CERGE
Center for Economics Research and Graduate Education
Charles University Prague

Essays on Interbank Interest Rates

Kamil Kovář

Dissertation

Prague, February 10, 2022

Kamil Kovář

Essays on Interbank Interest Rates

Dissertation

Prague, February 10, 2022

Dissertation Committee

Sergey Slobodyan (CERGE-EI; chair)

Filip Matějka (CERGE-EI)

Veronika Selezneva (CERGE-EI)

Referees

Lorenzo Burlon (European Central Bank)

Michal Franta (Czech National Bank)

This dissertation is dedicated to my GG. Without her love and patience, I would not have completed it. And to my parents, thanks to whom I can always look into the future and not worry about the present.

Contents

List of Figures

List of Tables

Abstract

This thesis studies the behavior of interbank interest rates in the aftermath of the global financial crisis. This crisis and its macroeconomic consequences led to a sharp break in how monetary policy is conducted, with unconventional tools such as quantitative easing (QE) programs gaining prominence. One consequence of such policies is a change in the behavior of interbank interest rates. This behavioral change is due to the emergence of excess reserves, which are a side effect of many unconventional policies including quantitative easing.

The first chapter explores the nexus between the QE program conducted by the European Central Bank, its policy of negative policy rates and the interbank interest rates. It starts with data analysis that demonstrates two salient features of the behavior of interbank interest rates in the presence of excess reserves. First, when excess reserves are present interbank interest rates are anchored by the deposit rate rather than by the main refinancing rate, as was the case before emergence of excess reserves. Second, the amount of excess reserves is negatively correlated with the level of interbank interest rates whenever excess reserves are present. The chapter proposes a structural time series model that links interbank interest rates to the two policy rates in a two-regime structure with a threshold variable - the amount of excess reserves - determining the prevalence of each regime. The resulting model provides a very good fit for the observed historical time series and confirms that the amount of excess reserves is a statistically significant factor influencing interbank interest rates. I then use the model to answer the policy question of the chapter: What was the effect of the QE program on interbank interest rates? Since the QE program led to a large increase in excess reserves, which in turn led to a decrease in interbank interest rates, the effect on these rates was substantial. Quantitatively speaking, the effect is on par with changes in policy rates of standard size, and the same order of magnitude as a typical estimate of the effect that the announcement the QE program by the ECB had on bond yields.

Given the combination of its size and the absence of academic investigation, I call this effect 'the neglected channel of quantitative easing'.

The second chapter analyzes the importance of the factors identified in the first chapter, as well as the value of the structural model, in the context of (pseudo) out-of-sample forecasting. It starts by comparing the forecasting performance of the structural model with forecasting performance of all plausible benchmark models encompassed by the universe of linear reduced form models. This comparison reveals that, in a ex-post setting, the structural model substantially outperforms all benchmark models, especially at longer horizons. While the ex-ante forecasting exercise shows that this substantially better performance is true only when excess reserves can be forecast well, this condition is satisfied during period when a QE program is operational. The overall conclusion is that the structural model and the factors embedded in it are important for forecasting interbank interest rates. This result then serves as a basis for more general discussion of which model features are important for forecasting interbank interest rates in the presence of excess reserves. The heterogeneity in forecasting performance is then explained by the econometric nature of eurozone interbank interest rates, which are both nonstationary and cointegrated with monetary policy variables. Models that forecast interbank rates in either their stationary or nonstationary transformation are unable to account for this nature and hence produce forecasts that are problematic.

The third chapter presents the Eviews add-in - that is, user created package - I have developed as part of the work on the second chapter and subsequently published for public use under name SpecEval. The chapter is an excerpt from a longer document provided with the package which demonstrates how the package can be leveraged in developing and evaluating time series models used for forecasting. It does this by following the model development process illustrated in one of the applications presented in the full version of the document.

The fourth chapter of this thesis is dedicated to a separate research topic of experiment in rational inattention. The chapter presents experimental design and empirical results on rationally inattentive behavior of subjects in an interactive environment. Specifically, the experiment features pairs of subjects playing a simple game in which one player - the Sender - exerts costly attention effort to collect information, which is then communicated to a second player; meanwhile, the second player needs to exert costly effort to process the information provided by the Sender and then takes action that will determine the payoff of both players. The empirical analysis of the experimental data confirms the theoretical predictions. The main result is that subjects in both roles react to information about a proxy for the attention costs of their partners in the expected direction: when partner's attention costs are higher, subjects exert effort that is on average significantly smaller, both statistically and economically. Apart from attention effort the chapter also presents results indicating that subjects reflect on information about the likely precision of the signal they have at their disposal when they take action. This can be seen in the fact that Receivers take actions that are closer to their prior belief if they know that the information communicated by the Sender is likely to be imprecise. Taken together, these results provide the first experimental evidence of rationally inattentive behavior in response to the inattention of other players.

Acknowledgments

I express my sincere gratitude to my supervisor, Sergey Slobodyan, who always supported me despite my complicated research endeavors and without whom I would not be defending this dissertation. He was more than just a supervisor, which is exactly what I needed. I am also grateful to my dissertation committee members, Filip Matejka and Veronika Selezneva, and to my referees, Lorenzo Burlon and Michal Franta, who provided useful comments that led me to improve this dissertation considerably. In addition, the fourth chapter greatly benefited from lengthy discussions with Daniel Martin during my research stay at Northwestern University, to whom I am grateful. Finally, I am also grateful to Deborah Novakova, who took on the gargantuan task of doing the first edits to the text, to Gray Krueger, who took over from her, and to the CERGE-EI SAO staff members, who have always made me wonder why all study affairs offices cannot be like the one at CERGE-EI.

Introduction

The global financial crisis of 2007-2008 and the Great Recession that followed marked a sharp break in how monetary policy is conducted. While before central banks relied almost exclusively on steering market interest rates using adjustments to their policy rate instruments, afterwards many central banks expanded their toolkit by adding unconventional policy instruments. Primary among these were large-scale asset purchases, more commonly known as quantitative easing (QE) programs. These were complemented by explicit forward guidance, and in the case of the European Central Bank (ECB), also long-term refinancing operations (LTROs) and negative policy rates.

This shift in the conduct of monetary policy has not been lost on academic literature, with a large body of articles studying the effects of unconventional monetary policy. Most of the existing literature focuses on the direct effects of the policies as standalone instruments. In the case of quantitative easing programs, the focus is primarily on their effects on government bond yields, given that the programs entail purchases of medium- and long-term government bonds. Meanwhile, negative policy rates are studied mostly in terms of their pass-through to bank deposit/lending rates and the effect on loan origination. However, relatively little attention has been paid to links between QE programs and interbank interest rates, or to the interaction between quantitative easing and negative policy rates, and the role this interaction plays in the emergence of negative interbank interest rates. This thesis aims to fill this gap.

The first chapter explores the nexus between the QE program conducted by the European Central Bank, its policy of negative policy rates, and the interbank interest rates. It starts by analyzing the key data series, focusing on how the link between the policy and interbank interest rates is influenced by the presence of excess reserves. Relying on event and correlation analyses, it shows that the emergence of excess reserves leads to a relationship that is fundamentally different from the relationship that prevailed before the emergence of excess reserves. The event analysis shows that, while prior to the emergence of excess reserves the interbank interest rates depended on the main refinancing rate, afterwards the interbank interest rates were anchored by the deposit rate. I demonstrate this by using high-frequency data on the behavior of interbank interest rates around dates on which the ECB changed its policy rates asymmetrically. This event analysis shows that, in the presence of excess reserves, changes in the main refinancing rate do not influence interbank interest rates, while changes in deposit rate do.

Meanwhile, I complement the event analysis by conducting a correlation analysis that highlights a second role played by excess reserves: while the presence of excess reserves changes the policy rate anchoring interbank interest rates, the amount of excess reserves influences the spread between interbank interest rates and the deposit rate. This is what one would expect based on a model of the market for excess reserves: the interbank interest rates are the price for lending or borrowing (excess) reserves, so that an exogenous increase in excess reserves should lead to a decrease in their price.

The data analysis raises an interesting question: What was the overall effect of the QE program conducted by the ECB on interbank interest rates? While this question is analogical to the question on which most academic literature has focused - the effect of the QE program on bond yields - the effect on interbank interest rates has not been explored. To answer this question, I propose a semi-structural time series model that reflects the conclusions drawn from the previous data analysis. The model links interbank interest rates to the two policy rates in a two-regime structure with a threshold variable - the amount of excess reserves - determining the prevalence of each regime. In addition to playing a role in determining the regime, the amount of excess reserves also influences the level of interbank interest rates in the excess reserve regime. The resulting model provides a very good fit for the observed historical time series and confirms that the amount of excess reserves is a statistically significant factor influencing interbank interest rates.

I then use the model to create a path for interbank interest rates under the observed path for excess reserves and under the counterfactual path for excess reserves that could be expected to prevail in the absence of the QE program. Since the QE program was the main driver of changes in excess reserves during the period of its operation, the difference in the amount of excess reserves with and without the program is large. Correspondingly, the interbank interest rates in the absence of the QE program would have been significantly different, with a peak effect of between 20 and 30 basis points depending on the maturity in question. This effect is on par with changes in policy rates of standard size, and the same order of magnitude as - even if smaller than - the typical estimate of the announcement effect of the QE program on bond yields. Despite its apparent size, and despite the fact that this impact channel of the QE program is known in policy circles, this effect has not been previously quantified, and hence I call it 'the neglected channel of the QE'.

The first chapter focuses on explaining the observed behavior of interbank interest rates, and shows that it can be explained very well with the relevant factors, while also demonstrating that ignoring the relevant structure leads to a substantially worse fit. This raises the question of whether the same factors and structure are also valuable in forecasting interbank interest rates in out-of-sample context, in addition to just explaining the behavior within given sample.

The second chapter of this dissertation addresses this question by performing an extensive set of forecasting exercises. These exercises compare the out-of-sample forecasting performance of the structural model proposed in the first chapter with the performance of all linear reduced-form models that could be considered relevant benchmarks. This includes the standard univariate ARMA models, as well as single equation multivariate models of the ARDL family and multi-equation models of the VAR family.

I then analyze the performance of all considered models both in terms of ex-post forecasts - where forecasts are made using observed values of independent variables - and ex-ante forecasts, where independent variables are also forecasted. The results of ex-post forecasting exercises demonstrate the value of the factors included in the structural model and the structure imposed by this model: the forecasting performance of the model is substantially better than any class of the reduced form models or any individual model when all horizons are taken into consideration. This is especially true at longer forecast horizons, with performance being an order of magnitude better, while for shortest horizons, some individual reduced form models can match or even exceed the performance of the structural model. However, the ex-ante forecasting exercise shows that this better performance depends on knowledge of the future path of policy variables, and especially excess reserves. When this is not the case, the forecasting performance of the structural model is either only slightly better or becomes worse than the performance of best reduced form models. The overall conclusion is that the structural model and the factors embedded in it are valuable for forecasting as long as path for excess reserves is known with some degree of certainty. The chapter shows that this is indeed the case during period when a quantitative easing program is operational.

The comparison of the forecasting performance of the structural model and plausible benchmark models forms the basis for the main value added of the second chapter, which is its analysis of what factors are linked to good forecasting performance for interbank interest rates in general. This analysis shows that variations in forecasting performance of individual models can be explained by the specific statistical nature of the series under consideration. The key insight is that the interbank interest rates at medium frequency are both nonstationary and cointegrated with policy variables. Since reduced form models cannot capture the complex nature of the data, they are effectively modelling the series in terms of its non-stationary transformation or in terms of its stationary transformation. Of course, neither of these choices is fully satisfactory and hence both lead to problematic forecasts in particular parts of the sample. In contrast, the structural model is able to account for the complex nature of the data and hence avoids pitfalls faced by other types of models.

The brief third chapter of the thesis is in essence a technical appendix to the second chapter. It presents the Eviews add-in - a user created package - developed as part of the work on the second chapter and subsequently published for public use under the name SpecEval. The SpecEval package is focused on creation, visualization, and evaluation of forecasts from equations or VAR models estimated in Eviews software. Correspondingly the package can be used as a tool for performing forecasting exercises such the one presented in the second chapter. Nevertheless, the third chapter is not motivated by the use of SpecEval for forecasting exercises aimed at the comparing performance of alternative models. Instead, it offers a short demonstration of how the package can be leveraged in *developing* and evaluating time series models used for forecasting. Of course, forecasting exercises are a bedrock of such evaluations, where they are called backtesting exercises. However, academic forecasting exercises typically focus on reporting numerical results. In contrast, SpecEval is not limited to numerical measures of forecasting performance, but also includes numerous graphical representations of the forecasting performance, some of which appeared in chapter 2. The idea behind this is that graphical information often provides more useful information for model development than information summarized in a single summary statistic.

The functionality of the SpecEval package, though, is not limited to backtesting exercises. There are two other types of exercises SpecEval can perform. First, SpecEval has the functionality to create and compare scenario forecasts, assuming that the information about scenario values of some exogenous variables is available. Second, SpecEval can create graphs of shock responses for multiple types of shocks. While this is standard feature of the evaluation of VAR models, it is less common for single equation models and statistical packages typically do not include any such functionality. This is a hard-to-justify omission and hence SpecEval provides a valuable service.

The fourth chapter of this thesis is dedicated to a separate research topic of experiment in rational inattention. Rational inattention has received increasing attention in economic literature. While this literature has been primarily focused on theoretical advances, more recently there has also been interest in providing validation for the theoretical models through empirical and, especially, experimental work. The fourth chapter contributes to this.

The chapter presents an experimental design and empirical results on rationally inattentive behavior of subjects in an interactive environment. The experiment features pairs of subjects playing a simple game in which one player - the Sender - exerts costly attention effort to collect information that is then communicated to a second player; meanwhile, the second player needs to exert costly effort to process the information provided by the Sender and then to take action that will determine the payoff of both players. This game has several important features that allow us to study whether the experimental subjects behave in accordance with theoretical predictions based on rational inattention theory. First, the attention effort of each player is a strategic complement for the attention effort of the other player, but neither player can shirk in hope that the other player will pick up the slack. Second, both players agree on value of information and the desired action given collected information. Thus the game is fully co-operative.

The empirical analysis of collected experimental data confirms the theoretical predictions. The main result is that subjects in both roles react to information about the proxy for attention costs of their partners in the expected direction: when the partner's attention costs are higher, given subject exerts effort that is on average significantly smaller, both statistically and economically. In addition to attention effort, the chapter also presents results indicating that subjects reflect on information about the likely precision of the signal they have at their disposal when they choose an action. This can be seen in the fact that Receivers take actions that are closer to their prior belief if they know that the information communicated by the Sender is likely to be imprecise. Taken together, these results provide the first experimental evidence on rationally inattentive behavior in response to the inattention of other players.

Chapter 1

Negative policy rates and interbank interest rates: The neglected channel of Quantitative Easing

"[C]entral bankers and academics have only started to systematically capture and quantify the various impact channels... [M]ore analysis is certainly needed and forthcoming to better understand the instruments' transmission channels at work, ***and in particular their interaction****."* (Altavilla, Carboni, and Motto 2021)

"...many of the favourable ***cross externalities among tools ... are largely ignored in available analyses****, but are in fact what has made them so powerful and probably indispensable within the ECB's multidimensional easing strategy."* (Rostagno et al. 2019)

1.1 Introduction

The global financial crisis of 2007-2009 and macroeconomic developments since then have led central banks to dramatically change the tools they use to influence the economy. While the focus before was on changes in monetary policy rates, since then the focus has shifted towards unconventional monetary policy instruments. These instruments include large-scale asset purchases (commonly referred to as the policy of quantitative easing), forward guidance, and in the case of the European Central Bank (ECB), also long-term refinancing operations (LTROs) and negative policy rates. This chapter studies the effect of these policies on financial markets, and specifically on interbank interest rates (IIRs). Its key focus is on the interaction between the two most significant unconventional policies, quantitative easing (QE) and negative policy rates, and what role this interaction plays in the emergence of negative IIRs.

This chapter contributes to existing literature in two fields: empirical modeling of IIRs, and analysis of unconventional monetary policy. On the modeling front, I propose and estimate a novel semi-structural model for IIRs. The main innovation of the model is to suggest that the relationship between policy rates and IIRs is fundamentally altered by the presence of excess reserves in a way that causes there to be two different regimes of operation for IIRs. In the normal regime - which prevails when there are no excess reserves and hence had been prevailing in period before the global financial crisis - the IIRs depend only on the interest rate at which banks can regularly obtain liquidity from the central bank, the main refinancing rate. In excess reserves regime - which prevails in the presence of excess reserves and hence has applied since the global financial crisis - the IIRs depend only on the rate at which commercial banks can deposit their excess reserves with the central bank: the deposit rate (DR). Apart from changing the regime and hence the policy rate anchoring the IIRs, excess reserves also affect the IIRs through their effect on the spread between the IIRs and deposit rate. Specifically, in excess reserves regime, the *spread* between IIRs and the DR decreases as the *amount* of excess reserves increases.

This relationship between policy rates, excess reserves and the IIRs suggests that two roles are played by the policy of QE, which directly leads to the emergence of and/or increases in excess reserves. First, if prior to the QE there are no excess reserves, the program can lead to switch from the normal regime to the excess reserves regime as the QE instigates the emergence of excess reserves. Second, as excess reserves increase due to continued asset purchases, the IIRs are pushed further towards the deposit rate, something that has been acknowledged in policy circles (e.g. Altavilla, Carboni, and Motto (2021) or Rostagno et al. (2019)), but never quantified.[1]

This leads to the policy contribution of this chapter. I use this novel semi-structural model to answer following policy question: What was the effect of the QE program of the ECB on IIRs? The estimates of the model suggest that there is statistically and economically significant effect of the QE program on the IIRs. If the ECB had not implement its the QE program, the Eonia rate would be, on average, 6bps higher since the start of the program, with a peak effect of 16-22bps, depending on the model used. Moreover, the effect on longer-maturity IIRs is possibly even larger.

