criteria, the model is estimated with one lag. Whereas AIC criteria and HQIC recommended two and seven respectively, given the evolution of the variables observed in Figure 3.1 and 3.3, and the inherent characteristics of the "reserves-cost" theory, it is unrealistic to use more than one lag (one quarter lagged). The reason is that banks will decide at each period the quantity of excess reserves to be held and later, the loans rate depending on the reserves costs at the current period, so as to not incur in losses. For the next period, the costs will be analyzed again, and new decisions will be taken regarding excess reserves and the loans rate. In any case, alternative models are estimated with further lags, new variables measuring factors such as risk or uncertainty, and finally, the structural identification is altered given the problems exposed above.

4.4 Results

In this section, I analyse the impulse response functions and the forecast error variance decomposition for the model developed above. Figures 3.5.1 to 3.5.5 displays the impulse response functions for the period 1922:I-2017:III with the identification scheme described in the previous section.

Figure 3.5.2 and 3.5.3 gather the relevant results for the "reserves-cost" theory. Figure 3.5.2 shows the responses to a positive RC-STR shock. The first panel displays that banks increase excess reserves (ER) about 0.018 percentage points after one quarter when the spread RC-STR grows one percentage point, and the effect slowly decreases for nine quarters. The opposite is found for NBR (panel 2), where the response is negative and significant until quarter seven. The explanation is that greater costs are caused by higher rates. In order to increase rates, the Fed diminishes nonborrowed reserves. Panel 4 shows that banks increase the spread LR-STR by 0.05 percentage points immediately after RC-STR is one percentage point larger. The response is significant and positive for two quarters. Regarding the responses to a positive MAX shock, that is, when the maximum rate at which

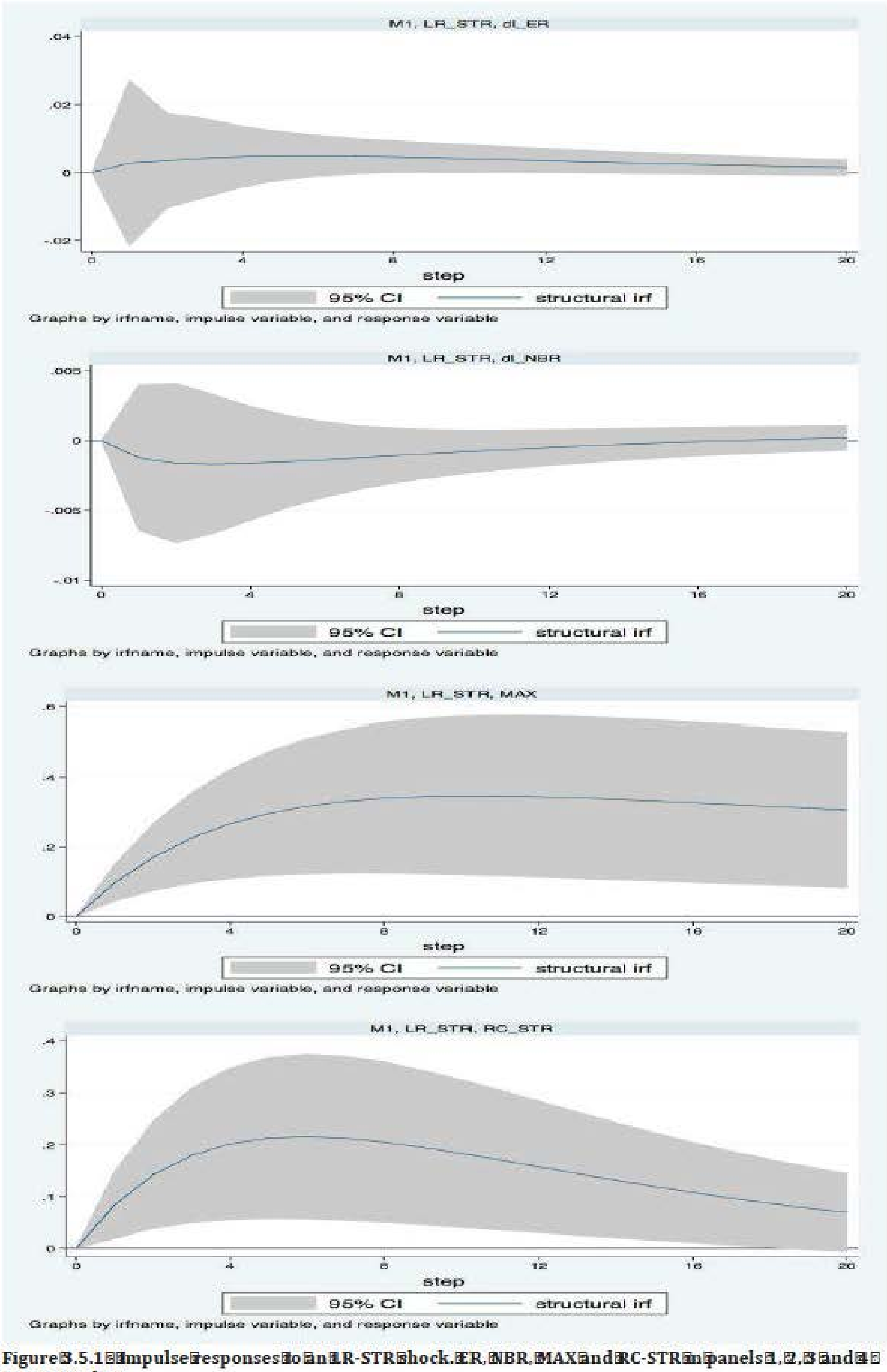

Figure 3.5.1 Impulse responses to an LR-STR shock. ER, NBR, MAX and RC-STR in panels 1, 2, 3 and 4 respectively.

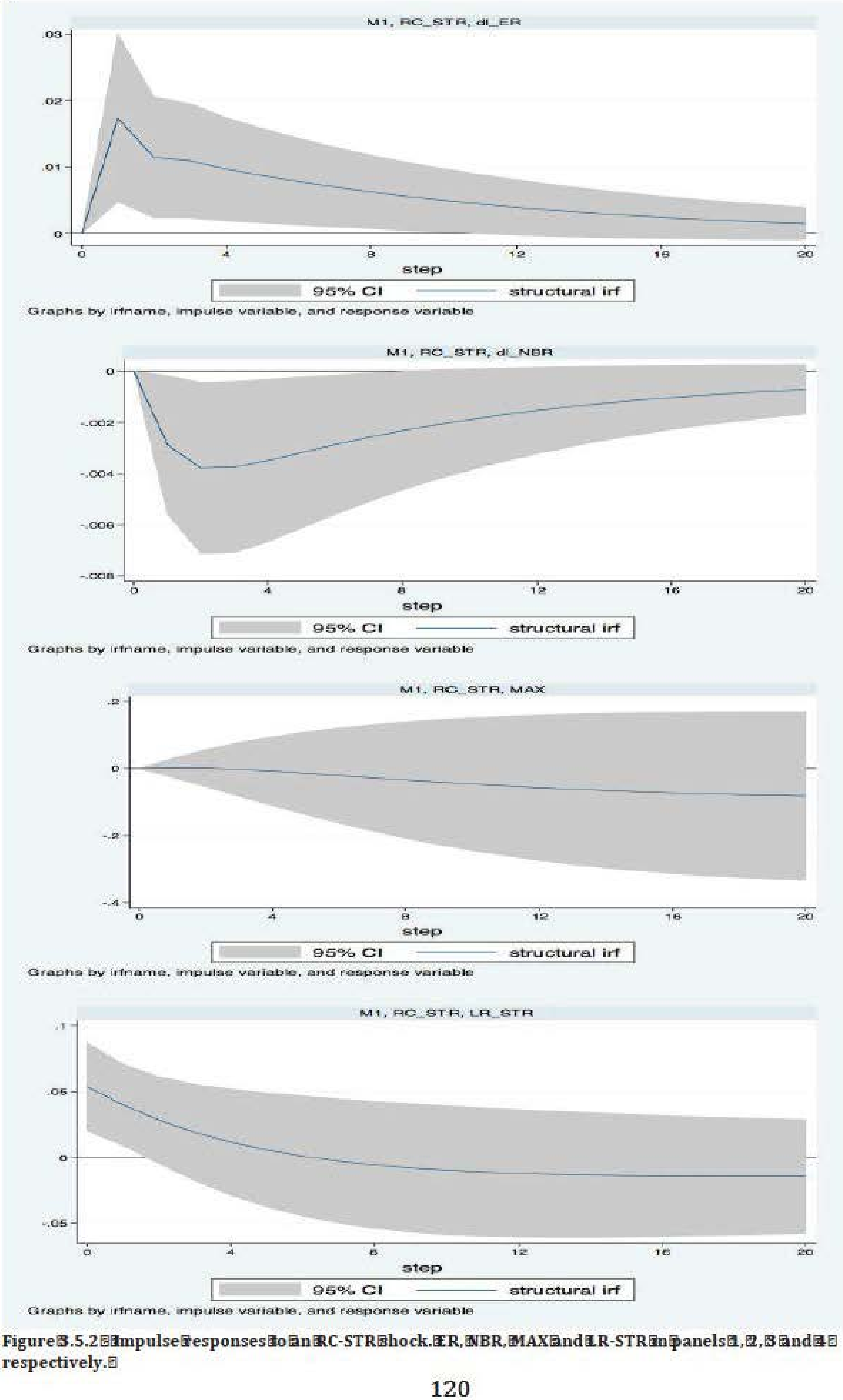

Figure 3.5.2 Impulse responses to an RC-STR shock. ER, NBR, MAX and LR-STR in panels 1, 2, 3 and 4 respectively.

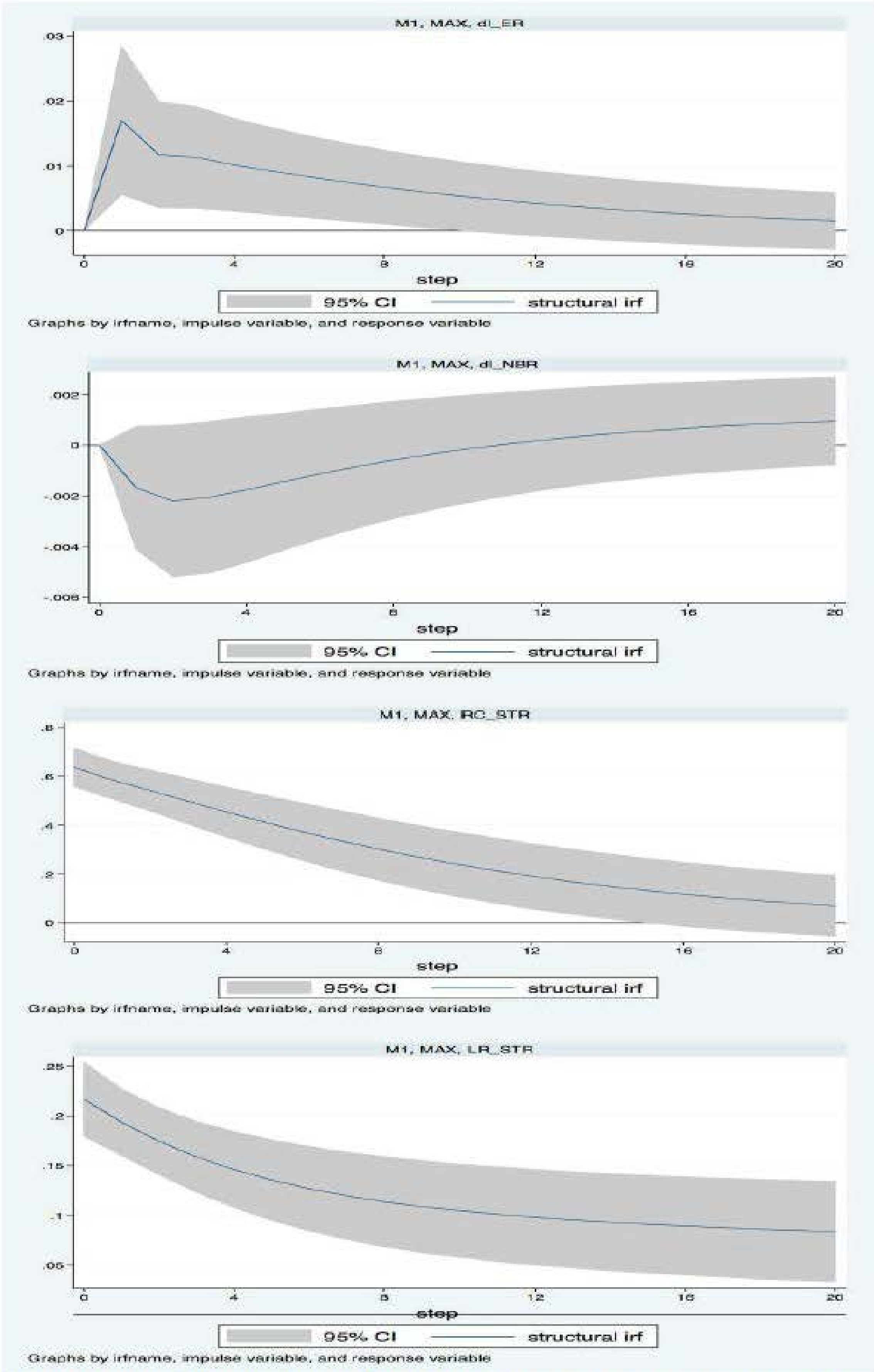

Figure 3.5.3 Impulse responses to a MAX shock. ER, NBR, RC-STR and LR-STR in panels 1, 2, 3 and 4 respectively.

banks are trapped with expensive reserves is one percentage point higher than the short-term rate of reference, Figure 3.5.3 (panel 1) displays that banks increase ER also by 0.018 percentage points. As in the previous figure, the response is positive and significant for nine periods. However, the NBR response is no significant in this case (panel 2). In panel 4, it is seen that a shock to MAX triggers an immediate increase in the loans rate around 0.2 percentage points more than the short-term rate. The effect lasts around 20 quarters.

In Figure 3.5.4, ER responds (panel 1) positively to a NBR shock, but surprisingly it is not significant. The second and fourth panel have positive and significant MAX and LR-STR responses respectively. The reason is that when the Fed intends to lower interest rates, more nonborrowed reserves are provided. Consequently, banks get trapped with reserves at higher costs. Thus, when nonborrowed reserves increase and banks accumulate excess reserves, the spread MAX becomes larger, and banks set a higher loans rate in relation to the short-term rate (LR-STR).

Looking at Table 1, the forecast error variance decomposition shows that the variation in ER is mostly explained by itself. While NBR is the second variable with more relevance in the first quarters, the reserves cost variables are ahead since quarter eight. As explained in section 2, the reserves cost will have effect on excess reserves only after the quantity of non-borrowed reserves provided in relation to the demand for credit is accounted. If the demand is relatively low, nonborrowed reserves will have a higher explanation power. NBR presents similar results, except that the reserves cost variables remains in the third and fourth place for the 20 quarters. MAX can be mostly forecasted by itself, followed by NBR and ER, although LR-STR takes over since quarter eight. For RC-STR, apart from its own explanatory power, MAX displays similar percentages, and is followed by ER. Last, regarding LR-STR, approximately 65% of its variation is explained by itself at the beginning, reaching about 50% after 20 quarters. It is followed by MAX, which explains around 27% in quarter one, but end ups with almost a 40%. ER and NBR can explain together between 4-6% of its variation, while RC-STR is left behind with around 1-2%.

Summing up, the results confirm the mechanism proposed by the "reserves-cost" theory, because a higher reserves cost increases ER and LR-STR. Whereas for ER, either the average cost or the maximum cost seems to have similar effect and explanatory power, the loans rate in LR-STR, seems to be more influenced by the maximum cost at which banks are trapped with expensive reserves.

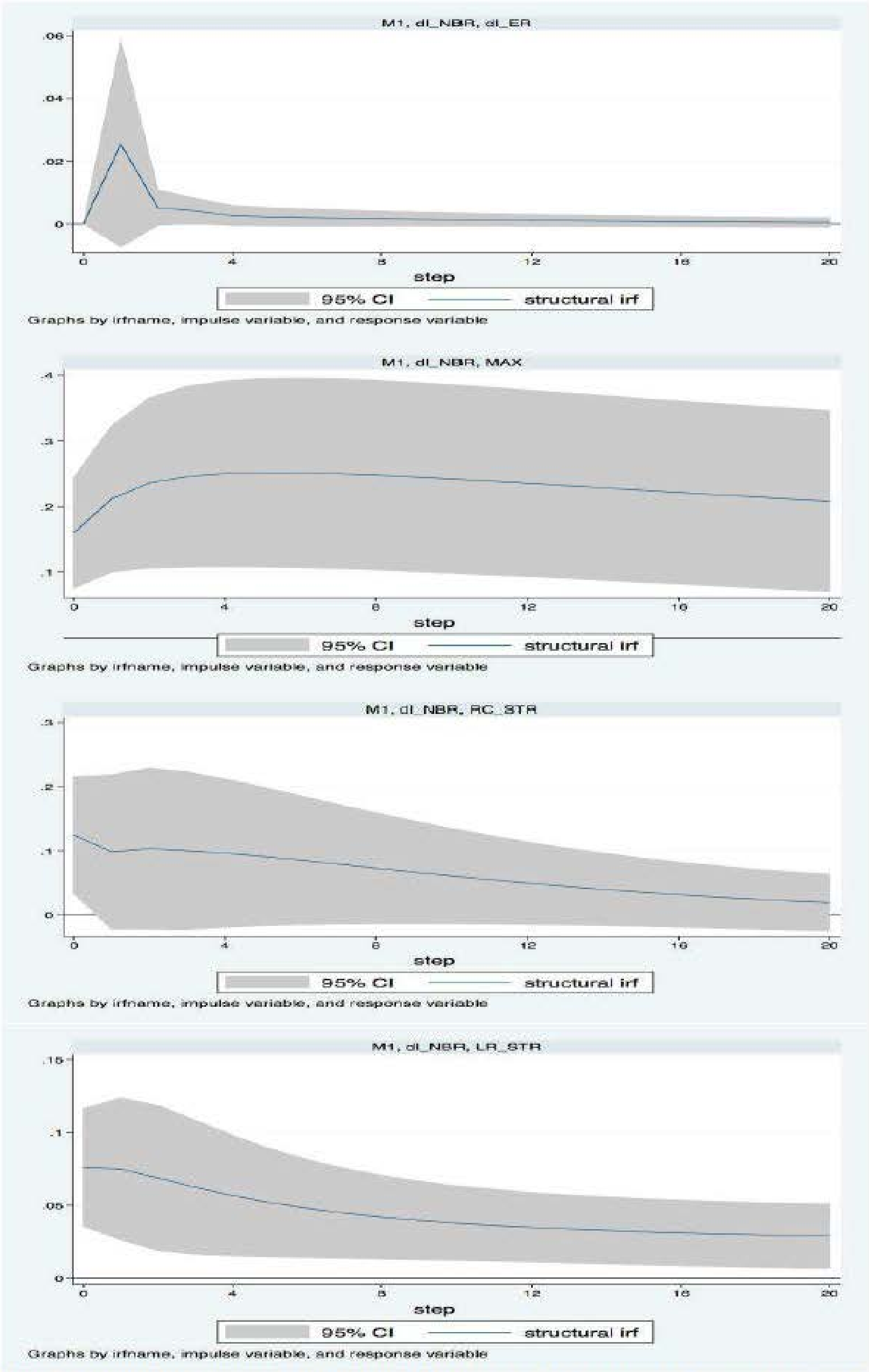

Figure 3.5.4 Impulse responses to an NBR shock. ER, MAX, RC-STR and LR-STR in panels 1, 2, 3 and 4 respectively.

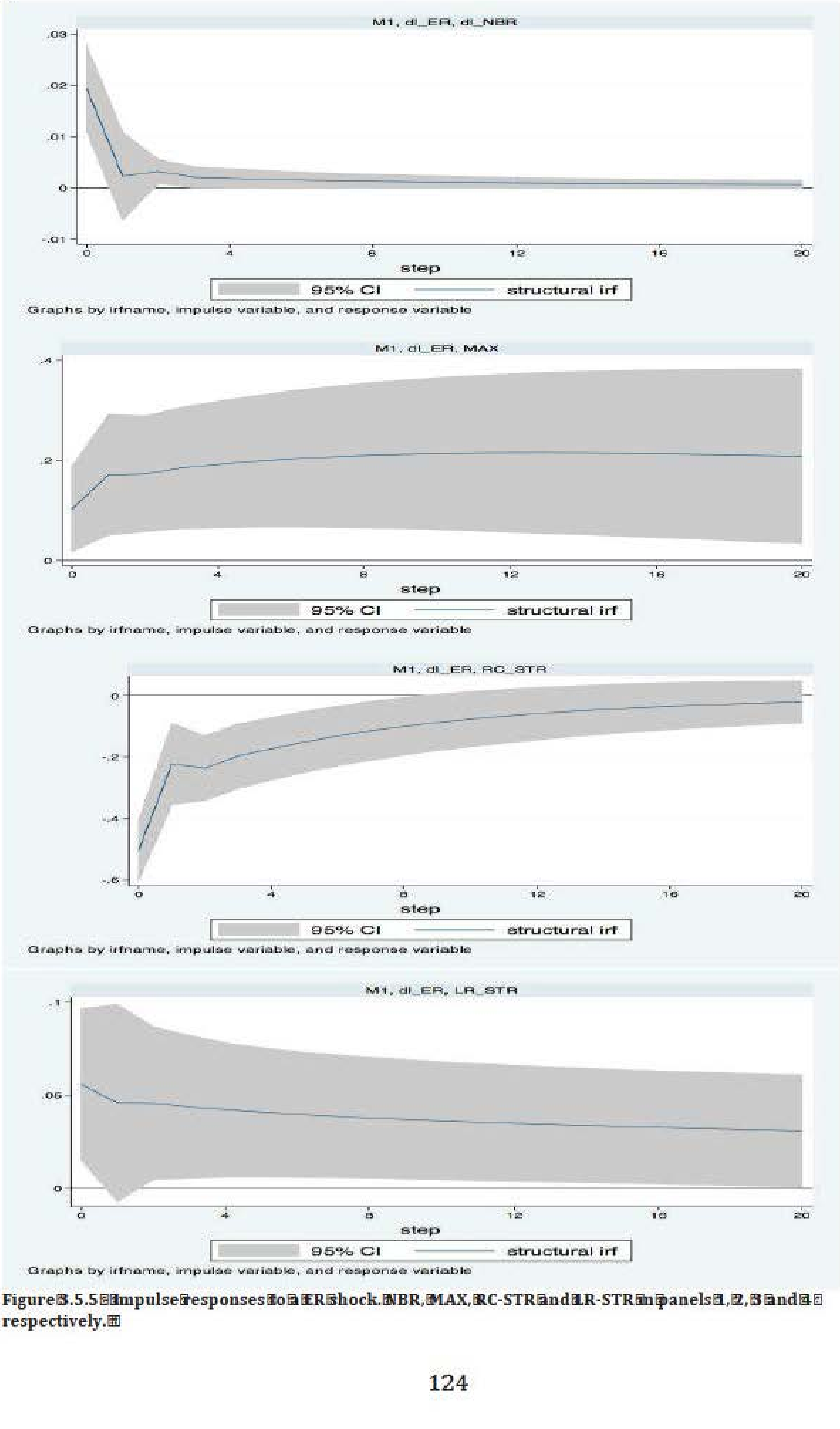

Figure 3.5.5 Impulse responses to a ER shock. NBR, MAX, RC-STR and LR-STR in panels 1, 2, 3 and 4 respectively.

Forecast error	*Forecast horizon*	*Proportion of forecast error variance h periods ahead accounted for by innovations in*				
in	*h*	**ER**	**NBR**	**MAX**	**RC-STR**	**LR_STR**
ER	1	1	0	0	0	0
	2	0,9923	0,0039	0,0017	0,0018	0
	3	0,9904	0,0041	0,0026	0,0026	0,0001
	4	0,9887	0,0042	0,0033	0,0033	0,0002
	8	0,9845	0,0043	0,0052	0,005	0,0007
	12	0,9825	0,0043	0,006	0,0057	0,0012
	20	0,9814	0,0043	0,0065	0,0061	0,0015
NBR						
	1	0,0524	0,9475	0	0	0
	2	0,048	0,9503	0,0003	0,001	0,0001
	3	0,0487	0,9469	0,0009	0,0028	0,0005
	4	0,049	0,9439	0,0014	0,0045	0,0008
	8	0,0499	0,9365	0,0023	0,0092	0,0018
	12	0,0504	0,9336	0,0024	0,0111	0,0022
	20	0,051	0,9313	0,0028	0,0124	0,0023
MAX						
	1	0,0146	0,0351	0,9501	0	0
	2	0,0264	0,0466	0,9206	0	0,0062
	3	0,0301	0,0542	0,899	0	0,0165
	4	0,0329	0,059	0,8796	0	0,0283
	8	0,0406	0,0672	0,8214	0,0002	0,0704
	12	0,0462	0,0701	0,7857	0,0009	0,0969
	20	0,0545	0,0724	0,7484	0,0033	0,1212
RC-STR						
	1	0,2322	0,0142	0,3727	0,3808	0
	2	0,165	0,0136	0,4079	0,4094	0,0038
	3	0,1437	0,0143	0,4177	0,4133	0,0108
	4	0,1301	0,0149	0,4224	0,4107	0,0193
	8	0,1061	0,0168	0,4229	0,4018	0,0521
	12	0,0976	0,0178	0,4187	0,3932	0,0724
	20	0,093	0,0183	0,4143	0,3872	0,087
LR-STR						
	1	0,0183	0,0342	0,2771	0,0168	0,6532
	2	0,0177	0,039	0,2881	0,0151	0,6398
	3	0,0189	0,042	0,2982	0,0136	0,6272
	4	0,0202	0,0441	0,3078	0,0123	0,6155
	8	0,0254	0,0491	0,3403	0,0093	0,5757
	12	0,0299	0,0518	0,3646	0,0086	0,5448
	20	0,0367	0,055	0,398	0,0093	0,5007

Table 1 Forecast error decomposition

4.5 Robustness tests

The figures corresponding to the following sections are available in Appendix F.

4.5.1 Alternative orders

When NBR is located below LR-STR, so that it responds instantaneously to all variables, the results are almost identical (not presented). However, the results change slightly when RC-STR and MAX are interchanged. Figures F.6.1 and F.6.2 display the response of each variable to an RC-STR and MAX shock respectively, as their responses are the only ones that vary. The differences in Figure F.6.1 regarding F.5.2 are found in the non-significant NBR response (panel 2), the positive and significant MAX response beyond 20 quarters (panel 3), and the larger LR-STR response (11 quarters instead of 2, panel 4). In Figure F.6.2 (compared to Figure 3.5.3), ER lost its significance (panel 1). NBR response remains no significant (panel 2), but it changes its sign. Last, RC-STR response turns no significant to a MAX shock (panel 3). Thus, with this order, RC-STR has a greater impact on LR-STR and ER, while MAX loses its importance in determining ER. Despite this variation, the logical order is the one applied in the baseline model.

4.5.2 Risk and uncertainty

I have also introduced two variables in the model as proxies for risk and uncertainty (two different models), ordered after ER. For the former, I have used the growth rate of public debt (DEBT), while for the latter I have calculated the IPI standard deviation for an interval of four years (16 observations) and rolling it one period ahead (GDPD). The results are almost identical to those obtained in the original model (not presented). A shock to DEBT increases instantaneously ER around 0.05 percentage points but the effect disappears after one quarter. It also increases the spread LR-STR by the same amount. In this case, however, the effect is permanent. The impact of a GDPD shock on LR-STR and ER is no significant. Therefore, it seems that banks take into account risk, but not uncertainty, in line with Chang, Contessi and Francis (2014). On the contrary, NBR increase slightly with a shock to GDPD. Hence, the Fed seems to inject more reserves for periods of uncertainty.

4.5.3 Subperiods

Other source that potentially could modify the results obtained is the different episodes over the period under analysis. From 1922 to 2017, there are two episodes when the

Fed flooded banks with reserves and could be distorting the impact of the variables under consideration. Consequently, I analyse the entire period divided in three subperiods using as criteria the level of ER displayed in Figure 3.1. The first one, from 1922 to 1960 (156 observations), includes the first great reserves injection into the banking sector and the consequent increase in excess reserves. Around 1960, ER levels are near the levels before the Great Recession. Therefore, it signals the end of the first subperiod and beginning of the second one. The second subperiod, from 1960 to 2006 (188 observations), can examine the period when the reserves supplied by the Fed can be considered as normal and the relationship between variables could be clearer. Last, the third subperiod analyses the same window as the second subperiod, but including the QE episode (1960-2017, 231 observations). For the last subperiod, as the results are quite similar to those of the baseline model, they are not presented.

4.5.3.1 Subperiod 1922-1960

When looking at Figure F.7.2, the shock to RC-STR shows variations in NBR and MAX responses (panels 2 and 3 respectively), which are now positive and not significant. Figure F.7.3 shows the responses to a MAX shock. Unlike the baseline model, ER response (panel 1) is initially positive and no significant for the first quarters, but negative and significant after eight quarters. This result is likely due to the fact that MAX remains equal for the entire subsample (Figure 3.1).

4.5.3.2 Subperiod 1960-2006

Regarding the subperiod 1960-2006, a RC-STR shock (Figure F.8.2) offers similar results to the original model. The only difference is that NBR response is not significant. For the shock to MAX (Figure F.8.3), results are similar. Figure F.8.4 presents quite a different picture in relation to Figure F.5.4. ER response to an NBR shock (panel 1) is negative and significant after one quarter, but positive and significant in quarter three. Probably, the ER low levels led to a more heterogeneous relationship between both variables, depending on the demand for loans. Last, Figure F.8.5 shows that the shock to ER has a no significant NBR response (panel 1), unlike the original model. Thus, the Fed seems to not have paid attention to the reserves held during that period. The other panels show similar results.

4.5.4 Further lags

Estimating the model with more lags (results no presented), specifically until seven, from lags two to four, the impulse responses keep their shape in general, despite becoming more jagged the more lags are included. However, most of the significant responses obtained in the original model disappear. From lags five to seven, the jagged shapes remain but the responses to MAX and RC-STR shocks recover their significance as in the original model. More lags trigger that RC-STR has a longer positive and significant impact on LR-STR.

4.5.6 Monthly data

The same baseline model was estimated changing the periodicity from quarterly to monthly (results not presented). The sample starts in 1947, as the data for loans (in the denominator of NBR) is available with monthly data only since then. Following the original model, it has been estimated with one lag and the results obtained are similar. As I increase the number of lags, the results present the same behaviour as commented in the previous section.

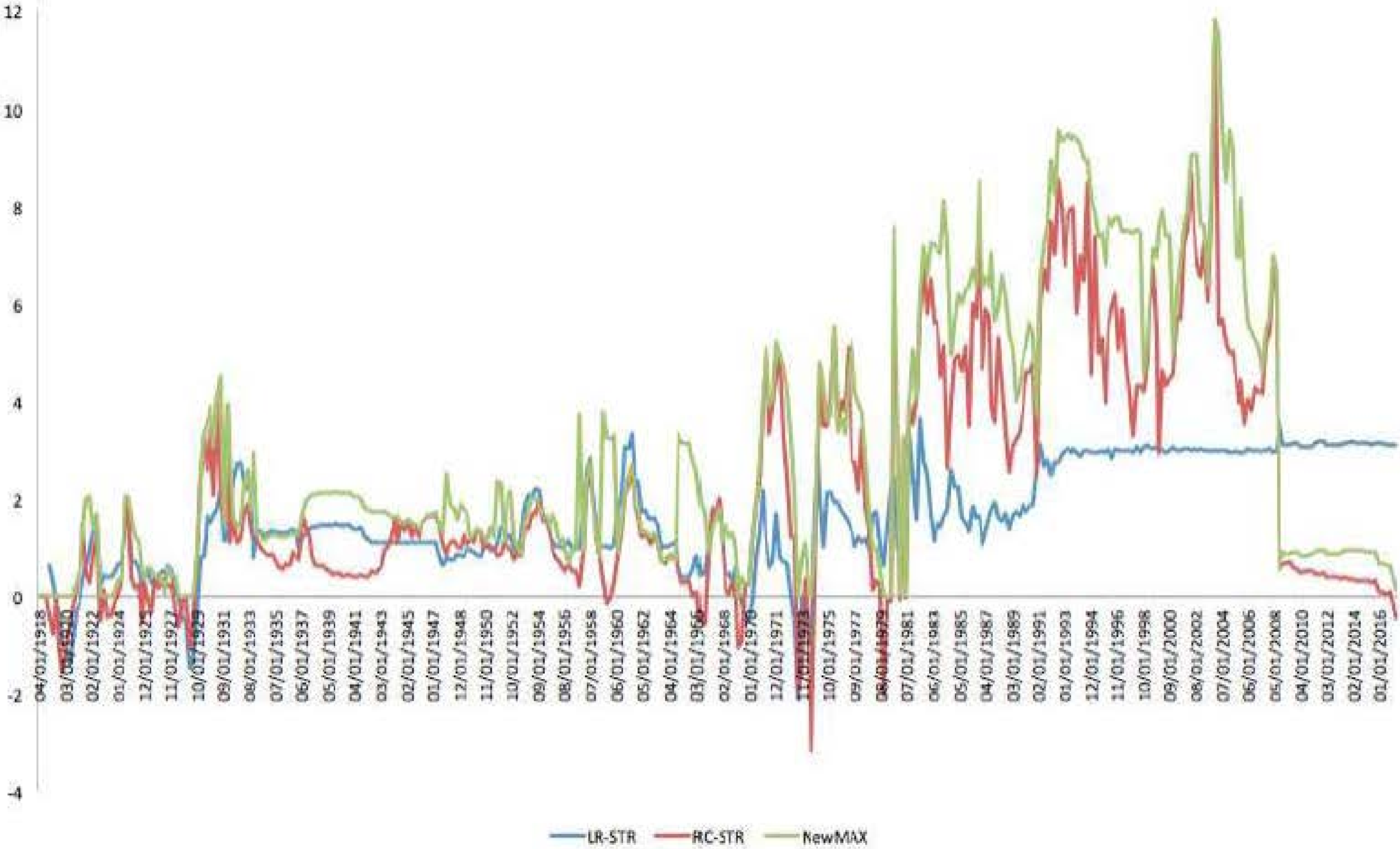

Figure 3.9 New MAX variable

4.5.7 New MAX variable

Given the variable MAX only measures the maximum rate at which banks obtained reserves, it may be the case that it is not representative of the cost of those reserves held above current short-term rates. The reason is that as the amount of reserves hold at those maximum rates could be small in relation to the expensive reserves held, banks could consider them as insignificant when setting the loans rate. To solve that issue, using the same method as for the creation of RC, I have calculated the average cost of the reserves, but this time accounting only for those held above short-term rates. Even though the idea may seems adequate, the outcome does not. Figure 3.9 displays again Figure 3.3, but using this new variable (NewMAX) instead of MAX. As easily seen, it evolves almost parallel to the variable RC (above 93% correlation). As one can expect, including this variable into the model does not add much information. Even so, the model was estimated again with it and the only variation obtained is that ER does not respond significantly to a NewMAX shock, and an RC shock double its impact on ER.

4.6 Interest on reserves and its potential future consequences

Nowadays, the Federal Reserve pays interest on required reserves (IORR) and excess reserves (IOER). With this tool, central banks intend to control separately interest rates and excess reserves. During and after the QE, the federal funds rate has been zero or nearby and without interest paid on reserves, the Fed would have had to operate massively in the open market to control the federal funds rate, given the excessive liquidity in the banking sector. Since January 2016, when the Fed began to normalize monetary policy, excess reserves have decreased around 25%. Thus, whereas the federal funds rate has increased from 0.35 to 2-2.25 percentage points since then, excess reserves have decreased from 2279 billion dollars to around 1700 billion dollars until October 2018. Approximately, the same decrease is observed in non-borrowed reserves. While in January 2016 the loans growth rate was about 8%, at the end of 2018 was near 5%. Those rates are around ordinary levels, and as predicted by Bindseil (2004), we have not seen any inflation outburst.

Todd (2013) and Ennis and Wolman (2015) argued that interests paid on reserves are encouraging banks to hold reserves. It is important to highlight that the federal funds rate and the IOER or IORR are almost the same. Then, why have I not included those rates in the models estimated? There are two reasons. First and most important, the interest paid on reserves would appear only in the last part of the sample, coinciding with the QE. Excess reserves started to decline only in 2016 and therefore, there would be only 8 observations for which excess reserves declined once interest rates started to be paid on reserves. There are not data for interest rates paid on reserves during "ordinary" periods. Hence, avoiding sample selection bias I decided to not use a variable capturing this new tool. Second, while IOER and IORR are increasing along with the federal funds rate, excess reserves are decreasing anyhow, as nonborrowed levels are declining in relation to the demand for credit. This supports the argument exposed previously by Moore (1998), Lindley, Clifford and Mounts (2001), Dwyer (2010), Calomiris, Mason and Wheelock (2011) among others, that banks lend when there is demand for credit and is profitable. Therefore, those interests paid on reserves, despite being a good tool to control the federal funds rate, seems to not influence banks' decisions between reserves held and lending. The reason is that once banks can lend because there is demand, as the loans rate is higher than the federal funds rate and the interest on reserves, they will lend and use their excess reserves.

Beyond that the interest rates paid on reserves are effective to isolate the federal funds rate from the amount of reserves and cannot avoid that banks use their excess reserves when the demand for credit is higher, it is not clear whether they could be used to control the impact that the cost of past accumulated reserves has on the loans rate, as exposed in the "reserves-cost" theory. There are two ways under which they could be used for that task:

-First, the Federal Reserve, being able to control the federal funds rate regardless of the amount of reserves, could drain whatever level of excess reserves is held on the banking sector and allows banks to borrow the same amount of reserves drained or that desired by banks, at the discount window. This would suppose the elimination of the influence of past reserves on the loans rate. Consequently, as the new reserves would have been obtained at the current discount rate, the Fed would have again an impact ration 1:1 over the loans rate with raises in the discount rate, and setting the same rate for the federal funds rate. This process would have to be undertaken every time that banks accumulate past reserves and that accumulation modifies the impact ratio of Fed's interest rates on the loans rate. The feasibility of this procedure is beyond my knowledge, but if possible, it would be effective.

-Second, and depending totally on banks decisions, as they are receiving interest on their reserves, at some point, they may price those past reserves with the interest rate that they are receiving, so that the impact ratio would be readjusted again (1:1) just because of that fact.

In case that none of these situations happens, the interest rate paid on reserves would be useless to control the impact of past reserves on the loans rate, and those fearing an inflation outburst given the current level of reserves could be right, although the mechanism working there is more complex than simply more excess reserves leading to higher inflation, and the consequences could also be different. That mechanism can only be explained from the "reserves-cost" theory and three scenarios are considered:

-For the first scenario, we have as example the 1960s and 1970s, where excess reserves were scarce. Under those circumstances, the Fed is able to have direct impact on the loans rate when using the federal funds rate (and interest rates on reserves). Therefore, one percentage point increase should force banks to increase the loans rate by the same quantity. The problem by then was that as the Fed set the discount rate below the federal funds rate, banks could borrow cheaper reserves at the discount window and increase the loans rate relatively less than the federal funds rate was raised. This led to more credit and inflation as showed in Chapter 3. Thus, the "reserves-cost" theory establishes that, under a scenario of low excess reserves, if the discount rate is at the same level or above the federal funds rate, the only factor able to cause bubbles (given the higher demand for credit) or high inflation would be a faulty model and policymakers setting lower rates than required. Otherwise, the Fed loses part of its power by allowing banks to borrow cheaper reserves at the discount window and increase the loans rate less in relation to the federal funds rate.

-For the second scenario, the present can be taken as example. Given high excess reserves levels, a positive demand shock could lead banks to use their reserves without necessity for obtaining them from another source. Bubbles or high inflation could come from two sources in this case. The first source is again a faulty model and policymakers setting lower rates than required. In this case, banks could lend almost as much as the demand side wishes to borrow at that low interest rate, either because banks can use their excess reserves (as long as their cost is below the federal funds rate) regardless of the interest paid on reserves, as the loans rate is above them, or because under an interest rate target the Fed would provide banks with all the necessary reserves to maintain the aimed low interest rate target. The second source would emerge if, under current levels and prices for excess reserves, the Federal Reserve increases the federal funds rate faster than banks can get rid of their cheaper reserves. In that case, banks could increase the loans rate below the ratio 1:1 in relation to the federal funds rate, by using their cheaper reserves.

However, the Federal Reserve may be lucky under a particular situation and this second source would be cancelled. The results have proved a positive relationship between the spread RC-STR and LR-STR, and MAX and LR-STR, the last one being more important and durable. Thus, while banks are trapped with reserves at higher rates, the loans rate tends to increase more than the federal funds rate. In Figure 3.1 is observed that when the federal funds rate declined after its maximum in the early 1980s, banks got trapped with expensive reserves. Consequently, they increased the loans rate more in relation to the federal funds rate (impact ratio above 1:1), but the spread reached a maximum of three percentage points. This spread has been constant since 1990 and is the largest of the entire period. It is the case that the interest rates reached in the 1980s are far from the current ones and while this condition remains, the loans rate will likely move one to one with federal funds rate as after 1990 (even if the Fed increases the federal funds rate faster than banks can get rid of their cheaper reserves), because there seems to be a ceiling for that spread[28]. That is, when the rate at which banks are trapped is too far from current rates, banks seem unable to widen more the spread and the loans rate is anchored to the federal funds rate.

-The process until that ceiling is reached, the third scenario, implies an impact ratio above one. Contrary to the inflation outburst and asset bubbles, in this case the Federal Reserve may find problems to increase inflation, as every increase in the federal funds rate will suppose a greater hardening in its policies through a further increase in the loans rate. This hardening will provoke a posterior decrease in interest rates, triggering an impact ratio above one with a spread even higher for the next rise of the federal funds rate. When this process is extended without reaching the last global maximum, as it has been the case since the early 1980s, interest rates will evolve in a descending trend, with a local maximum lower than the previous one. Lower interest rates for such a long time can lead to debt accumulation. Under those circumstances, the next time that the interest rate is raised, as the loans rate will continue increasing above the ratio 1:1 and the spread will be larger, probabilities for recessions and defaults will be higher, making even more difficult to

[28] The answer to the email reproduced in the introduction ended as follow: "…It is important to know that some states have laws that set a ceiling on the amount of interest that banks can charge for certain loans, but these ceilings are not set by the Federal Reserve or any other federal agency."

increase interest rates in the future. This situation can be extended to the exceptional case commented for the second scenario, namely, the spread ceiling. When the loans rate is finally anchored to the federal funds rate, if the process of the impact ratio above one has been long enough to accumulate high debt levels as a consequence of interest rates being pushed to the floor, what was considered a lucky situation in the second scenario will become an unfortunate episode of low inflation, recessions and defaults.

In order to have an idea of the banking sector' reserves inventory nowadays and its potential consequences for the next years or even decades, a snapshot has been taken (Figures 3.10 and 3.11) for the last period available for RC (June 2018). The Y-axis displays the amount of excess reserves (billion), while the X-axis represents the interest rate at which they were obtained. Figure 3.10 shows that banks are holding most of their reserves at 0.25%, followed by reserves at 0.5%, 1.25% and 2.25%. Overall, those four costs account for approximately 1,860 billion dollars of reserves held by then. If those points are removed, the quantity of reserves remaining is insignificant in relation to the previous figure (Figure 3.11). Summing all of them, the result is around 3,1 billion and yet, most of them are below 5%.

Therefore, if the American economy were to face an increasing credit demand in the near future and interest rates surpassed five percentage points, dangers of inflation and/or bubbles would be likely if: excess reserves with lower cost than the federal funds rate (and therefore, IORR and IERR) remains in the banking sector, so that the federal funds rate does not have an impact ratio 1:1 on the loans rate, or, evidently, if the federal funds rate is set sufficiently low to not restrain the demand for credit. For the first case, however, as banks seems to have fixed the spread between the federal funds rate and the loans rate because of the much higher cost of past reserves, the impact ratio 1:1 is likely to continue. The situation may change once short-term rates reach the zone near 7%, because banks started to peg the loans rate to the federal funds rate around those levels. In the meantime, during the process of normalizing interest rates, given the maximum spread LR-STR of 3%, the Fed may find difficulties not only to increase inflation above certain levels, but also interest rates, due to the higher probabilities of defaults, given debt accumulation during low interest rates periods.

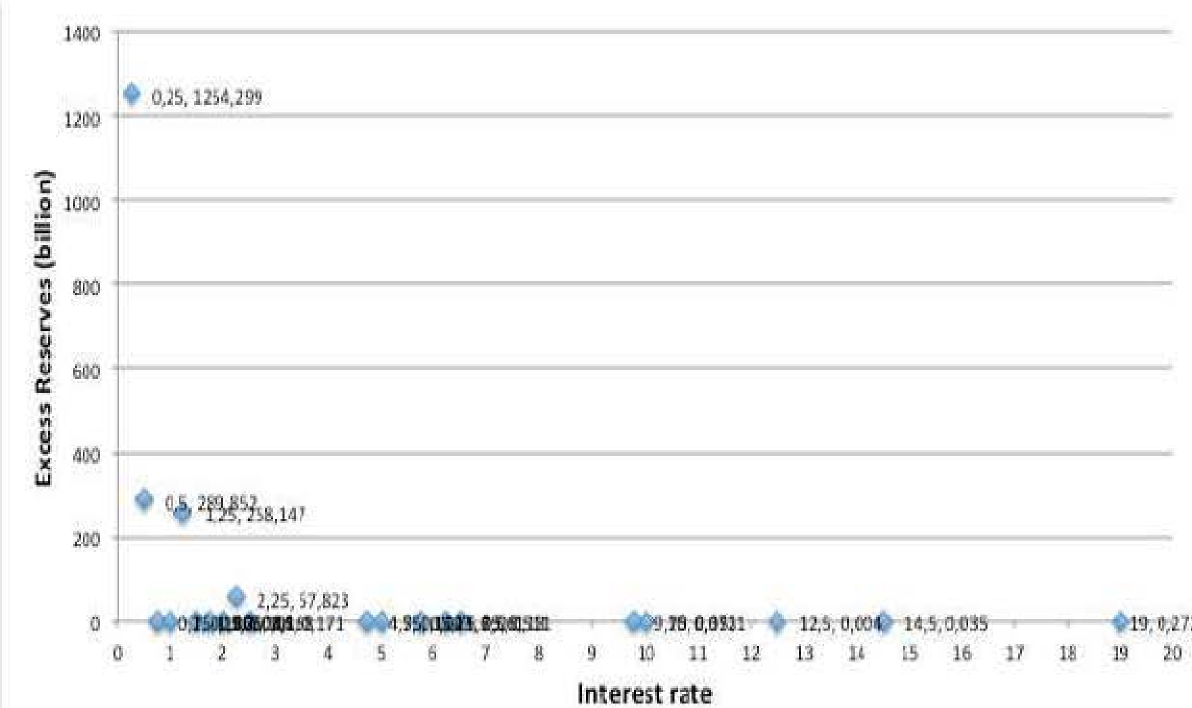

Figure 3.10 – Reserves cost inventory by June 2018

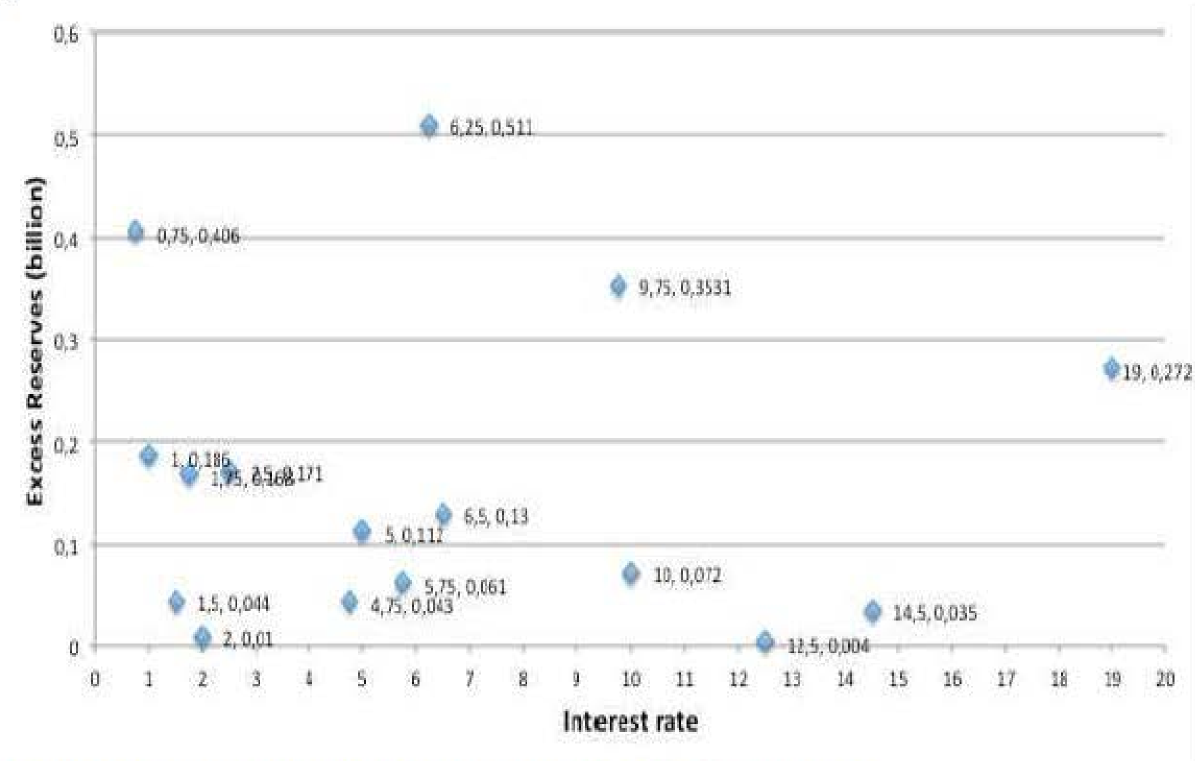

Figure 3.11 – Reserves cost inventory by June 2018 without the cheapest reserves.

Having this in mind, as policy advise, a more efficient monetary policy would imply to keep an eye on the federal funds rate impact ratio on the prime loans rate. Thereby, when the ratio is below one, the Fed must increase further the federal funds rate (keeping the discount rate at the same level or above it), so that the loans rate has the desired impact on the real economy. Besides, after relative high interest rates, the Federal Reserve should pay attention to how banks accumulate excess reserves in the process of lowering rates, namely, to prices and quantities. The reason is that, on the one hand, if banks hold excessive levels of reserves, they may able to modify the impact ratio in the future, when interest rates are raised again. On the other hand, an impact ratio above 1:1 will increase the probabilities of remaining in a low interest rates scenario for too long. Deflation and increasing debt may be some of the consequences. For this last case, only a strong and increasing demand that wipes out the high excess reserves levels and is able to endure higher interest rates (what entails the reduction of debt) would be the solution. Again, these would be the possible scenarios and advices, just in case that the IOER and IORR are not able to control the cost of past accumulated reserves through the two ways explained above.

4.7 Conclusions

Periods of great accumulation of excess reserves as during the Great Depression and the Great Recession have attracted the interest of researchers, who have tried to unveil the reasons behind that hoarding and its consequences. However, periods with normal levels have fallen into oblivion. Being those periods of vital importance to understand the mechanisms determining the evolution of excess reserves and its consequences, and driven by the need to complete the "reserves-cost" theory initiated in Chapter 3, a SVAR is estimated for the period 1922-2017. The renovated "reserves-cost" theory proposes, in summary, that banks consider the different costs at which reserves are obtained, so that, in relation to market rates, those costs will determine the accumulation of excess reserves and the evolution of the loans rate. Thus, the only way whereby the Fed can influence the real economy is directly by the impact of its policies on the reserves cost, and indirectly by the influence of that cost on the loans rate, which in the end is the rate that the real economy endures. The results have supported this theory by showing that when the average cost of excess reserves is above short-term rates, banks accumulate more excess reserves and the loans rate increase in proportion more than the short-term rate. The same is observed when short-term rates are below the maximum rate/cost at which banks obtained reserves. For this last case, it seems that this factor is stronger in explaining the evolution of the loans rate in relation to the short-term rate.

Therefore, the main lesson is that the accumulation of reserves (quantity) does not cause by itself increasing lending and inflation levels. It is the interest rate (price) that the demand side of the money market faces. Accordingly, the Federal Reserve should include as target the tracing of banks reserves, what implies the price at which those reserves are obtained. The lack of attention over this matter is likely to grant a significant part of monetary policy management to the banking sector. This factor can create either, inflation outbursts or asset bubbles when the short-term rate has an impact ratio below 1:1 on the loans rate, or deflation and higher debt levels for remaining with low interest rates for too long, as a consequence of enduring an extended impact ratio above 1:1.

-- Final conclusions --

Discussion

This thesis has unveiled a fundamental mechanism for monetary policy management in the United States, which has been either ignored or unknown for the Federal Reserve. This mechanism, described in the “reserves-cost” theory, implies that the banking sector is able to modify Federal Reserve’s policies and hence, lessen its power to conduct monetary policy. This, however, does not exempt it from the responsibility for the American economy’s performance. This mechanism operates through the cost of reserves. As banks can obtain borrowed or nonborrowed reserves from different sources, if the Federal Reserve aims to be the only agent driving the path of the economy, the same cost must be targeted for both types of reserves, starting from a scenario where excess reserves are scarce. By having reserves at the same cost at the discount window or in the money market, banks will set the loans rate according to the evolution of the unique reserves cost available. Under these circumstances, an increase in short-term rates aimed by the Fed will have the same impact on the loans rate and hence, the economy will absorb that impact directly as the Federal Reserve intended. When banks accumulate excess reserves or the Federal Reserve provides arbitrage opportunities for the reserves cost, the impact of short-term rates on the loans rate will be different from the ratio 1:1. When this ratio is below 1, the economy will endure a loans rate that will restrain the demand for credit insufficiently in relation to Fed’s desires, triggering higher inflation levels. When the ratio is above 1, the consequences will be just the opposite, with the risk that if these periods are long lasting, the economy will face a situation of low interest rates, where debt accumulation is likely. Under these circumstances, the future raise of interest rates will find difficulties given those debt levels. In this case, deflation can be also probable. In addition to the impact ratio of short-term rates, its variability is also important. Regardless of the ratio being below or above one, if it is constant over time, the Federal Reserve will still have a great influence on the economy, as the spread between short-term rates and the loans rate will be constant, and the economy will accommodate every movement equally. On the contrary, when the ratio is varying over time, the Federal Reserve power will be weaker, and the variability of the spread will be transmitted into the economy, as the agents will be experimenting different impact ratios every time that the Fed is aiming its targets. Higher volatilities in inflation and output as those seen during the Great Inflation will be the result. It has also been showed that the reserves cost is able to explain the evolution of excess reserves on the banking sector. Thus, when banks hold reserves above the cost at which current reserves are offered in the market, they will obtain further reserves at current rates to maximize profits by using those cheaper reserves and maximizing the spread between the reserves cost and the loans rate. Also, they will accumulate reserves with the intention of not incurring in loses by using those reserves with a cost superior to the rate at which money is lent. On the other hand, when banks hold reserves with a cost below market rates, they will use their excess reserves for the opposite reasons just commented. Last, this theory in its early development also unveiled that the known price puzzle whereby inflation increases despite the federal funds rate is raised, is actually the consequence of those arbitrage opportunities previously commented. Thus, when banks could obtain cheaper reserves at the discount window given the positive spread between the federal funds rate and the discount rate, increases in the federal funds rate had an impact below 1 on the loans rate. Consequently, lending was insufficiently restrained and inflation increased as result of those bad policies that allowed positive spreads.

While this study is only focused on the American economy, posterior studies should ascertain whether this mechanism could be applied to other economies and central banks' policies.

Main contributions

While the "reserves-cost" theory is the main contribution of this thesis, several steps were necessary before reaching the completeness of that theory, each step being a minor contribution. First, the standard procedure whereby monetary policy stance has been measured during the last decades has been declared erroneous, as by using intermediate targets that measures supply and demand forces in the first submarket of the money market, bias is introduced into the models. Therefore, given that to measure monetary policy only the supply side must be captured, that procedure cannot be applied. Using the actual Fed's instruments and the spread between short-term rates and the discount rate, which is a hybrid instrument, was proposed as the right procedure to capture monetary policy stance, this being the first contribution. From there, another one was extracted, as that procedure unveiled the inexistence of the price puzzle. The increase in inflation when the interest rate was raised was not a puzzle, but the consequence of bad policies undertaken by the Fed when it allowed positive spreads between the federal funds rate and the discount rate, providing arbitrage opportunities for the cost at which reserves were obtained. The last minor contribution from the new procedure was found in those results showing that monetary policy is transmitted through prices and not quantities.

A deeper reflexion about the new procedure to measure monetary policy stance led to the next contribution. The first version of the “reserves-cost” theory suggested that to measure monetary policy stance the instruments are not necessary, and based on the premise that prices are the ones transmitting monetary policies, it proposed that just measuring the price at which reserves are obtained is enough to capture monetary policy stance. Thus, measuring Fed’s policies and the banking sector’s behaviour with the spread between the federal funds rate and the discount rate, and the spread between the loans rate and the federal funds rate respectively, it was discovered that the extent to which monetary policies can reach the real economy depends on banks’ decisions regarding the loans rate, which is previously conditioned on the cost of reserves obtained, that cost being determined by Fed policies. This mechanism is proposed as the one explaining the high inflation levels, as well as the higher volatilities in inflation and output during the Great Inflation, and the lower volatilities and inflation witnessed during the Great Moderation.