Insofar as it is IIRs that matter for the wider economy, this is a sizeable effect, comparable to a standard-sized decrease in policy rates. It is even significant relative to various estimates of the effect of other unconventional monetary policy tools. For example, the announcement of the QE program by the ECB in January 2015 led to a similar decrease in long-term interest rates, even though the cumulative effect of the program was likely several times larger (see e.g. Altavilla, Carboni, and Motto (2021); Greenlaw et al. (2018) provide a more skeptical view of the effects of QE programs). The effect is also similar to the peak effect of forward guidance (Rostagno et al. 2021). Despite its apparent size, and despite

[1]Note that this channel did not initially feature prominently in the discussions of the effects of the QE program. For example, early assessment of the QE program by Hartmann and Smets (2015) does not mention it at all.

the fact that this impact channel of the QE program is known in policy circles, this effect has not been previously quantified, and hence I term it 'the neglected channel of QE'.

Turning to negative interest rates, the effect of the QE program on IIRs also suggests that it plays a role in the emergence of negative *interbank* interest rates (as opposed to negative *policy* rates, which are purely under the control of the ECB). Since the model suggests that IIRs would be significantly higher in the absence of the QE, it also suggests that they would either remain positive or only mildly negative for all of 2015. While they would turn negative at the end of 2015 irrespective of the QE program, they would reach substantially less negative values in the following years.

Moreover, this discussion potentially understates the importance of the QE in the emergence of negative the IIRs, as the ECB's the QE program was initiated when excess reserves were already present thanks to other policies, and even increased during the period of the QE program. Since only the deposit rate was lowered into negative territory, the emergence of negative IIRs crucially depends on the presence of excess reserves: in the absence of excess reserves IIRs would have remained slightly above the main refinancing rate and hence positive throughout the last decade. If the QE program had been the only policy causing excess reserves during the relevant period, we would not observe negative IIRs in the absence of the QE program. Hence, negative policy rates and the QE policy (or an alternative policy leading to excesss reserves) are complementary to each other - as only their combination leads to negative IIRs.

Corresponding to its two contributions, this chapter is related to two strands of literature. First, there is a large literature studying the behaviour of IIRs, especially overnight rates. Green et al. (2016) provide a recent survey of this literature, most of which focuses on the high-frequency behaviour of overnight interest rates and its relationship to policy interest rates and liquidity management by central banks. In contrast, this chapter focuses on medium- and long-term movements in IIRs. The closest work is that of Marquez, Morse, and Schlusche

(2013). They focus on the exit strategies of the Federal Reserve (Fed) and hence model the overnight rate as a function of excess reserves, among other things. The paper finds a presence of the same negative effect of excess reserves on overnight rate. Nevertheless, they focus on the somewhat different institutional environment of the United States and hence their model is very different. Specifically, due to the institutional differences, they do not postulate the existence of different regimes for IIRs.

Second, a more closely related strand of literature is the burgeoning literature on unconventional monetary policy, and specifically on the ECB's multiple novel instruments introduced since 2014. Altavilla et al. (2021b) provide a detailed overview of these instruments and survey of literature analyzing them. This literature is mostly organized around the instruments and their primary impact channel. A large number of articles studies various QE programs and their impact on long-term interest rates, see D'Amico et al. (2012) and Gagnon et al. (2011) in the case of the US, and Altavilla, Carboni, and Motto (2021) and Eser et al. (2019) in the case of euro zone.[2] These studies focus on the two main channels of QE programs, the portfolio-rebalancing channel induced by a decrease in supply of long-term bonds, and the signaling channel. A more closely related study is that by Christensen and Krogstrup (2016), who highlight that portfolio rebalancing and the corresponding decline in bond yields can also be induced by an increase

[2]There is also literature on central banks intervention during the global financial crisis, be it liquidity infusions or (indirect) purchases or risky assets. Examples include Adrian, Kimbrough, and Marchioni (2010), Berger et al. (2015), Boyson, Helwege, and Jindra (2015) and Duygan-Bump et al. (2013). This literature is separate, though, because the policy instruments were addressing malfunctioning financial markets, which likely makes their effects conditional on the state of financial markets. In contrast, the ECB's unconventional instruments discussed in this chapter were activated to address the combination of deteriorating macroeconomic outlook and the absence of space for conventional monetary policy tools, not to address stress in financial markets (Rostagno et al. 2019).

Yet another strand of literature studies the macroeconomic impacts of unconventional policies, using either the approach of shadow interest rate, e.g. Wu and Xia (2016), or VAR approach augmented by information from high-frequency analysis, e.g. Rostagno et al. (2021). Here I focus only on the impact on IIRs, and do not address the implications for the wider economy. As such, the literature on the estimation of shadow rates is not related to this paper: for observed IIRs it is the actual policy rates that are relevant.

in excess reserves caused by purchases of other assets than long-term bonds. Like this chapter, they highlight that excess reserves created by asset purchases impact market interest rates, an impact that is independent and in addition to the impact of asset purchases. In other words, the liability of the central bank balance sheet also plays an important role in determining the impact of asset purchases. However, the focus of these studies is firmly on long-term bond yields and they do not discuss IIRs.

Another unconventional monetary policy tool receiving substantial attention is the policy of negative interest rates.[3] Jackson (2015), Bech and Malkhozov (2016) and Jobst and Lin (2016) provided early discussions of the policy around the world, including basic discussions of pass-through to market interest rates and how it relates to excess reserves. Nevertheless, none of these papers develops a model for the behaviour of IIRs, which constitutes the first contribution of this chapter, or discuss quantitatively the role played by the QE program, what constitutes the second contribution of this chapter. Later studies focus on the transmission of negative policy interest rates to bank funding and its effect on lending behavior (Altavilla et al. 2021a; Bottero et al. 2019; Eggertsson et al. 2019; Heider, Saidi, and Schepens 2019). However, these studies focus on bank-level data and hence do not aim to model the aggregate IIRs or to quantify the effect of the QE program on those rates. Insofar as bank funding/lending partly reflects the aggregate level of IIRs, this chapter studies the earlier link in the chain of effects leading to lending.[4]

There is also slowly emerging literature on interaction among the unconventional tools and their impacts through secondary channels, to which this chapter is most closely related. The importance of this interaction of tools is emphasized

[3] The other two unconventional policies were (targeted) long-term refinancing operations (LTROs) and forward guidance. See for example Andreeva and Garcia-Posada (2019) for LTROs and Altavilla et al. (2019) for the effects of forward guidance. Altavilla et al. (2021b) provides an overview of relevant literature.

[4] The discussion of negative policy rates also links this chapter to the issue of zero lower bound and literature focusing on existence of such bound (e.g. Brunnermeier and Koby (2016, Rognlie (2015)). However, the empirical facts of negative interest rates addressed here make it clear that the chapter is more related to the *absence* of zero lower bound, than to its existence, similarly to Altavilla et al. (2021a).

by Rostagno et al. (2019) and Altavilla, Carboni, and Motto (2021) (see quotes at beginning of this chapter), who also provide a helpful classification of all channels and interactions. Among these studies Rostagno et al. (2019) and Rostagno et al. (2021) demonstrate that decreases in policy rates into negative territory were associated with abnormally large decreases in expected future IIRs and bond yields with medium and long maturity, and relate this to, among other things, the interaction between negative policy rates and forward guidance; see also Wu and Xia (2020). Meanwhile, Ryan and Whelan (2021) discuss the interaction between the QE program and negative policy rates that arises thanks to the emergence of excess reserves caused by the QE program. However, these studies either focus only on *expected* IIRs, or do not discuss IIRs at all, and hence are silent on the connection between *current* IIRs and the QE program. While Demiralp, Eisenschmidt, and Vlassopoulos (2021) highlight the importance of excess reserves in the transmission of negative policy rates to bank behavior, they are only interested in the cross-section of bank lending and not in IIRs.

Finally, the most closely related paper is by Arrata et al. (2020), who study the effect of the QE program on money market rates (a broader category than IIRs). Their investigation shows the link between money market rates and asset purchases under the QE program. However their focus is on the role of scarcity induced by asset purchases and hence study special repo rates. This leads them to perform their empirical analysis at daily frequency and to exploit cross-sectional variation, both features that make their analysis quite different from mine. While their results also point towards the role played by excess reserves, they downplay the role and hence do not explore it in more detail.

The remainder of the chapter is organized as follows. Section 1.2 describes three key aspects of the data. Section 1.3 describes and discusses the modelling approach and section 1.4 presents the results of the estimation. Section 1.5 presents the estimates of the effects of the QE program. Finally, section 1.6 concludes.

1.2 Data analysis

In this section I document and discuss three main features of the data which serve as motivation for my latter formal econometric analysis. First, I discuss the emergence of negative interest rates in the euro zone, focusing both on policy rates and, more importantly, interbank interest rates (IIRs). Second, I present and analyze the evolution of excess reserves of commercial banks and outline the role played by quantitative easing (QE) policy in their evolution. Finally, I will show how these two phenomena are connected.

1.2.1 Negative interest rates

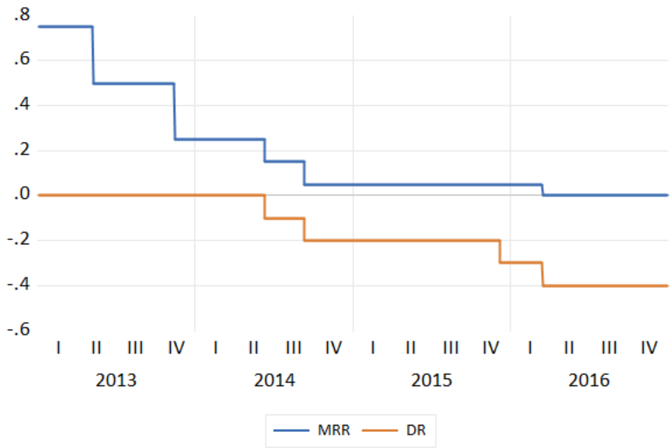
The European Central Bank (ECB) introduced a negative deposit rate (DR) in June 2014 in response to falling inflation and inflation expectations (Draghi 2014a; Rostagno et al. 2019). Initially, the DR was set at -10 basis points and was lowered to -20 basis points in September 2014 (see Figure 1.1), which was at that time considered the effective lower bound (Draghi 2014b; Draghi 2014c; Rostagno et al. 2019). Nevertheless, following the experience of other countries throughout 2015, the ECB lowered the DR to -30 basis points in December 2015 and to -40 basis points in March 2016, where it remained until September 2019.[5] Meanwhile, the other policy rate, the main refinancing rate (MRR), was decreased from 0.75% to 0% in a series of steps spanning 2013-2016. This points towards aspect of the data that I will exploit later and which also visible in Figure 1.1: during this period we can observe several asymmetric changes to the two policy rates.

Turning to interbank interest rates[6] , the left panels of Figure 1.2 show that

[5]In this section and throughout the chapter, I limit the sample to end in August 2019. This is because in September 2019 the ECB introduced a two-tiered system for remunerating excess, reserves and in beginning of October the Eonia rate, which will be my main variable of interest, was replaced by the Euro Short-Term Rate rate (ESTR); both of these changes potentially influenced the mechanisms described in this chapter.

[6]I use the term of "interbank interest rates" - rather than the term "money market rates" - to indicate that I focus on banks on the interbank lending market specifically. Money market is broader term than intebank market, and money market rates encompass greater universe of rates than just interbank interest rate. For example, they also include repo rates.

Figure 1.1: ECB policy rates



these slowly trended downwards, following the DR into negative territory. Nevertheless, both the Eonia rate and the Euribor rates went negative with a substantial delay compared to the DR. The Eonia rate and Euribor rates with shortest maturity became negative for the first time only at the beginning of September 2014, after the second decrease in the DR, and remained relatively stable for the next six months, before decreasing dramatically after the start of QE purchases in March 2015. Meanwhile, Euribor rates with longer maturities remained positive throughout 2014 and turned negative only in March and April 2015, after QE purchases started. Later in this section and throughout the chapter I will argue that this timing reflects a causal relationship between excesss reserves and the IIRs.

It is instructive to translate the observed the behavior of levels of the IIRs into behavior of spreads from the DR (see right hand side panels of Figure 1.2). Before the start of QE purchases, the spreads were volatile, but did not display any clear upward or downward trend. In contrast, all the IIR spreads decreased dramatically following the start of QE purchases, halving in span of three months between March and June 2015. Afterwards, the spreads continued their downward path, but the magnitude of the decrease was much smaller. Finally, in last years of the sample the spreads were stable at levels of around one third of their value prior to the start of QE purchases.

1.2.2 Excess reserves

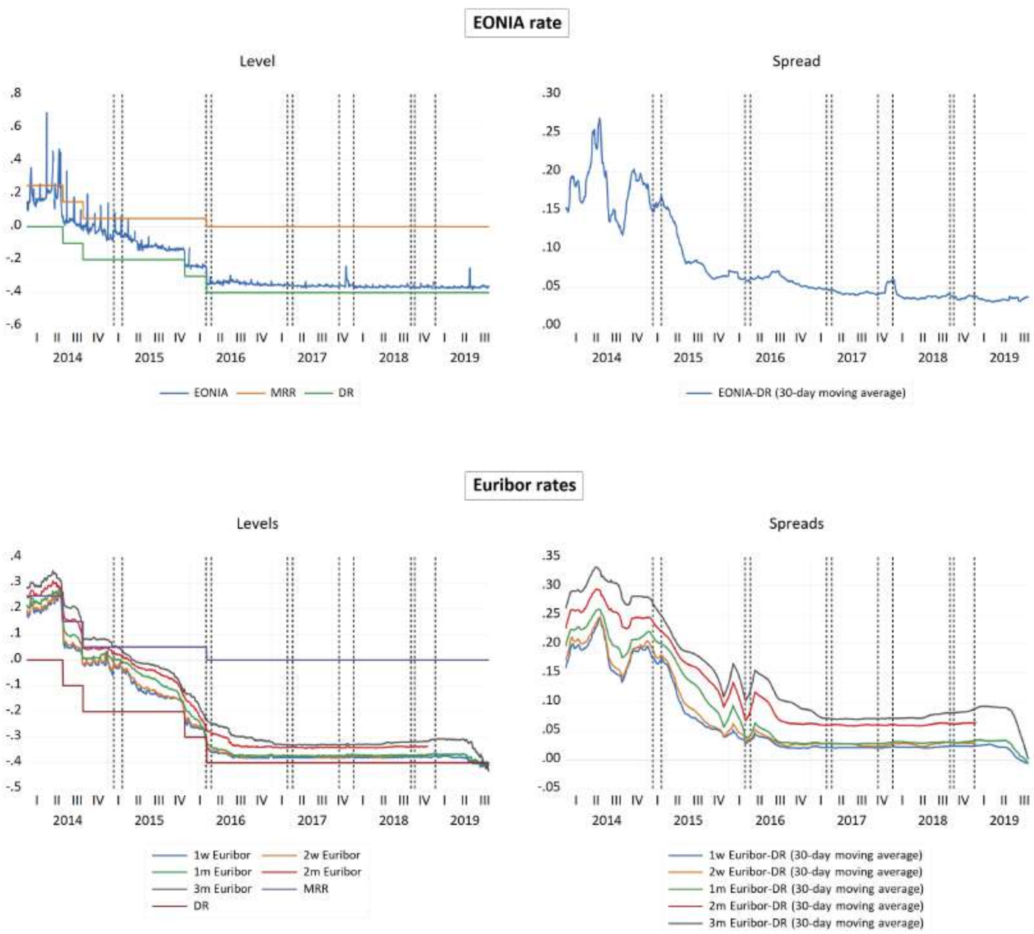
The autumn of 2008 marks the emergence of the phenomenon of excess balances of the euro zone commercial banks at their accounts with the ECB (I will refer to these as *excess reserves in the interbank market*, or simply as *excess reserves*), defined as the sum of current and deposit account balances of commercial banks with the ECB minus the reserve requirements. The rest of this subsection briefly describes the evolution and sources of excess reserves over past decade. More detailed discussion is relegated to Appendix 1.A.

Substantial excess reserves in the euro zone interbank market were observed for the first time in October 2008 (see Figure 1.3), right after the ECB switched from fixed allotment to full allotment tender procedures. This switch meant that the ECB was no longer determining or limiting the amount of liquidity commercial banks could obtain in refinancing operations. After jump in fall of 2008, excess reserves were fluctuating throughout 2009 and 2010, staying in between 40 and 300 billion euro until the beginning of 2011. During that year, excess reserves first almost disappeared, then again increased substantially in the autumn. Finally, at the end of 2011 and the beginning of 2012, excess reserves recorded several large jumps which left them at almost 800 billion. The main reason for the increases in excess reserves were the two Very Long-Term Refinancing Operations (VLTROs), executed in December 2011 and February 2012. Concurrently with the VLTROs the ECB lowered the reserve ratio from 2% to 1%, effective from January 18th 2012, causing a decrease in required reserves and a corresponding jump in excess reserves.[7]

Implementation of the second VLTRO marked a local peak in excess reserves, after which they began gradual decline lasting until the middle of 2014, when they

[7] The effect of a decrease in reserve requirements can be seen in a decrease in the current account holdings and an offsetting increase in deposit account holdings around the effective date.

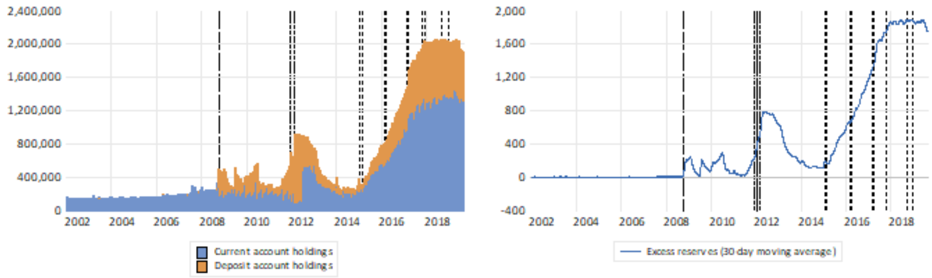
Figure 1.2: Interbank interest rates



Notes: In this and other figures in this section, the vertical lines indicate the timing of relevant policy interventions by the ECB. The interventions are:

- Large dashed lines indicate the announcement of the shift to full allotment tender procedures (October 2008).
- Dashed-dot lines indicate announcement and implementation of VLTROs in 2011-2012.
- The solid line indicates change in reserve requirements in 2012.
- Small dashed lines indicate announcements and implementations of changes in asset purchase programs, startingin 2014.