The last step and contribution that led to the final “reserves-cost” theory was the explanation of why banks hold reserves and the impact of those reserves in the loans rate. Introducing excess reserves levels and their price into the theory shed a clearer light on how monetary policy is transmitted into the real economy through the banking sector, namely, the loans rate, and how it determines the evolution of inflation, output, and their volatilities, depending on different scenarios. Those scenarios are the amount and price at which excess reserves are held in relation to short-term rates and the spread between the loans rate and those short-term rates. Thus, when banks hold reserves above short-term rates, they accumulate more reserves and use the ones available in the market to maximize their profits. At the same time, the loans rate is raised in relation to short-term rates to lower the probability of incurring in loses in case of having to use their expensive reserves.

Therefore, the final lesson and contribution is that the intermediate target that the Federal Reserve should aim to achieve its final targets is the cost at which banks obtain their reserves. As long as this mechanism is not under control, banks will modify Fed’s policies, weakening its power, but without exempting it from the responsibility for the American economy’s performance.

Open gates for future research

The most logic step that should be taken after discovering this mechanism is to ascertain whether it is applicable to other economies. If it were not the case, it should be considered only for the American economy and even, re-tested under different approaches to discover whether the original intention of this thesis, namely, to discover missing or misunderstood knowledge of monetary policy at some stage, should be also applied to this theory. Additionally, a more accurate accounting measure for the cost of reserves should be developed, so that the impact ratio of the short-term rates on the loans rate is exactly quantified. That knowledge would facilitate the analysis regarding how increases in interest rates drive the evolution of variables such as output, inflation, money supply or lending.

The last line of research that this thesis should open is related to the low inflation levels, which may also unveil further factors determining the periods of high inflation. When banks are trapped with expensive reserves, as it has been happening from the early 1980s, the impact ratio of the federal funds rate on the loans rate was above one until the spread reached a hypothetical ceiling of three percentage points, being almost the maximum of the period under analysis. A ratio above one implies stronger restrictions on the economy that little by little push inflation, and subsequently, interest rates, to lower levels. If during that descending path and once interest rates reach a floor, the accumulation of debt is significant, more difficulties will be found to raise interest rates in the future without causing a recession or a worse episode. Thus, given the current situation where banks are holding reserves at a cost above market rates, an urgent analysis is necessary to solve how to wipe out the significant amount of excess reserves in the banking sector. Thereby, the ceiling seeing on the spread could be reduced and the loans rate could increase in the future along with short-term rates without damaging output growth and employment, and allowing a smooth recovery in debt and inflation levels. The interest rate paid on required and excess reserves may provide some options to achieve it.

Bibliography

Adrian, T. and Shin, H. S. (2008). Financial Intermediaries, Financial Stability, and Monetary Policy. *Federal Reserve Bank of New York. Staff Reports* 346.

Ahmadi, P. A., and Ritschl, A. (2009). Depression Econometrics: A FAVAR Model of Monetary Policy during the Great Depression. *SFB 649 Discussion Paper 2009-054*.

Altunbas, Y., Gambacorta, L. and Marqués, L. (2007). Securitisation and The Bank Lending Channel. *European Central Bank Working Paper Series*, No. 838.

Ashcraft, A., McAndrews, J. and Skeie, D. (2011). Precautionary Reserves and the Interbank Market. *Journal of Money, Credit and Banking*, 43(2): 311-348.

Barth III, M. J. and Ramey, V. A. (2001). The Cost Channel of Monetary Transmission. *NBER Macroeconomics Annual 2001* Volume 16, Bernanke and Rogoff 2002.

Benati, L., and Surico, P. (2009). VAR Analysis and the Great Moderation. *American Economic Review.* 99 (4): 1636-1652.

Bernanke, B.S. (1983). Non-monetary Effects of the Financial Crisis in the Propagation of the Great Depression. *American Economic Review* 73, 257-276.

Bernanke, B.S. (1993). Credit in the Macroeconomy. *Federal Reserve Bank of New York, Quarterly Review. Spr*., 50-70.

Bernanke, B.S. (1995). The Macroeconomics of the Great Depression: A Comparative Approach. *Journal of Money, Credit and Banking* 27 (1): 1-28.

Bernanke, B. S. (2010). Federal Reserve's exit strategy. *Testimony of Federal Reserve Official. U.S. House of Representatives*, Washington, D.C.

Bernanke, B.S. and Blinder, A.S. (1992). Federal Funds Rate and the Channels of Monetary Transmission. *The American Economic Review,* 82 (4): 901-921.

Bernanke, B.S., Boivin, J. and Eliasz, P. (2005). Measuring The Effects Of Monetary Policy: A Factor-Augmented Vector Autoregressive (FAVAR) Approach. *Quarterly Journal of Economics* 120, 387-422.

Bernanke, B.S. and Gertler, M. (1995). Inside the Black Box: The Credit Channel of Monetary Policy Transmission. *The Journal of Economic Perspectives*. 9 (4): 27-48.

Bernanke, B.S. and Gertler, M. (1998). The Liquidity Effect and Long-Run Neutrality. *Carniege-Rochester Conference Series on Public Policy*, 49 (1): 149-194.

Bernanke, B.S., and James, H. (1991). The Gold Standard, Deflation and Financial Crisis in the Great Depression: An International Comparison. R.G. Hubbard, ed., *Financial Markets and Financial Crisis*, Chicago: University of Chicago Press, IL 19, 33-68.

Bernanke, B. S. and Mihov, I. (1998). Measuring Monetary Policy. *Quarterly Journal of Economics* 108 (3): 869-902.

Bindseil, U. (2004). The Operational Target of Monetary Policy and The Rise and Fall of Reserve Position Doctrine. *European Central Bank Working Paper Series*, No. 372.

Bindseil, U., Camba-Mendez, G., Hirsch, A. and Weller, B. (2006). Excess reserves and the implementation of monetary policy of the ECB. *Journal of Policy Modeling*, 28: 491-510.

Blanchard, O. and Simon, J. (2001). The Long and Large Decline in U.S. Output Volatility. *Brookings Papers on Economic Activity* 2001 (1): 135-164.

Boivin, J. (2000). The Fed's Conduct of Monetary Policy: Has it Changed and does it Matter? Essays on the Analysis of Changes in the Conduct of Monetary Policy. *Princeton University, PhD Dissertation,* Chapter 3.

Boivin, J., and Giannoni, M. P. (2006). Has Monetary Policy Become More Effective? *Review of Economics and Statistics* 88, 45-462.

Boivin, J., and Giannoni, M. P. (2002). Assessing Changes in the Monetary Transmission Mechanism: A VAR Approach. *Economic Policy Review,*8(1):97-111.

Bolton, P. and Freixas, X. (2006). Corporate Finance and the Monetary Transmission Mechanism. *The Review of Financial Studies,* 19 (3): 829-870.

Bordo, M. B. (1993). The Gold Standard, Bretton Woods and Other Monetary Regimes: An historical appraisal. *NBER Working Paper* No. 4310.

Bordo, M. D. and Schwartz, A. J. (1999). Monetary Policy Regimes and Economic Perdormance: The Historical Record. In: *Taylor, J.B. and Woodford, M. (Eds.), Hanbook of Macroeconomics. Elsevier, Amsterdam*, pp. 149-234.

Borio, C. and Disyatat, P. (2009). Unconventional Monetary Policies: An Appraisal. *BIS Working Papers* 292.

Boschen, J.F. and Mills, L.O. (1995). The Relation between Narrative and Money Market Indicators of Monetary Policy. *Economic Inquiry,* 33 (1): 24-44.

Bowman, D., Gagnon, E. and Leahy, M. (2010). Interest on Excess Reserves as a Monetary Policy Instrument: The Experience of Foreign Central Banks. *International Finance Discussion Papers* 996.

Brunner, K., and Meltzer, A. H. (1968). What Did We Learn from the Monetary Experience of the United States in the Great Depression? *Canadian Journal of Economics*1, 334-348.

Cagan, P. (1969). Interest Rates and Bank Reserves—A Reinterpretation of the Statistical Association. *In Jack M. Guttentag and Phillip Cagan, eds. Essays on Interest Rates* 1, 223-271.

Calomiris, C.W., Mason, J. and Wheelock, D. (2011). Did Doubling Reserve Requirements Cause the Recession of 1937-1938? A Microeconomic Approach. *NBER Working Paper No.* 16688.

Canova, F. and Gambetti, L. (2010). Do Expectations Matter? The Great Moderation Revisited. *American Economic Journal* 2 (3): 183-205.

Canova, F and Forero, F.J. (2015). Estimating Overidentified, Nonrecursive, Time-Varying Coefficients Structural Vector Autoregressions. Quantitative Economics 6 (2): 359-384.

Carlson, M. and Burcu D. (2016). The Tools and Transmission of Federal Reserve Monetary Policy in the 1920s. FEDS Notes. *Washington: Board of Governors of the Federal Reserve System,* https://doi.org/10.17016/2380-7172.1871.

Carpenter, S.B., Demiralp, S. (2012). Money, Reserves, and the Transmission of Monetary Policy: Does the Money Multiplier Exist? *Journal of Macroeconomics*. 34 (1): 59-75.

Cecchetti, S.G. (1992). Prices during the Great Depression: Was the Deflation of 1930-1932 Really Unanticipated? *American Economic Review* 92, 141-156.

Chang, S., Contessi, S. and Francis, J. L. (2014). Understanding the accumulation of bank and thrift reserves during the U.S. financial crisis. *Journal of Economic Dynamics and Control*, 43: 78-106.

Christiano, L. J., Eichenbaum, M., Evans, C. L. (1998). Monetary Policy Shocks: What have We Learned and to What End? In: *Taylor, J.B. and Woodford, M. (Eds.), Hanbook of Macroeconomics. Elsevier, Amsterdam*, 65-148.

Clarida, R., Galí, J., and Gertler, M. (2000). Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory. *Quarterly Journal of Economics* 115 (1): 147-180.

Clouse, J. A. (1994). Recent Developments in Discount Window Policy. *Federal Reserve Bulleting*, November, 965-977.

Cogley, T. and Sargent, T. J. (2001). Evolving Post-World War II U.S. Inflation Dynamics. *NBER Macroeconomics Annual* 16.

Cogley, T. and Sargent, T. J. (2005). Drifts and Volatilities: Monetary Policies and Outcomes in the Post WWII US. *Review of Economic Dynamics* 8, 262-302.

Cook, T. and Hahn, T. (1988). The Information of Discount Rate Announcements and Their Effect on Market Interest Rates. *Journal of Money, Credit and Banking,* 20 (2): 167-180.

Cottarelli, C. and Kourelis, A. (1994). Financial Structure, Bank Lending Rates, and the Transmission Mechanism of Monetary Policy. *Palgrave Macmillan Journals on behalf of the International Monetary Fund*. 41(4): 587-623.

Disyatat, P. (2010). The Bank Lending Channel Revisited. *BIS Working Papers 297.*

Dow Jr, P. J. (2001). The Demand for Excess Reserves. *Southern Economic Journal,* 67(3): 685-700.

Dudley, W.C. (2009). The Economic Outlook and the Fed's Balance Sheet: The Issue of "How" versus "When". *Federal Reserve Bank of New York*, *Speech*.

Dwyer, G. P. (2010). Excess Reserves in the 1930s: A Precautionary Tale. Federal Reserve Bank of Atlanta. Center for Financial Innovation and Stability. Notes from the Vault.

Eichengreen, B. (1990). International Monetary Instability Between the Wars: Structural Flaws or Misguided Policies? Published as: The Evolution of the International Monetary System, edited by Suzuki Y., Miyake J. and Okabe, M. Universityof Tokyo Press, 71-116.

Ennis, H. M. and Keister, T. (2008). Understanding Monetary Policy Implementation. *FRB Richmond Economic Quarterly*, 94(3): 235-263.

Ennis, H. M. and Wolman, A. L. (2015). Large Excess Reserves in the United States: A view from the Cross-Section of Banks. *International Journal of Central Banking*, 11(1): 251-289.

Favero, C. A. and Rovelli, R. (2003). Macroeconomic Stability and the Preferences of the Fed: A Formal Analysis, 1961-98. *Journal of Money*, *Credit and Banking* 35 (4): 545-556.

Fedpoint (2015). The Discount Window. Federal Reserve Bank of New York. *https://www.newyorkfed.org/aboutthefed/fedpoint/fed18.html*

Fernández-Villaverde, J., Rubio-Ramírez, J. F., Sargent, T. J. and Watson, M. W. (2007). ABCs (and Ds) of Understanding VARs. *American Economic Review*, 97 (3): 1021-1026.

Fforde, J. S. (1954). The Federal Reserve System 1945-1949. *Oxford University Press*.

Forni, M., Gambetti, L., and Sala, L. (2014). No News in Business Cycles. *The Economic Journal* 124, 1168-1191.

Forni, M. and Gambetti, L. (2010). The Dynamic Effects of Monetary Policy: A Structural Factor Model Approach. *Journal of Monetary Economics* 57 (2): 203-216.

Forni, M. and Gambetti, L. (2014). Sufficient Information in Structural VARs. *Journal of Monetary Economics* 66, 124-136.

Friedman, M. (1968). The Role of Monetary Policy. *The American Economic Review,* 58 (1).

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

Frost, P.A. (1971). Banks' Demand for Excess Reserves. *Journal of Political Economy*, 79 (4): 805-825.

Frost, P.A. and Sargent, T.J. (1970). Money-Market Rates, the Discount Rate, and Borrowing from the Federal Reserve. *Journal of Money, Credit and Banking,* 2(1): 56-82.

Galí, J. and Gambetti, L. (2009). On the Sources of the Great Moderation. *American Economic Journal*: Macroeconomics 1 (1): 26-57.

Gertler, M. and Gilchrist, S. (1993). The Role of Credit Market Imperfections in the Monetary Transmission Mechanism: Arguments and Evidence. *The Scandinavian Journal of Economics,* 95 (1): 43-64.

Geweke, J. (1992). Evaluating the Accuracy of Sampling-Based Approaches to the Calculation of Posterior Moments. *Bayesian Statistics,* Oxford University Press 4.

Giannone, D., Lenza, M. and Reichlin, L. (2008). Explaining the Great Moderation: It is not the Shocks. *Journal of the European Economic Association* 6, 621–633.

Goldfeld, S.M. and Kane, E.J. (1966). The Determinants of Member-Bank Borrowing: An Econometric Study. *Econometric Research Program, Research Memorandum* 81.

Goodfriend, M. (1993). Interest Rate Policy and the Inflation Scare Problem: 1979-1992. *FRB Richmond Economic Quarterly* 79 (1): 1-23.

Goodfriend, M. (2005). Interest on Reserves and Monetary Policy. *Economic Policy Review*, 8(1).

Goodhart, C. (2007). Whatever became of the Monetary Aggregates? *LSE Financial Markets Group Paper Series*. Special Paper 172.

Goodhart, C. (2010). Money, credit and bank behaviour: need for a new approach," *National Institute Economic Review*, 214(1).

Hamdani, K. and Peristani, S. (1991). A Disaggregate Analysis of Discount Window Borrowing. *Federal Reserve Bank of New York Quarterly Review.* 16 (2): 52-62.

Hamilton, J.D. (1987). Monetary factors in the Great Depression. *Journal of Monetary Economics* 19, 145-170.

Hanson, M. S. (2006). Varying Monetary Policy Regimes: A Vector Autoregressive Investigation. *Journal of Economics and Business* 58, 407-427.

Hetzel, R.L. (1985). The Rules Versus Discretion Debate over Monetary Policy in the 1920s. *Federal Reserve Bank of Richmond Economic Quarterly* 71, 3-14.

Hetzel, R. L. (2008). The Monetary Policy of the Federal Reserve: A History. *Cambridge University Press*.

Hetzel, R. L. and Leach, R. F. (2001). The Treasury-Fed Accord: A New Narrative Account. *Federal Reserve Bank of Richmond Economic Quarterly* 87.

Hetzel, R. L. (2008). The Monetary Policy of the Federal Reserve: A History.Cambridge University Press.

Higgins, B. (1980). Free Reserves and Monetary Policy. Federal Reserve Bank of Kansas City *Economic Review*. July-August, 15-26.

Holmes, A.R. (1969). Operational Constraints on the Stabilization of Money Supply Growth. *Money and Finance: Readings in Theory, Policy and Institutions*. 391-400.

Jakab, Z. and Kumhof, M. (2015). Banks Are Not Intermediaries of Loanable Funds— And Why This Matters. *Bank of England Working Paper* No. 529.

Jones, D.S. (1981). Contemporaneous vs. Lagged Reserve Accounting: Implications for Monetary Control. Federal Reserve Bank of Kansas City. *Economic Review. November, 3-19.*

Kashyap, A.K. and Stein, J.C. (1994). Monetary Policy and Bank Lending. *The University of Chicago Press,* 221-261.

Kashyap, A.K., Stein, J.C., and Wilcox, D.W. (1993). Monetary Policy and Credit Conditions: Evidence from the Composition of External Finance. *The American Economic Review,* 83 (1): 78-98.

Kashyap, A.K., Stein, J.C., and Wilcox, D.W. (1995). The Impact of Monetary Policy on Bank Balance Sheets. *Carnegie-Rochester Conference Series on Public Policy* 42, 151-195.

Kashyap, A.K., Stein, J.C., and Wilcox, D.W. (2000). What Do a Million Observations on Banks Say about the Transmission of Monetary Policy. *The American Economic Review* 90 (3): 407-428.

Kasriel, P. L. and Merris, R.C. (1982). Reserve Targeting and Discount Policy. *Economic Perspectives*, 6 (3).

Keister, T., Martin, A., and McAndrews, J. (2008). Divorcing Money from Monetary Policy. *FRBNY Economic Policy Review*, September.

Keister, T. and McAndrews, J. (2009). Why Are Banks Holding So Many Excess Reserves. Federal Reserve of New York. *Current Issues in Economics and Finances*, 15.

Koch, C. (2015). Deposit Interest Rate Ceilings as Credit Supply Shifters: Bank Level Evidence on the Effect of Regulation Q. *Journal of Banking and Finance* 61, 316-326.

Leeper, E. M. (1997). Narrative and VAR Approaches to Monetary Policy: Common Identification Problems. *Journal of Monetary Economics* 40 (3): 641–657.

Leeper, E. M., Walker, T. B., and Yang, S. S. (2013). Fiscal Foresight and Information Flows. *Econometrica, Econometric Society*, 81 (3): 1115-1145.

Leroy, S.F. (1979). Monetary Control Under Lagged Reserve Accounting. *Southern Economic Journal*. 42 (2): 460-470.

Levin, F.J and Meek, P. (1981). Implementing the New Operating Procedures: The View from the Trading Desk. In Stephen H. Axelrod, ed., *New Monetary Control Procedures.* Washington: Board of Governors of the Federal Reserve System.

Lindley, J. T., Sowell, C. B. and Stewart Mounts Jr. W.M. (2001). Excess Reserves during the 1930s: Empirical Estimates of the Cost of Converting Unintended Cash Inventory into Income-Producing Assets. *Journal of Economics and Finance,* 25(2).

Lippi, M., and Reichlin, L. (1994). VAR Analysis, Non-fundamental Representations, Blashke Matrices. *Journal of Econometrics* 63 (1): 307-325.

Lombra, R.E. (1992). Understanding the Remarkable Survival of Multiplier Models of Money Stock Determination. *Eastern Economic Journal,* 18 (3): 305-314.

Lütkepohl, H. (2005). New Introduction to Multiple Time Series Analysis. *Springer-Verlag*, Berlin.

Martin, A., McAndrews, J. and Skeie, D. (2011). A Note on Bank Lending in Times of Large Bank Reserves. Federal Reserve Bank of New York, *Staff Reports* 497.

Mayer, T. (1999). Monetary Policy and the Great Inflation in the United States: the Federal Reserve and the Failure of Macroeconomic Policy, 1965-1979. *Edward Elgar Publishing*.

McCallum, B.T. (1990). Could a Monetary Base Rule have Prevented the Great Depression? *Journal of Monetary Economics* 26 (1): 3-26.

McCallum, T.B and Hoehn, J.G. (1982). Money Stock Control With Reserve and Interest Rate Instruments Under Rational Expectations. *Journal of Money, Credit and Banking* 15 (1): 96-101.

McConnell, M. M. and Perez-Quiros, G. (2000). Output Fluctuations in the United States: What Has Changed since the Early 1980's. *American Economic Review* 90 (5): 1464-1476.

McLeay, M., Radia, A. and Thomas, R. (2014a). Money Creation in the Modern Economy: An Introduction. *Bank of England, Quarterly Bulletin*.

McLeay, M., Radia, A. and Thomas, R. (2014b). Money Creation in the Modern Economy. *Bank of England, Quarterly Bulletin.*

Meltzer A.H. (2003). A History of the Federal Reserve: 1913-1951. Volume I. The University of Chicago Press.

Meltzer A.H. (2009a). A history of the Federal Reserve: 1951-1969. Volume II.1. The University of Chicago Press.

Meltzer A.H. (2009b). A history of the Federal Reserve: 1970-1986. Volume II.1. The University of Chicago Press.

Meulendyke, A. (1988). A Review of Federal Reserve Policy Targets and Operating Guides in Recent Decades. *Federal Reserve Bank of New York Quarterly Review*. Autumn, 6-17.

Miron, J. A. (1988). The Founding of the Fed and the Destabilization of the Post-1914 Economy. *NBER Working Paper* 2701.

Moore, J.B. (1983). Unpacking the post Keynesian black box: bank lending and the money supply. *Journal of Post Keynesian Economics*. 5 (4): 537.

Moore, J. B. (1988). The Endogenous Money Supply. *Journal of Post Keynesian Economics*, 10(3).

Moore, C. H. (1990). The Federal Reserve System: A History of the 75 first years. McFarland & Co.

Moreno, A. (2004). Reaching Inflation Stability. *Journal of Money, Credit and Banking*, 36(4): 801-825.

Ogawa, K. (2007). Why Commercial Banks Held Excess Reserves: The Japanese Experience of the Late 90s. *Journal of Money, Credit and Banking*, 39(1): 241-257.

Oliner, D.S. and Rudebusch, G.D. (1995). Is There a Bank Lending Channel for Monetary Policy? *Federal Reserve of San Francisco Economic Review* 2, 3- 20.

Oliner, D.S. and Rudebusch, G.D. (1996). Monetary Policy and Credit Conditions: Evidence from the Composition of External Finance: Comment. *The American Economic Review*. 86 (1): 300-309.

Orphanides, A. (2003). Monetary policy in deflation: the liquidity trap in history and practice. *The North American Journal of Economics and Finance*, 15: 101-124.

Orphanides, A. (2004a). Monetary Policy in Deflation: the Liquidity Trap in History and Practice. *North American Journal of Economics and Finance* 15, 101-124.

Orphanides, A. (2004b). Monetary Policy Rules, Macroeconomic Stability, and Inflation: A view from the Trenches. *Journal of Money, Credit and Banking* 36, 151-175.

Orphanides, A. (2003). Historical Monetary Policy analysis and the Taylor Rule. *Journal of Monetary Economics* 50, 983-1022.

Pearce, D.K. (1993). Discount Window Borrowing and Federal Reserve Operating Regimes. *Economic Inquiry* 31, 564-579.

Peristani, S. (1991). The Model Structure of Discount Window Borrowing. *Journal of Money, Credit and Banking 23 (1): 13-34.*

Pivetta, F. and Reis, R. (2007). The Persistence of Inflation in the United States. *Journal of Economic Dynamics and Control* 31 (4): 1326-1358.

Plosser, C. I. (2011). Exit. *Shadow Open Market Committee*, New York. Speech.

Polakoff, M.E. (1960). Reluctance Elasticity, Least Cost, and Member-Bank Borrowing. A Suggested Integration. *The Journal of Finance* 15 (1): 1-18.

Polakoff, M.E. and Silber, W.L. (1967). Reluctance and Member-Bank Borrowing: Additional Evidence. *The Journal of Finance*. 22 (1): 88-92.

Primiceri, G. E. (2005). Time Varying Structural Vector Autoregressions and Monetary Policy. *Review of Economic Studies* 72, 821-852.

Raftery, A.E. and Lewis, S.M. (1992a). Comment: One Long Run with Diagnostics: Implementation Strategies for Markov Chain Monte Carlo. *Statistical Science*, 7 (4), 493-497.

Raftery, A.E. and Lewis, S.M. (1992b). The Number of Iterations, Convergence Diagnostic and Generic Metropolis Algorithm. In: *Gilks, W.R. and Spiegelhalter, D.J (Eds), Practical Markov Chain Monte Carlo,*115-130.

Ramey, V. (1993). How Important is the Credit Channel in the Transmission of Monetary Policy. *Carnegie-Rochester Conference Series on Public Policy* 39, 1-45.

Ramey, V.A. (2011). Identifying Government Spending Shocks: It's all in the Timing. *The Quarterly Journal of Economics*, Oxford University Press 126 (1): 1-50.

Rasche, R.H., Meltzer, A.H., Sternlight, P.D. and Axilrod, S.H. (1982). Is the Federal Reserve's Monetary Control Policy Misdirected? *Journal of Money, Credit and Banking* 14 (1): 119-147.

Ritschl, A. and Woitek, U. (2000). Did Monetary Forces Cause the Great Depression? A Bayesian VAR analysis for the US economy. *Discussion Paper No. 2547*, CEPR.

Rodrigues, A.P. (1993). Government Securities Investments of Commercial Banks. Federal Reserve Bank of New York *Quarterly Review*. Summer, 39-53.

Roley, V.V. and Troll, R. (1984). The Impact of Discount Rate Changes on Market Interest Rates. Federal Reserve Bank of Kansas City, *Economic Review. January, 27-39.*

Romer, C. D. and Romer, D. H. (1989). Does Monetary Policy Matter? A New Test in the Spirit of Friedman and Schwartz. *NBER Macroeconomics Annual* 4.

Romer, C.D. and Romer, D.H. (1990). New Evidence on the Monetary Transmission Mechanism" *Brookings Papers on Economic Activity*. 41 (1): 149-213.

Romer, C. D. and Romer, D. H. (2004). A New Measure Of Monetary Shocks: Derivation And Implications. *American Economic Review* 94, 1055-1084.

Rudebusch, G. D. (1998). Do Measures of Monetary Policy in a VAR Make Sense? *International Economic Review* 39 (4): 907-931. Symposium on Forecasting and Empirical Methods in Macroeconomics and Finance.

Sellon Jr., G.H. (1980). The Role of the Discount Rate in Monetary Policy: A Theoretical Analysis. Federal Reserve Bank of Kansas City *Economic Review.* June, 3-15.

Sellon, Jr. G.H. and Seibert, D. (1982). The Discount Rate: Experience Under Reserve Targeting. Federal Reserve Bank of Kansas City *Economic Review* September-October, 3-18.

Shapiro, M.D. (1994). Federal Reserve Policy: Cause and Effect. *Monetary Policy, Mankiw.*

Sims, A. C. (1998). Comment on Glenn Rudebusch's "Do Measures of Monetary Policy in a VAR Make Sense? *International Economic Review* 39 (4): 933-941, Symposium on Forecasting and Empirical Methods in Macroeconomicsand Finance.

Sims, A. C. and Zha, T. (2006). Were There Regime Switches in U.S. Monetary Policy? *American Economic Review* 96 (1): 54-81.

Stock, J. H. and Watson, M. W. (2003). Has the Business Cycle Changed and Why? *NBER Macroeconomics Annual* 2002, 17.

Strongin, S. (1995). The Identification of Monetary Policy Disturbances. Explaining the liquidity puzzle. *Journal of Monetary Economics* 35, 463-497.

Sternlight, P.D. (1980). Monetary Policy and Open Market Operations in 1979. Federal Reserve Bank of New York *Quarterly Review*. Summer, 50-64.

Taylor, H. (1983). The Discount Window and Money Control. *Business Review* 3-12.

Temin, P. (1976). Did Monetary Forces Cause the Great Depression? *New York: Norton.*

Thornton, D.L. (1982). The Discount Rate and Market Interest Rates: What's the Connection? *Federal Reserve Bank of St. Louis*. June-July, 3-14.

Todd, W. F. (2013). The Problem of Excess Reserves, Then and Now. *American Institute for Economic Research*, Working Paper No. 763.

Toma, M. (1989). The Policy Effectiveness of Open Market Operations in the 1920s. *Explorations in Economic History* 19, 99-116.

Toma, M. (1997). Competition and Monopoly in the Federal Reserve System, 1914-1951. *Cambridge University Press*.

Triffin, R. (1964). The Evolution of the International Monetary System: Historical Reappraisal and Future Perspectives. *Princeton Studies in International Finance* 12.

Uhlig, H. (2005). What are the Effects of Monetary Policy on Output? Results from an agnostic identification procedure. *Journal of Monetary Economics* 52, 381-419.

Urich, T. and Wachtel, P. (1981). Market Response to the Weekly Money Supply Announcements in the 1970s. *The Journal of Finance* 36 (5): 1063-1072.

Velde, F. R. (2004). Poor Hand or Poor Policy? The Rise and Fall of Inflation in the U.S. Federal Reserve Bank of Chicago. *Economic Perspectives* 28 (1): 34-51.

Wheelock, D.C. (1989). The Strategy, Effectiveness and Consistency of Federal Reserve Monetary Policy 1924-1933. *Explorations in Economic History* 26:453-476.

Wheelock, D.C (1990). Member Bank Borrowing and the Fed's Contractionary Monetary Policy during the Great Depression. *Journal of Money, Credit and Banking* 22 (4): 409-426.

Wheelock, D.C. (1991). The Strategy and Consistency of Federal Reserve Monetary Policy 1924-1933. Cambridge University Press.

Wicker, E. (1965). Federal Reserve Monetary Policy, 1922-33: A Reinterpretation. *Journal of Political Economy* 73, 325-343.

Wilcox, J. A. (1984). Excess Reserves in the Great Depression. *NBER Working Paper Series* No. 1374

Appendices

Appendix A - Impulse response functions

In this section the reader can find the remaining impulses responses from section 5. Even though they are not essential to understand the results obtained, they offer some details about the mechanisms working between instruments and final targets.

A.1 - Interwar period

Paying attention to an M1 shock (Figure A.4.7), the response of the inflation shows no clear pattern. Between 1925 and 1927 an increase in the money supply have a mostly positive response, although only after one month. During the period when gold inflows were accelerating but the Fed was offsetting them (1927-1929), increases in M1 would have increased the inflation more than in the rest of the period given the levels of deflation. From 1930 to 1933, the response is mostly negative, except for a delayed positive response between the second half of 1931 and first half of 1932. The stagnation of business, bank failures, the increase in currency and reserves holdings, and the gold outflow, could have contributed to this mostly negative response, without forgetting the negative spread. The delayed positive response between 1931 and 1932 could correspond to gold inflows and the pressures that the Fed received to purchase in the open market. Since 1933 the response of the inflation is initially negative and becomes slightly positive after three-five months. Again, accumulation of excess reserves, increases in reserve requirements and the negative spread could have contributed.

A shock to M1 (Figure A.4.8) shows mostly a negative response of the IPI between 1928-1929, years of deflation and sterilization of gold inflow. From 1925 to 1927, the response is mostly positive, except for the initial negative response in 1925 and part of 1926, again, periods of deflation or low inflation. Thus, deflation could have restrained output growth. Contrary to those periods of deflation, the response is positive from 1930-1932, (brief open market purchases and gold inflows for those years) becoming negative in the second half of 1932 and having a negative peak in 1933, likely due to banking failures. The negative response continues until the end of 1935, slightly after the devaluation and the purchase of silver and gold. It turns positive since 1936, when the Treasury desterilized the gold flows from previous years.

Regarding the impulse responses of M1, a shock to both, the discount rate (Figure A.4.9) and to C-D (Figure A.4.11) has a negative response, although for the last

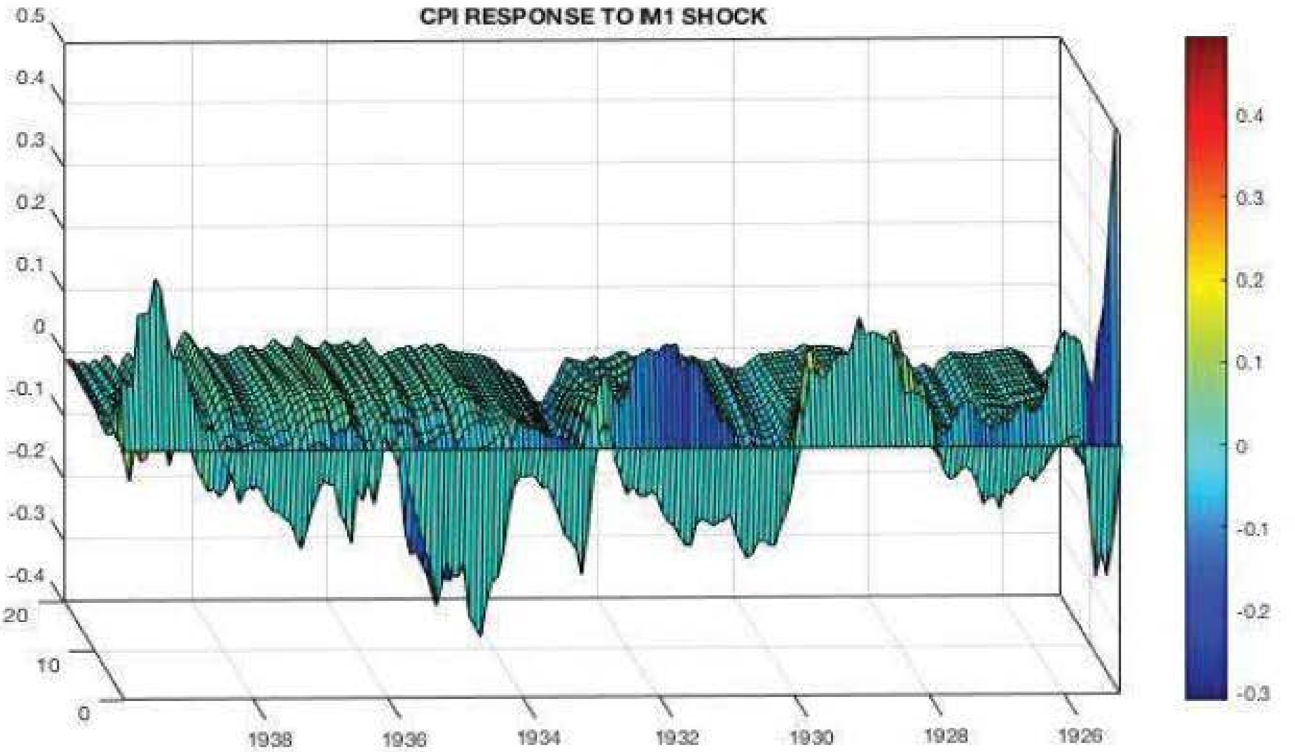

Figure A.4.7 - CPI inflation impulse response to an M1 shock. Note: Posterior means.

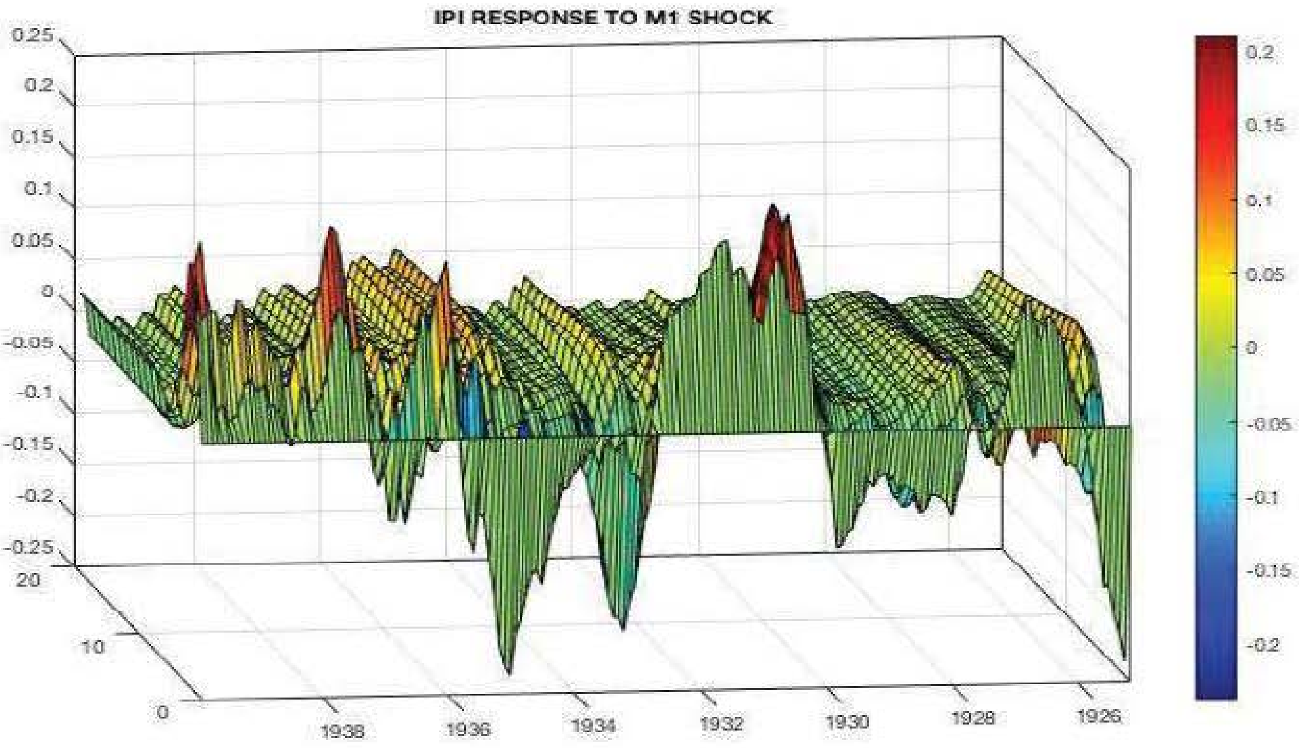

Figure A.4.8 - IPI impulse response to an M1 shock. Note: Posterior means.

figure, it becomes positive after two months and it is not significant (Figure 1.5.3). When the shock is to OMO (Figure A.4.10) the response is positive as expected, but not significant (Figure 1.5.2). There is a positive peak in 1929 belonging to the moment when the Fed perceived indications of a recession. Accordingly, the Fed purchased in the open market, the money supply increased, but it was not enough (as observed in 1.4.2) to increase inflation. Only the discount rate has a significant impact on M1 for virtually the entire period (Figure 1.5.1), except for 1930-1932 and 1937, corresponding to the sterilization of gold and deflation periods. For the other cases, the responses are significant mostly after the second month and last beyond the impulse response horizon. The response of C-D to a discount rate shock (Figure A.4.12) has a positive response that becomes negative or zero after approximately nine or ten months, from 1925 to the end of 1929. This is the period when the discount rate was below the call loans rate. From 1930 to the end of 1937 the initial response is negative, becoming positive after approximately six to eight months between 1930 and 1934, and one to two months from 1934 to the end of 1937. This happens while the discount rate is above the call loans rate. After 1937, when both rates were at the same level, the response is negative. Thus, the increase in the discount rate led to a relative higher increase in the call loan rate in the first part of the sample, and this pattern reversed since then, when the Fed and the Treasury tried to keep short-term rates low to finance government spending. The response of C-D to an OMO shock (Figure A.4.13) is negative for the entire period, although for 1928 the response becomes positive after two to four months. Therefore, purchases in the open market decreased short-term rates and consequently, the spread. Last, the response of OMO to a discount rate shock (Figure A.4.14) is negative for the whole period, as it would be expected.

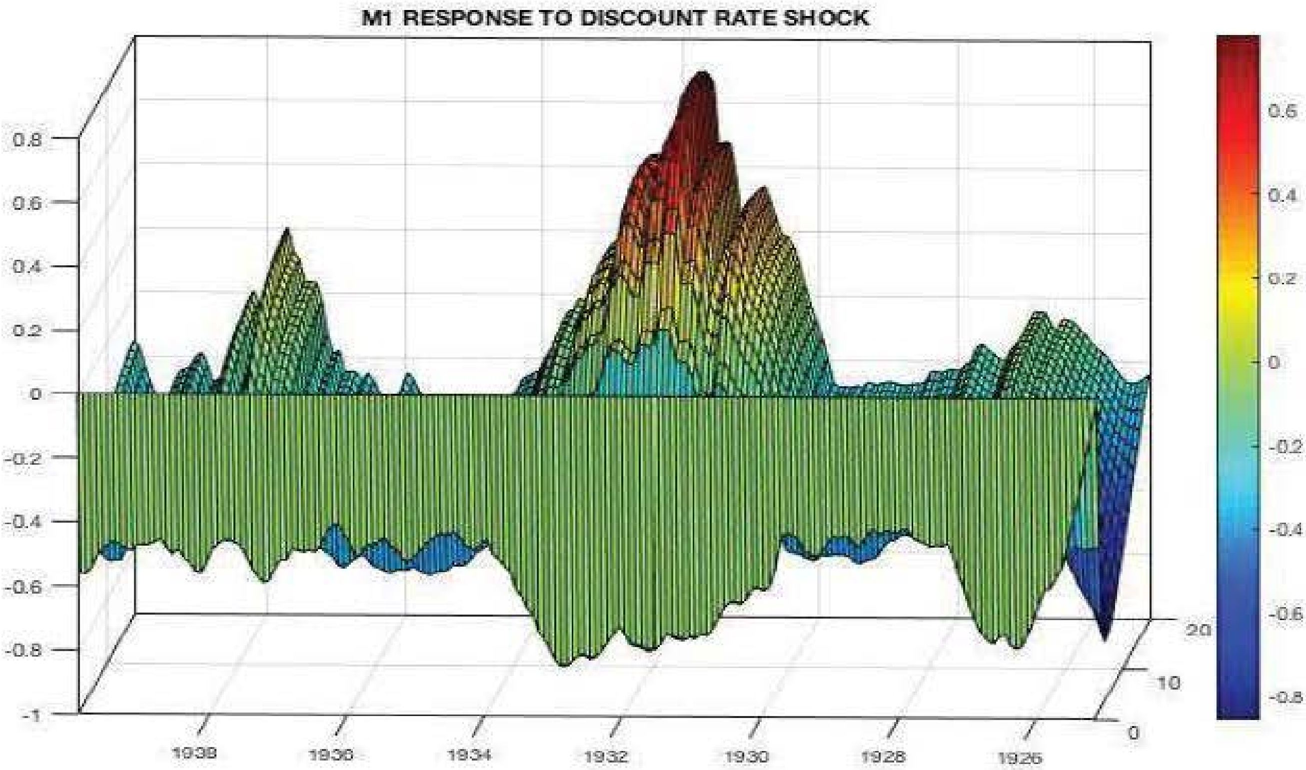

Figure A.4.9 - M1 impulse response to a discount rate shock. Note: Posterior means.

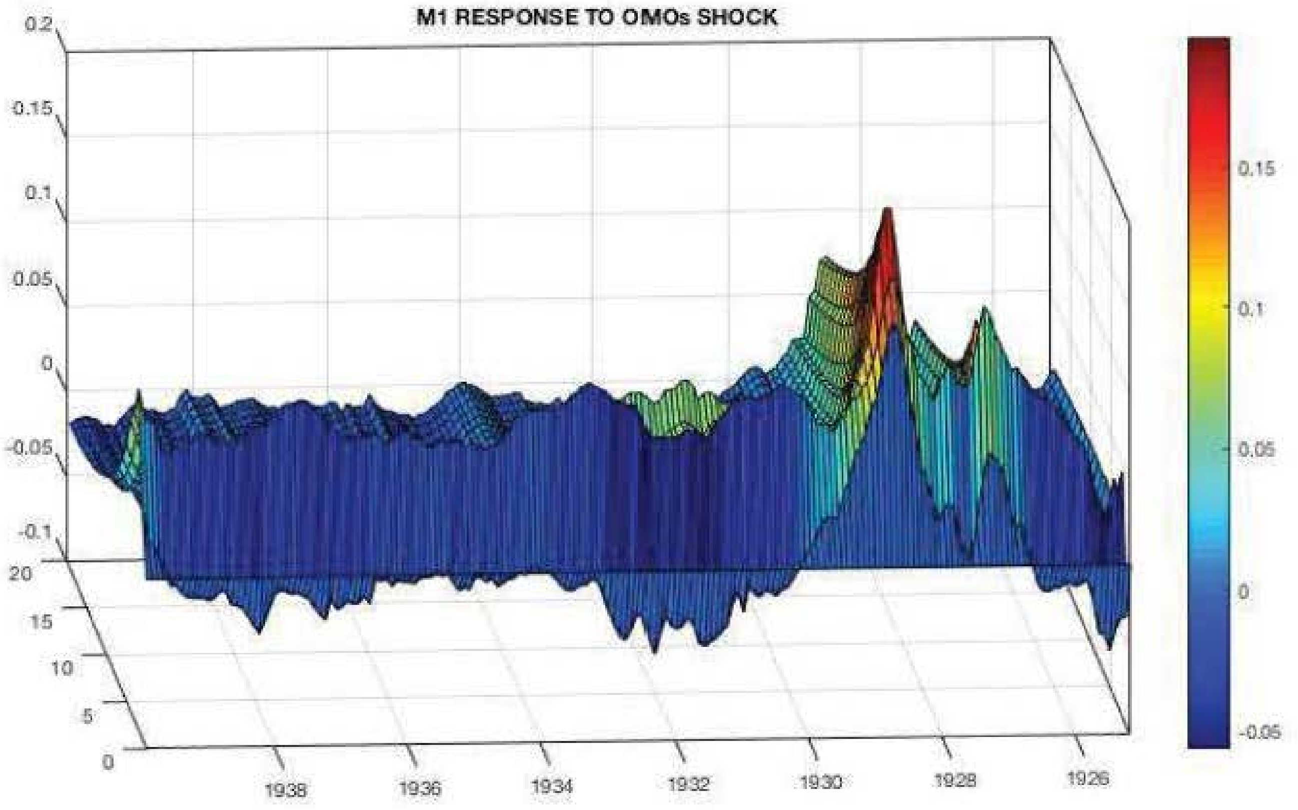

Figure A.4.10 - M1 impulse response to an OMO shock. Note: Posterior means.

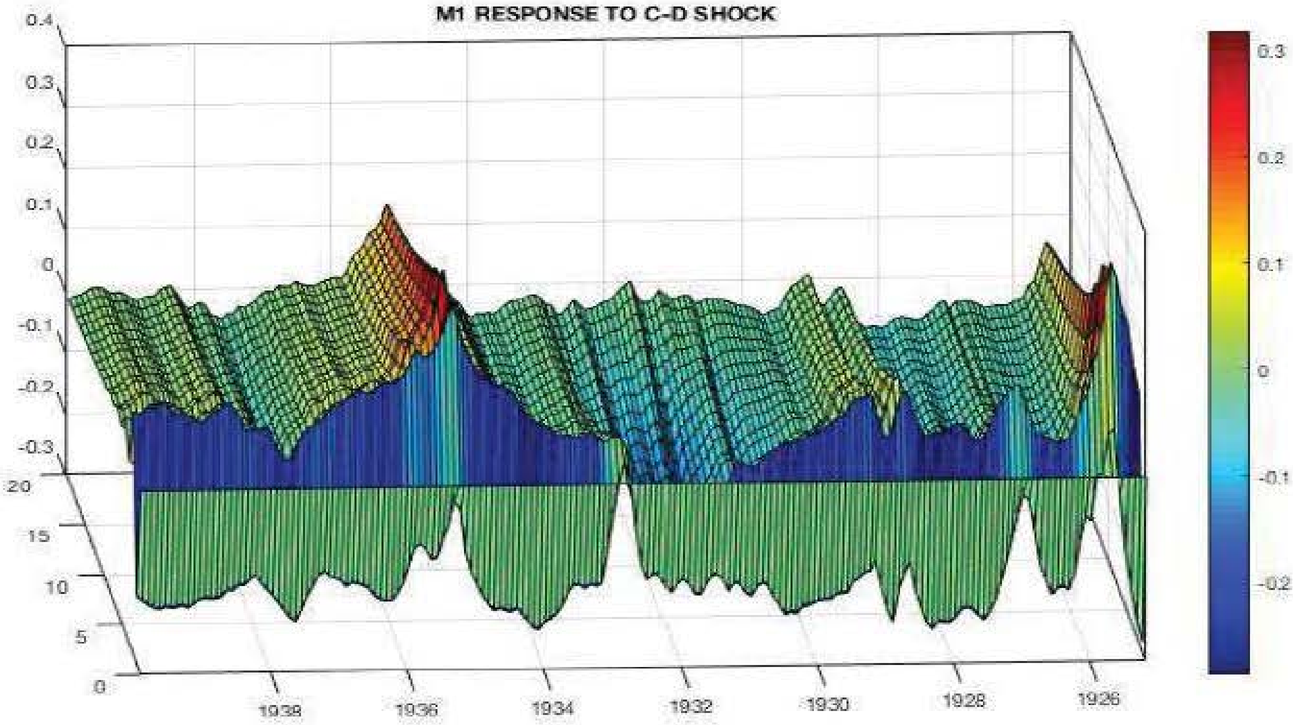

Figure A.4.11 - M1 impulse response to a C-D shock. Note: Posterior means.

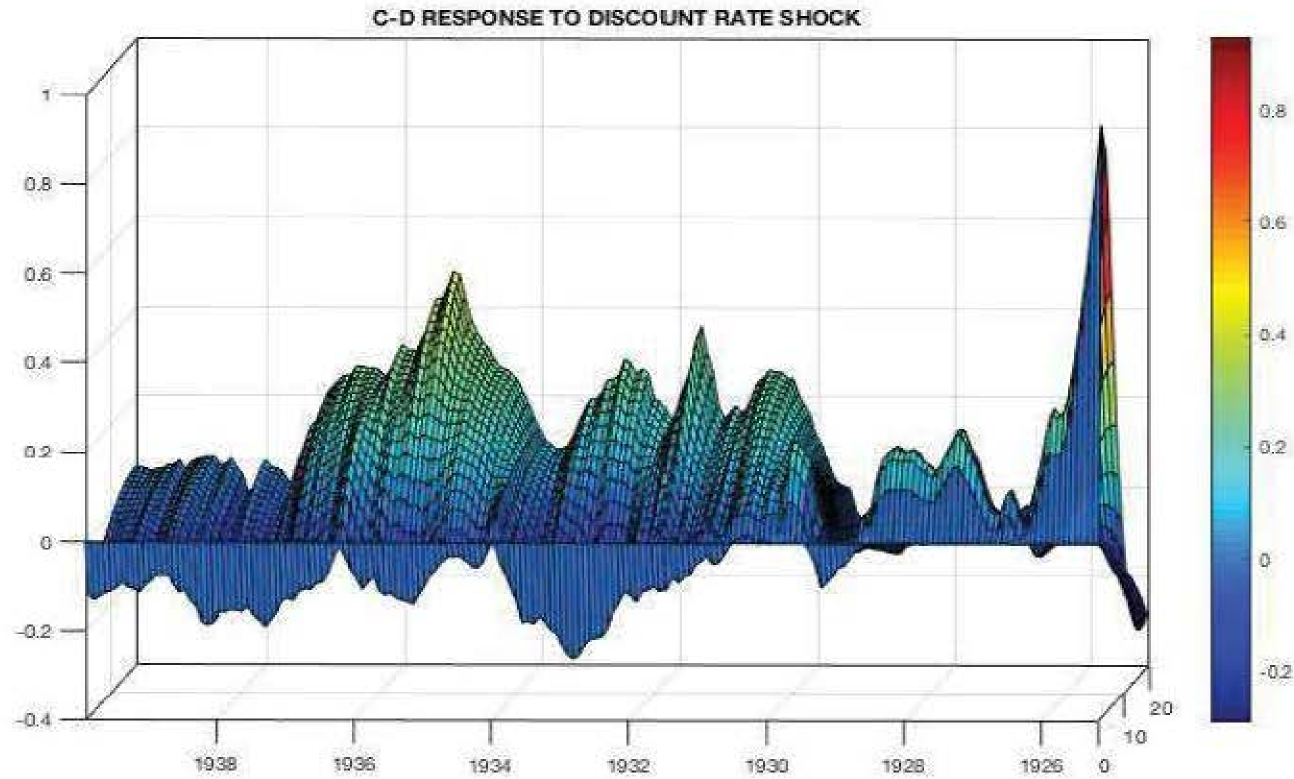

Figure A.4.12 - C-D impulse response to a discount rate shock. Note: Posterior means.

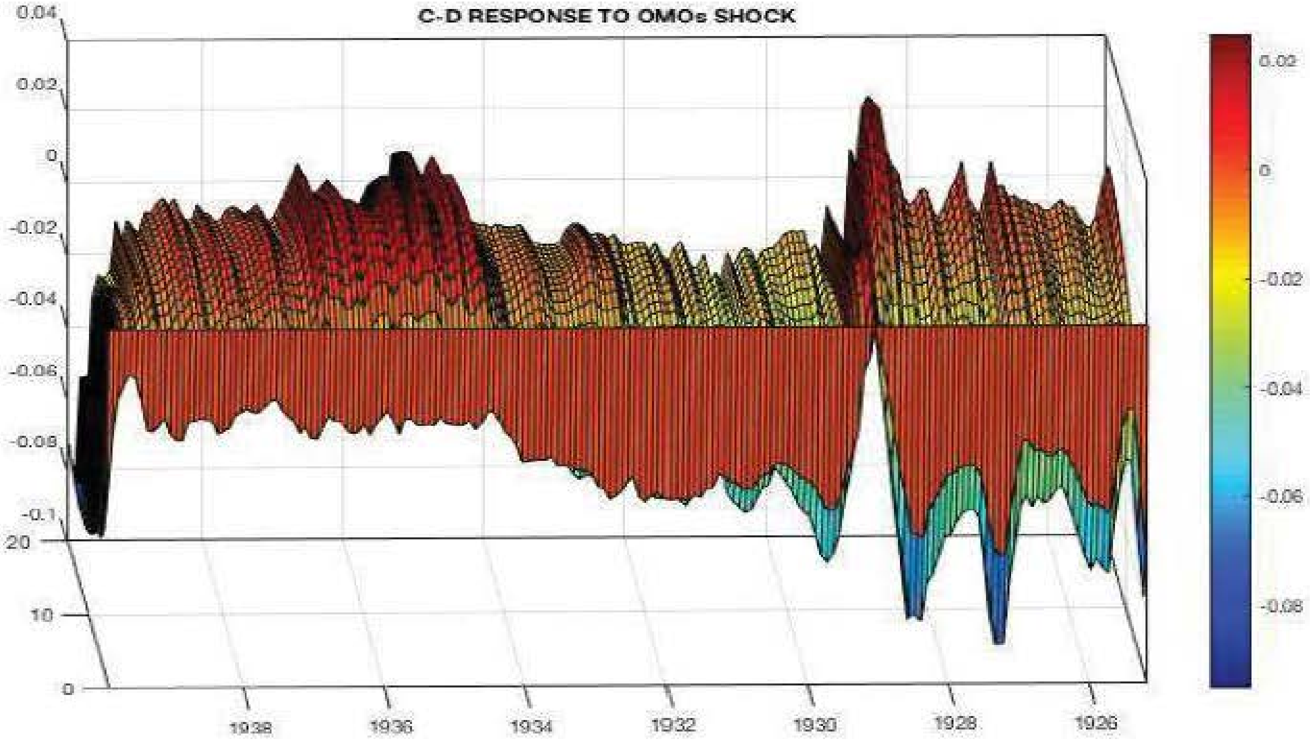

Figure A.4.13 – C-D impulse response to an OMO shock. Note: Posterior means.

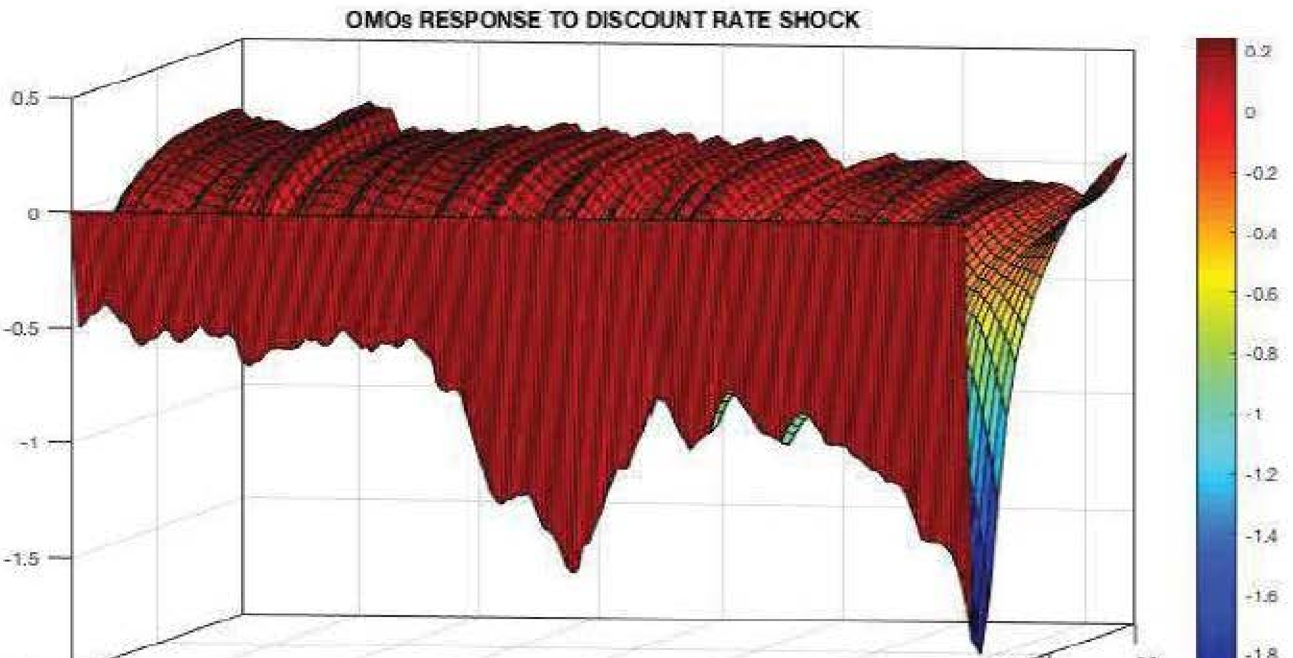

Figure A.4.14 – OMO impulse response to a discount rate shock. Note: Posterior means.