Figure 1.3: Excess reserves



Left panel shows the current and deposit account balances of commercial banks with ECB, as captured in the ECB balance sheet statements. Right panel shows excess reserves as defined in text.

briefly dropped below 100 billion. April 2014 then marks a local minimum of excess reserves, which first jumped back to above 100 billion for the rest of 2014, and then started increasing rapidly from March 2015.

The figure highlights that the dramatic break in the trend of excess reserves observed in March 2015 coincides with the start of asset purchases under the QE program. Under this program, the ECB purchased large quantities of financial assets, mainly government bonds (see Claeys, Leandro, and Mandra (2015) or Hammermann et al. (2019) for more details). Crucially, these purchases were unsterilised, meaning that their effect on the size of the ECB balance sheet was not reversed through other ECB's actions, which was one of the motivations for launching the program (Rostagno et al. 2019). Effectively, these purchases replaced private sector holdings of government bonds with commercial banks' balances at the ECB and hence are causally related to excess reserves (Boucinha and Burlon 2020); see a detailed discussion in Appendix 1.A.2, which illustrates this point using accounting techniques. As Figure 1.3 shows, since the start of the QE program, excess reserves continued to increase at a steady pace, reaching almost two trillion in 2017, where they remained until the end of the sample.

Before discussing the connection between negative IIRs and excess reserves, it is useful to highlight an important aspect of the evolution of excess reserves over the QE period: the fact that the changes in excess reserves (i.e. asset purchases under the QE program) are exogenous with respect to the behavior of commercial banks (ECB 2015) and with respect to determinants of short-term interbank interest rates in general. Not only is the primary goal of these asset purchases their effect on inflation, but the primary channel through which they are meant to affect the economy is their effect on medium-to-long-term bond yields (Rostagno et al. 2019), not short-term IIRs.[8] Even more importantly, the monthly increases are almost deterministic, with the amount of assets to be purchased each month specified ahead of time for prolonged periods. This is in contrast to changes in excess reserves in the pre-QE period, which were determined to a large degree by the commercial banks themselves, and hence are potentially endogenous to IIRs. Correspondingly, the exogeneity of QE-related excess reserves will be a key aspect of my empirical strategy. Ryan and Whelan (2021) offer similar arguments with respect to the exogeneity of excess reserves during the QE program.

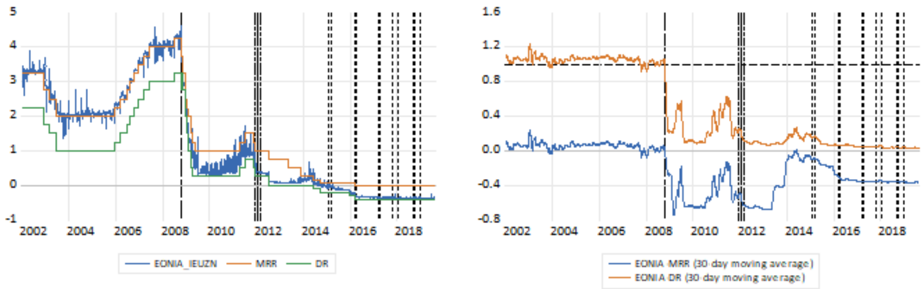
1.2.3 Policy rates, excess reserves and the IIRs nexus

The previous two subsections discussed the two key aspects of data relevant for this chapter: negative interest rates and the presence of excess reserves. Before turning to the developing model that links these two aspects together, this subsection provides preliminary data analysis that highlights how the relationship between policy rates and interbank interest rates is fundamentally altered by the presence of excess reserves. To do this, it presents an analysis of the relationship between the Eonia rate and policy rates before and after the emergence of excess reserves.

Figure 1.4 shows that, before the emergence of excess reserves in October 2008, the Eonia rate was consistently just above the MRR, which is equivalent to a small positive the spread between these two rates. This has changed starting in October 2008, when the Eonia rate dropped below the MRR where it has remained ever

[8]Note that most of the empirical analysis of effect of QE program on financial markets *assumes* that the program does not affect short-term IIRs. See for example Altavilla et al. (2019) or Rostagno et al. (2021).

Figure 1.4: Eonia rate - full sample



since, which translates into a negative spread between these two rates over the whole period after 2008. In contrast, the Eonia rate never dropped below DR, so that the spread between the Eonia rate and the DR remained positive at all times. In terms of the volatility of the spread between the Eonia rate and either the ECB rate, the period after 2008 witnessed much higher volatility than the period before 2008. These observations clearly suggest a change in regime in October 2008, which affected both the level and the volatility of these spreads, something I confirm formally in Appendix 1.F.1.

The higher volatility of the spread of the Eonia rate from either policy rate masks one important fact: in the period after 2008, the spread between the Eonia rate and the DR was visibly more stable than the spread between the Eonia rate and the MRR. This chapter will argue that the reason for this greater stability in the DR spread is that the Eonia rate (and other IIRs) are anchored by the DR rather than by the MRR whenever excess reserves are present. To support this hypothesis, it is natural to analyze the asymmetric changes in the ECB policy rates, as those can shed light on how IIRs change when the ECB changes just one of policy rates, and hence whether both rates influence the IIRs in presence of the excess reserves. Since September 2008, we have witnessed 5 such asymmetric changes in policy rates: 2 decreases in the MRR without a change in the DR, 1 decrease in the DR without a change in the MRR, and 2 decreases in both rates but of different size. Figure 1.5 shows policy rates together with the Eonia rate during these 5 changes, while Table 1.1 shows the change in the Eonia rate following the changes in policy rates.

Table 1.1: Assymetric changes in policy rates

Date	Change in given interest rate				
	MRR	DR	Eonia rate 1-day	Eonia rate 3-day	1w Euribor
21/01/2009	-50	-100	-64.9	-94.3	-57.4
08/05/2013	-25	0	-1.1	-0.1	0
13/11/2013	-25	0	-0.4	0.5	-0.5
09/12/2015	0	-10	-8.8	-8.4	-6.1
16/03/2016	-5	-10	-9.1	-9.7	-6.9
22/01/2015	–	–	4.9	4.8	2.6

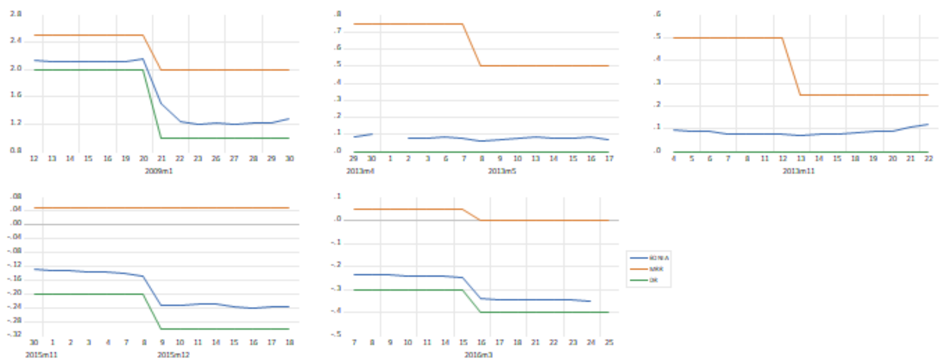
Notes: Changes in policy rates and the IIRs, in basis points, corresponding to policy changes effective on specified date[a]. For the Eonia rate the changes are calculated as difference between interest rate prevailing 1 and 3 days after the change in policy rates minus the interest rate prevailing on the day before the change in policy rates. For 1-week Euribor rate the changes are calculated as difference between interest rates prevailing 1 day after the change in policy rates minus the interest rate prevailing 5 days prior to the policy decision. Last row captures change around the first announcement of the QE program.

[a]Note that the changes in monetary policy rates are effective with few days delay. Throughout the chapter I use the effective date as the relevant timing for the policy changes.

Both the figure and the table clearly establish that, in a period with excess reserves, changes in the MRR do not influence the Eonia rate, while changes in the DR cause an almost equivalent decrease in the Eonia rate. When the MRR decreased twice in 2013 by 25 basis points, the Eonia rate reacted only minimally in either 1-day or 3-day window (2nd and 3rd panel/row). In contrast, the decrease in the DR in December 2015, at which time the MRR was left unchanged, resulted in an almost equivalent decrease in the Eonia rate (4th panel/row). Finally, when both the MRR and the DR decreased by different magnitudes, the decrease in the Eonia rate was substantailly closer to the decrease in the DR (1st and 5th panel/row).

I complement these results with a brief analysis for the 1-week Euribor, numbers

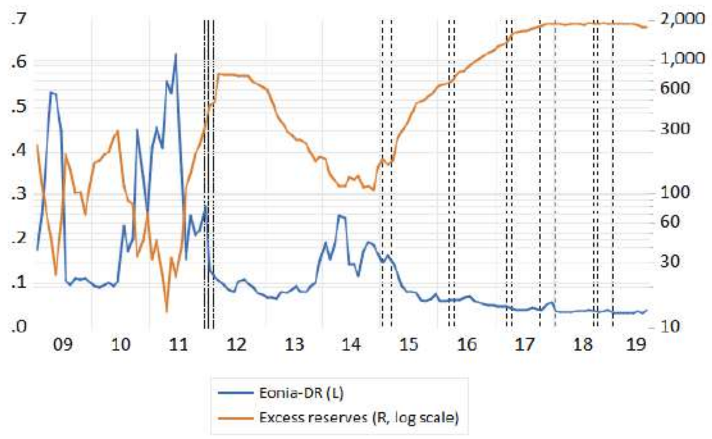
Figure 1.5: Effects of assymetric changes in policy rates



for which are in the last column of Table 1.1. Broadly speaking, the results for the 1-week Euribor indicate the same relationship as with the Eonia rate, with the Euribor rate not reacting to changes in the MRR, and reacting strongly to changes in the DR. Overall, the data in Table 1.1 and Figure 1.5 establish the first effect of excess reserves on relationship between the IIRs and the ECB policy rates: for both overnight and longer-maturity the IIRs it is the level of DR, rather than the level of MRR, that influences the level of various interbank interest rates *in the presence of excess reserves.*

The change in anchor - from the MRR to the DR - is not the only effect of excess reserves. Additionally, the *amount* of excess reserves also plays a role in the *size* of the spread between IIRs and the DR. There are two ways to establish this claim. First, Figure 1.6 combines time series data from Figures 1.2 and 1.3 to suggest a clear negative correlation between the evolution of excess reserves and level of the spread between the IIRs and the DR. For example, as discussed in section 1.2.1, the level of different IIRs turned substantially negative only with the onset of the QE program, which led to dramatic increase in excess reserves. In Figure 1.6 this is captured by large decrease in the spread between the Eonia rate and the DR after the start of the QE program. This clearly suggests relationship between the amount of excess reserves and the spread between the IIRs and the DR.

Figure 1.6: Excess reserves and the Eonia rate the spread

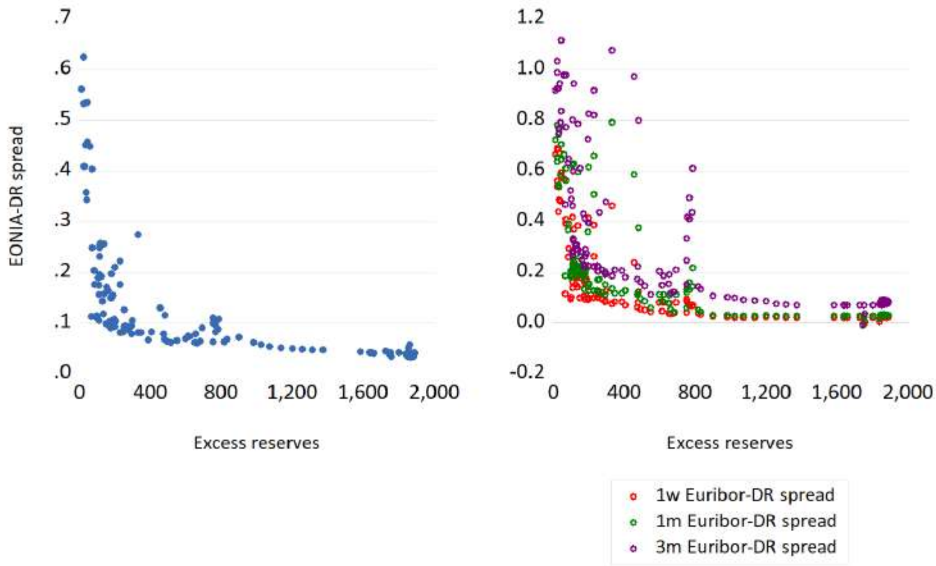


Second, this relationship and its strength can best be seen in the scatter plot of the spread between the IIRs and the DR plotted against the amount of excess reserves, presented in Figure 1.7. Not only is the relationship clearly negative, but it seems to be very strong, especially at low levels of excess reserves.[9] Moreover, the relationship is also very close in the case the Eonia rate, which is uncontaminated by additional factors, and reasonably close in the case of the IIRs with longer maturity.

Does negative correlation between excess reserves and IIRs reflect causation? Before turning to the modeling strategy based on these empirical observations, it is worthwhile to highlight that the negative relationship between

[9]Similar graph can be found in Valiante (2015). However, he misinterprets the causality between excess reserves and the spread. Specifically, sometimes it is (wrongly) argued that excess reserves mean that banks are hoarding reserves. This is incorrect for following reason: While individual banks can decrease their amounts of reserves, they can do so only by lending/transferring them out to other banks, and hence the banking system as a whole cannot decrease the amount of excess reserves (apart from collectively forcing the public to hold more deposits). Therefore, the amount of excess reserves is, to a large degree, determined by the actions of the central bank. See Appendix 1.A for an explanation.
Corresponding to this misconception that the banking system as such decides how much excess reserves it holds, Valiante (2015) concludes that excess reserves increase as the spread decreases, thus reversing the causality. In contrast, Marquez, Morse, and Schlusche (2013) interpret identical graph in the same way as I do.

Figure 1.7: Effect of excess reserves on the IIRs



Notes: The horizontal axis corresponds to excess reserves in billion euros at a monthly frequency. Vertical axes are spreads in basis points. The sample covers 2009M01-2019M08.

excess reserves and IIRs is very likely of a causal nature. I support this by two arguments, one based on theory and one based on observed historical data.

First, from a theoretical perspective, it is not surprising that excess reserves *cause* lower IIRs. At the most basic level, IIRs measure the price of lending/borrowing reserves between banks. Standard microeconomic theory clearly indicates that the supply of (excess) reserves should play a key role in determining such price. Therefore, the basic argument explaining why excess reserves *cause* movements in IIRs is the standard supply and demand argument, which is indeed at the center of exposition of modern central banking; see for example Ennis and Keister (2008).

These general considerations have found their expression in theoretical micro-founded models starting with Poole (1968). Section 4 of Green et al. (2016) provides a recent review of this literature. The most relevant work with respect to the current analysis is Bech and Monnet (2016), who depart from similar data analysis and then provide a micro-founded theoretical model of IIRs, explicitly accounting for excess reserves. Their conclusion is that *"when the surplus of reserves is large then the [Eonia] rate tends to the deposit facility rate"* ((Bech and Monnet 2016), p. 50), supporting the empirical analysis I perform here.

Meanwhile, from an empirical perspective, the most likely reservation about the causality of the relationship rests on the role played by the QE program: one could consider the relationship in Figure 1.7 to be a side-effect of the effect that asset purchases have on other market interest rates; lower bond yields caused by the QE program might have also pushed down IIRs, with no role played by excess reserves. There are two reasons such a conclusion is hardly warranted. First, the negative relationship is present also before asset purchases under the QE program, reflecting the fact that the source of excess reserves is of second order importance from the perspective of the theory. This suggests that excess reserves play a role in determining IIRs.

Second, there is important difference between the effect of the QE program on bond yields and the effect of excess reserves on IIRs. The effect on bond yields is immediate thanks to the forward-looking nature of the pricing of long-term bonds and hence occurs at announcement of the policy (Altavilla et al. 2019). In contrast, the effect on IIRs, which have short maturity, occurs only once excess reserves are actually infused into the interbank system. That is why we do not observe any decreases in the IIRs at the announcement of the QE program (see last row of Table 1.1). Instead, the declines in IIRs occur as the excess reserves increase, something I document in more detail in section 1.4.[10] I return to this difference in section 1.5.

Finally, if excess reserves *causally* impact IIRs, while the QE program *causes* an increase in excess reserves, then the QE program *causes* a decrease in IIRs (see also Boucinha and Burlon (2020)). This effect of the QE program on IIRs then constitutes a separate channel through which the QE program affects market

[10]Another way to see this is by observing that the co-movement between IIRs and short-term bond yields is rather weak and likely reflects the impact of IIRs on bond yields, see the discussion in Appendix 1.C.6. There, I also show empirically that the movements in IIRs are not driven by the movements in short-term bond yields.

interest rates, which goes above and beyond the effect on bond yields: while the duration channel of the QE affects *medium- and long-term* interest rates through decreasing the supply of corresponding bonds (Hammermann et al. 2019), the effect on excess reserves affects the *short-term* IIRs. This mirrors discussion in Christensen and Krogstrup (2016) and Christensen and Krogstrup (2019), who show that the creation of excess reserves via asset purchases leads to a decrease in bond yields *independently* of the effect of asset purchases themselves, which they call the reserve-induced channel of the QE program.

1.3 Modelling approach

Following the conclusions from previous section, this section proceeds by proposing a regime switching model for the IIRs. The model has two regimes, which are determined according to a threshold-switching mechanism, with excess reserves playing the role of an (exogenous) threshold variable. In both regimes the IIRs are anchored by two main policy rates: the main refinancing rate, and the deposit rate, but which policy rate acts as an anchor differs across regimes.

The remainder of this section is separated into three parts: a detailed part dedicated to the Eonia rate, a brief part dedicated to rates with longer maturities, and a concluding discussion.

1.3.1 Eonia rate

The key idea for modelling IIRs in a normal regime is the substitutability between funds obtained from the ECB and funds obtained from other commercial banks. Specifically, commercial banks always have an option to obtain funds from the ECB through regular refinancing operations, the price of which is given by the main refinancing rate (MRR). At the same time, banks can in principle always borrow from other commercial banks through overnight interbank market, with averaged prices of such transactions recorded as the Eonia rate. This suggests that the Eonia rate should not be significantly higher than the MRR, as otherwise banks would profit from an arbitrage opportunity by borrowing from the ECB and lending to other commercial banks. Therefore, the Eonia rate should be equal to the MRR plus the spread to compensate banks for the associated risk of lending funds to other commercial bank. Denoting the Eonia rate in the given period with the general label IIR_t we obtain the following simple model:

$$IIR_t = \beta_0 + MRR_t + \epsilon_t \tag{1.1}$$

Though this equation seems quite innocuous, it embodies three important assumptions: (1) the absence of other policy rates and excess reserves, (2) the absence of future (or past) values of the MRR, (3) the assumption that the IIR move one-for-one with the MRR, i.e., that the spread does not vary with the level of MRR. The first and third assumption reflect theoretical and empirical considerations;[11] I discuss the second assumption in greater detail in Appendix 1.B.3.