A.2 - 1958-2007

The response of M1 to a F-D shock (Figure A.7.7) is very similar to Figure A.7.1 Thus, for those years with positive spread banks increased borrowing and lending, increasing M1. After 1986, as already commented, there was another large spread. However, as in Figure 1.7.1, now the response is not as positive as in the other cases[29]. From 1970 to 1990 the positive response is significant (Figure 1.8.1). After 1995, when the spread was positive again, the response becomes more positive. Surprisingly, after 2003, although the discount rate was above the federal funds rate, and even above inflation after 2004, the response of M1 becomes increasingly positive. This time, however, it is not significant (Figure 1.8.1). The response of M1 to an OMO shock (Figure A.7.8) has a positive effect for the entire period but is not significant (Figure 1.8.2). The positive peaks occur when the federal funds rate and discount rate are at similar levels after their local maximum. This could indicate that when the Fed observed low borrowing at the discount window, possibly a signal of low credit growth, it purchased more in the open market to boost lending with lower rates.

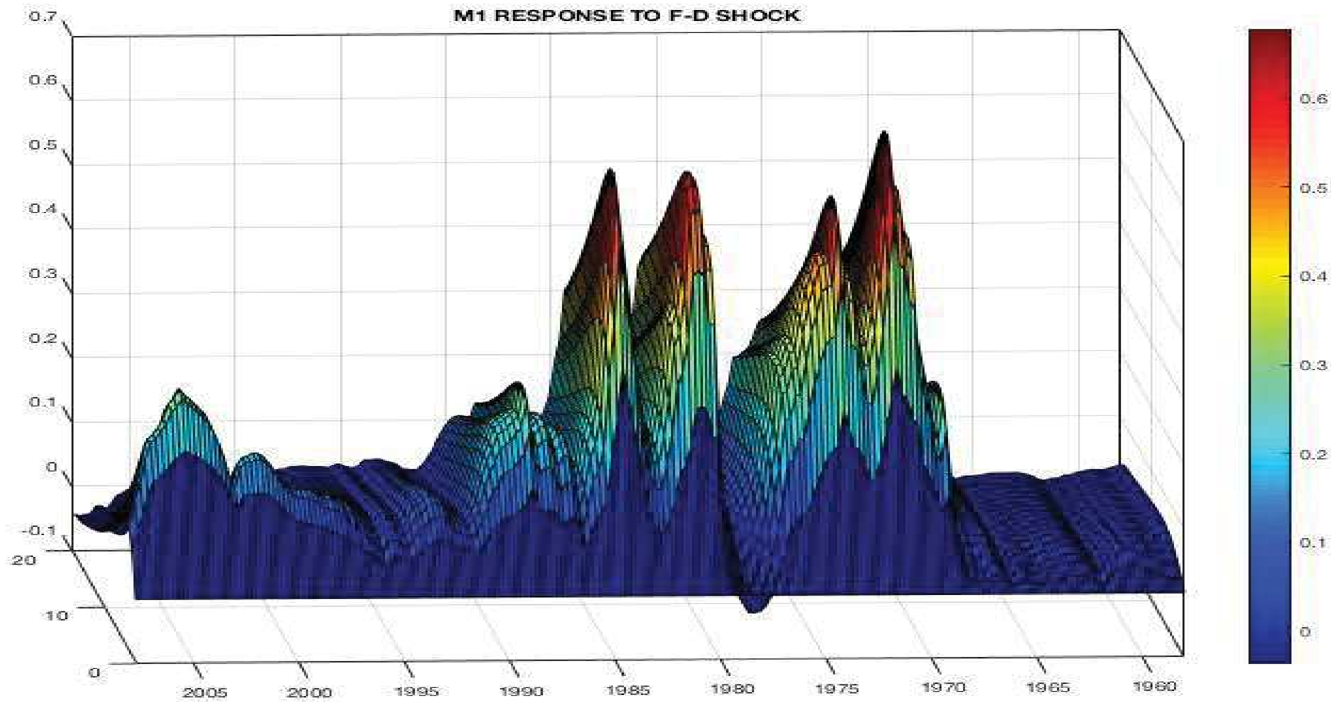

Figure A.7.7 – M1 impulse response to a F-D shock. Note: Posterior means.

[29] Using K-P algorithm and two lags, the impulse response shows even an initially negative response after 1985, although it turns positive shortly after.

This hypothesis is reinforced by the response of F-D to an OMO shock (Figure A.7.9), where the positive peaks coincide with periods of a small (positive or negative) or zero spread between rates, what occurs just after those rates have been in local maximum[30]. It seems that despite this pattern, the Fed extended its purchases a little more during the years of positive spread between 1969-1971. In 1990 the almost-zero or zero spread did not avoid Fed's purchases due to the recession. This is in line with Figure A.7.8, as after the 1980s is observed that despite having positive spreads in the periods 1986-1990 and 1994-2000, it seems that, likewise, the Fed purchased significantly in the open market, as the response of M1 to an OMO shock is positive. This could be related to the decrease in borrowing after the 1980s. Figure A.7.10 confirms the results of Figure 1.7.3, as the response of M1 to a discount rate shock is positive for the entire period, pointing out that despite the increase in the discount rate, M1 still was raising because of the purchases in the open market and increases in borrowing given the positive spreads. Both figures share significance for almost the same periods (Figure 1.8.3). Analysing the inflation response to an M1 shock (Figure A.7.11), it is negative almost for the whole period, not being the expected result. The negative peaks (except for that between 1960-1962) occur some time after the maximums of the federal funds rate. Hence, it seems to be capturing the following: when there were maximum in the federal funds rate (Figure 1.1), which normally coincided with large spreads, it was when M1 was in its highest levels. Subsequently, rates were lowered, because the economy could not endure those high rates and lending decreased. However, the higher supply of money was already in the market and it was a matter of time that prices increased. When it happened, the money supply had already begun to decrease (this is what the model is capturing here), as positive spreads had disappeared, together with the decline in borrowing and loans. Thus, the relationship M1-CPI may need more lags to capture the right effect. Beyond that possible explanation, the behavior changes after mid-1998. Inflation responds positively initially, although after two quarters, it goes to zero or slightly negative.

[30] Although this effect is not clear, as the model estimated with two lags show a negative response, OMO impact is more likely to be captured with one lag. Hence, I consider the results with 1 lag more reliable in this case.

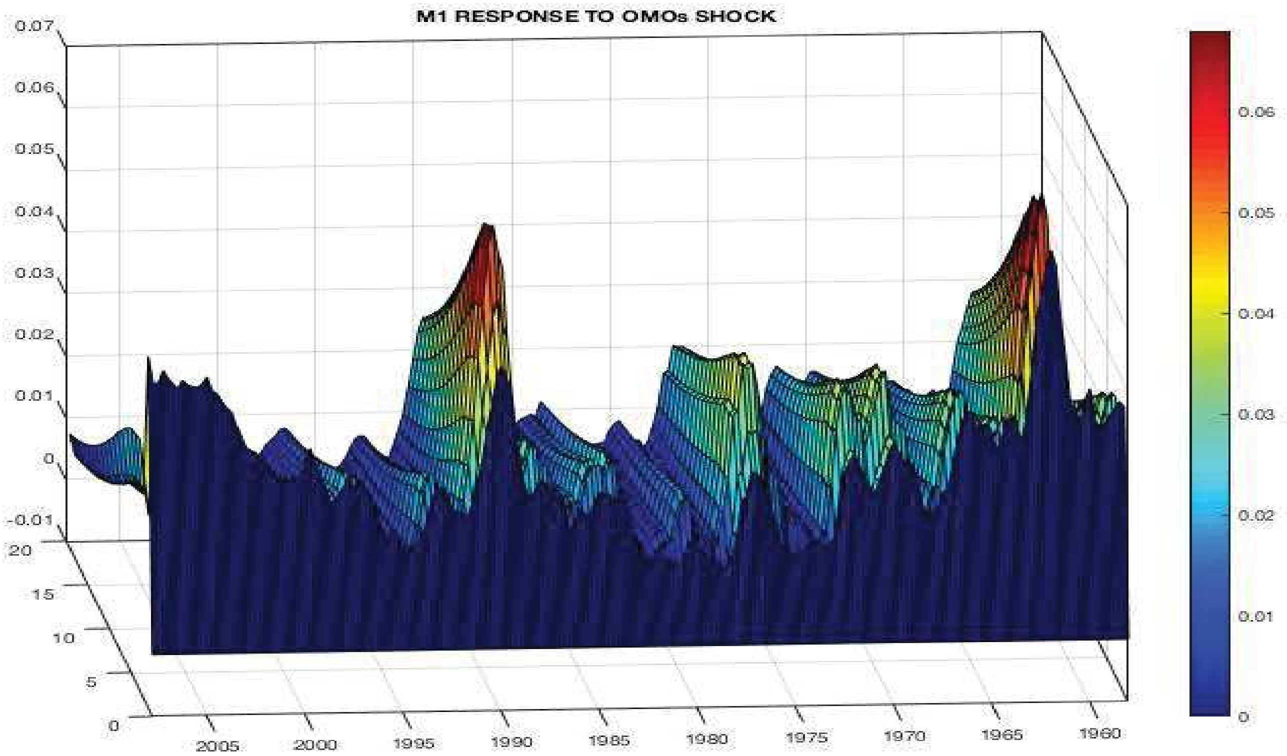

Figure A.7.8– M1 impulse response to an OMO shock. Note: Posterior means.

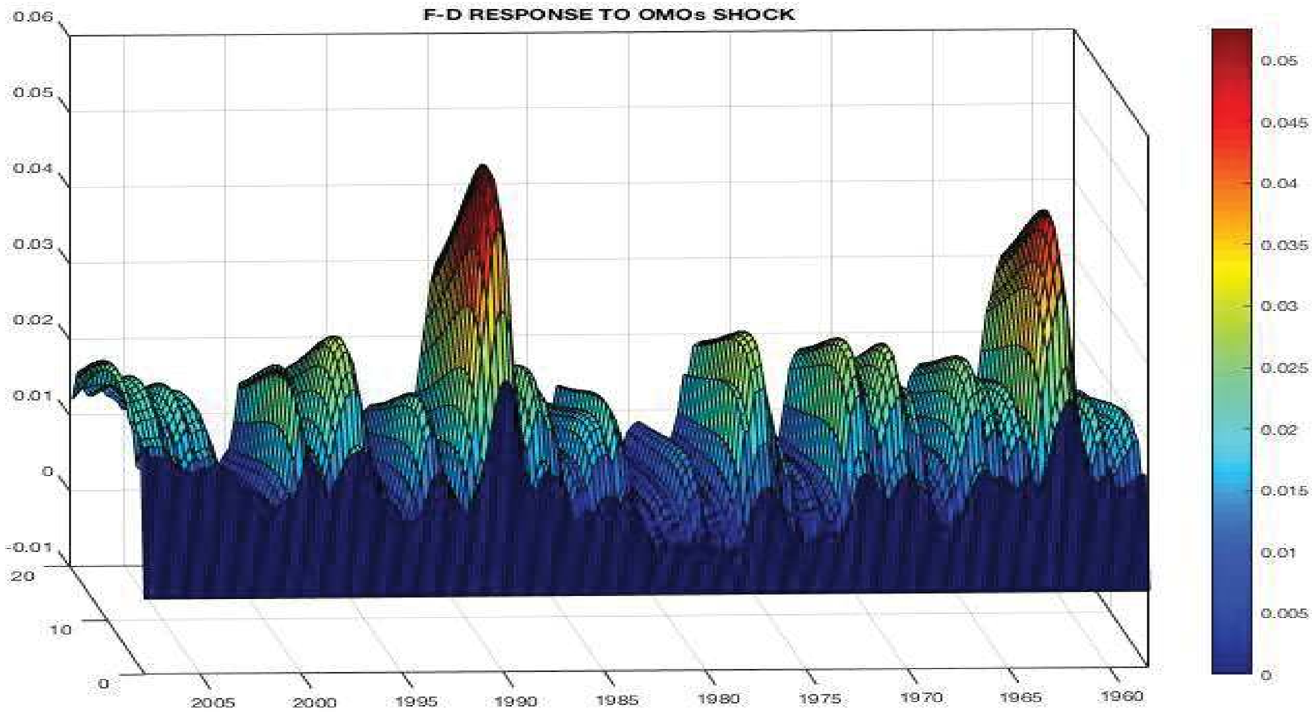

Figure A.7.9– F-D impulse response to an OMO shock. Note: Posterior means.

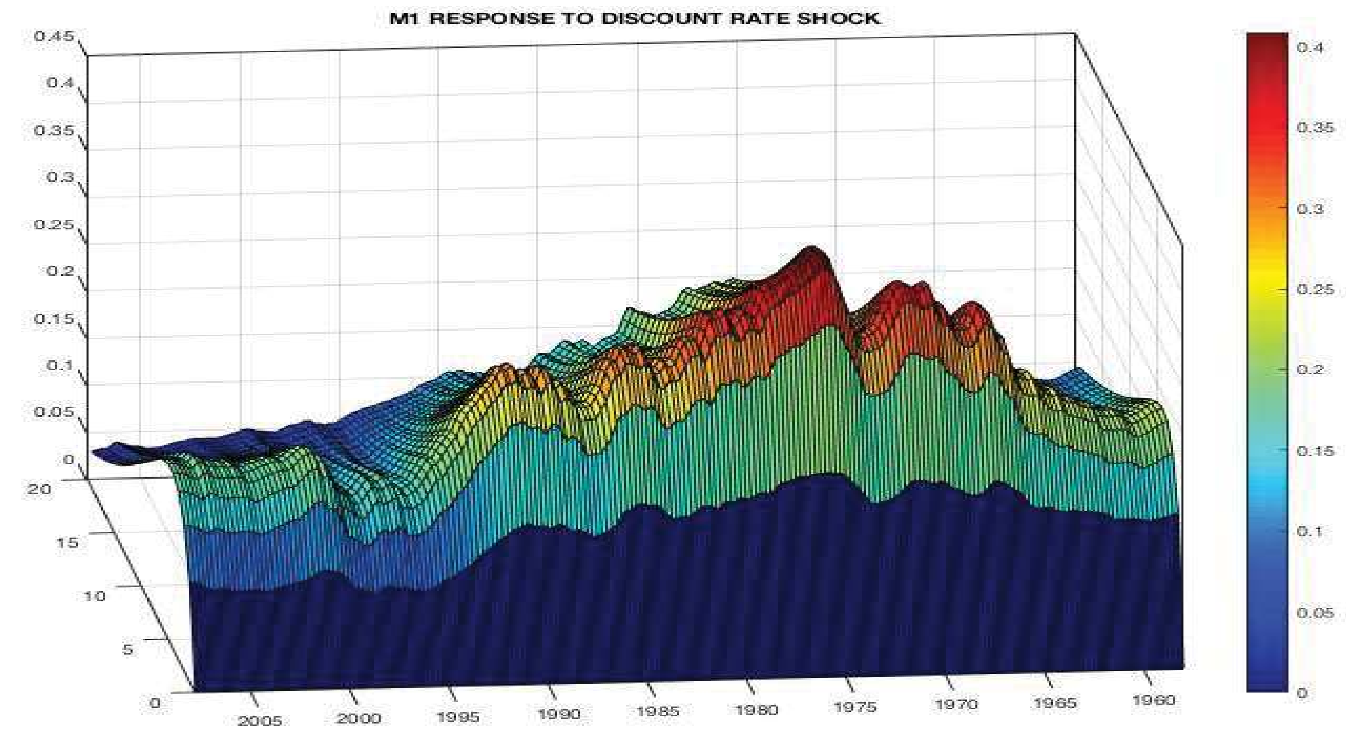

Figure A.7.10– M1 impulse response to a discount rate shock. Note: Posterior means.

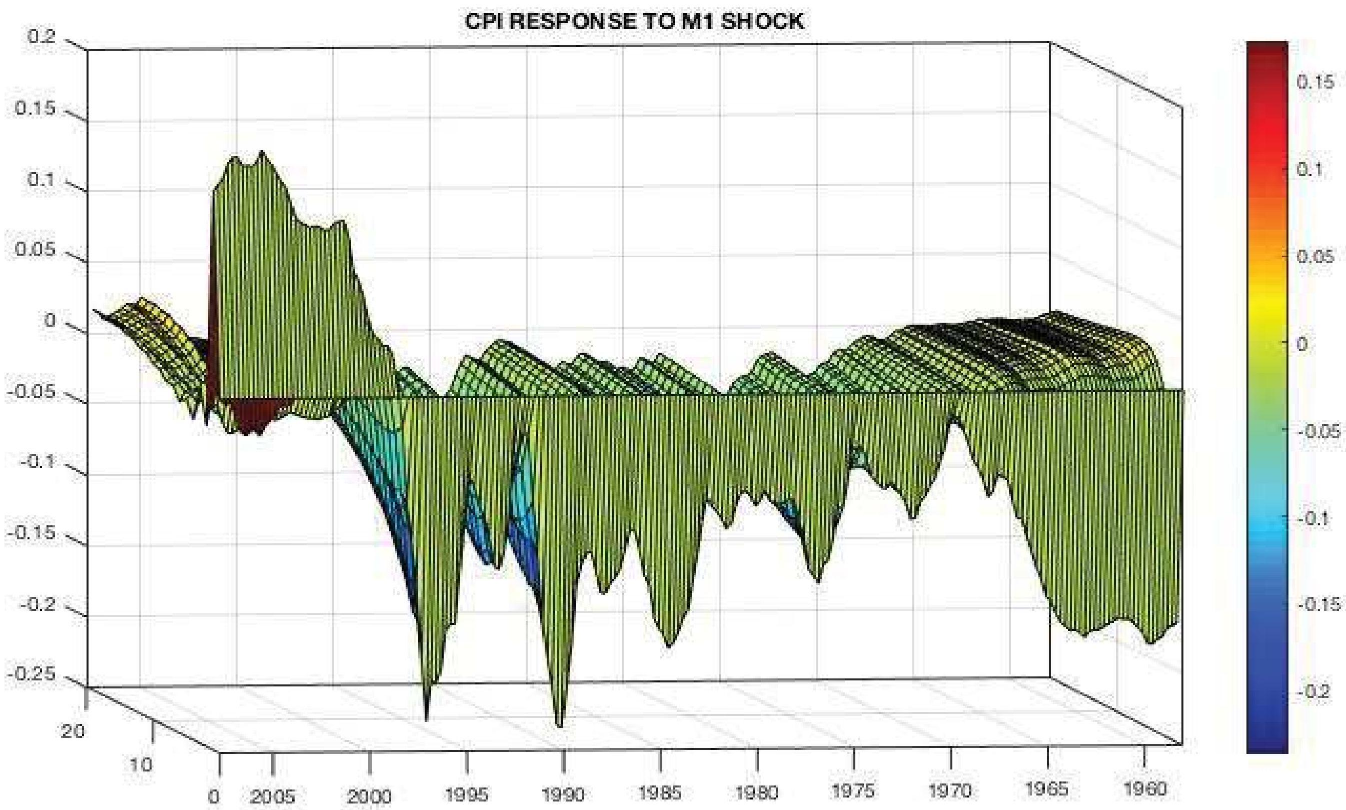

Figure A.7.11– CPI Inflation impulse response to an M1 shock. Note: Posterior means.

The response of the IPI to an M1 shock (Figure A.7.12) is initially negative, but after three or four quarters the response becomes positive, although this only starts to happen after approximately 1966. Subsequently, the initial negative response diminishes until disappearing at the end of 1987. After 1995, those initial or belated positive peaks occur when the spread was almost zero or negative. Again, the initial negative response could be a delayed effect, pointing out the necessity of more lags.

Last, the discount rate shock to F-D (Figure A.7.13) has a positive response for the whole period, showing that when the Fed increased the discount rate, the federal funds rate increased more in relation to the discount rate. The response of OMO to a discount rate (Figure A.7.14) shock is negative, as expected.

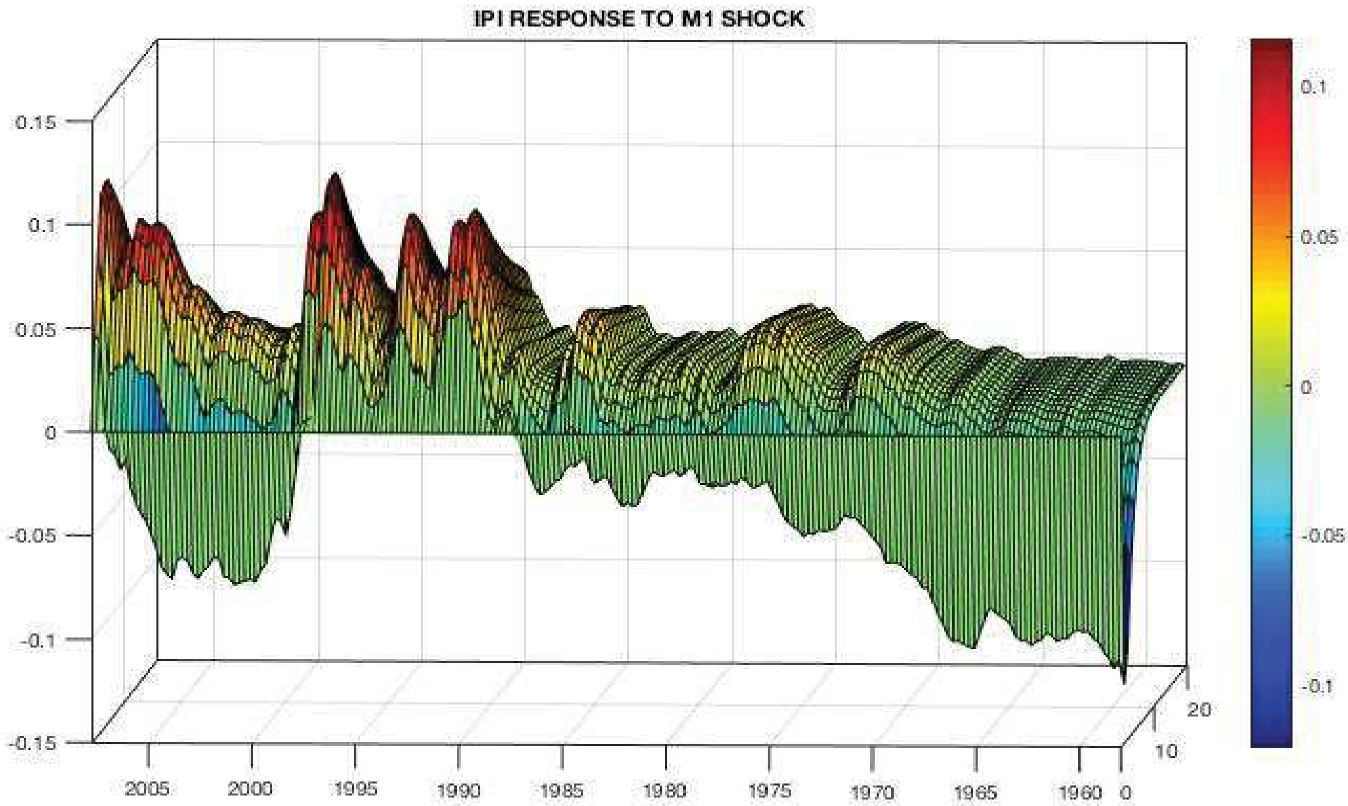

Figure A.7.12– IPI impulse response to an M1 shock. Note: Posterior means.

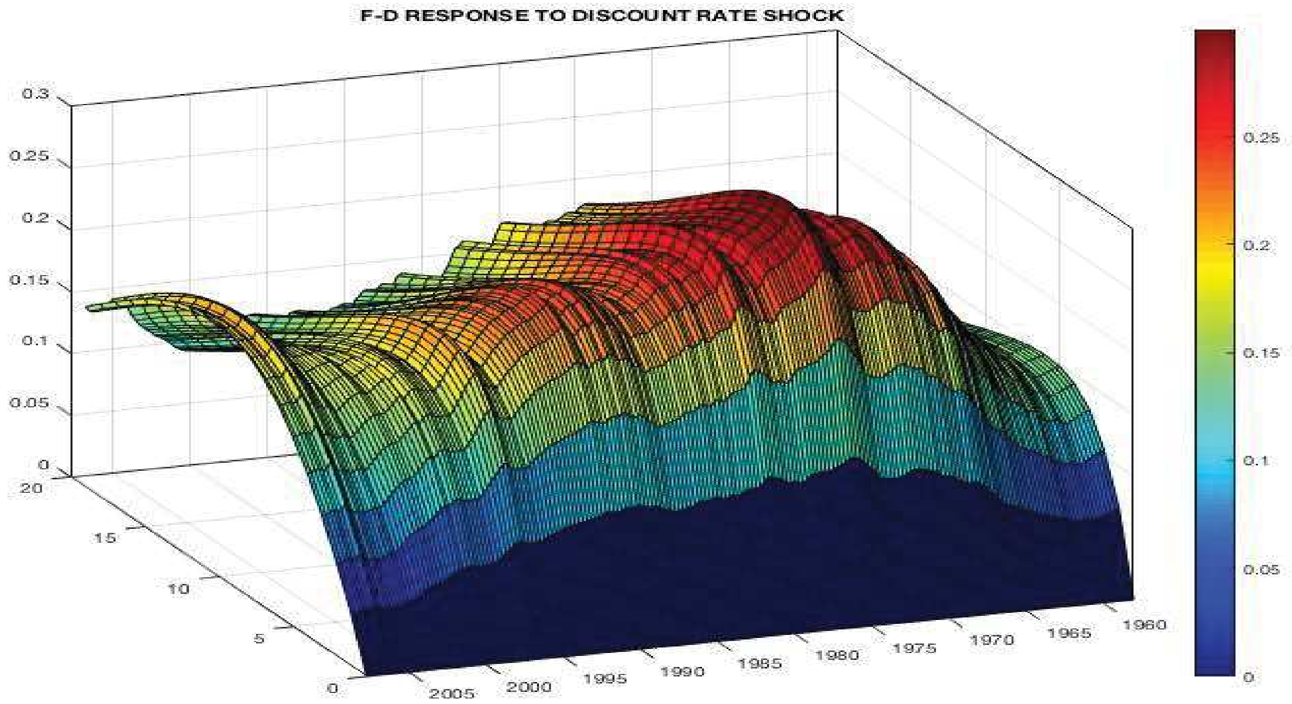

Figure A.7.13– F-D impulse response to a discount rate shock. Note: Posterior means.

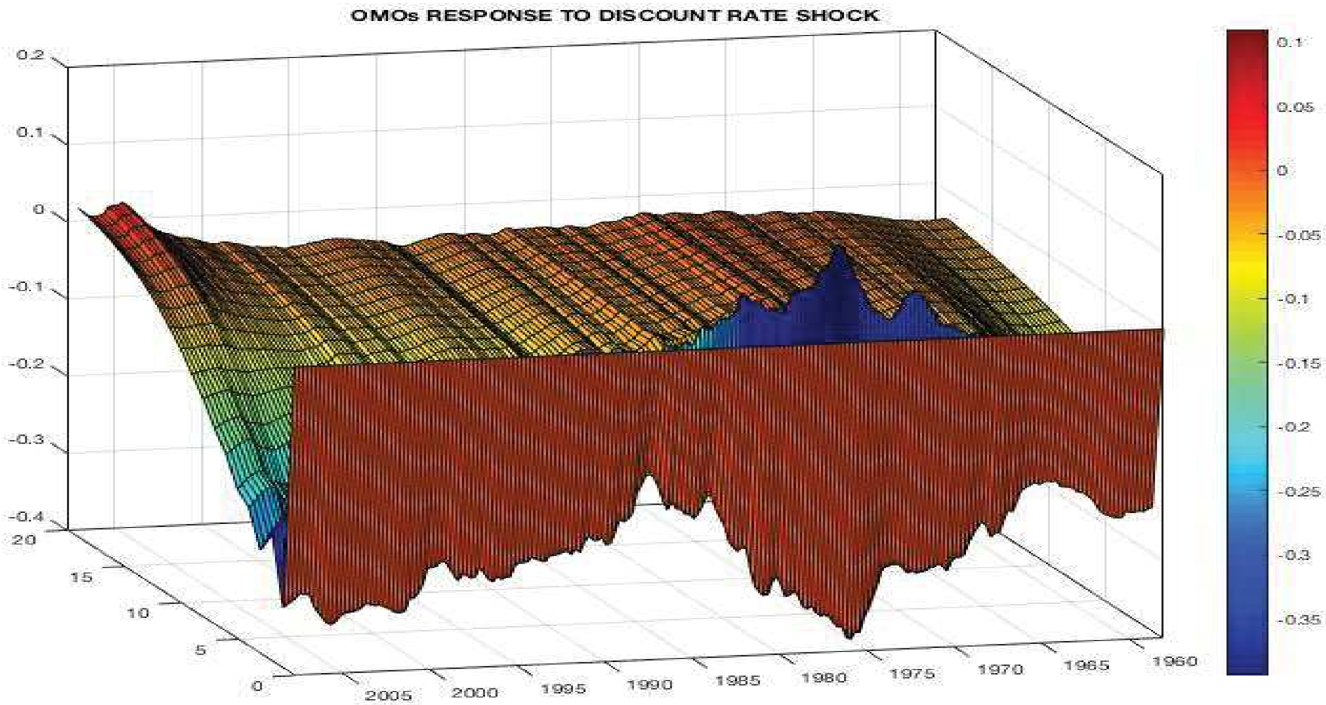

Figure A.7.14– OMO impulse response to a discount rate shock. Note: Posterior means.

Appendix B - Counterfactual methodology

The impulse response analysis for the second period shows that Fed's instruments, OMO and the discount rate, only have a regime switch around 1965. Moreover, the new variable F-D seems to present quite a dynamic behavior in relation to inflation, the IPI and M1. While during the 1970s the difference between the discount rate and the federal funds rate was not managed adequately as it increased inflation, around 1990, the sign of the response in Figure 1.7.1 is the opposite, negative. Hence, it seems that policies did not change and the banking sector modified its behavior. To support this hypothesis and discover whether different policies were applied, in this section I carry out a counterfactual analysis.

For that task, I have used the posterior mean of the average value of the parameters between 1995 and 1999, representing Greenspan's policies, and between 1971 and 1977 for Burns' policies. These values are used to simulate new ones for the rest of the parameters and for the other periods of the sample. Thus, the new values obtained for the parameters can be interpreted as those that would have been observed, had those policies been applied to the rest of the sample. In this case, unlike other works that draw the average of the posterior distribution from the monetary policy rule equation, meaning, the federal funds rate equation, here I draw the average values from the OMO, discount rate and F-D equations. In this type of analysis, the Lucas' critique arises as expectations and the private agents' behavior could have changed, had policies been modified at some point. However, given the Bayesian framework, in which policy is random and the model presents stochastic time variation of policy, the issue is hugely mitigated. Apart from that, I have included a new variable in the model, which is intended to capture inflation expectations. A new posterior distribution will be created for this variable every time that I introduce the new averaged values from Greenspan or Burns' policies. Thus, new expectations about inflations will be created for each counterfactual exercise. This new variable is inspired by Goodfriend (1993, pp.5-6), where he explained that the long-term yields should be a sum of the short-term rates with a variation, perhaps between two or three percentage points, plus the expected inflation. Thus, when the long-term rates increase more than the short-term rates, it is because the inflation expectations are higher. Consequently, the variable for the inflation expectations is built as follow: first, I take the difference of the short-term rate (3-months Treasury Bill rate) between period t and period t-1, and the same is done for the long-term rate (10-years government bonds yields). Once I have the difference of both rates, I use the difference of those two values. Then, the inflation expectations are captured when the long-term rate has increased or decreased more than the short-term rate. This variable is located in the fourth position of the VAR, before M1 and after F-D, as during Volcker and Greenspan's mandate, they targeted long-term rates to control inflation. The results obtained are displayed in Figures B.10.1 (Greenspan's counterfactual) and B.10.2 (Burns' counterfactual). The CPI response to an F-D shock under Burns and Greenspan's policies are very similar. Under Burns' policies, the positive values are more positive than under Greenspan's policies, and for those periods when the response is negative, Burns' policies would have decreased inflation less. Apart from that, there is not significant difference between both figures and I cannot claim a change in Fed's policies after the 1980s or 1990s, despite Figure 1.7.1 shows a regime change for those years. For the rest of the impulse response analysis, figures are all almost identical for both counterfactuals. The same counterfactual was carried out with the 1960-1965 parameters, when Martin was the chairman of the Fed, but the results are alike. This casts doubts on the possible regime change observed for that period in Figures 1.7.2 and 1.7.3.

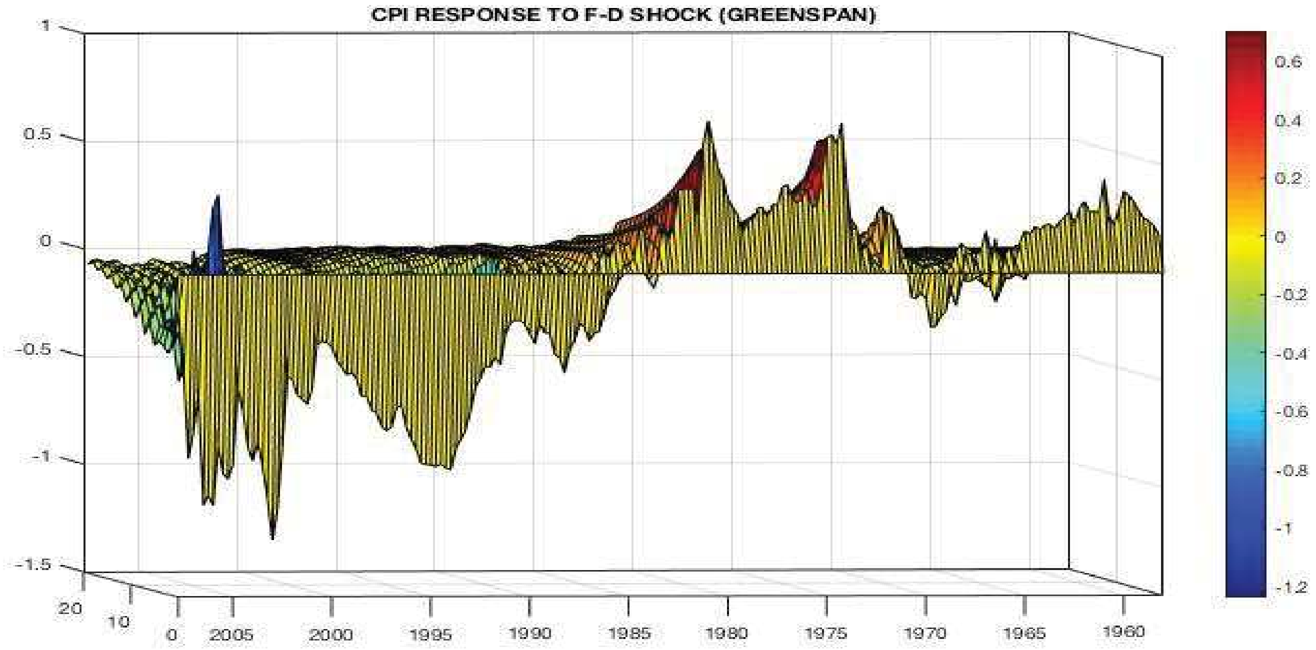

Figure B.10.1– CPI Inflation impulse response to a F-D shock (Greenspan). Note: Posterior means.

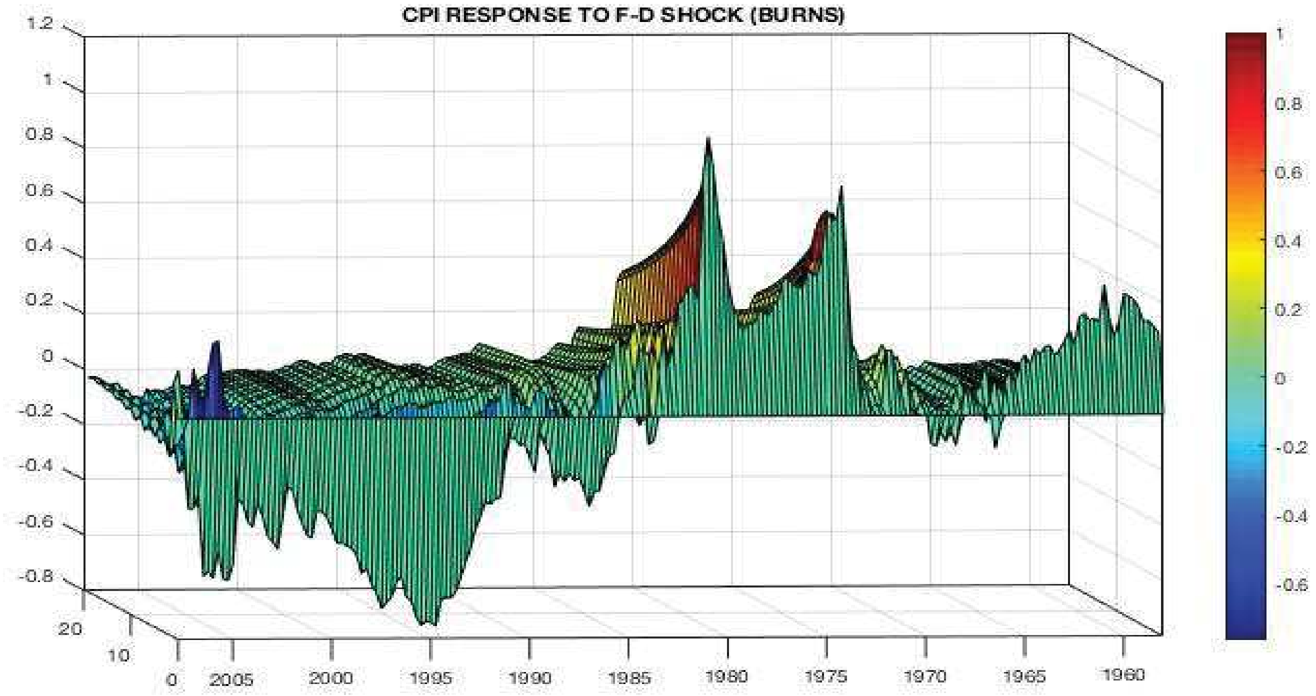

Figure B.10.2– CPI Inflation impulse response to a F-D shock (Burns). Note: Posterior means.

Appendix C – Convergence tests I

In this section, convergence of the Markov chain Monte Carlo algorithm is assessed for the baseline models. First, I present the inefficiency factor (IF), which is the inverse of the relative numerical efficiency (RNE) measure developed in Geweke (1992), for the posterior estimates of the parameters. RNE is a function of the serial correlation characteristics of the chain. The estimate (IF) is performed using a 4% tapered window for the estimation of the spectral density at frequency zero. Values below or around 20 can be considered as satisfactory. For space reasons results are not presented for the hyperparameters. Also, because they behave better than the parameters and all the values are below 20.

Second, to reassure that after the initial discarded sample and thinning of the chain the sample generated adequately represent the posterior distribution of interest, I calculate the I-statistic from Raftery and Lewis (1992b) that measures "the increase in the number of iterations due to dependence in the sequence" (Raftery and Lewis, 1992a). It is obtained from the formula $(M+N)/N_{min}$, where M is the initial number of iterations that should be discarded, N the number of iterations stored and required to achieve certain precision, and N_{min} the minimum number of iterations to reach convergence (Raftery and Lewis, 1992a). In this case, I apply it to the sample already "cleaned", so that I can evaluate if more iterations are needed, burned or it needs more "thinning". Those numbers are calculated for the quantile 0.025 of the posterior distribution of the parameters, estimated to within ± 0.005 with probability 0.95. That is, 95% intervals with posterior probability between 0.94 and 0.96. Values greater than 5 indicate dependence problems. Again, and for the same reasons as for the IF, the figures are presented only for the parameters. Figure C.11.1 and Figure C.11.2 display the results for the parameters B, α, and σ for the interwar period, and Figure C.12.1 and Figure C.12.2 for the second period.

In general, all the parameters reach convergence for the number of iterations indicated in section 4.1, as their values are below 20 for the IF estimate and below 5 for the I-statistic. Although the convergence of the Bs for the interwar period with K-P algorithm is not as good as for the other parameters, the IF estimate has a mean of 28.062, which can be considered as satisfactory. Regarding the I-statistic, only a few values are higher than 5.

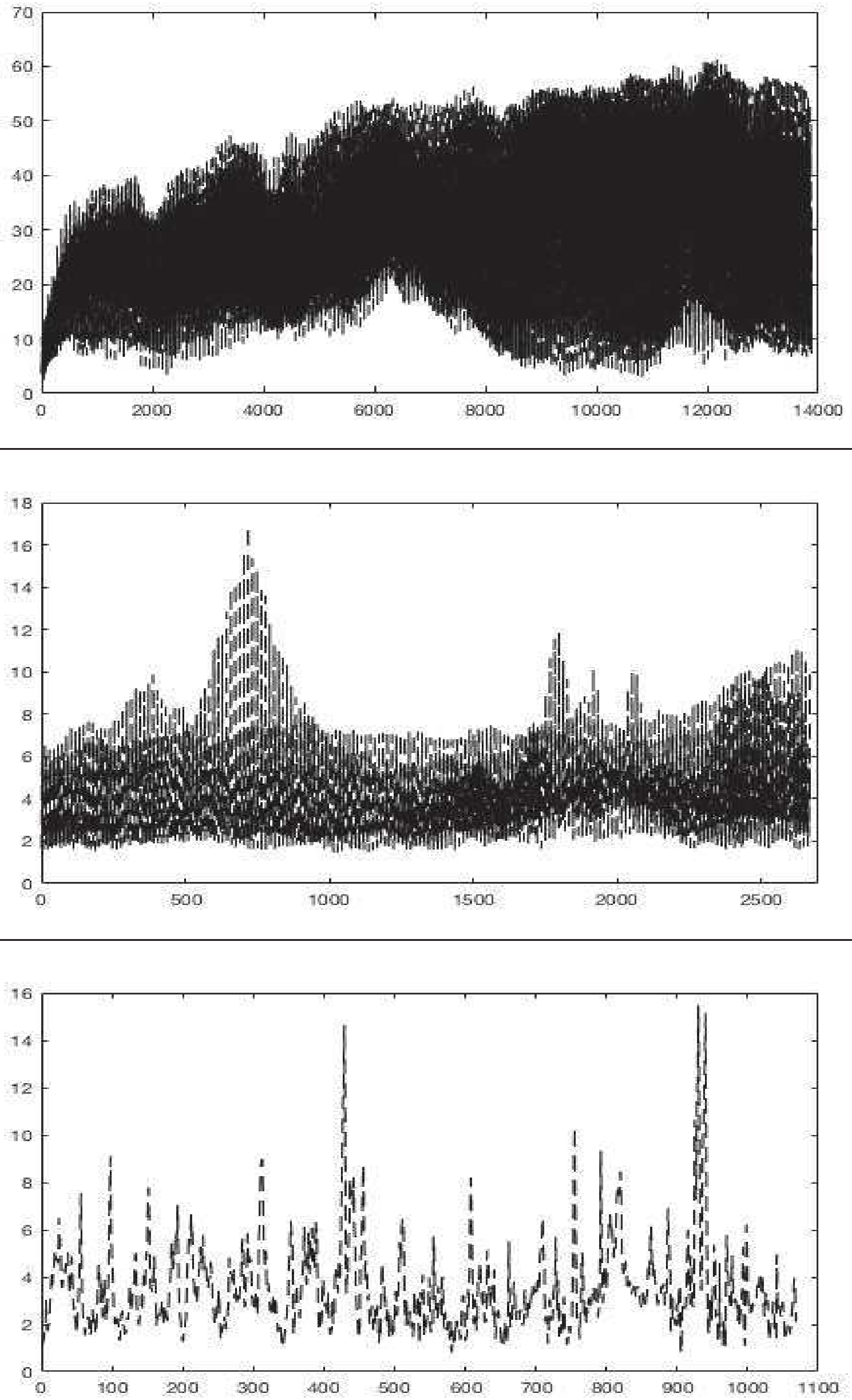

Figure C.11.1 – IF estimate for parameters B (Panel 1), α (Panel 2), and σ (Panel 3)

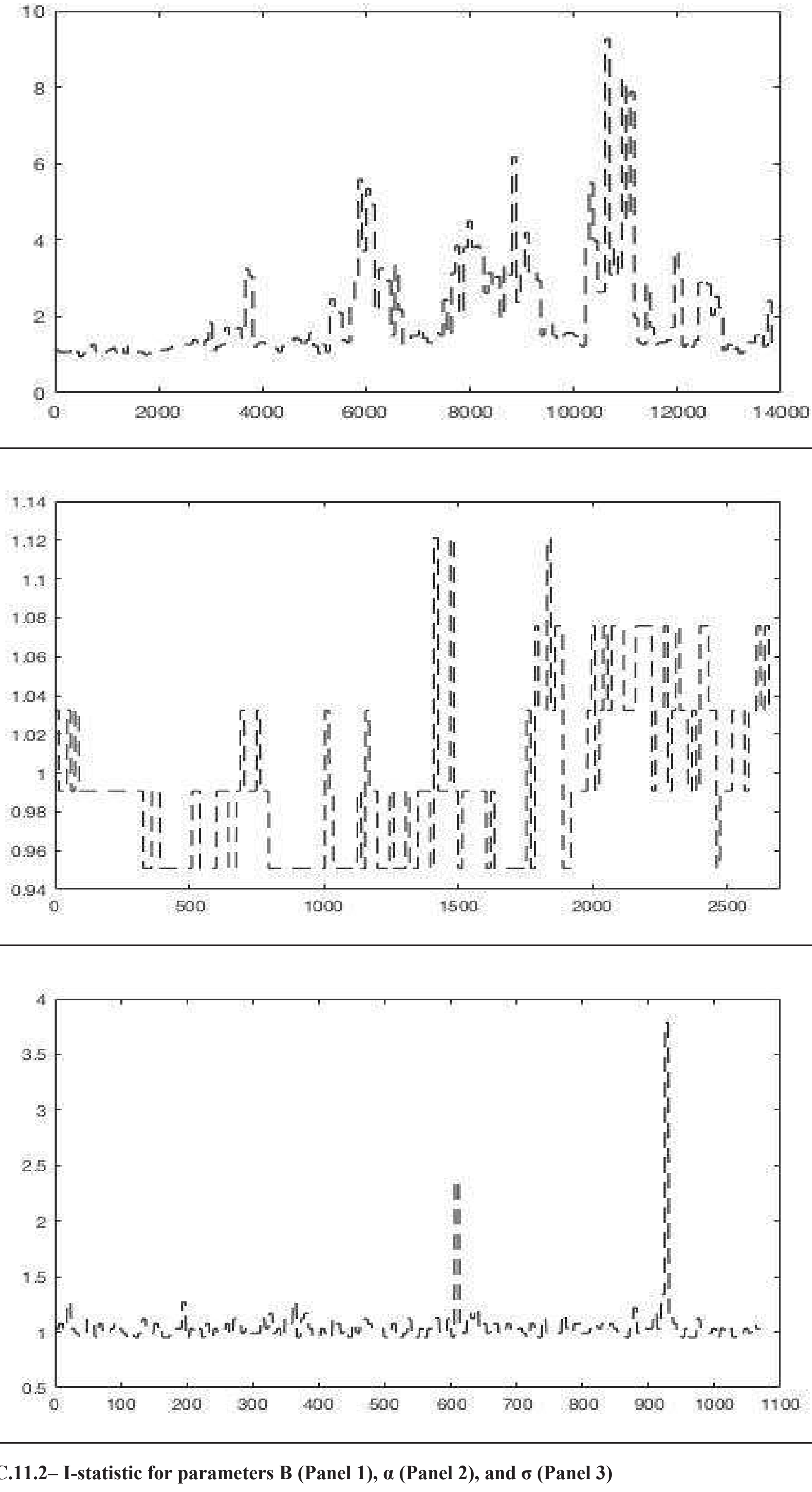

Figure C.11.2– I-statistic for parameters B (Panel 1), α (Panel 2), and σ (Panel 3)

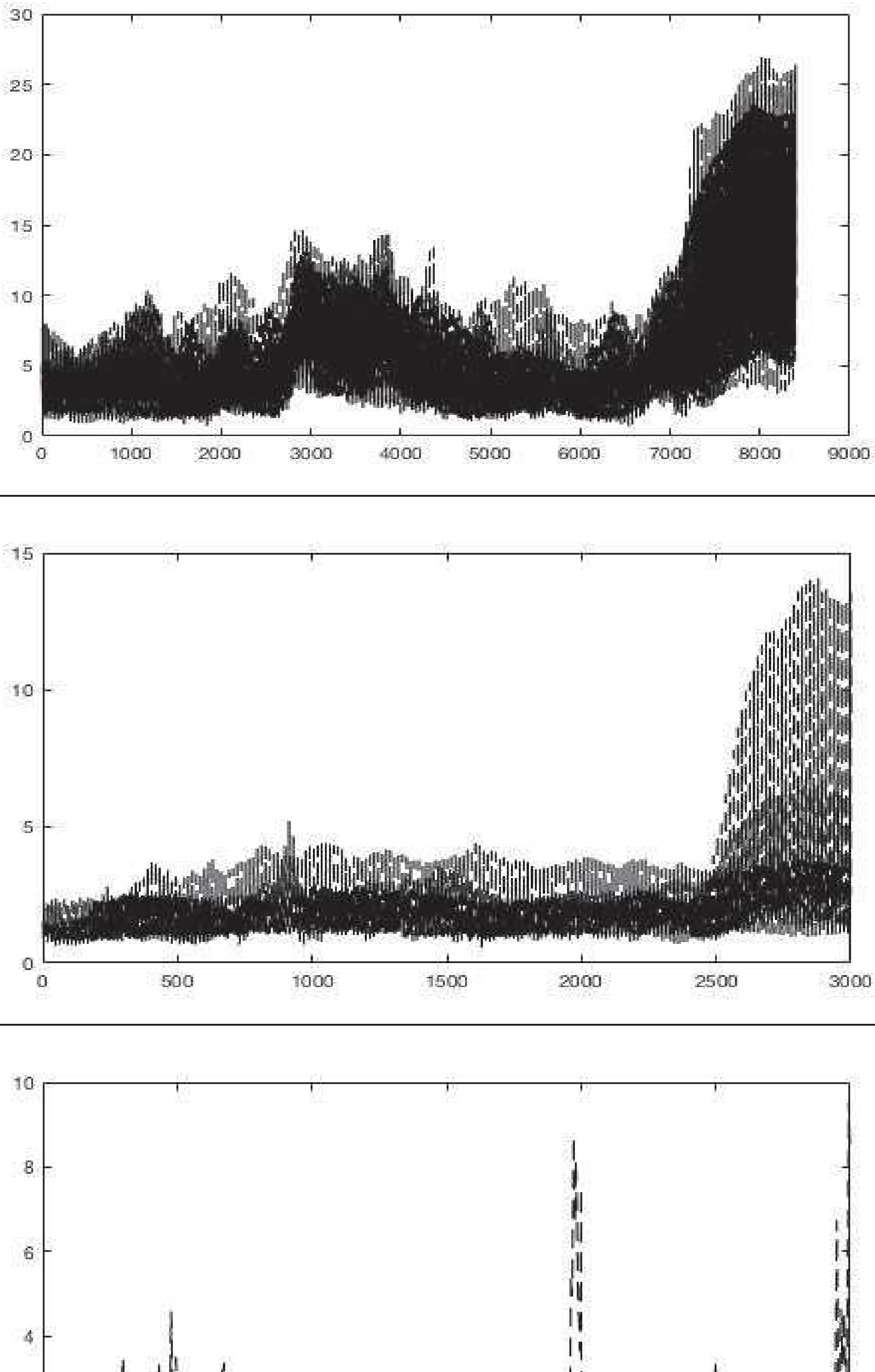

Figure C.12.1– IF estimate for parameters B (Panel 1), α (Panel 2), and σ (Panel 3)

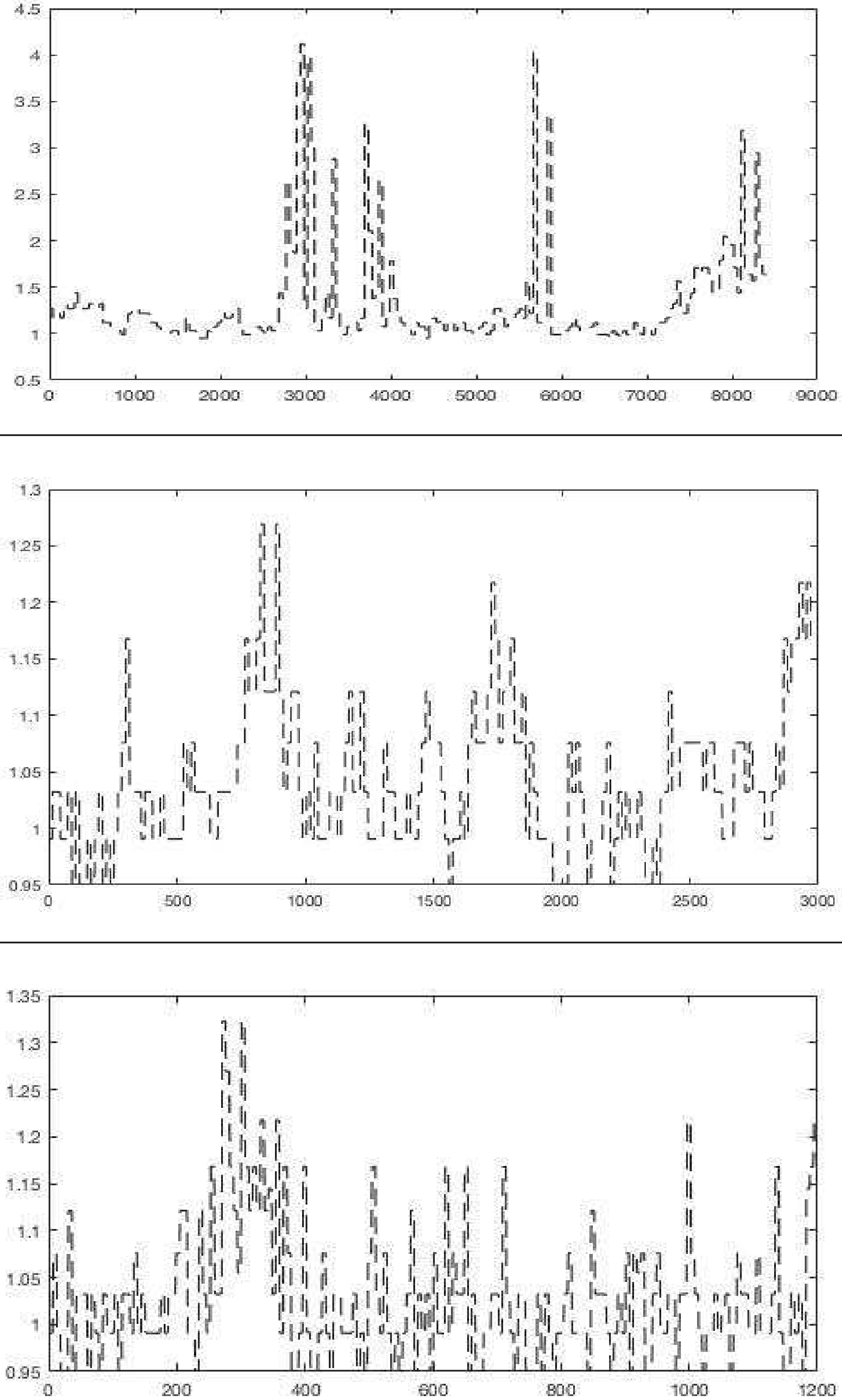

Figure C.12.2 – I-statistic for parameters B (Panel 1), α (Panel 2), and σ (Panel 3)

Appendix D – Convergence tests II

As in the previous appendix, convergence of the Markov chain Monte Carlo algorithm is assessed for the baseline model. Again, for space reasons results are not presented for the hyperparameters and because they behave better than the parameters. Figures D.20 and D.21 display the results for the parameters B, □, and □

In general, all the parameters reach convergence for the number of iterations indicated in section 3.3.2, as their values are below 20 for the IF estimate and below 5 for the I-statistic.