The relationship captured by equation (1.1), however, should be expected to hold only in situations when the ECB is adjusting the amount of liquidity to ensure that the Eonia rate is close to the MRR and, therefore, when there are no excess reserves. The arbitrage argument above explained why the Eonia rate should not be significantly *above* the MRR. However, since excess reserves are remunerated by the DR, and since the DR is always below the MRR, there is potential for the Eonia rate to be below the MRR when there are excess reserves. Simply, the opportunity cost of lending to other commercial banks is the DR, not the MRR. This means that when commercial banks have excess reserves, they are willing to lend to other commercial banks at rates below the MRR, as long as

[11]Specifically, prior to the switch to a full-allotment policy, the ECB had been continuously adjusting its liquidity provisions to ensured that the Eonia rate is close to the MRR, which ensures both the irrelevance of other policy rates (assumption 1) and the one-for-one co-movement with the MRR (assumption 3). It also ensured that there were no excess reserves during the normal regime. See Figure 1.4 for empirical support for these assumptions. Appendix 1.F provides formal tests of the first and third assumptions, showing that it is indeed the case that other policy rates and excess reserves do not influence IIRs in the normal regime, and that in that regime IIRs move one-for-one with the MRR.

they receive a rate above the DR. This suggests that in the presence of excess reserves, the Eonia rate is anchored by the DR rather than the MRR, something I have empirically established in the previous section. Overall, we are left with a two-regime relationship between the Eonia, the MRR and the DR:

$$IIR_t = \begin{cases} \beta_{10} + MRR_t + \epsilon_t & \text{if } D_t^{ER} = 0 \\ \beta_{20} + \beta_{21} DR_t + \epsilon_t & \text{if } D_t^{ER} = 1 \end{cases} \tag{1.2}$$

Here D_t^{ER} is regime dummy variable indicating whether excess reserves are present or not, so that the equation (1.2) is an exogenous regime-switching model, with the presence of excess reserves determining the prevalent regime. Correspondingly, I label the regimes *normal* regime and *excess reserves* regime. Postulation of this regime-switching specification is the first main contribution of this chapter in terms of modelling approach.

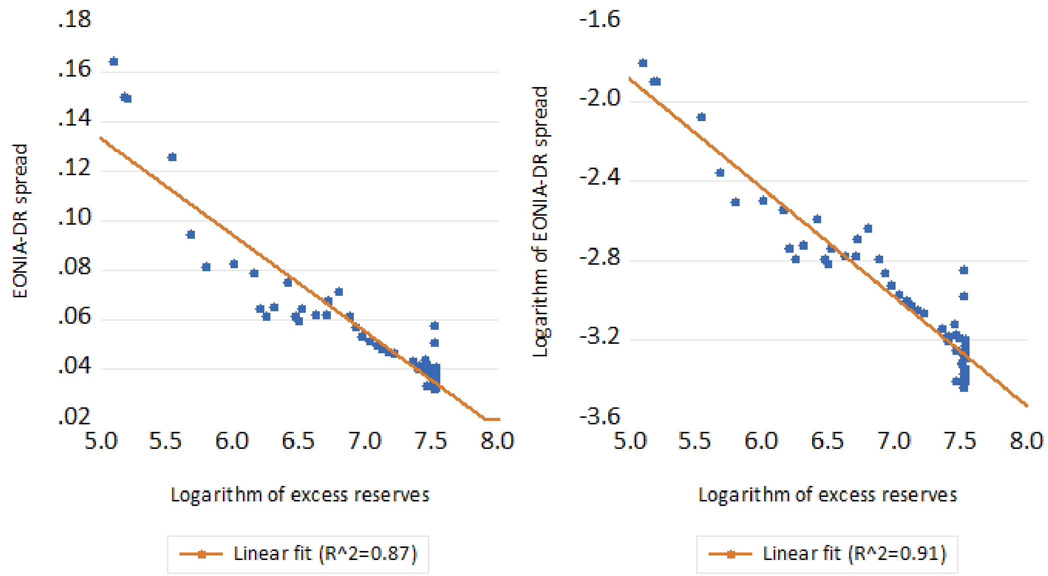
Two comments are in order with respect to equation (1.2). First, note that I allow for the spread between the IIR and the DR to depend on the level of DR, in contrast to the normal regime captured in equation (1.1). Second, as in equation (1.1), I assume that other policy rates do not influence the IIR in excess reserves regime, a restriction that is tested in Appendix 1.F.

Equation (1.2) omits the second channel of excess reserves stressed in section 1.2.3: the effect of the *amount* of excess reserves on the *size* of the spread between the IIR and the DR. This means that equation (1.2) is somewhat incomplete: we need to include amount of excess reserves as one of the regressors. Assuming a general functional form linking excess reserves and the IIR, we obtain the following specification:

$$IIR_t = \begin{cases} \beta_{10} + MRR_t + \epsilon_t & \text{if } D_t^{ER} = 0 \\ \beta_{20} + \beta_{21} DR_t + \beta_{22} g(ER_t) + \epsilon_t & \text{if } D_t^{ER} = 1 \end{cases} \tag{1.3}$$

where ER_t is the amount of excess reserves. To estimate (1.3) we need to specify

Figure 1.8: Log excess reserves and the Eonia rate the spread



Notes: The horizontal axes corresponds to a logarithm of excess reserves in billion euros at monthly frequency. The vertical axes are spreads in basis points (left panel) and a logarithm of this the spread (righ panel). The linear fit corresponds to linear regression with a constant. The sample covers 2009M01-2019M08.

the exact functional form of function $g(\cdot)$ linking IIR_t and excess reserves. Figure 1.7 suggests that the effect of additional excess reserves should be decreasing in the amount of already accumulated excess reserves, indicating a concave function. A logarithmic function is a natural choice. Indeed, the left panel of Figure 1.8 shows that when we display a scatter plot of the Eonia rate the spread vs the logarithm of excess reserves, we achieve a reasonably good linear fit.

While using a logarithmic function ensures that the effect is concave, it does not ensure that the IIR does not go below the DR. Both theory and international experience reviewed by Bowman, Gagnon, and Leahy (2010) strongly suggests that IIRs in the eurozone should not decrease below the DR. This would suggest that the specification above is globally invalid and hence potentially misspecified as it allows for such possibility for high enough values of excess reserves. To obtain a fully correct specification we would have to allow the effect of additional excess reserves to approach zero not only with the level of excess reserves, but also with the size of the spread between the IIR and the DR. One simple way to do this, suggested by the superior fit of the right panel of Figure 1.8, is to use a logarithmic transformation of the spread between the IIR and the DR in excess reserves regime as the dependent variable, yielding following specification

$$f(IIR_t) = \begin{cases} IIR_t - MRR_t = \beta_{10} + \epsilon_t & \text{if } D_t^{ER} = 0 \\ log(IIR_t - DR_t) = \beta_{20} + \beta_{22} g(ER_t) + \epsilon_t & \text{if } D_t^{ER} = 1 \end{cases} \tag{1.4}$$

In this specification the dependent variable is either the spread between the Eonia rate and the MRR, or *log-spread* between the Eonia rate and DR. I use both equation (1.3) and (1.4) in the following sections.

1.3.2 Euribor rates

While the primary focus of this chapter is on the Eonia rate, I also provide indicative results for IIRs with longer maturity, i.e. the Euribor rates. There are two main differences between the Eonia rate and the Euribor rates that should be reflected in any model. First, while the overnight Eonia rate should (mostly) reflect only the current value of policy rates, rates with longer maturity should reflect *expectations* about *future* policy rates (and other policy instruments), in accordance with the expectational hypothesis. Second, the longer maturity of these rates also means that their risk component explains a much larger fraction of the variation in these rates: over longer periods, there is a greater chance that liquidity or credit risk will materialize.

In light of these considerations, can models (1.3) and (1.4) be used for the Euribor rates? If one were to use these models also for Euribor rates, one would be omitting the two expectations and risk components from the model. Effectively, one would be modeling the long-run (steady state or equilibrium) component of the Euribor and ignoring the transitory components related to expectations and risk components. From the perspective of estimation, there is a risk that this will lead to biased coefficients due to the omitted variable bias. Nevertheless, this is the approach I follow in the main text; meanwhile, in Appendix 1.B.3, I show that when I use proxy variables for the two missing components, the results are virtually unchanged.[12] This is not surprising given that in the main sample of interest there were no fluctuations in the risk component and almost no fluctuations in the component related to expectations of future changes in policy rates. Moreover, even if these components were present, not including them in the model does not have to be problematic from the perspective of the main empirical question of this chapter; only if these components were systematically correlated with excess reserves would this cause bias in relevant coefficients. I discuss this issue further in the following subsection.

1.3.3 Potential sources of bias

The modelling approach presented above warrants further discussion of potential sources of bias in coefficient estimates. Here, I address two such potential sources, (i) endogeneity of excess reserves and (ii) endogeneity of policy rates. Meanwhile, in Appendix 1.B.3, I also discuss two more potential sources: (iii) the absence of expectations of the future path of policy rates and (iv) the use of the proxy for financial market stress.[13] These other two concerns are relevant for IIRs with longer maturity but unlikely to *significantly* affect the Eonia rate given its 1-day maturity.

Before the discussion it is important to point out that the interest of this chapter lies in effect of changes in the IIR due to changes in excess reserves resulting from the QE program. This means that even if *some* coefficients are biased, the conclusions would remain intact. To see this, consider defining the effect of the

[12]The appendix also explains why using overnight index swaps as proxy for policy rate expectations is not viable option. Correspondingly I use actual realized future policy rates as a proxy.

[13]The appendix first argues that these issues are unlikely to invalidate the empirical results and then shows that when I control for them the results are indeed unchanged.

QE (EQE) as the difference in the IIR prevailing with and without QE:

$$
\begin{aligned}
EQE_t \equiv IIR_t^{QE} - IIR_t^{NOQE} = & \qquad (1.5)\\
= \Big\{ (1 - D_t^{ER,QE}) \left[\beta_{01} + \beta_{11} MRR_t\right] + D_t^{ER} \Big[\beta_{02} + \beta_{12} DR_t + \beta_{22} f(ER_t^{QE})\Big] \Big\} - &\\
- \Big\{ (1 - D_t^{ER,NOQE}) \left[\beta_{01} + \beta_{11} MRR_t\right] + D_t^{ER} \Big[\beta_{02} + \beta_{12} DR_t + \beta_{22} f(ER_t^{NOQE})\Big] \Big\} &\\
= \beta_{02} + \beta_{12} DR_t + \beta_{22} f(ER_t^{QE} - \beta_{02} + \beta_{12} DR_t + \beta_{22} f(ER_t^{NOQE} &\\
= \beta_{22} \Big(f(ER_t^{QE}) - f(ER_t^{NOQE}) \Big) &
\end{aligned}
$$

where QE indicates values of variables under QE, while $NOQE$ indicates values of variables without QE. The first line follows from defintion of the model in equation (1.3). The second-to-last equation follows from assuming that $D_t^{ER,QE} = D_t^{ER,NOQE}$, which amounts to assuming that a change in excess reserves due to the QE did not lead to a change in regime. This is indeed the case for all the results presented in following sections. Finally, the last equation follows from assuming that the coefficients capturing the relationship between IIRs and monetary policy variables is unchanged by the sole presence of the QE program, which is supported both by theoretical considerations and empirical evidence presented in Appendix 1.F. The main takeaway is that all coefficients except for β_{22} are eliminated in the comparison, so that we should be concerned only about factors that bias β_{22}; bias in any other coefficient will not influence our estimate of the effect of the QE on IIRs.

Endogeneity of excess reserves. Equations (1.3) and (1.4) include excess reserves as an independent variable, indicating that excess reserves affect interbank interest rates. However, it is plausible that there is also a reverse relationship: the amount of excess reserves could also respond to the level of IIRs. This is because excess reserves are *partly* determined by the actions of commercial banks via their borrowing from a central bank in refinancing operations. Since obtaining liquidity from central bank or from other commercial banks are partly substitutes, then a decrease in the Eonia rate should lead to lower demand for refinancing from the

ECB and hence lower amount of excess reserves. In econometrics terminology this implies that the IIRs and excess reserves form a simultaneous equation system, creating the possibility of biased coefficient estimates.

This is the main reason why in estimation I primarily focus on the QE program period: in this period it was the actions of the ECB that were clearly the main driver behind the evolution of excess reserves, weakening the potential for simultaneity bias driving my results. Moreover, when estimating the equations I eliminate this possible source of endogeneity so that I eliminate this as a potential driver behind my results. Finally, note that the simultaneity bias in the QE period is likely to be positive - excess reserves should be lower when the IIRs are lower - which means that if true coefficients in (1.3) and (1.4) are negative then the estimated coefficients will be attenuated towards zero. Hence, finding negative coefficients is likely to occur *despite* the estimation bias.

Endogeneity of policy rates. Another potential source of estimation bias is the reverse causality between IIRs and the ECB policy rates: It is possible that outside shocks to IIRs could prompt the central bank to adjust its monetary policy rates. In such a situation, the coefficient estimated from the model linking IIRs and policy rates could be biased. However, as argued above, the only coefficient that will be important in our empirical analysis is the coefficient on excess reserves, which is unlikely to be influenced by this possible endogeneity of policy rates with respect to IIRs, especially in the main estimation sample covering the period of the QE program. During that period, excess reserves were continuously increasing more-or-less according to the pre-determined profile of asset purchases, making correlation with any non-trending factor statistically impossible. Moreover, if such correlation is present it is most likely positive, and hence would cause upward bias in our coefficient of interest and correspondingly smaller estimated effects of excess reserves. Hence, as above, finding negative coefficients is likely to occur *despite* the estimation bias.

There is another possible, more complex relationship between IIRs and policy rates, resting on the endogeneity of monetary policy instruments to macroeconomic developments. The stance of monetary policy, and hence the values of its instruments, responds to macroeconomic developments, and specifically to the inflation outlook. Insofar as it is IIRs that influence these macroeconomic developments, one could imagine that shocks to IIRs influence monetary policy instruments, including the size of balance sheet, indirectly. Further, in this case, the resulting correlation and bias are unlikely to be strong, and if anything more likely to be positive.

1.4 Estimation

In this section, I discuss estimation results for models (1.3-1.4) together with their fit. I discuss results from two different estimation exercises, corresponding to two estimation samples. First, I present the estimates based only on the period of the QE program, during which the variation in excess reserves was, to a large degree, reflecting actions of the ECB, rather than the actions of commercial banks. Hence excess reserves are closer to being exogenous with respect to the dependent variable, as I discussed above. Then I present results estimated on the whole sample. Here I focus only on the non-linear model (1.4) for brevity. The discussion is more detailed in the case of the Eonia rate, as that forms my primary focus, and relevant conclusions also apply to Euribor rates.

In the main text I use monthly frequency of the data for expositional convenience.[14] There are two reasons for this choice of frequency. First, the purchases of assets under the QE program were targeted to be constant at monthly, not weekly, frequency, leading to higher volatility of excess reserves when viewed at weekly frequency. Second, the focus of this chapter is on the effect of medium- and long-term developments in excess reserves, and hence monthly frequency corresponds more closely: in addition to more volatile values for excess reserves, the IIR also feature significantly more volatility at weekly frequency. It is unlikely that the

[14]The results for other frequencies are presented in Appendix 1.B. They are qualitatively and quantitatively not different from results at monthly frequency.

short-term volatility in these series is related to the QE program.

1.4.1 Robust estimation sample

I start with estimation results for a sample in which excess reserves varied mostly for exogenous reasons and hence coefficient estimates should be most robust. This is the sample covering the period of asset purchases under the QE program, March 2015 until August 2019, yileding 56 observations. Since the QE program was initiated in a period when there was already a large amount of excess reserves, estimation of models based only on this period have a single regime, excess reserves regime. During this regime equations (1.3) and (1.4) simplify to the following:

$$IIR_t = \beta_0 + \beta_1 DR_t + \beta_2 log(ER_{t-1}) + \epsilon_t \tag{1.6}$$

$$log(IIR_t - DR_t) = \beta_0 + \beta_2 log(ER_{t-1}) + \epsilon_t \tag{1.7}$$

I refer to equation (1.6) as a linear model and to equation (1.7) as a non-linear model. Two comments are in order. First, I used the lagged value of excess reserves. Since theory does not provide strong suggestions about the timing of the effect, I experimented with three different specifications of the timing of this regressor - concurrent, lagged, and both - and chose the lagged timing as it led to the overall best model fit across various estimation methods. Appendix 1.C shows that the results are broadly unchanged when different timing is used. Second, I allow ϵ_t to follow a general ARMA process when applicable.[15]

I use several different estimation methods to address potential issues with the estimation strategy and to highlight the robustness of the results. I start with simple OLS, therefore ignoring the (potential) endogeneity of excess reserves and

[15]The ARMA structure is selected based on model fit criteria and formal statistical tests for the presence of autocorrelation in residuals; the reported results are virtually unchanged when alternative specifications are used: see Appendix 1.C. The appendix also contains results for an alternative to allowing ϵ_t to follow an ARMA process: inclusion of a lagged dependent variable. The model with ARMA errors always had a better fit and the conclusions when using lagged dependent variable are unchanged.