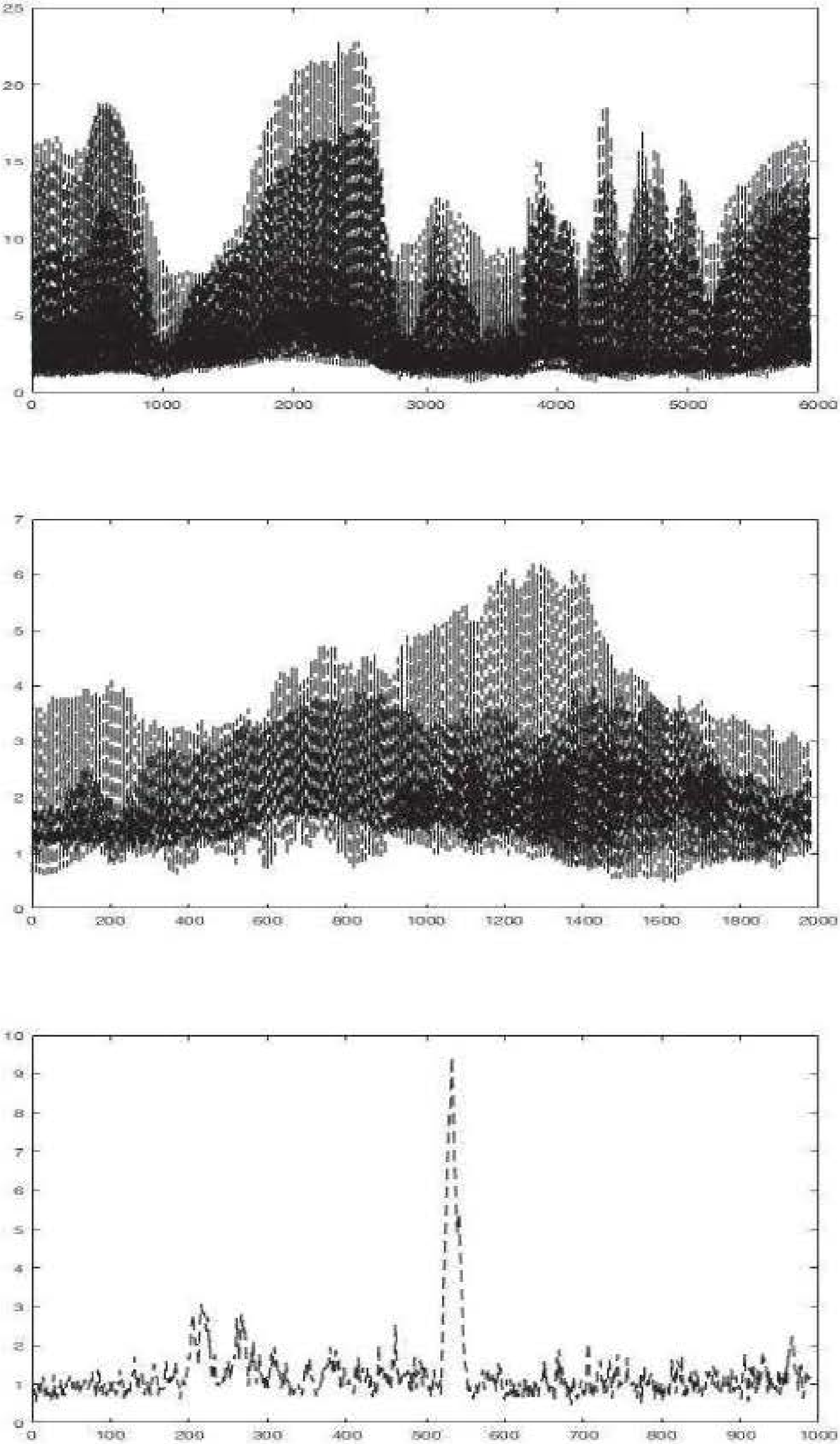

Figure D.20 IF estimate for parameters [illegible] (Panel 1), α (Panel 2), and σ (Panel 3).

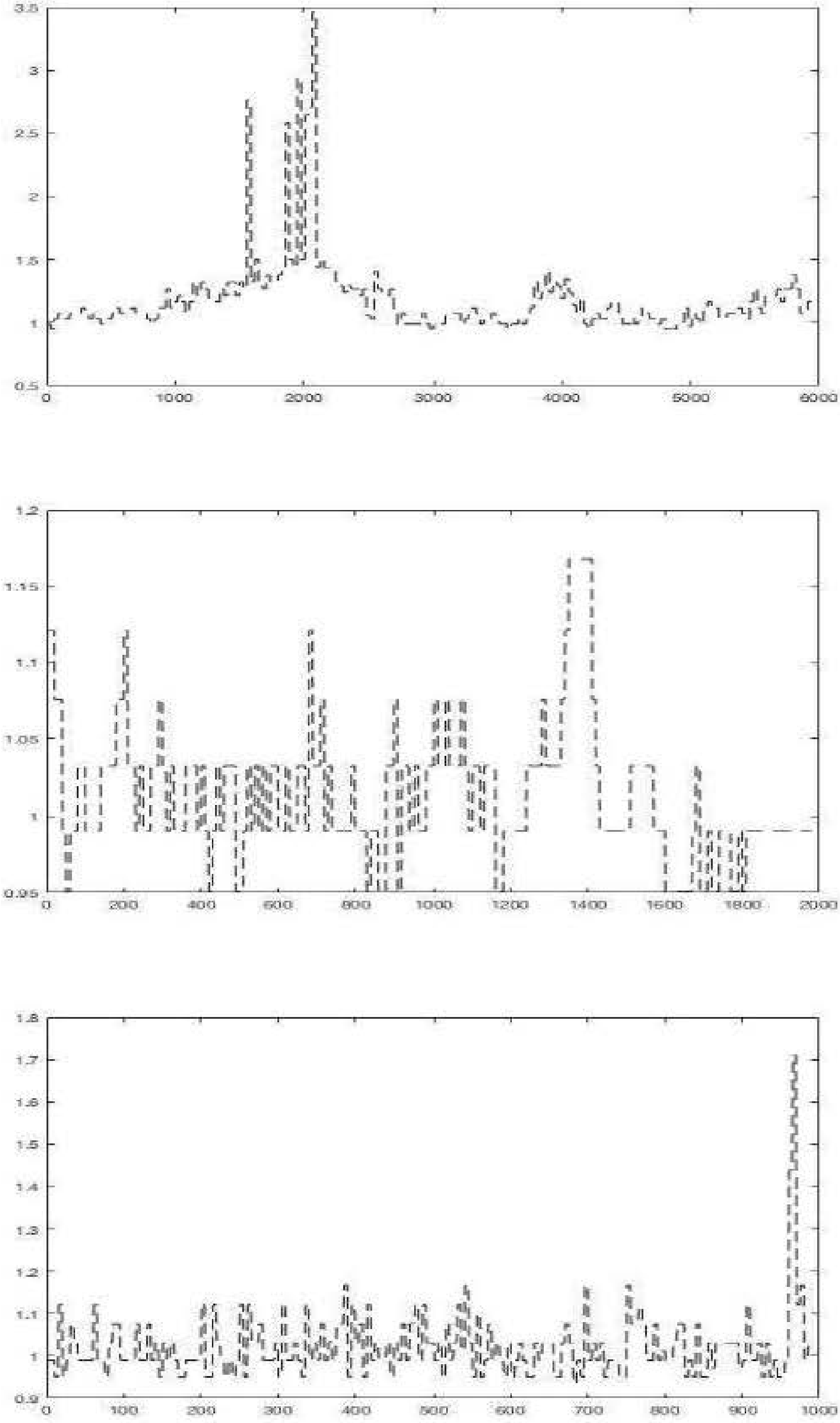

Figure D.21 – R-statistic for parameters β (Panel 1), α (Panel 2), and σ (Panel 3).

Appendix E – Data

The data sources for the variables are "Banking and Monetary Statistics 1914-1941", "Banking and Monetary Statistics 1941-1970" available in FRASER, and last, the FRED database.

- Nonborrowed reserves are those reserves that are not borrowed from the central bank at the discount window. It is measured in billion of dollars.

- Required reserves are the reserves that a bank must hold in their vaults or at the Federal Reserve, depending on amount of deposits. It is measured in billion of dollars.

-Excess reserves are those reserves held in excess of what is required by regulators. It is measured in billion of dollars.

- The prime loans rate is the base rate used by banks to price short-term business loans, posted by a majority of 25 insured U.S.-chartered commercial banks. The prime loans rate appears officially in 1934 defined as "the rate that banks charge their most creditworthy business customers on short-term loans. It is the base from which rates charged on loans to other business customers are scaled upward. Generally speaking, the prime rate has not been considered a sensitive rate that fluctuates daily in response to short-terms changes in demand and supply as measured by a national market." Further, " Primer rates are "formally" posted only by largest banks. A nationally publicized and uniform prime rate did not emerge until the depression of the 1930's. The rate in that period -1.5 per cent- represented a floor below which banks were said to regard lending as unprofitable" (Banking and Monetary Statistics 1941-1970, p. 642). Thus, from 1934 to 1949, the data for this variable is collected from the source just mentioned. Since 1949, the data is collected directly from FRED. For the periods 1919-1929 and 1930-1933 the rates on customers' loans and commercial loans in New York are collected respectively. These rates, despite not being the prime loans rate, are the most similar once there are data available for it.

- The short-term rate of reference has been changing through the period under analysis, depending on the sources from which banks obtained funds. For the period 1919-1933 an average of the most important short-term open-market instruments is used. These instruments are: 4- to 6-month commercial paper and prime 90-day bankers' acceptances (loans based on commercial transactions) and 90-day Stock Exchange time loans and Stock Exchange call loans, new and renewal (loans based on security collateral). Given that after 1931 there is a decline in commercial and bankers' acceptances holding and an increase in short-term Treasury bills, for the period 1934-1954 rates on 3-months T-bills are collected instead. Last, after 1954 the federal funds rate is considered the short-term rate of reference.

- The "Loans" variable is not homogeneously available for the entire period and different data needed to be plugged and extrapolated. From 1919 to 1947 the quarterly amount of loans from "member banks" have been collected. In some cases, some quarters are missing. To fill those quarters, I have calculated the average of loans between t-1 and t+1. However, "member banks" does not represent all commercial banks, for which there are only semi-annual data. Thus, for those periods when data is available for both series, I have calculated the proportion of commercial banks' loans regarding member banks' loans. The resulting number is multiplied for the amount of member banks' loans for the quarters missing in the commercial banks' loans series, until the next datum in the semi-annual series for commercial banks is found, when the proportion is calculated again. This series comprises loans on securities, real state, banks and other loans. Since 1948, the data has been collected from FRED "Loans and Leases" series, which includes commercial, industrial, consumer and real state loans and other loans and leases. It is measured in billion of dollars.

RC, MAX, GDPD and DEBT have been already described in sections 4.2, 4.3 and 4.4.

Appendix F – Robustness tests

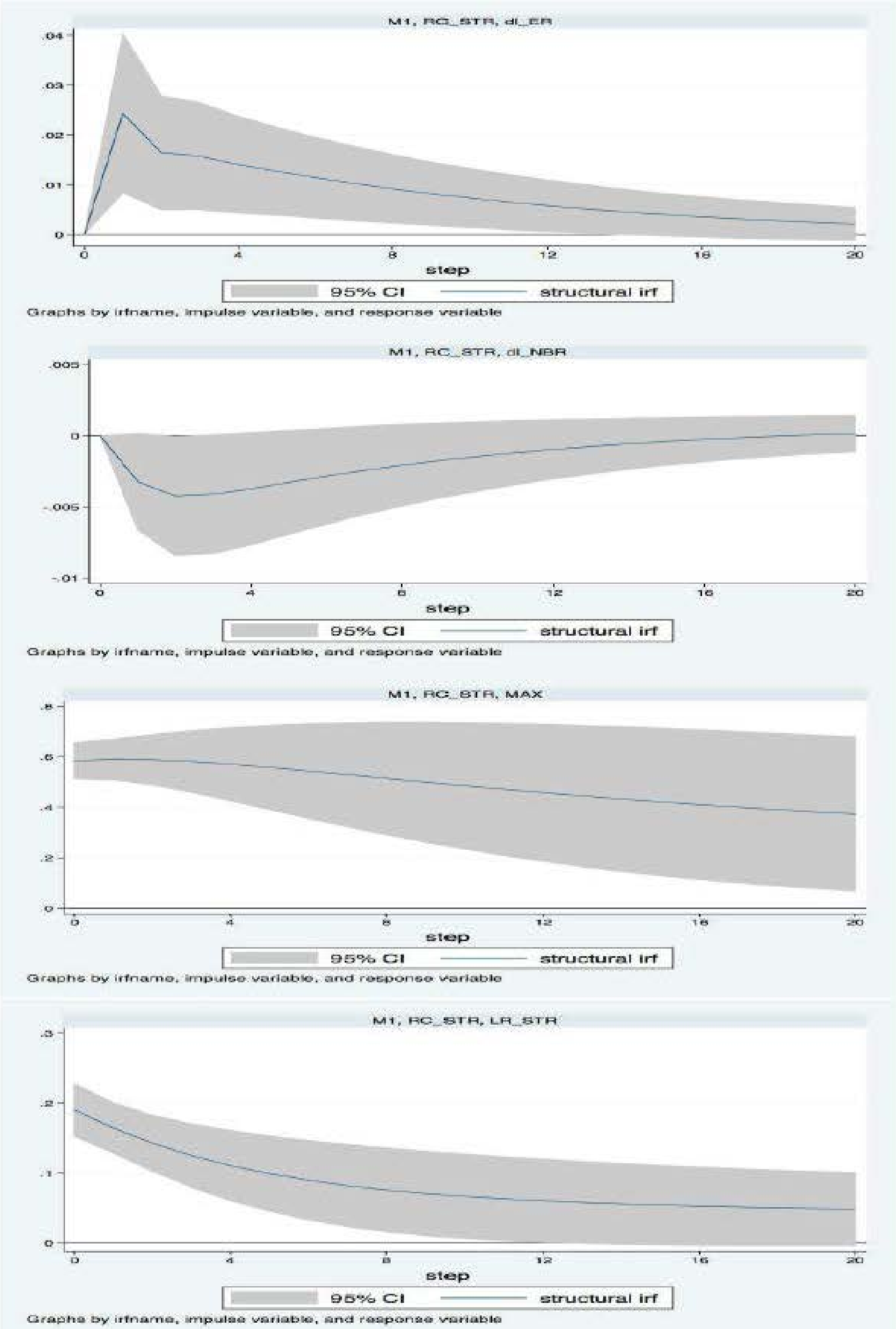

Figure F.6.1 Impulse responses to an RC-STR shock. ER, NBR, MAX, and LR-STR in panels 1, 2, 3 and 4 respectively.

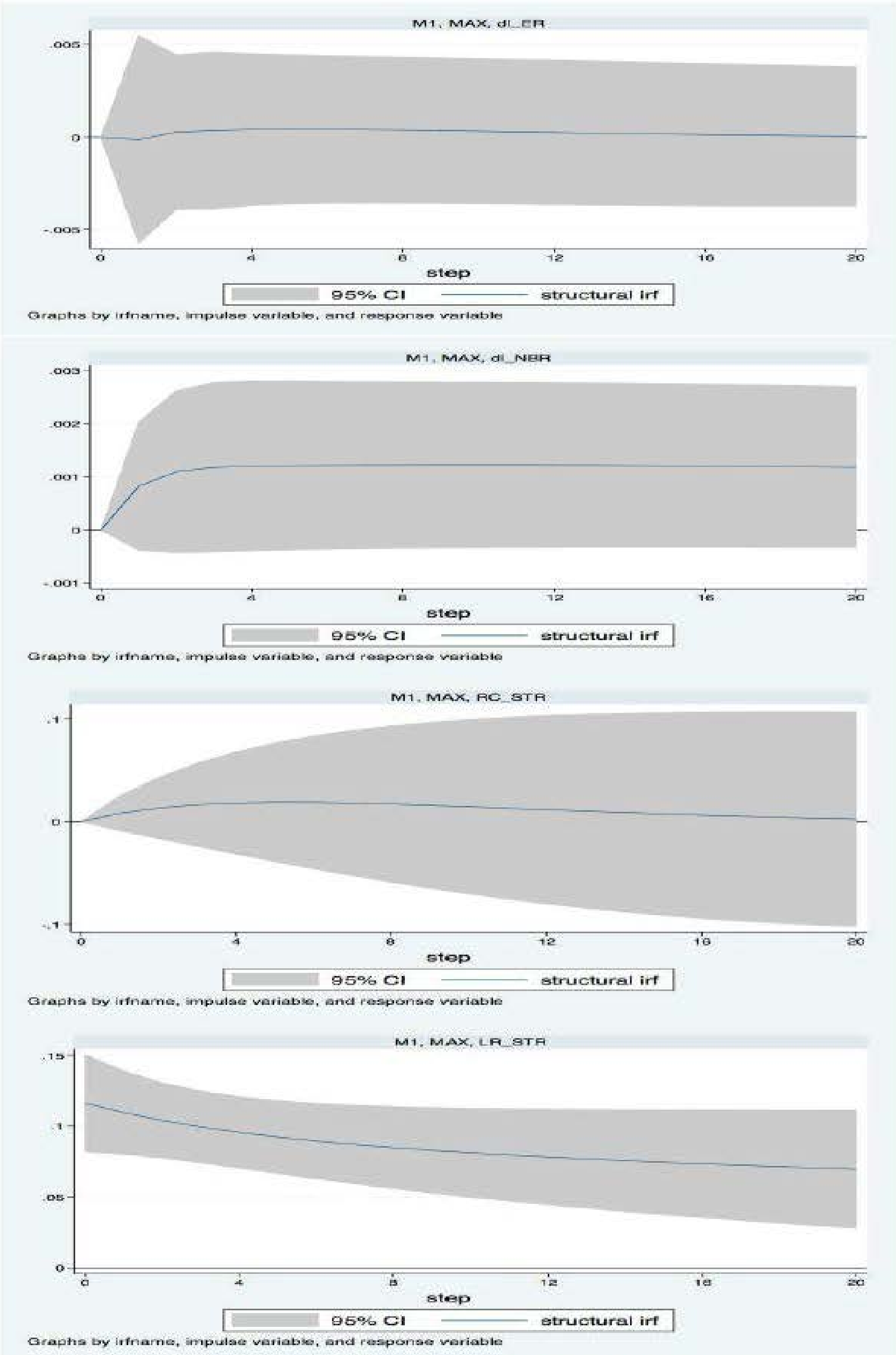

Figure F.6.2 Impulse responses to a MAX shock. ER, NBR, RC-STR, and LR-STR in panels 1, 2, 3 and 4 respectively.

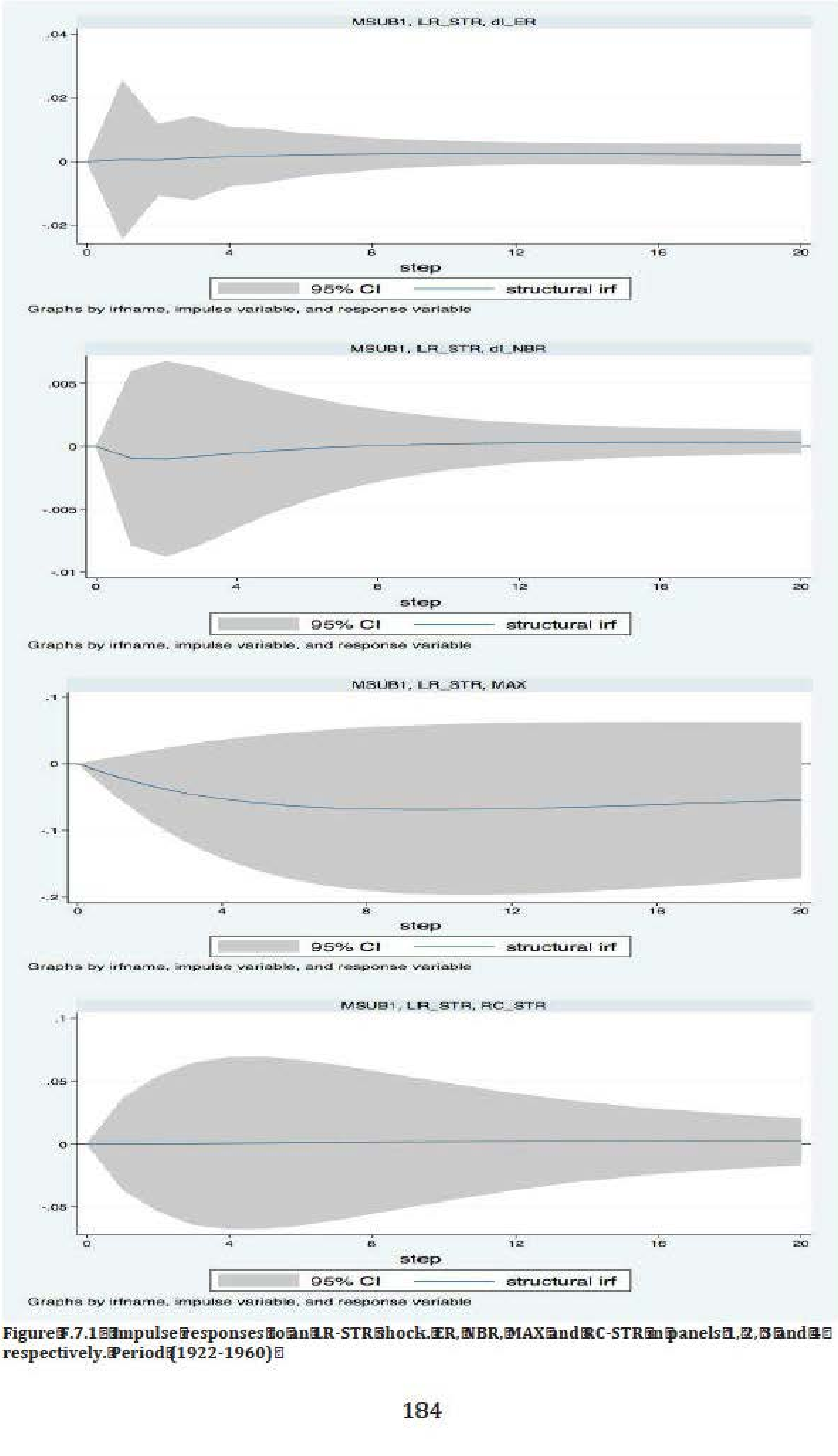

Figure F.7.1 Impulse responses to an LR-STR shock. ER, NBR, MAX and RC-STR in panels 1, 2, 3 and 4 respectively. Period (1922-1960)

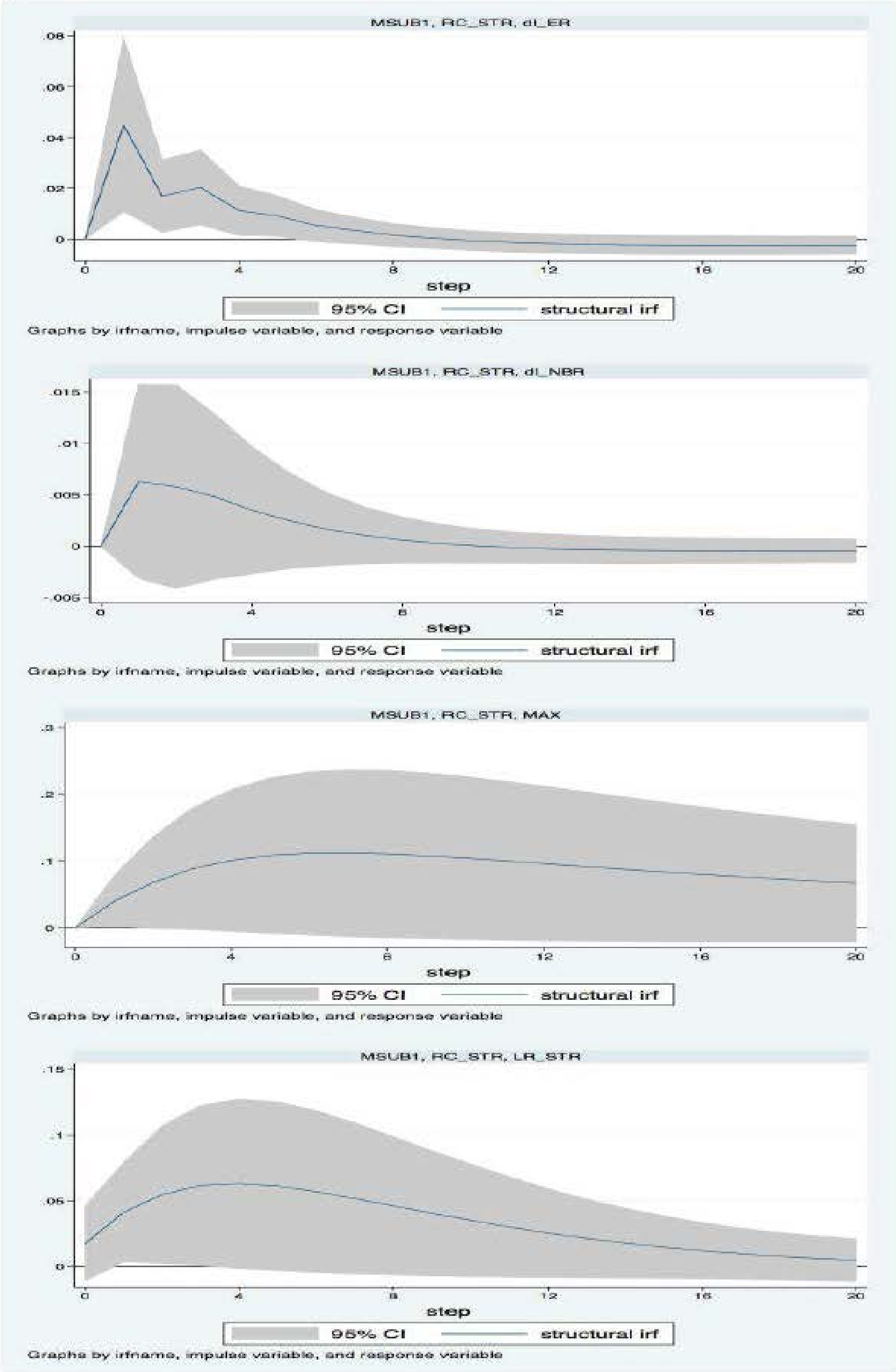

Figure F.7.2 Impulse responses to an RC-STR shock. ER, NBR, MAX and LR-STR in panels 1, 2, 3 and 4 respectively. Period (1922-1960)

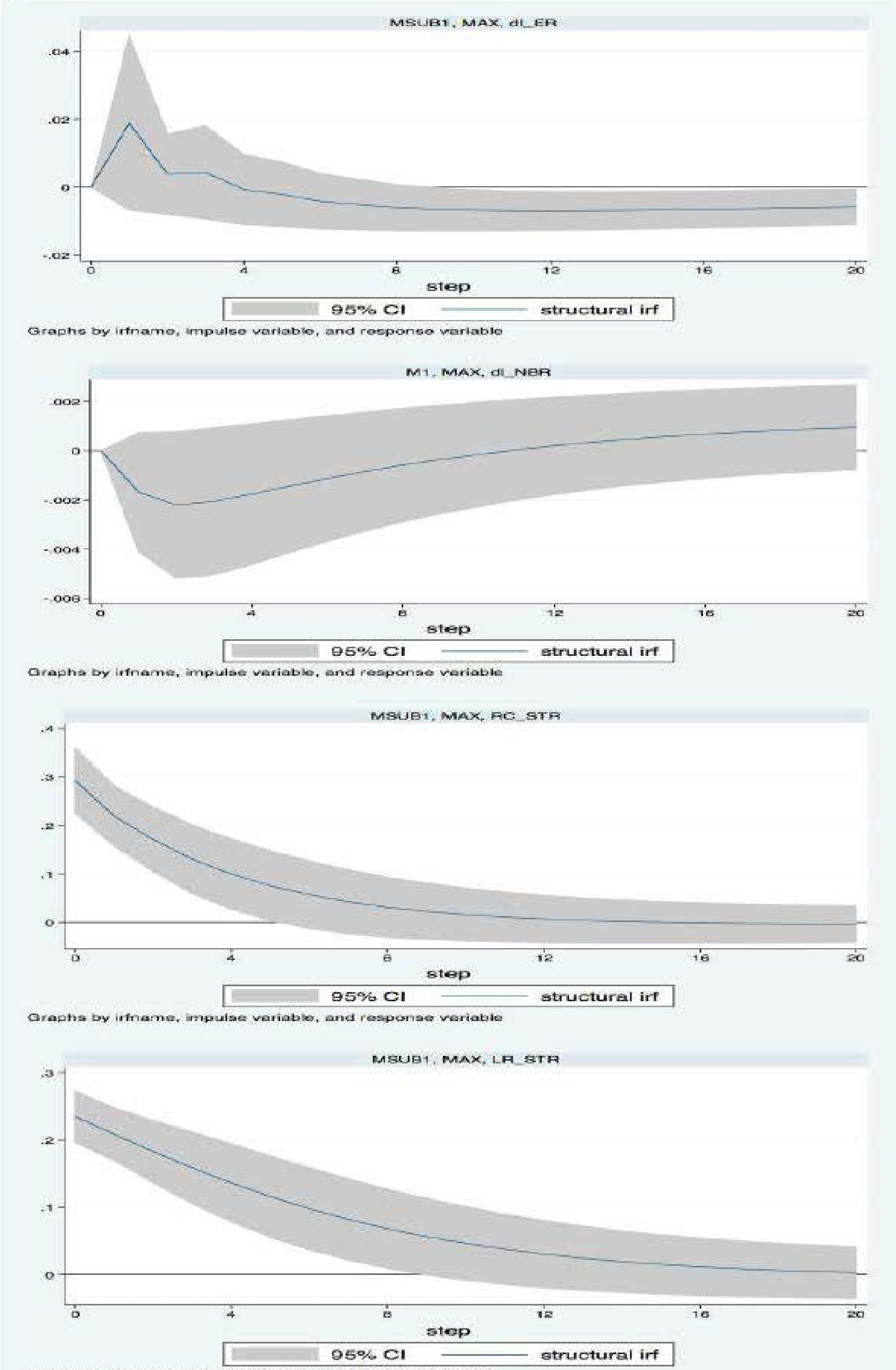

Figure 7.3 Impulse responses to a MAX shock. ER, NBR, RC-STR and LR-STR in panels 1, 2, 3 and 4 respectively. Period (1922-1960)