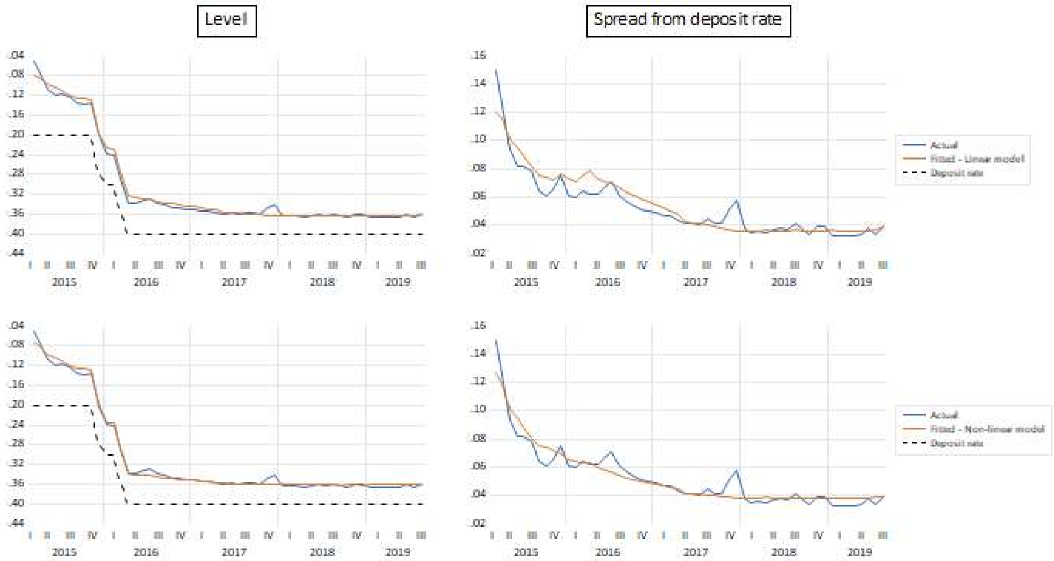
non-stationarity of the series in the regression. Next, given the potential bias in OLS coefficient estimates due to endogeneity of excess reserves, I turn to a TSLS estimation. Specifically, thanks to the fact that QE purchases are pre-determined to occur at fixed rates over prolonged periods of time, I can use the balance of assets purchased under the QE program as an instrument for the amount of excess reserves. Moreover, since asset purchases are the main driver of excess reserves over the period, they are very strong instruments.

Both OLS and 2SLS effectively ignore the time-series nature of the series under analysis. Most importantly, the methods ignore that all three series are likely to be non-stationary. Indeed, key added value of this chapter is to argue that the spread between the Eonia rate and the deposit rate (DR) depends on excess reserves, a series that is clearly non-stationary. This on its own implies non-stationarity of the Eonia rate, which is confirmed by formal statistical tests, see Appendix 1.F.2. At the same time, the model postulates that, after accounting for excess reserves, the spread between the Eonia rate and the DR is stationary, or in other words, that excess reserves are the only source of non-stationarity. In econometrical terminology, the Eonia rate, the DR and excess reserves form a co-integrating relationship. Again, I confirm this using formal statistical tests in Appendix 1.F.2. Given this conclusion, I next turn to estimation methods based on cointegration analysis. I present the results from 2 alternative cointegration estimators:[16] First, I use the autoregressive distributed lag approach of Pesaran and Shin (1998). Second, I use the non-parametric approach of fully modified OLS (FMOLS) proposed by Phillips and Hansen (1990).

Before turning to results it is worth emphasizing that, in the view of the cointegration, the OLS estimates discussed before are valid even if excess reserves are endogenous: since equations (1.6) and (1.7) are effectively cointegration relationships, then the associated coefficients are super-consistent, which ensures consistency even in the presence of endogeneity.

[16]Results from other estimators are in Appendix 1.C. The estimated coefficients are virtually unchanged.

Figure 1.9: Model fit - single-regime models



Notes: The fitted values are based on equations (1.6) and (1.7), respectively referred to as linear and non-linear models. Both equations have been estimated using the FMOLS method with results reported in last column of Table 1.2 and 1.3, respectively.

Table 1.2 presents the estimation results from all estimation methods for equation (1.6) and Table 1.3 presents results for equation (1.7). Meanwhile, figure 1.9 shows the fit of the equations in terms of the level of Eonia rate and in terms of the spread from the DR. Since the estimated coefficients are very similar across the methods presented in Table 1.2, as well as other methods presented in Appendix 1.C, I show figures and provide discussion of results only for one representative estimation method, FMOLS, which, out of the two robust cointegration regression methods, has a better model fit as demonstrated by higher R^2. The qualitative conclusions are unchanged and quantitative conclusions are only mildly different.

Both linear and non-linear model have a very good fit, which can be seen both in the tables and the figures. First, the models are able to explain almost all of the variation in the level of the Eonia rate, as evidenced by the left panels in Figure 1.9. The Eonia rate is predicted to decrease rapidly over the first year of

Table 1.2: Estimation results for the Eonia rate - linear model

	Method			
Coefficient	OLS	TSLS	ARDL	FMOLS
β_0 (constant)	0.300*** (0.021) [14.436]	0.240*** (0.019) [12.381]	0.226*** (0.042) [5.391]	0.337*** (0.018) [18.484]
β_1 (DR_t)	0.947*** (0.059) [16.161]	0.917*** (0.025) [36.389]	0.938*** (0.043) [21.742]	0.902*** (0.034) [26.346]
β_2 (ER_{t-1})	-0.038*** (0.006) [-6.654]	-0.031*** (0.003) [-9.785]	-0.028*** (0.007) [-3.953]	-0.045*** (0.004) [-11.305]
Observations	56			
ARMA/Lag structure	MA(2)	MA(1)	(3,1,2)	-
Model R^2	0.996	0.997	0.983	0.991
MRR-spread R^2	0.992	0.995	0.973	0.986
DR-spread R^2	0.907	0.946	0.707	0.844

Notes: The table shows point estimates, with standard errors in round brackets and corresponding t-statistics in square brackets. The standard errors are heteroskedasticity and autocorrelation consistent due to use of Newey-West standard error estimator, where applicable. *,**,*** indicate significance at 10%, 5% and 1%. ARMA/lag structure indicates the order of ARMA errors or the lag structure of the ARDL model (number of lags of the dependent variable, followed by the number of lags of independent variables). For the ARDL model I report results for the implied long-run form corresponding to the cointegration relationship. Order of ARDL model is selected according to Bayesian information criterion with maximum 4 lags for both dependent and independent variables.
See footnote 17 for details about R^2 measures.

Table 1.3: Estimation results for the Eonia rate - nonlinear model

	Method			
Coefficient	OLS	TSLS	ARDL	FMOLS
β_0 (constant)	0.538*** '(0.241) [2.232]	0.581*** '(0.237) [2.449]	0.069*** '(0.337) [0.206]	0.390*** '(0.256) [1.520]
β_1 (ER_{t-1})	-0.506*** '(0.035) [-14.618]	-0.513*** '(0.034) [-15.072]	-0.444*** '(0.047) [-9.447]	-0.487*** '(0.036) [-13.337]
Observations	56			
ARMA/Lag structure	MA(1)	MA(1)	(1,0)	-
Model R^2	0.888	0.888	0.985	0.886
MRR-spread R^2	0.991	0.991	0.977	0.99
DR-spread R^2	0.9	0.901	0.762	0.89

Notes: See notes under Table 1.2 for explanation of the values in the table.

the sample and more gradually after. The majority of the initial decrease can be explained by the two decreases in the DR that occurred in December 2015 and March 2016. While this is, in hindsight, obvious effect, a model which would include only the MRR, as is customary, would not be able to explain this effect, which highlights the first contribution of this chapter (see discussion below for more detail). Meanwhile, the right panels in Figure 1.9 show the importance of the second contribution of this chapter: the role played by the *amount* of excess reserves. While without excess reserves the spread between the Eonia rate and the DR would be (almost) constant, which is at odds with reality: the actual model predicts a large gradual decrease in the spread, corresponding to the increase in excess reserves. This increase in excess reserves is driven by the QE program, a topic of the next section.

I also document the very good model fit by three quantitative measures: model R^2, as well as R^2 in terms of the spread from the MRR and the DR, respectively.[17]

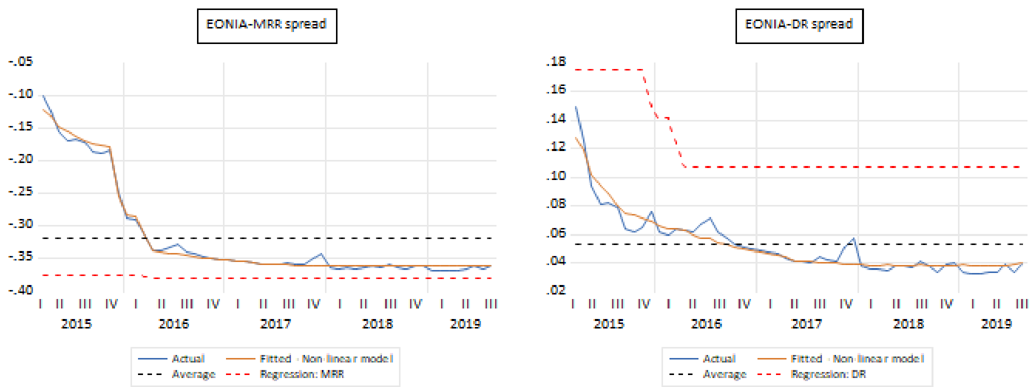
[17]The three different measures of model fit are calculated as follows:

Not only is the linear model able to explain 99.1% of the variation in the level of the Eonia rate, but it is also able to explain 98.6% of the variation in the spread of the Eonia rate from the MRR and 84.4% variation in the spread from the DR.

To highlight both contributions of this chapter in terms of modelling the Eonia rate, Figure 1.10 compares the fit of the non-linear model in terms of the spread between the Eonia rate and policy rates with simple (and common) alternatives. The left panel demonstrates that using a simple model based on the MRR (rather than the DR) estimated on the whole available sample would lead to almost no variation in the predicted Eonia rate and correspondingly to a bad fit. Another alternative - using the average value computed over given sample - would of course lead to a better fit, but not substantially. Meanwhile, the right panel shows that using the DR (rather than the MRR) would lead to some variation in the Eonia rate and a significantly better fit, as argued by this chapter and hence highlighting the first modelling contribution. Still, the fit would be rather poor since the model does not include excess reserves and, as such, the spread from the DR changes

1) The model R^2 is the coefficient of determination of the estimated equation. This can be interpreted as a share of the variance of the *dependent variable* explained by the model; both its structural and non-structural parts (Briefly, the structural part refers to the equation without ARMA errorrs.) This is not comparable between linear and non-linear models, as they have different dependent variables, or between models with and without dynamics terms, as those artificially increase R^2.
2) The MRR-based R^2 is normalized so that it is comparable across models with different dependent variables. Specifically, for each model I compute the predicted values of the Eonia rate based on the structural part of the model. I then compute the implied spread from the MRR in order to eliminate the variation originating in the MRR. This is done for two reasons: First, variation in policy rates can explain the vast majority of variations in the Eonia rate, and hence not accounting for it would result in very high R-squared, lowering its information value. Second, constant spread between the Eonia rate and the MRR would be a reasonable baseline model. Finally, I compute the total sum of squares (TSS) and residual sum squares (RSS) as well as the corresponding coefficient of determination: $R^2 = 1 - \frac{RSS}{TSS}$. This can be interpreted as the fraction of the variation in the spread between the Eonia rate and the MRR explained by the model.
3) The DR-based R^2, which is calculated analogically to the MRR-based one, but uses the DR instead of the MRR. Correspondingly, the value can be interpreted as the fraction of the spread between the Eonia rate and the DR explained by the model. This metric can highlight the role played by the *amount* of excess reserves, as opposed to the role played by a regime switch from the MRR to DR. Note that ordering of the models is the same irrespective of whether the MRR- or DR-based R^2 is used, but potentially not when model-based R^2 is used.

Figure 1.10: Model fit - Comparison with simple alternatives

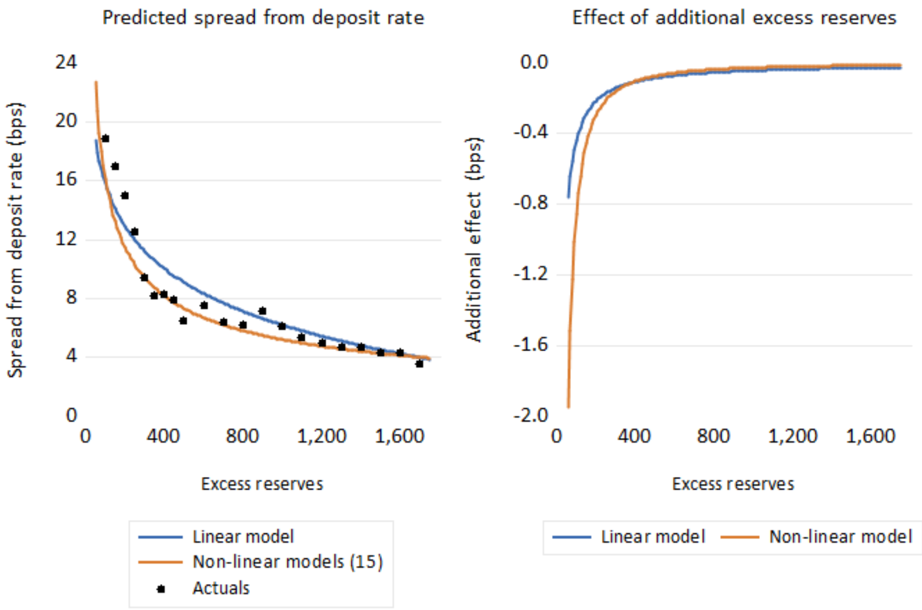


Notes: The fitted values are based on equation (1.7) estimated using the FMOLS method. "Average" refers to the average of the spread between the Eonia rate and the respective policy rate over 2015M03-2019M08 sample. The regressions are simple linear regressions of the Eonia rate on respective policy rate, based on the 2002M01-2019M08 sample, with the following coefficient estimates: $Eonia_t = -0.38 + 1.12 * MRR_t$ and $Eonia_t = 0.24 + 1.34 * DR_t$.

only with changes in the DR. This highlights the second modelling contribution of this chapter.

Turning to individual coefficients, the estimation results confirm the main hypothesis of this chapter: there is negative relationship between excess reserves and the spread between the Eonia and policy rates. The relationship is both highly statistically significant and economically important. To help interpret the coefficients on excess reserves, Figure 1.11 shows the predicted values of the spread from the DR (left panel) and the effect of an extra 10 billion euros in excess reserves on this the spread (right panel) for both models. The left panel makes it clear that the effect of excess reserves is substantial: in both models, the spread between the Eonia rate and the DR decreases from around 20 basis points (bps) to around 4bps as excess reserves increase from 50 billion to 1800 billion euros. Meanwhile, the right panel makes clear that the effect of excess reserves quickly diminishes: at 50 billion in excess reserves the estimated effect of extra 10 billion in excess reserves is almost 1bps for the linear model and almost 2bps for the non-linear model, but at 200 billion in excess reserves, this becomes 0.2bps and 0.3bps, respectively. At

Figure 1.11: Implied effects of excess reserves



Notes: The left panel shows the predicted the spread between the Eonia rate and the DR corresponding to different values of excess reserves. In case of model (1.7), the DR is set to -0.4%, as that was the prevailing rate for most of the sample under consideration. In addition, the graph includes several observed values of the spread taken from Figure 1.7.
The right-hand panel shows the effect of additional excess reserves on this spread: the predicted change in the Eonia rate as a consequence of 10 billion euro increase in excess reserves.

levels of excess reserves observed in 2018-2019 period, the effect is almost zero. The figure also highlights the difference between the two models: the non-linear model implies a larger effect at lower values of excess reserves and a smaller effect at higher values. This is what one would expect given the differences in the structure of the models.

As a final note, notice that the estimated coefficients are very stable across different estimation methods that take into account the possible endogeneity of excess reserves, which is true even for the significance of the coefficients. This suggests that the results are not driven by this potential endogeneity bias. Appendix 1.C.6 additionally shows that the results are also *not* driven by a potential alternative channel: the bond purchases under the QE program could lead to decreases

Table 1.4: Estimation results for Euribor rates

	Linear model		**Non-linear model**	
Coefficient	1W Euribor	3m Euribor	1W Euribor	3m Euribor
β_0 (constant)	0.316*** (0.024) [12.951]	0.599*** (0.054) [11.107]	3.069*** (0.633) [4.848]	2.843*** (0.745) [3.815]
β_1 (DR_t)	1.109*** (0.045) [24.478]	1.111*** (0.100) [11.106]	-	-
β_2 (ER_{t-1})	-0.034*** (0.005) [-6.315]	-0.064*** (0.012) [-5.413]	-0.926*** (0.090) [-10.234]	-0.724*** (0.107) [-6.801]
Observations	56	53	56	53
Model R^2	0.974	0.972	0.698	0.565
MRR-spread R^2	0.96	0.96	0.972	0.778
DR-spread R^2	0.627	0.79	0.738	-0.159

Notes: See notes under Table 1.2 for explanation of the values in the table. The sample is restricted to end in 2019M05 due to negative value of the spread in months after, likely reflecting expectations of policy rate changes that actually occurred in September.

in bond yields, which in turn could lead to decreases in interbank rates.

EURIBORs. To complement the results for the Eonia rate, I also briefly discuss results for Euribor rates with two selected maturities, 1-week and 3-month. Table 1.4 shows coefficient estimates using the FMOLS estimation method for both models.[18] The main takeaways are the same as for the Eonia rate: the models are able to explain a large share of the variation in the rates. Unsurprisingly, the R^2 for the spread from either policy rate is lower than in the case of the Eonia rate, reflecting the presence of expectations and stress components that I do not control for. Still, even for the 3-month Euribor the model explains more than three quarters of the variation in the spread from the MRR and more than half in the spread from the DR.

[18]Results for remaining methods and maturities are presented in Appendix 1.B and 1.C.

In terms of our main coefficient of interest, the estimates provide support to the main hypothesis: coefficients on excess reserves are negative and large as well as strongly statistically significant in all cases. Therefore, it is safe to conclude that the effect of excess reserves on interbank interest rates is not limited to risk-free overnight rates, but is also present for the more important longer-maturity risky rates. This is what one would expect, given that loans with different maturities are substitutes for each other.

1.4.2 Full estimation sample

The single regime models provided ample evidence to support the main hypothesis of this chapter: the presence of a negative relationship between excess reserves and the Eonia rate. However, the model does not apply for the whole sample, as argued in the previous section. Specifically, the relationship between excess reserves and the IIRs is present only when there is a large enough amount of excess reserves in the interbank markets. To determine the threshold and provide model estimates that are valid for the whole sample, I now turn to estimation of the two-regime model. For the sake of brevity I will focus only on the non-linear model, relegating results for the (worse-fitting) linear model to Appendix 1.B, and presenting results based only on the FMOLS estimation method. As before, the results are qualitatively and quantitatively unaffected. I present results for the Eonia and Euribor rates together, again focusing on the former in the discussion.

Before jumping to estimation results, it is important to point out that the estimation in case of samples spanning multiple regimes is complicated by the presence of multiple regimes. This means that one needs to specify which regime prevails in each period, which is further complicated by the fact that the threshold value is unknown and must be estimated. Fortunately, in the present case, characterized by single exogenous threshold variable, this turns out not to be a major complication: for the given threshold value, the estimation can proceed as if the threshold is known. Therefore, I follow Chan (1993) and estimate the equation for all plausible threshold values of excess reserves and select the threshold value leading to the smallest residual sum of squares (RSS).

Eonia. The first column of Table 1.5 presents the estimation results for the Eonia rate. The main takeaway is that the coefficient on excess reserves is broadly consistent with the value of the coefficient presented in Table 1.3: while based on the robust sample I obtain the coefficient -0.487, based on the full sample I obtain the coefficient -0.575, with a slightly increased standard error. Therefore, as in the case of single-regime models, the estimation results from the general model confirm the main hypothesis of this chapter.

As before, I report various measures of model fit, which are complemented by Figure 1.12. Its top row shows the fit of the model for the Eonia rate, both in terms of levels and the spread from the two policy rates, with yellow shading indicating the prevalence of an excess reserve regime. As before, the model is able to explain the vast majority of variation in the Eonia rate, as well as variation in the spreads between the Eonia rate and policy rates. While the former finding is not surprising, the ability to explain such a large share of variation in the Eonia rate *after accounting for variation in policy rates* is a strong endorsement of the model. Apart from the decrease in the spread from the DR in the QE program period discussed in the previous subsection, the model is also able to partly explain the variation of the spread from the two policy rates in the period from 2008M10 until the start of the QE program in 2015. Focusing on the spread from the MRR, the behavior in 2013-2014 nicely illustrates the first contribution of this chapter - the importance of anchoring the IIRs using the DR rather than the more common the MRR. With the MRR decreasing but the DR remaining unchanged during this period, the model predicts no change in the Eonia rate and, hence, its convergence to the MRR, almost exactly matching the actual behavior.

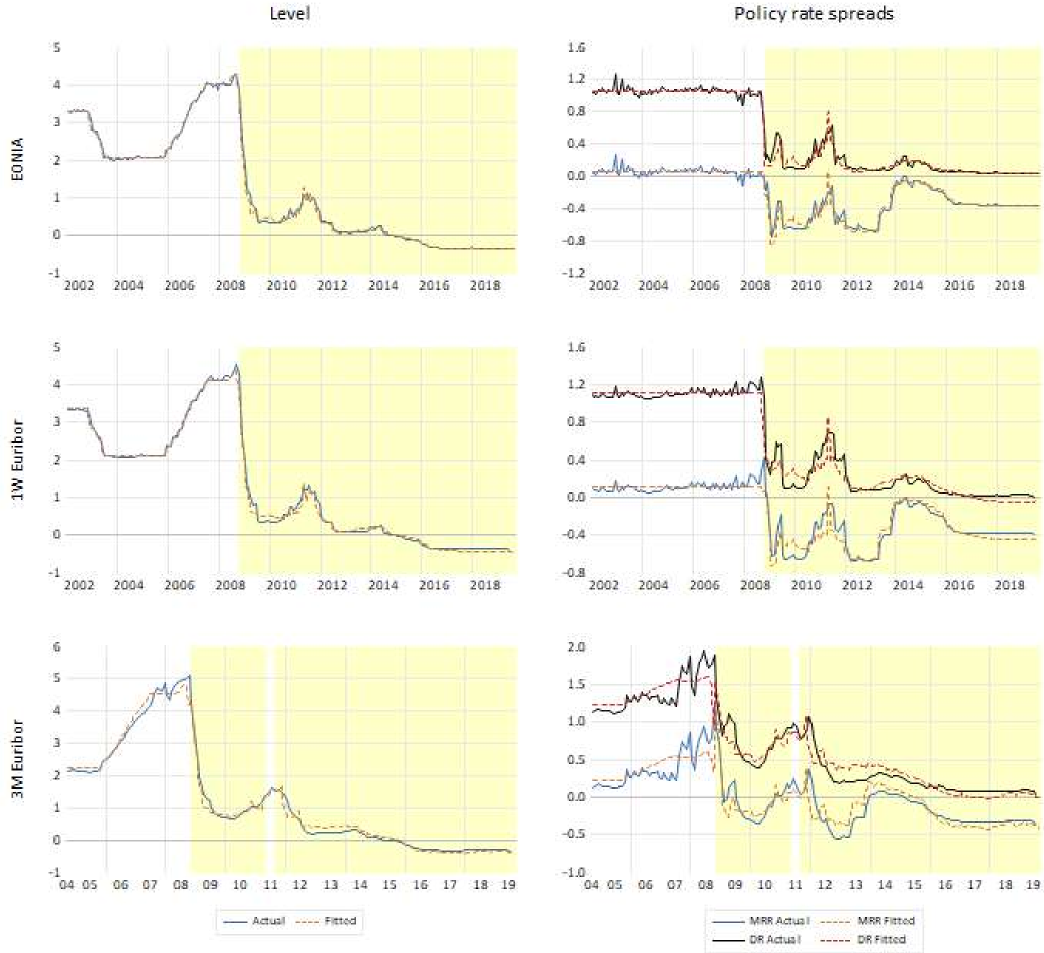
Meanwhile, looking at the behavior of the spread from the DR, the period from 2013 until the start of the QE program is another demonstration of the second contribution of this chapter - the link between the amount of excess reserves and the spread between the Eonia rate and the DR. During the first half of this

Table 1.5: Estimation results - 2 regime models

	Dependent variable		
Coefficient	Eonia rate	1W Euribor	3M Euribor
α_{01}	0.051***	0.102***	0.238***
(constant)	'(0.008)	'(0.012)	'(0.045)
(NR)	[6.702]	[8.211]	[5.296]
α_{02}	1.044***	2.804***	3.808***
(constant)	'(0.365)	'(0.840)	'(1.159)
(ERR)	[2.860]	[3.339]	[3.287]
α_{02}	-0.569***	-0.883***	-0.866***
(ER_{t-1})	'(0.060)	'(0.138)	'(0.185)
(ERR)	[-9.416]	[-6.395]	[-4.678]
Observations	212	209	209
Level R^2	1.00	1.00	0.98
the MRR-spread R^2	0.94	0.91	0.42
DR-spread R^2	0.98	0.97	0.78
Threshold	22.85	31.61	87.10

Notes: Estimation results, sample 2002M01-2019M08. See notes under Table 1.2 for the explanation of all values. NR refers to normal regime and ERR refers to excess reserves regime. Note that the MRR-based R^2 is in general lower than that of the DR-based R^2, reflecting smaller volatility of the former the spread in my sample due to higher values of the DR the spread in the pre-2008 period.
Note that the FMOLS method is applicable only to the excess regime equation, since the normal regime equation contains only stationary regressors. Therefore, the two regime equations are estimated separately, with the normal regime equation estimated using OLS with HAC standard errors. The two equations are then used to create combined fitted values for the IIRs on which the resulting R^2 is calculated.

Figure 1.12: Model fit (two-regime models)



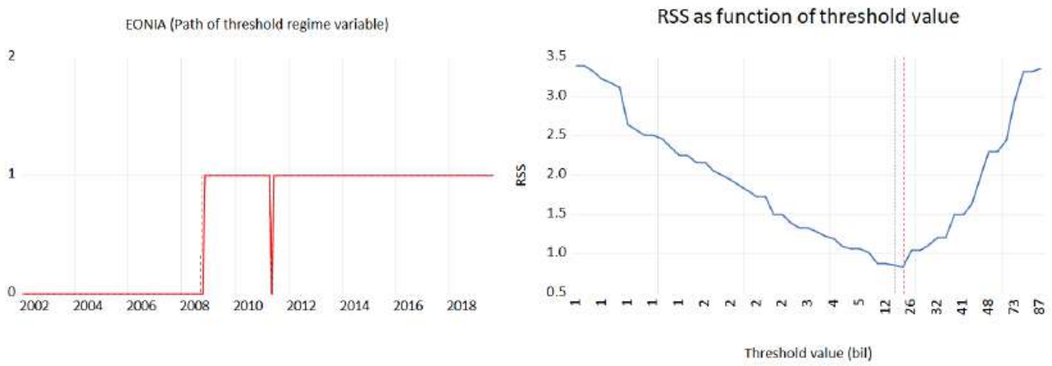
period, the amount of excess reserves was decreasing, as multiple policies that increased excess reserves were winding down. Meanwhile, during the second half of this period, the ECB started multiple (small-scale) asset purchase programs, leading to gradual increases in excess reserves. As a result, the model predicts first increase and then decreased in the spread from the DR, again capturing the actual developments.

To complement Table 1.5 and Figure 1.12, I also present basic results with respect to the regime estimates for the Eonia rate. The left panel of Figure 1.13 shows the estimated paths of the regime dummy variable, D_t^{ER}, while the right panel shows the RSS for different values of the threshold. The regime estimates are intuitive for the most part: The excess reserves regime prevails more-or-less continuously from October 2008, when the ECB switched to a full allotment policy and initiated other liquidity infusions, leading to large increase in excess reserves. The exception are a few months during 2011 when the normal regime prevailed, corresponding to almost complete evaporation of excess reserves due to winding down of various liquidity-infusing policies of th ECB. Meanwhile, the right panel shows that the RSS as a function of threshold values is concave almost everywhere, as desired: the fit first improves as the threshold for the excess reserve regime increases, but then it starts to worsen rapidly.

EURIBORs. As before, I complement the results for the Eonia rate with the results for Euribor rates based on models capturing only the equilibirum component. Columns 2-3 of Table 1.5 show estimation results for 1-week and 3-month Euribor rates, while the second and third rows of Figure 1.12 show the fitted values. As in the case of single regime models, the model fit for Euribor rates is somewhat worse than for the Eonia rate, but still very good. Similarly, the coefficient on excess reserves is statistically and economically significant.

In addition to confirming my conclusions from the Eonia rate analysis, the coefficient estimates for rates with longer maturity also shed light on one additional aspect of the data which was noticeable in Figure 1.2: there seems to be compression of the spread between the IIRs with different maturity, with rates with

Figure 1.13: Regimes estimates



Notes: The left panel shows the values of D_t^{ER} in each month corresponding to the value of the threshold corresponding to the lowest RSS (red line) and the second lowest RSS (dashed gray line). The right panel shows values of RSS for all considered threshold values, with the positions of the lowest and the second lowest indicated by vertical lines.

longer maturity converging closer to rates with shorter maturity. This is mirrored in my results in the fact that the coefficient on excess reserves is higher with longer maturity. This suggests that the increase in excess reserves not only lowers all the IIRs, but it also leads to compression of the spread between the IIRs with different maturity. Finally, having estimates for the IIRs with longer maturity (and hence a non-negligible risk component) allows me to highlight one additional feature of the data: that most of the variation in the spread between Euribor rates and the DR after 2009 can be explained by factors related to the ECB policy influencing excess reserves and hence does not reflect stress in financial markets. This can be clearly seen in the right panels for both Euribors, which show that the model is able to explain the variation in the spread from the DR throughout most of the sample. Since this variation reflects only excess reserves, then I can conclude that the ECB policy explains most of this variation. This is especially interesting in the case of period covering the height of Eurozone sovereign debt crisis in 2011-2012, a topic I discuss at greater length in an unpublished companion paper (Kovar 2020).

1.4.3 Contribution and links to existing literature

Before turning to the main empirical question of this chapter, it is useful to highlight the contribution of this chapter so far. The first top line contribution focuses on the link between interbank interest rates and excess reserves. While the link is well known in policy circles - see for example various post-2014 issues of Economic Bulletin from the ECB as well as other analysis from the ECB such as Boucinha and Burlon (2020) - the link had a much smaller imprint on academic literature. Among other things, this is reflected in misunderstanding the nature of the link, leading to problematic model specifications. This can take several forms, from relatively benign to very serious. At the relatively benign end of the spectrum Corradin et al. (2020) and Arrata et al. (2020) use excess reserves in their linear rather than logarithmic transformation. In contrast, the model of Bech and Monnet (2016) implies a logarithimic relationship similar to the one presented here.

A more serious misspecification centers on the time series transformation of the dependent and independent variables. The presented model links levels of IIRs to levels of independent variables, something that is justified by the cointegration nature of the relationship between the variables. In contrast, some authors use different time series transformations. Arrata et al. (2020) link *changes* in money market rates (MMRs) to *changes* in excess reserves, which amounts to model misspecification given that the variables are cointegrated.[19] Even more problematic is the model of Corradin et al. (2020), which links the *level* of MMRs to *changes* in excess reserves. Unsurprisingly, the authors do not find a significant relationship between the MMRs and excess reserves, which is of course because MMRs depend on the *level* of excess reserves.

The most problematic is the empirical analysis of Ogawa (2007), who estimates a regression linking banks' excess reserves holdings and overnight IIRs on a sample of the original Japanese QE program, claiming to show that banks increased their

[19] However, in their case this treatment might be partially justified by the focus on money market rates, which might not be fully cointegrated with the monetary policy variables, and by the focus on the cross-section dimension of their dataset.

holdings of excess reserves because of low IIRs. Similarly, critiques of interest rate on reserves set by the Fed often argue that it lead banks to demand excess reserves, see e.g. Selgin (2016) or Beckworth (2018). Of course, this chapter argues that the relationship runs the other way around, in line with the notion that the aggregate amount of excess reserves are determined by the central bank, as is widely accepted in policy circles, e.g. Bernanke (2013).

The second key contribution focuses on the regime-switching nature of the relationship between IIRs and the two main policy rates, and specifically the role of excess reserves in changing the anchoring variable from the MRR to the DR. While this regime-switching nature has also been acknowledged in the literature - for example, Ryan and Whelan (2021) discuss the multiple regime nature of monetary policy in general, while Boucinha and Burlon (2020), Rostagno et al. (2019) and Altavilla et al. (2021a), among others, highlight that in presence of excess reserves IIRs are anchored by the DR - existing works do not focus on this nature and hence do not explore it in detail. In contrast, to my knowledge, this chapter is the first work to provide systematic evidence for the regime-switching nature in the form the event analysis presented in section 1.2.3. For example, while Boucinha and Burlon (2020) highlight the (almost) complete pass-through of the DR decreases into negative values of Eonia rate, they do not distinguish between symmetric and asymmetric changes in the DR and the MRR, as this chapter does. Similarly, none of the existing literature translates this notion into asemi-structural model for IIRs.

Both the role played by excess reserves, and the regime-switching nature, also have a bearing on studies that analyze pass-through of changes in policy rates to bank deposit and lending rates (see e.g. the recent example by Altavilla et al. (2021a)). First, insofar as deposit and lending rates reflect the cost of bank funding - i.e. among else also IIRs - the pass-through of changes in policy rates could be confounded by the variations in excess reserves. This is especially problematic when the analysis is based on relatively few observations due to the focus on pass-through of negative DRs. For example, when a decrease in the DR is followed by an increase in excess reserves - which was indeed the case in 12 months after the first and second decreases in the DR below zero - then IIRs decrease more than the policy rates. If deposit and lending rates also decrease in line with IIRs, then the analysis would wrongly ascribe the decrease to effect of the change in the DR, rather than to the increase in excess reserves.[20]

Moreover, this issue is not easily addressed by using a longer sample in the analysis of the pass-though. In such a situation, the regime-switching between the MRR and the DR anchoring IIRs will cause problems: throughout the whole sample not a single policy rate will be an appropriate measure of changes in policy stance. This discussion suggests that one needs to study pass-through from policy rates to IIRs, accounting for excess reserves, and then pass-through from IIRs to deposit and lending rates, something I plan to address in future research.[21]

The last key contribution lies in using the postulated and estimated model to explain the movements in IIRs over the decade since the start of the financial crisis. Insofar as understanding these developments is important, the explanation of their behavior should provide a valuable service to users of IIRs time series. To see this, consider two examples. First, among other things the model highlights how the decreases in the MRR during 2013 were largely irrelevant for IIRs. While this view is now accepted - for example Rostagno et al. (2019) states that the decreases were motivated by the link between MRR and the cost of *existing* loans of commercial banks from the ECB, not by their effect on IIRs - early analyses put too much focus on these changes (see for example Hartmann and Smets (2015)).

The second example concerns the quantitative importance of excess reserves in driving IIRs. Boucinha and Burlon (2020) make the observation that short-term (OIS) rates decreased by more than the DR between June 2014 and the end of 2019. The authors link this to the fact that *"the gap between the EONIA and*

[20] Similar considerations apply to the question of pass-through *speed*: the estimates of speed of pass-through could be heavily influenced by the changes in excess reserves.

[21] Altavilla et al. (2021a) control for excess reserves at the bank level when studying pass-through to deposit and lending rates. This, however, accounts for cross-sectional variation, and not for the aggregate effects of excess reserves discussed here.

the DFR [DR] was somewhat larger in mid-2014 than it is now [the end of 2019]". While this is a correct description, it misses the key point: the IIRs decreased more because of the increase in excess reserves. Section 1.2.2 has already highlighted how this increase in excess reserves was driven by the QE program, and thus the greater decrease in IIRs compared with the DR can be (partly) explained by the program, something I investigate quantitatively in the next section.

1.5 Quantitative easing and the IIRs

The model coefficient estimates presented in previous section firmly establish the negative relationship between excess reserves and the IIRs. In so far as quantitative easing policy leads to increase in excess reserves - something clearly visible from Figure 1.3 and explained in Appendix 1.A - this on its own establishes that the QE program led to decrease in IIRs. The remaining task that can be performed with these estimated models is to quantify this effect, in order to answer the main empirical question of the chapter: *What was the overall effect of the QE on the IIRs?*

To answer this question, equation (1.5) suggests one would need to proceed in three steps: (1) establish the effect of the QE program on amount of excess reserves, so that one can establish what would be would be the amount of excess reserves without the QE; (2) use this counterfactual amount of excess reserves to determine the counterfactul level of the IIRs; and (3) calculate the effect of the QE as the difference between predicted level of the IIRs with the observed level of excess reserves and the counterfactual level of excess reserves. However, it turns out that the first step - establishing the counterfactual amount of excess reserves without the QE - is in principle impossible to do with certainty given the available data. Fortunately, I can construct several plausible estimates that provide an indicative range of the effects of the QE on excess reserves and hence on the IIRs.