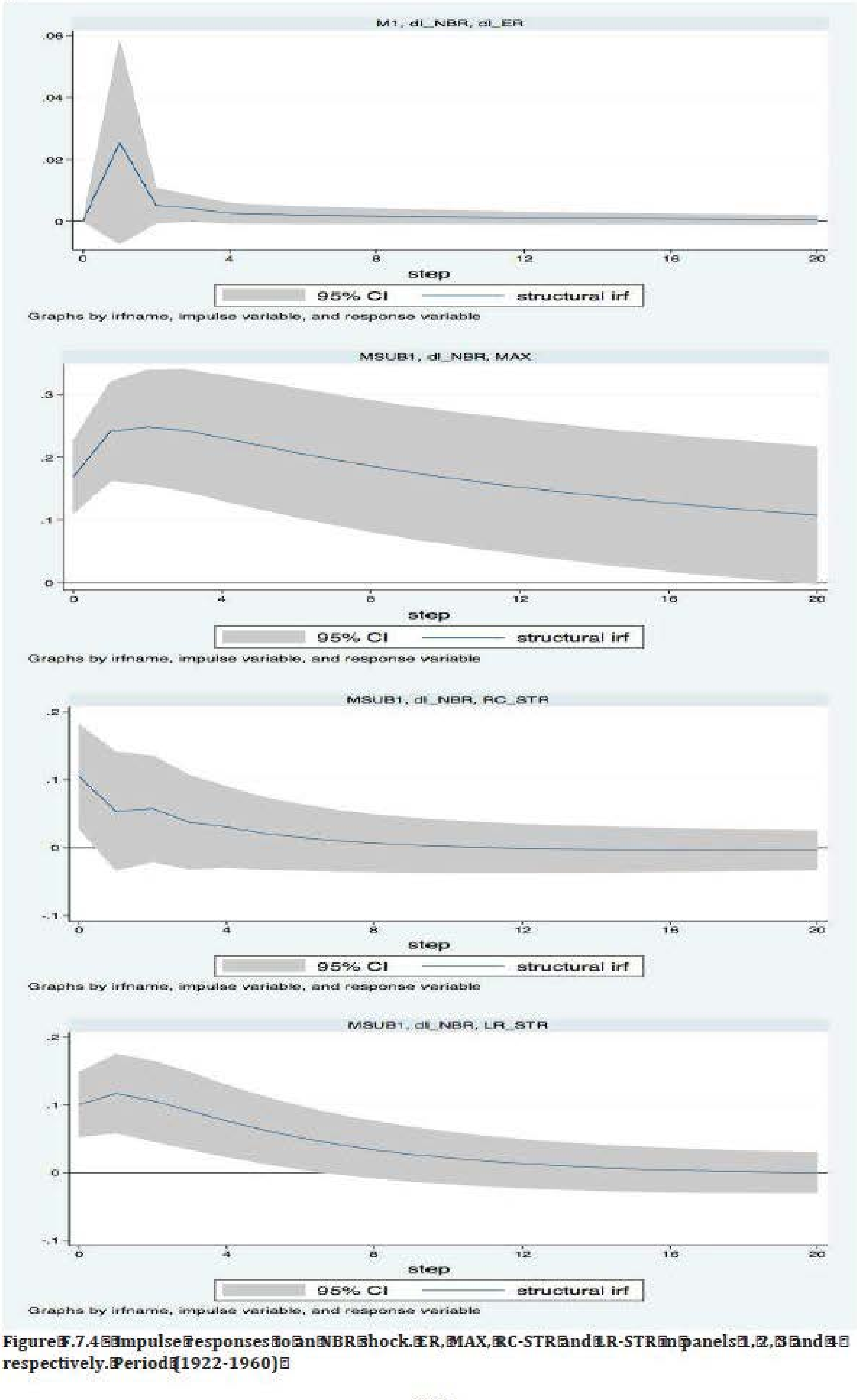

Figure F.7.4 Impulse responses to an NBR shock. ER, MAX, RC-STR and LR-STR in panels 1, 2, 3 and 4 respectively. Period (1922-1960)

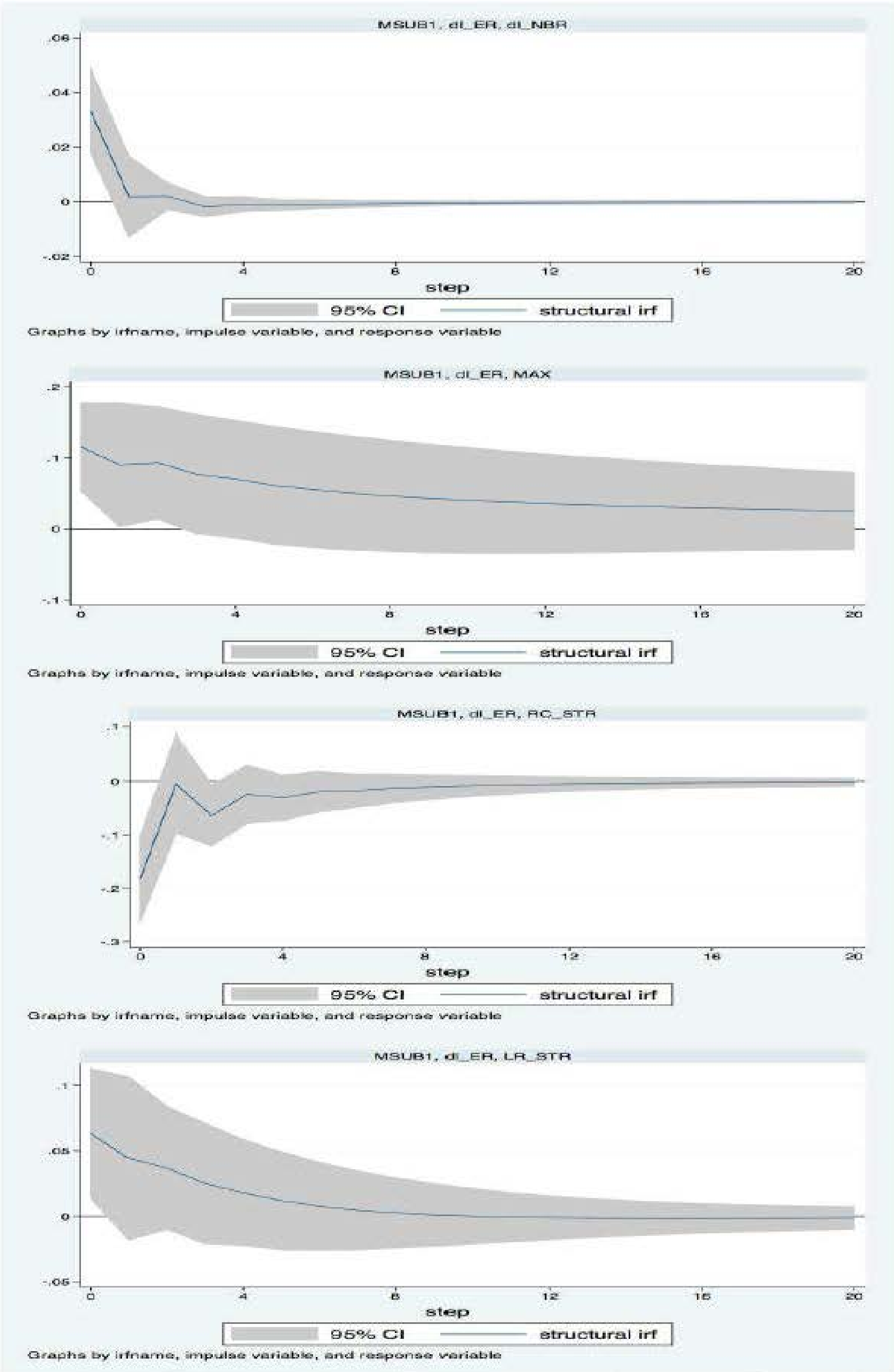

Figure [illegible].7.5 Impulse responses to an ER shock. NBR, MAX, RC-STR and LR-STR in panels 1, 2, 3 and 4 respectively. Period (1922-1960)

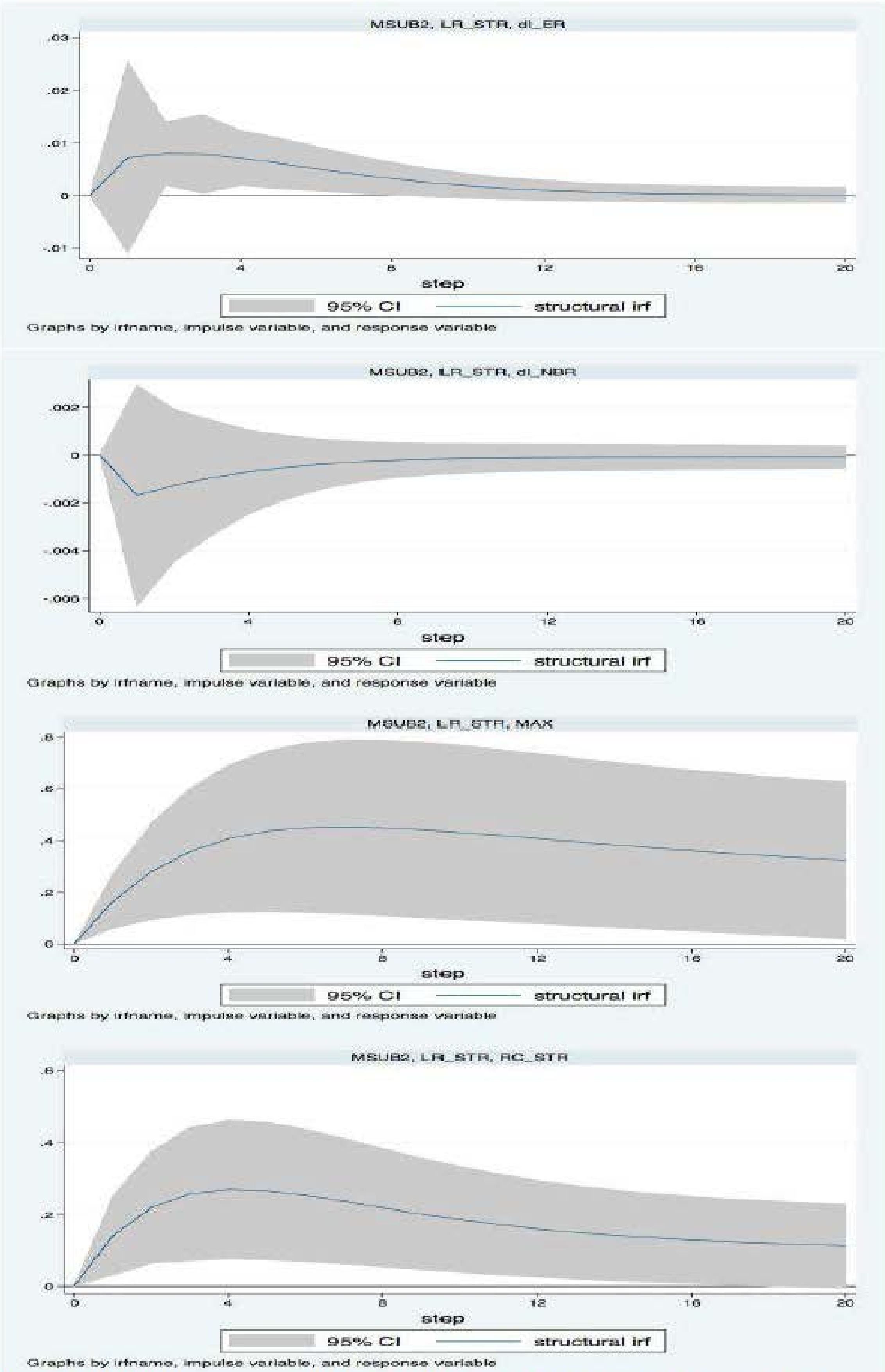

Figure F.8.1 Impulse responses to an LR-STR shock. ER, NBR, MAX and RC-STR in panels 1, 2, 3 and 4 respectively. Period (1960-2006)

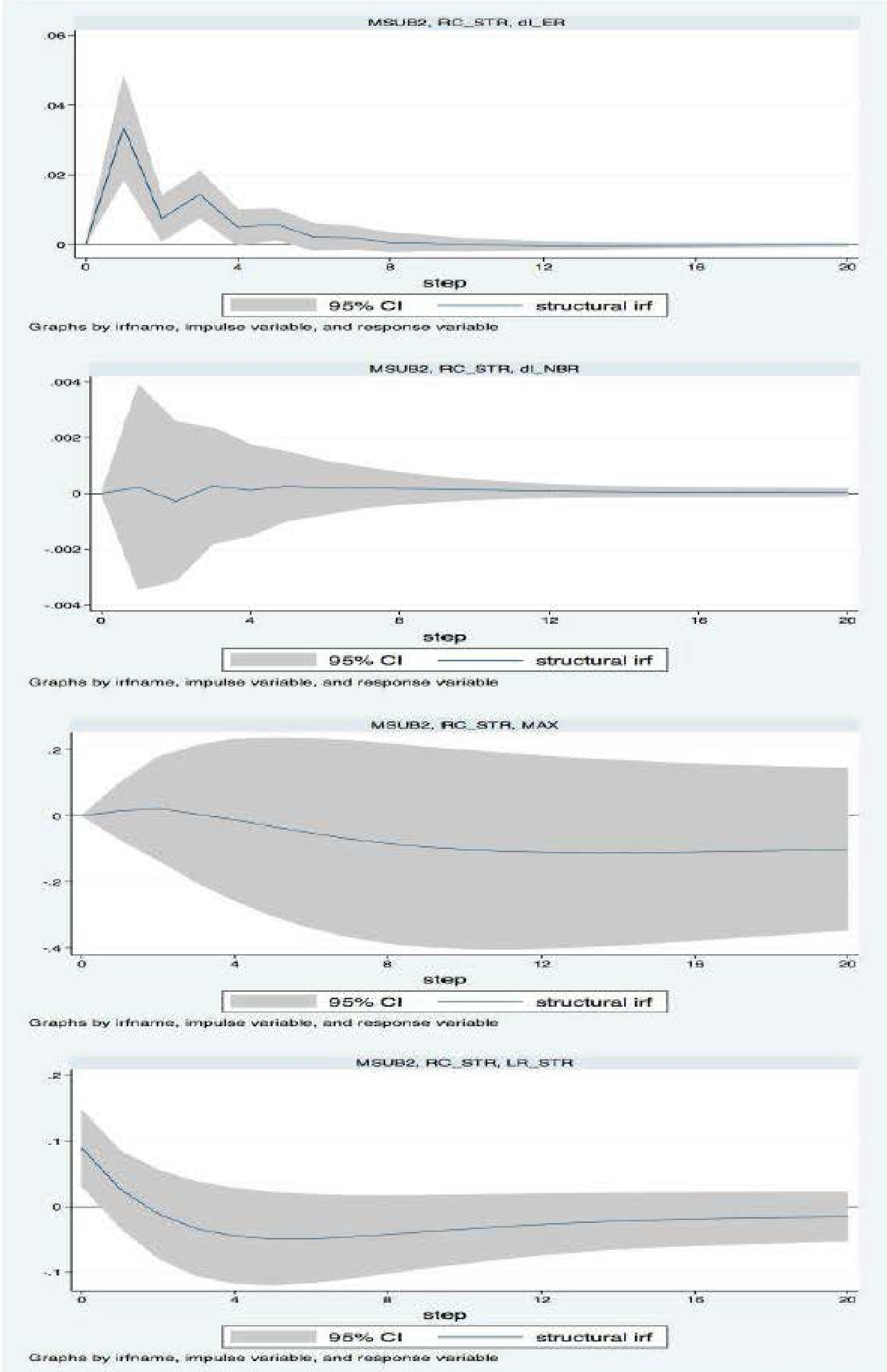

Figure F.8.2 Impulse responses to an RC-STR shock. ER, NBR, MAX and LR-STR in panels 1, 2, 3 and 4 respectively. Period (1960-2006)

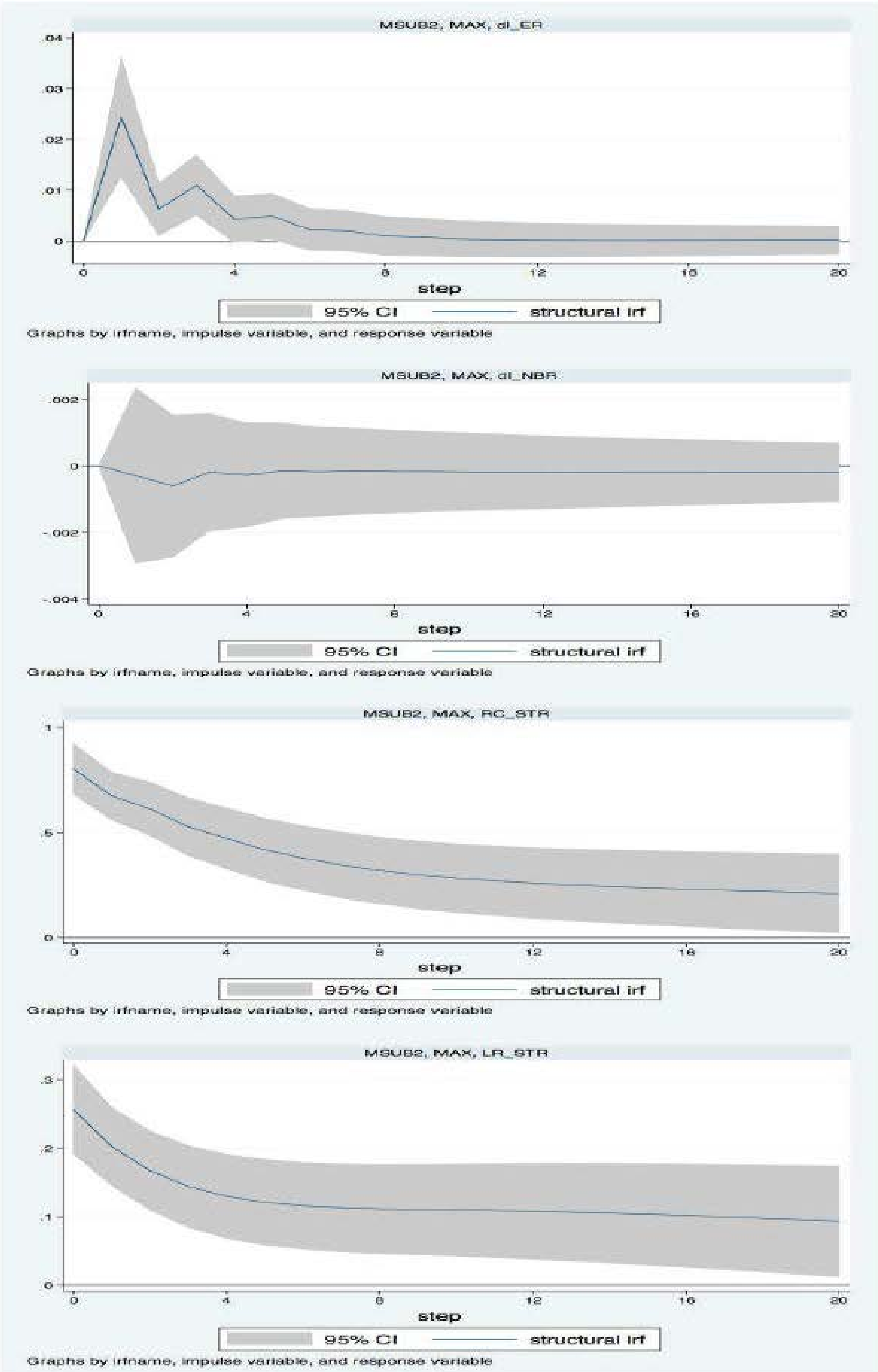

Figure F.8.3 Impulse responses to a MAX shock. ER, NBR, RC-STR and LR-STR in panels 1, 2, 3 and 4 respectively. Period (1960-2006)

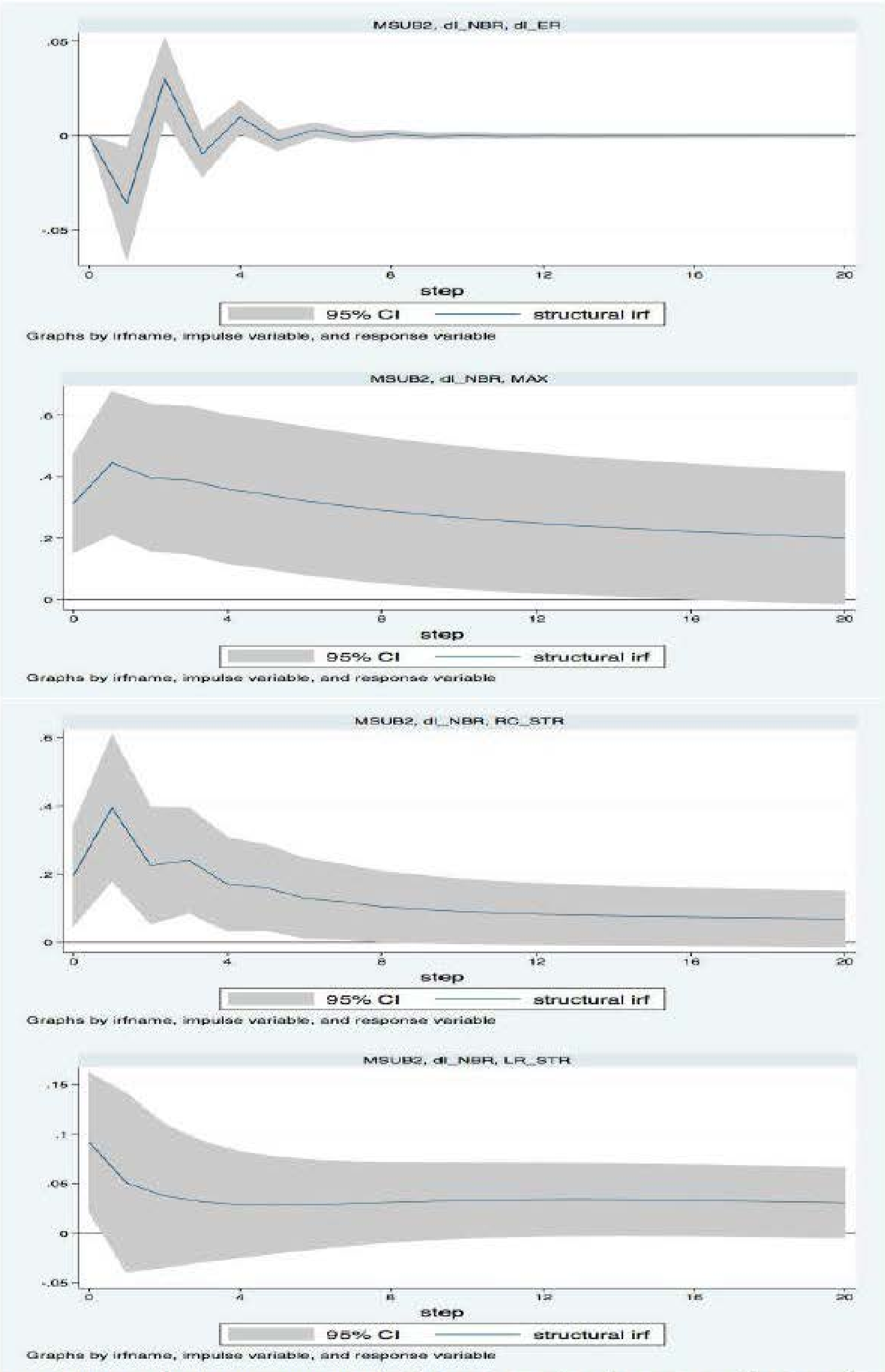

Figure F.8.4 Impulse responses to an NBR shock. ER, MAX, RC-STR and LR-STR in panels 1, 2, 3 and 4 respectively. Period (1960-2006)

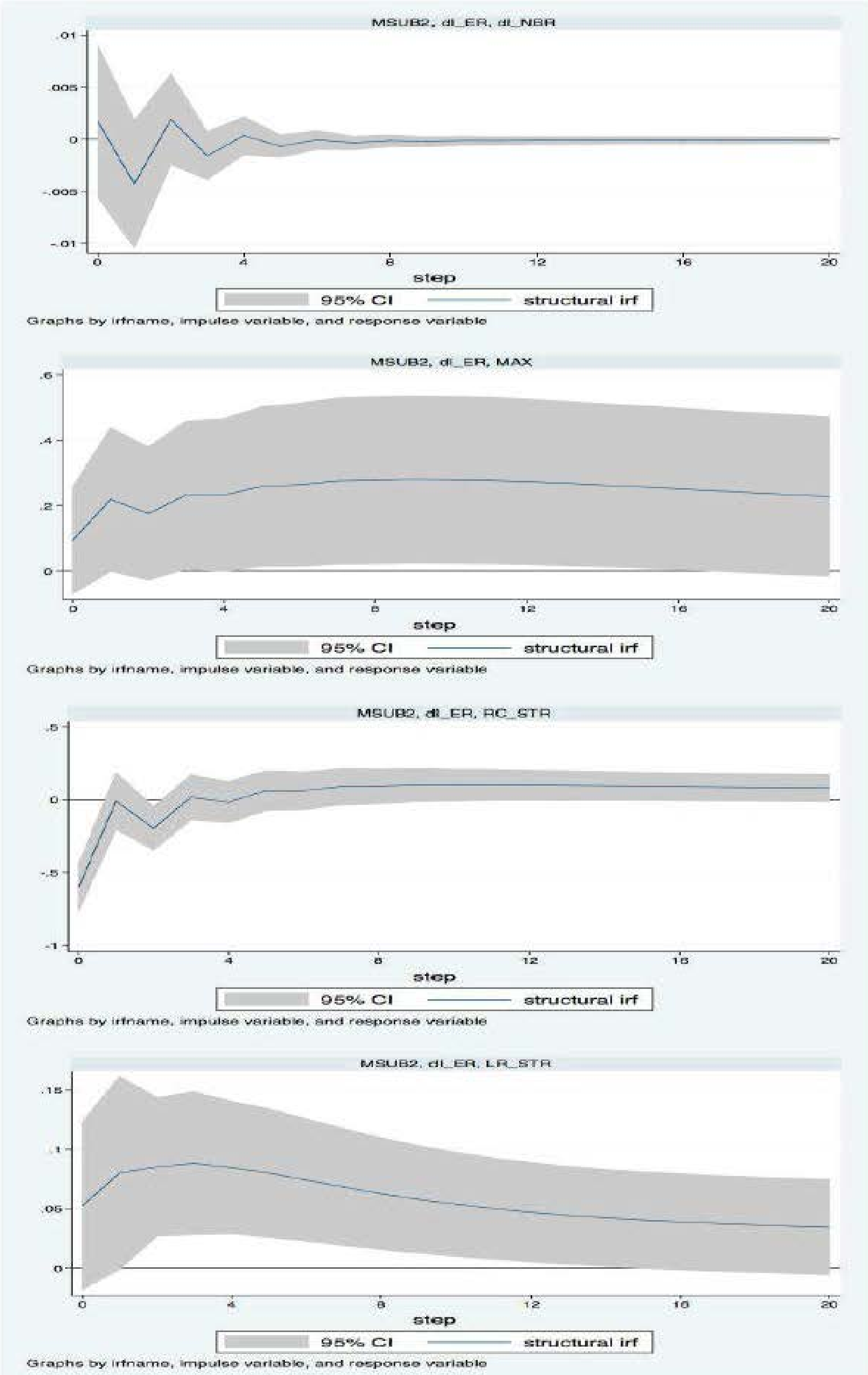

Figure F.8.5 Impulse responses to an LR-STR shock. NBR, MAX, RC-STR and LR-STR in panels 1, 2, 3 and 4 respectively. Period (1960-2006)

Appendix G - Primary Sources

- **National Bureau of Economic Research (NBER)**

  NBER Macrohistory Database:

  -Section XIV. Money and Banking

  - Section XIII. Interest rates

- **Federal Reserve Economic Data (FRED)**

- **Bureau of Labor Statistics**

- **Reserve Archival System for Economic Research (FRASER):**

  -Banking and Monetary Statistics 1914-1941

  -Banking and Monetary Statistics 1941-1970

  - Federal Reserve Bulletins from 1970 to 2003