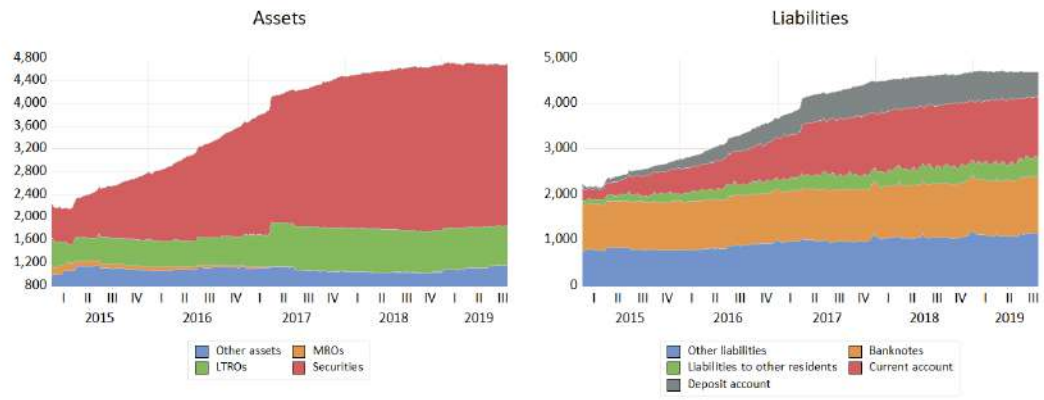
This section is divided into two subsections. The first subsection discusses the effects of the QE program on excess reserves and construction of a counterfactual path for excess reserves; the second then uses this couterfactual path to calculate the effects of the QE on the IIRs.

1.5.1 ECB balance sheet during the QE program

The starting point in establishing the effects of the QE on excess reserves is to take a closer look at developments in the ECB balance sheet during the QE program period. Figure 1.14 captures the evolution of the balance sheet, with the left panel capturing the asset side and the right panel capturing the liability side. In addition, Table 1.6 captures the size of the asset and liability components just before the start of the QE and at the end of the sample. In both cases the balance sheet is suitably simplified.

As can be seen, the ECB balance sheet has more than doubled over the period of the QE program, with most of the increase occurring between the beginning of 2015 and the end of 2017, after which the size of the balance sheet remained broadly unchanged. On the asset side, the vast majority of the change in the balance sheet is accounted for by the increase in securities held for monetary policy purposes, which leapt from 227 billion to 2.6 trillion, the result of various ECB asset purchase programs. The remainder is mostly accounted for by increases in lending to credit institutions. This development in top-level items masks an important aspect of its evolution: even though total lending to euro area credit institutions recorded only relatively small increases during this period, the two sub-components - main refinancing operations (MROs) and the longer-term refinancing operations (LTROs) - varied to a greater degree. Specifically, MROs became almost zero, while LTROs increased by 277 billion, resulting in an overall increase in lending. These two developments are likely related, since banks view liquidity obtained from MROs and VLTROs as substitutes (Vogel 2016). Nevertheless, the decrease in MROs can also be in response to the QE program - since the QE raises the reserves of commercial banks, there is overall less need for banks to obtain liquidity in the form of MROs. These conclusions are supported by the fact that MROs

Figure 1.14: ECB balance sheet developments



decrease both in discrete fashion around increases in LTRO balances, and also gradually over time. Finally, while other assets recorded some variations during this period, the developments were not out of ordinary.

The picture is more complicated in the case of liabilities, with several items recording substantial changes over the period. Most of the increase in total liabilities is accounted for by increases in the liabilities to credit institutions. This item effectively captures the deposits of financial institutions at the ECB and as such is equal to excess reserves. However, the increase in this balance sheet item is substantially smaller than both QE purchases and the overall increase of the balance sheet. In other words, this item explains only about two thirds of the increase in liabilities, in contrast to the asset side, where increases in securities account for 90% of the overall increase. This suggests that asset purchases under the QE program do not lead to equivalent increase in excess reserves.

The difference between the QE-implied increases in excess reserves and actual increases can be accounted for by the remaining liabilities equally. While increases in banknotes are not out of ordinary , the increase in liabilities to other residents and non-residents is likely related to the QE program. This can be seen in Figure 1.15, which shows that these items are fluctuating with excess reserves, and hence that the variation in other liabilities is causally related to the same forces, namely

Table 1.6: ECB balance sheet comparison

Assets

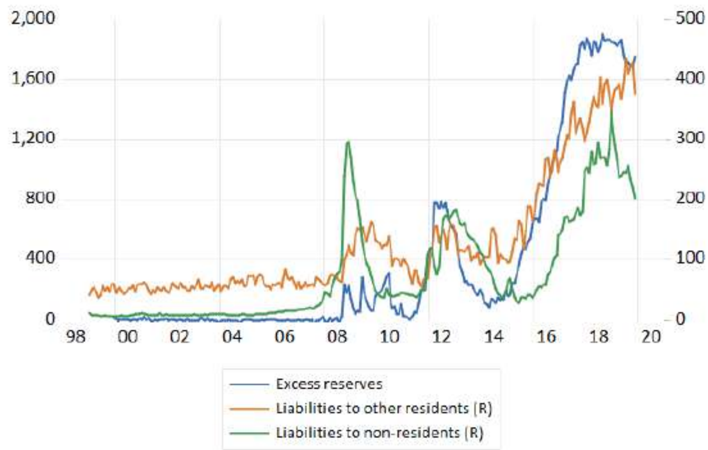
	30.1.2015	**30.8.2019**	**Difference**
Gold and gold receivables	**343,867**	**431,861**	**87,994**
Claims on non-residents	**273,726**	**347,900**	**74,174**
Claims on residents	**35,549**	**19,509**	**-16,040**
Lending to credit institutions	**579,646**	**695,654**	**116,008**
• Main refinancing operations	163,821	3,348	-160,473
• Longer-term refinancing operations	415,608	692,306	276,698
Other claims on credit institutions	**62,134**	**35,146**	**-26,988**
Securities of residents	**603,358**	**2,835,533**	**2,232,175**
• Securities held for mon. policy purposes	227,107	2,614,240	2,387,133

Liabilities

	30.1.2015	**30.8.2019**	**Difference**
Banknotes in circulation	**1,004,230**	**1,250,754**	**246,524**
Liabilities to credit institutions	**264,523**	**1,873,150**	**1,608,627**
• Current accounts	227,385	1,318,399	1,091,014
• Deposit facility	36,557	554,736	518,179
Liabilities to other residents	**111,448**	**415,267**	**303,819**
• General government	76,284	278,115	201,831
• Other liabilities	35,164	137,152	101,988
Liabilities to non-residents	**84,378**	**260,941**	**176,563**
Total assets/liabilities	**2,181,954**	**4,683,714**	**2,501,760**

Notes: Data from the weekly ECB balance sheet reports.

Figure 1.15: Liabilities to other residents



Notes: Data from the Statistical Data Warehouse.

the QE program. This is further supported by realizing that a negative DR also applies to these accounts, and so these accounts are treated the same as credit institutions' accounts.

What do these observations imply about the counterfactual path for excess reserves in the absence of a QE program? The discussion should make two things clear. First, the most straightforward approach - subtracting QE purchases from actually observed excess reserves - would lead to the nonsensical conclusion that excess reserves would be negative without a QE program, since excess reserves for most of the period are lower than QE purchases. Second, and as a result of this first point, it is obvious that the counterfactual path will have to be constructed based on assumptions on development in *multiple* components of the ECB balance sheets; this means that there are multiple plausible counter-factual paths for excess reserves which would prevail without the QE program, and that the final choice will always be arnitrary to some degree. I leave detailed discussion of the construction of alternative counterfactual paths to Appendix 1.D; here I will focus on two alternative counterfactual paths that are most plausible a-priori, and one of which provides reasonable upper bound for the counterfactual level of excess reserves and hence lower bound for the effect of the QE program.

One approach to construct a counterfactual path for excess reserves is to depart from the initial value of excess reserves and to simply assume that this amount of excess would change over the period of the QE program only in response to changes in LTRO balances:

$$ER_t^{NOQE} = ER_{2015W9} + (LTRO_t - LTRO_{2015W9})$$

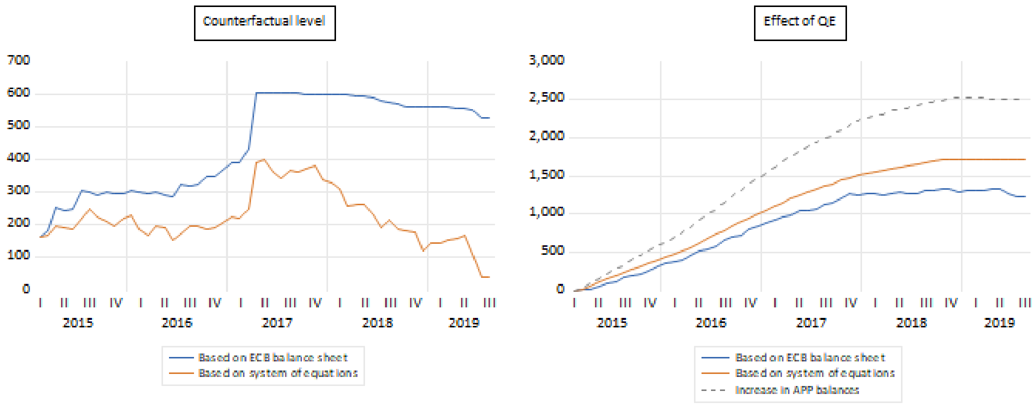
This alternative is appealing because it avoids the issues raised in the previous discussion, and especially by Figure 1.15: it does not require one to determine the value of other balance sheet items throughout the period, since it does not depart from observed excess reserves throughout the sample, but rather from the observed excess reserves before the start of the relevant sample. The exact opposite approach is to postulate and estimate amodel linking various balance sheet items in a multiple-equation framework.[22]

The two resulting counterfactual paths, and the corresponding effect of the QE program on excess reserves, are shown in Figure 1.16. Excess reserves under the first counter-factual path would generally increase in the first half of the sample, even though the increase is far from monotonous: there are several discrete jumps corresponding to successive rounds of TLTROs, especially the TLTRO program implemented in March of 2017, after which excess reserves are predicted to be just above 600 billion. The resulting estimated effect of the QE on excess reserves captured in the right panel increases gradually, reaching 1.2 trillion by the end of 2018, less than half of the increase in APP balances. The right panel also shows one drawback of this approach: the estimated effect of the QE program is quite volatile despite the fact that the increase in APP balances was smooth.

While relying on a multiple-equation model to produce the counterfactual path

[22]Briefly, the model has 4 endogenous variables - excess reserves, MRO balances, and the two types of other liabilities in Figure 1.15 - and 4 exogenous variables - APP balances, LTRO balances, banknotes, and other liabilities not-allocated. Excess reserves are treated as a residual category that responds to all other variables, while other endogenous variables respond to APP balances, LTRO balances and/or excess reserves.

Figure 1.16: Counterfactual excess reserves



Notes: See Appendix 1.D for the description of the construction of the time series.

for excess reserves leads to a path that is qualitatively similar, the path is substantially lower throughout and especially at the end of the sample. Under this alternative the excess would reach a maximum of 400 billion after the March 2017 LTRO, two thirds of the value of the first alternative. Moreover, after this jump, excess reserves would be expected to decrease gradually, eventually dropping below 100 billion. Correspondingly, the effect of the QE program on excess reserves reaches 1.7 trillion. In contrast to the first alternative, the alternative based on a multiple-equation model is supported by the smoothness of the effect of the QE on excess reserves throughout time.

What drives the substantial difference between the alternatives? The key realization is that the first alternative abstracts from developments in all other balance sheet items, which in their sum would likely imply lower excess reserves. For example, the alternative does not account for the observed decrease in MRO, which is at least partly related to LTRO balances, due to substitutability between these two categories. Even more importantly, multiple liability items increased over the period of the QE sample, and would do so even in the absence of a QE program, especially banknotes and other liabilities not allocated (as opposed to other liabilities to EU and non-EU insitutions). Their increase means that excess reserves would decrease, but by departing from pre-QE levels of excess reserves, the alternative effectively ignores this fact. All this together suggests that the alternative provides a conservative estimate of the effects of the QE on excess reserves. This is useful insofar as it can be used in the construction of a lower bound on the effects of the QE program on the IIRs. That said, most of the remainder of this section will focus on the second alternative based on a multiple-equation model.

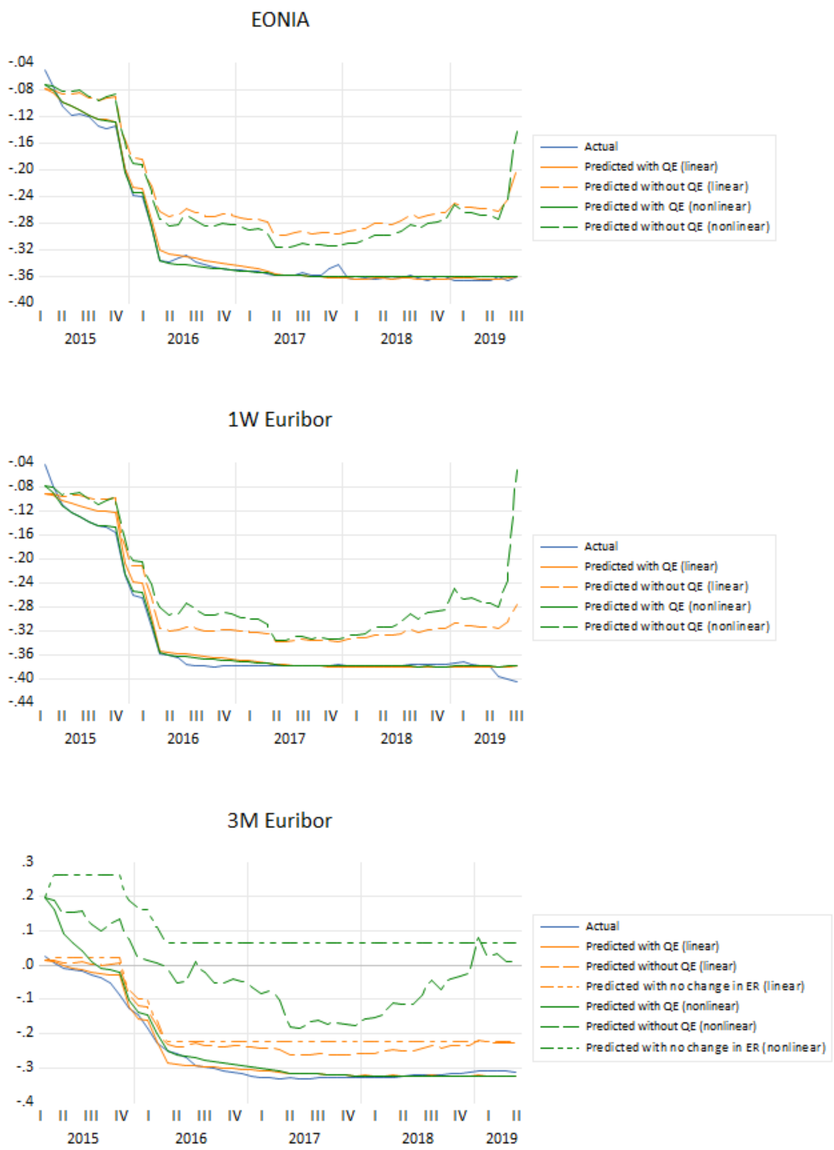
1.5.2 Effects of the QE program on the IIRs

In this subsection, I use the counterfactual path for excess reserves to forecast the counterfactual path for different IIRs and hence estimate the effects the QE policy had on these the IIRs. I focus on models estimated on a robust sample, and present results from both linear and non-linear models and for all three maturities.

Figure 1.17 provides the first look at the effect of the QE on the IIRs: solid lines are predicted values with observed excess reserves, while dashed lines are predicted values with counterfactual excess reserves; orange lines correspond to the linear model and green lines to the nonlinear model. As expected, in all cases the rates would be higher throughout the QE period in the absence of a QE program, reflecting the positive effect of the QE on excess reserves together with the negative relationship between excess reserves and the IIRs. This finding is even more significant in case of longer-maturity the IIRs: the third panel shows that not only would the longer-maturity the IIRs be higher, but the models predict that they would remain positive in the absence of a QE program during the first year of the program. The rate turned negative in April 2015, less than two months after QE purchases started; without these asset purchases, they would have remained positive until the further decreases in the DR at the end of 2015 and beginning of 2016. This suggests that, at mildly negative values of the DR, longer-maturity IIRs are negative only if a QE program (or other policy causing elevated excess reserves) is active.

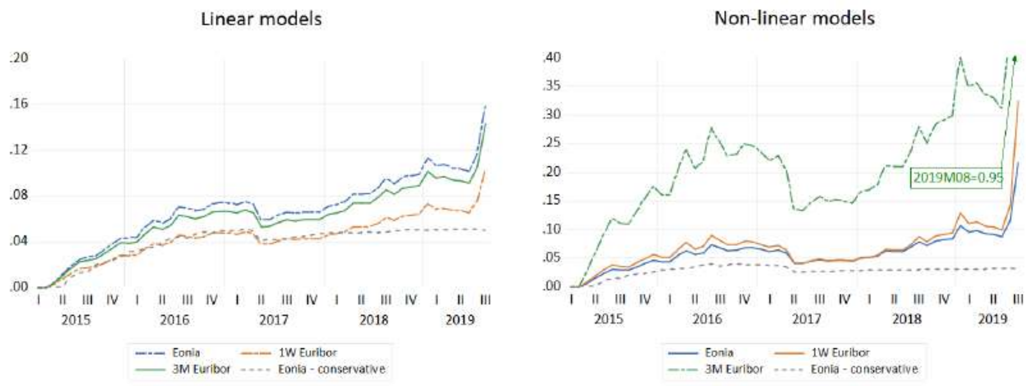
To explicitly quantify the effect, Figure 1.18 presents the full time path of the

Figure 1.17: Counterfactual forecasts



Notes: Predictions from models (1.3-1.4), coefficients of which are in tables 1.2-1.4. Predicted values with the QE are based on observed values of excess reserves, while those predicted without the QE are based on excess reserves shown in Figure 1.16.

Figure 1.18: Estimated effects



Notes: For details, see notes under Figure 1.17. Solid lines indicate better fit of the given (linear or non-linear) model.

differences between predicted values under the two alternative paths for excess reserves. As before I present results for both models and all 3 maturities, while indicating superior model fit with solid lines. In all cases, the effect gradually increases during the first 2 years of the program, before dropping in March 2017, corresponding to large increases in excess reserves due to VLTROs. However, since excess reserves were predicted to decrease afterwards, the effect starts to increase again. Moreover, as counterfactual excess reserves drop below 100 billion during 2019, the effect spikes. For the Eonia rate the peak effect is between 16-22bps, depending on the model, while the mean/median effect over the whole sample is 6-7bps. Figure 1.18 also shows the effect based on a conservative counterfactual path for excess reserves. Since under this counterfactual, excess reserves would be substantially higher, the effect, especially the peak effect, would be substantially lower, as suggested by the dashed gray line. That said, the effect is still somewhat economically significant, with an average of 4bps and peak of 5bps. Meanwhile, Appendix 1.E shows that the effects are also statistically significant .

The overall results are very similar for the 1-week Euribor rate, even though the peak effect based on the (better-fitting) non-linear model is estimated to be even higher than for the Eonia rate, at 33bps. Finally, in case of the 3-month Euribor, the non-linear model predicts a much higher effect on the average than for the other two maturities (22bps), and also very high peak effect; that said, these results should be taken with pinch of salt, since the likely-more-robust linear model predicts much smaller effects. Overall, one can conclude that the effect of the QE program on the IIRs was of the same order of magnitude as a reduction in policy rates of a standard 25 basis point size.

The drop in the effect size in March 2017 captured in Figure 1.18 suggests an important factor influencing the conclusions: the fact that, throughout the period, excess reserves were also increasing due to other the ECB policies, especially because of VLTROs. To illustrate this, dashed-dot lines in last panel of Figure 1.17 capture the level of 3-month Euribor if excess reserves would have remained at their average value for 2014. This shows that, in the absence of any change in excess reserves, the Euribor would be even higher than under the counteractual path for excess reserves in the absence of a QE program: under this alterantive counterfactual path the linear model predicts only mildly negative rates throughout the whole sample, while the non-linear model (implausibly) predicts positive rates all along. Therefore, it is only the combination of negative DR and the policy of large-scale liquidity infusions that produce (substantially) negative IIRs; without significant amounts of excess reserves the IIRs with longer maturity would likely remain either mildly negative or even positive. In terms of effect this would translate into an additional 5bps for the linear model. Taking the counterfactual path at face value, this suggests that the combined policies influencing excess reserves from the onset of the QE program lad to 21bps decrease in the 3-month Euribor by the end of the sample.

The above paragraphs demonstrate not only that the effect of the QE on the IIRs was substantial, but also that at the levels of DR actually observed, the IIRs turn substantially negative only because of such policy. This conclusion is further strengthened by realizing that the QE program was initiated in a situation in which excess reserves were already elevated due to multiple VLTROs and earlier small-scale asset purchase programs programs: not only would the effect be larger if initial excess reserves were lower, thanks to the concavity of the effect of excess reserves; but even more importantly, the model suggests that without excess reserves, the IIRs would remain positive irrespective of the value of the DR, since in a normal regime, the IIRs are anchored by the MRR, which remained non-negative.[23] Therefore, the two main policies of the ECB over the second half of previous decade - negative policy rates and large-scale asset purchase programs - are highly complementary as far as the effect on the IIRs is concerned.

1.5.3 Contribution and links to existing literature

There is a burgeoning literature dedicated to the effect of unconventional monetary policies on financial markets and the broader economy. Altavilla, Carboni, and Motto (2021) provide a recent broad overview of this literature and a helpful classification of impact channels into standalone channels and channels through which each instrument has influenced the transmission of the other channels. While the vast majority of existing literature focuses on the first category - and the majority of that part of the literature focuses on the standalone effect of only one policy, the QE program - authors such as Rostagno et al. (2019) and Altavilla, Carboni, and Motto (2021) stress that it is the interaction of channels that is at least as important as their standalone channels (see the quotes at beginning of this chapter).

How does this section contribute to this literature? The contribution can most clearly be seen in terms of he classification of Altavilla, Carboni, and Motto (2021): this section belongs squarely to the second category of channels, the interaction

[23]Arguably, in such a situation, the ECB could have achieved negative IIRs by setting the MRR to negative values and thus effectively paying commercial banks for borrowing from it, rather than charging them for excess reserves. Of course the difference is that, under the actual policy, the ECB achieved negative IIRs without incurring losses; the alternative might have been politically unfeasible: The policy of negative interest rates faced large opposition throughout the currency area, which would likely have been stronger if the policy were to entail losses for the ECB, and hence would in some sense entail transfers from the ECB to commercial banks.

channels, as it investigates the interaction between the QE program and negative policy rates. More specifically, this section is effectively quantification of the channel in cell 3.1 in Table 1 in Altavilla, Carboni, and Motto (2021), which is described as "[e]xtra liquidity [excess reserves] contributes to keeping overnight rate at DFR [DR]". The contribution of this section is then highlighted by the fact that Altavilla, Carboni, and Motto (2021) - or Rostagno et al. (2019), who also rely on this classification - do not provide any further discussion of this channel and correspondingly do not provide any references to existing research on this channel.

In a broader sense, this section also points to the limitations of existing empirical investigations of the unconventional monetary policy package, such as Rostagno et al. (2019) and Rostagno et al. (2021). In those works the authors construct couterfactual paths for IIRs in the absence of negative policy rates, the forward guidance and/or the QE program.[24] However, when investigating the effect of the QE program the authors rely on a large scale VAR model. Since this model does not include excess reserves, and hence the linkages between IIRs, excess reserves and asset purchases, then their analysis *assumes* that there is no effect of the QE program on the Eonia rate; see for example Figure 11.3 in Rostagno et al. (2021). Insofar as the results in this chapter (or the own words of the authors of that study) are taken at face value, this absence of the effect of the QE program on the Eonia rate is incorrect and hence their analysis underestimates the effect on the QE program on financial markets and the economy.

This limitation also points to the specific nature of the channel linking the QE program and IIRs. To see this, one can ask the question 'Why doesn't the empirical strategy of Rostagno et al. (2021), which relies on event analysis methodology, pick up the effect of the QE program on IIRs?'. The reason for this lies in what underpins the event methodology: the forward looking nature of financial markets

[24] To analyze the effects of negative policy rates and forward guidance, the authors rely on modifying distributions of *expected* IIRs teased from option prices for overnight index swap (OIS) rates. For example, for the counterfactual path of OIS rates in the absence of negative policy rates, they re-anchor the distribution at zero and reassign any negative values to zero. See further comments on this below.

(see e.g. Altavilla et al. (2019)). The way the effect of the QE program is typically studied is by analyzing the changes in yields on bonds with *medium- and long-term maturity.* If financial markets are forward looking then the announcement (or news) of the *future* purchases under the QE program will have an immediate effect on *current* bond yields, as is found in the literature (Altavilla et al. 2019; Rostagno et al. 2021).

The problem is that this methodology cannot capture the effect of QE on *short-term* IIRs. Since IIRs, and especially the Eonia rate, do not contain an expectations component, then one cannot rely on the *current* market reaction to identify the effect of *future* asset purchases.[25] The other way to see this is in terms of the classification of channels of QE programs into stock and flow channels. While the channel identified in this chapter is firmly stock channel, in its nature it is different from the regular understanding of the stock channel: it could be argued it is the *immediate* stock channel, as opposed to the *expected* stock channel, because it affects IIRs only once the asset purchases are made. Therefore, it lies in between the flow channel and the usual stock channel, which reflects the expected stock of purchases over the period of the QE program. Correspondingly, instead or relying on announcement effects one has to rely - equipped with the theoretical arguments discussed in 1.2.3 - on analysis of co-movement between asset purchase and excess reserves on one hand, and excess reserves and IIRs on the other hand.

Finally, while this chapter was focused on the effect of the QE program on IIRs, the estimated model could be used to answer related question: What would be the value of the Eonia rate in the absence of negative DR? Existing literature (Rostagno et al. 2019; Rostagno et al. 2021) typically assumes that in such a scenario the Eonia rate would be zero throughout the sample covering negative DR. Given that there is always small spread between the Eonia rate and the DR, this would be the case only if MRR (like the DR) would also be set to zero. Since

[25] Altavilla et al. (2019) use a related argument as part of their identification strategy for the QE channel. Note that this is also the reason one cannot use the monetary policy surprises from that work to analyze the effect of the QE on IIRs.

this is something the ECB was actively attempting to avoid (Rostagno et al. 2019), then such counterfactual scenario is unlikely. A much more likely scenario is that the MRR would remain somewhat above the DR, and hence the Eonia rate would remain above 0. I leave more detailed investigation of these issues for future research.

1.6 Conclusion

This chapter presents a novel model of interbank interest rates (IIRs), linking them to two main policy interest rates of the ECB and highlighting the role of excess reserves. The main feature of the model is its regime-switching nature, with the prevalance of each regime determined by the presence of excess reserves. Excess reserves not only influence which regime prevails and hence which policy rate acts as an anchor for the IIRs, but in the excess reserve regime, they also influence the spread between IIRs and the deposit rate. This suggests that policies that lead to large increases in excess reserves have the effect of lowering the IIRs, as I show for the case for the QE policy of the ECB. This policy caused the Eonia rate to be lower by up to 22 basis points, and other the IIRs by even more, and hence played a role in establishing negative IIRs.

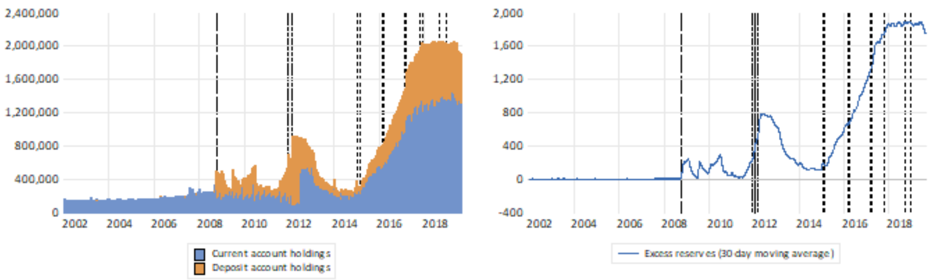
The focus of this chapter was on causal effects of excess reserves and of the QE policy. However, the model can be useful for other purposes, which is what I focus on in two companion papers. The second chapter of this thesis highlights the value of the model for forecasting the IIRs. Meanwhile, its unpublished companion paper focuses on disentangling developments in the Euribor rates during the 2010-2012 period. This period saw large movements in these rates, which are often mis-attributed to movements in the risk components. The paper shows that the majority of the movements in the Euribor rates during that period can be explained by fluctuation in excess reserves, which are to a large degree controlled by the ECB and hence not directly related to endogenous variations risk components. Future applications could focus on the role played by the regime-switch in the IIRs in forecasting other macroeconomic variables, such as GDP or the term structure of interest rates.

1.A Evolution of excess reserves

1.A.1 Excess reserves from 2008 onward

This appendix discusses in greater detail the evolution of excess reserves over the whole data sample.[26] Figure 1.19 reproduces Figure 1.3 from the main text displaying the current and deposit account balances at the ECB, and excess reserves over period from 2002 to 2019. It clearly shows that up until September 2008, the reserve holdings of banks were at all times roughly equal to the reserve requirements, with virtually all reserves being held in their current accounts and almost none in their deposit accounts. This is not surprising since the institutional set up of the euro zone monetary policy was forcing banks to economize on excess reserves.

Figure 1.19: Excess reserves



This changed dramatically after the fall of Lehman Brothers in September 2008. Following this event, the deposit account balances of commercial banks increased dramatically to more than 200 billion, while current account balances remained broadly unchanged. This increase was a combination of two factors. In reaction to ongoing stress in funding markets, the ECB switched from fixed allotment to full allotment procedure. This means that, instead of providing fixed amounts of financing in its refinancing operations, it allowed banks to borrow as much liquidity

[26]Rostagno et al. (2019) provide complementary discussion focused on monetary policy stance in general.

as they wished, to which banks responded by borrowing large amounts of extra liquidity. At the same time, the commercial banks were unwilling to lend this extra liquidity to other banks reflecting heightened liquidity and credit risks, and hence the banking sector overall had larger balances than before. This is reflected in the right panel, which shows a corresponding large jump in excess reserves in the fall of 2008.

During 2009 and 2010, excess reserves fluctuated wildly, with multiple decreases to near-zero values, followed by increases to 200-300 billion. These fluctuations mostly reflected changes in the total balance of liquidity obtained by banks in the form of LTROs and hence reflect the behavior of commercial banks. In addition to this asset side factor, excess reserves also changed, with fluctuations in other liability side items. For example, between the beginning and end of 2009, liabilities to non-euro area residents denominated in euro decreased by 250 billion. This highlights that excess reserves can fluctuate due to changes both on the asset side and the liability side of the EBC's balance sheet.

At the beginning of 2011, excess reserves stood at around 30 billion, decreasing all the way to 17 billion in April, which marks the lowest value in the period after October 2008. This reflected a return of refinancing to pre-crisis values: the sum of marginal refinancing operations and long-term refinancing operations stood at 421 billion, compared with a previous maximum of 843 billion reached in June 2010, and with 464 billion in August 2008. For the rest of 2011, excess reserves were increasing, first gradually and rapidly from August. The initial increase reflected higher demand from commercial banks during MROs. This was later complemented by increases in LTROs reflecting the reaction of the ECB to stress in funding markets related to re-intensification of stress in sovereign debt markets. The last factor behind the increase in excess reserves was increased asset purchases under the Securities Market Program (SMP), again in reaction to stress in government debt markets, even though corresponding to these, the ECB encouraged commercial banks to deposit some liquidity in the form of fixed term deposits.

As discussed in the main text, excess reserves recorded very large increases at the end of 2011 and the beginning of 2012, reflecting the two rounds of very long-term refinancing operations (VLTROs) and the lowering of reserve requirements. The first was executed in December 2011, and amounted to 489 billion euro, while the second was executed in February 2012, at 529 billion euro. Since these long-term loans to commercial banks were not sterilized in any way, and since the decrease in reserve requirements was not met by draining of excess reserves from the interbank market, these actions directly led to increases in excess reserves. Due to these factors, excess reserves reached almost 800 billion in May 2012, a multiple of the maximum values observed prior to that.

The following months and years were characterized by gradual decreases in excess reserves, reflecting the gradual repayment of the VLTRO loans by some of the commercial banks. Between its peak and the end of 2013, the LTRO balances decreased by more than half a trillion euro. On top of this decrease in LTRO balances, the ECB balance sheet also decreased due to other factors, such as decreases in securities purchased under SMP due to their maturing. Altogether, the ECB's balance sheet decreased by 700 billion over this period. This left excess reserves between 110 and 140 billion for most of 2014. In December that year, they again started to increase, reflecting jumps in LTRO balances due to the new program of targeted long-term refinancing operations (TLTROs) and thanks to small-scale security purchases under various asset-backed security purchase programs. These factors were then swamped by the large-scale asset purchase program announced in January and implemented from March, discussed in greater detail in the main text.

1.A.2 Effects of the ECB policies on excess reserves

This appendix discusses and illustrates the effects of the ECB policies on the ECB's balance sheet and on excess reserves, focusing on the VLTRO and QE programs (which had the largest influence on excess reserves as illustrated in Figure 1.3 in the text).

I illustrate the effects of these two policies by studying their effects on the accounting balance sheet of the ECB and of a generic financial institution (FI, most commonly a commercial bank).[27] For this purpose, Figure 1.20 shows a simplistic scheme of the normal state of affairs for the ECB and financial institutions. The ECB's assets include, among other things, securities (typically government bonds) and loans to financial institutions (typically main refinancing operations). The main assets of financial institutions are loans to the non-financial sector, but to a large degree also include various securities, typically government bonds. In addition to these two asset classes, financial institutions also always have deposits at the ECB, which in pre-2008 times were mostly held to satisfy reserve requirements. In terms of liabilities, financial institutions fund themselves mostly through commercial deposits, but can also issue their own bonds or use loans from the ECB.

Figure 1.20: Normal balance sheets

ECB

A	L
Securities	Banknotes
Loans to FI	FI deposits

Financial Insitution

A	L
Loans	Bonds
Securities	Deposits
ECB deposits	Loans from ECB

VLTRO program. The VLTRO program consisted of loans granted to FIs, that crucially were not sterilized, meaning that their effect on the ECB balance sheet was not offset by additional countervailing actions. While the loans were collaterilized by relevant securities, the securities remained on the balance sheets of FIs from an accounting perspective. The overall effect on the schematic balance sheets can be seen in figure 1.21.

The loans under VLTRO program have two connected effects on ECB's balance sheet under double-entry bookkeeping. First, on the asset side, the ECB increases the size of loans to FIs. Second, the loans are executed by depositing the accounts of relevant FIs, which from accounting perspective leads to an increase in liabilities.

[27]This approach is similar to that of McLeay, Radia, and Thomas (2014), who focus on money creation in a modern economy, and also apply this approach to the QE program.

Figure 1.21: Effect of VLTRO on balance sheets

ECB	
A	**L**
Securities	Banknotes
Loans to FI	FI deposits
+Loans to FI	+FI deposits

Financial Insitution	
A	**L**
Loans	Bonds
Securities	Deposits
ECB deposits	Loans from ECB
+ECB deposits	+Loans from ECB

Hence, both assets and liabilities increase and correspondingly the ECB balance sheet size increases. FI balance sheets also increase, with assets increasing due to changes in deposits at the ECB and liabilities increasing due to loans from the ECB.

QE program. The QE program consists of purchases of (mostly government) bonds by the ECB. The most common sellers are various FIs and especially commercial banks, and for ease of exposition I focus on this case in figure 1.22. The alternative situation results in the same outcome.

Figure 1.22: Effect of quantitative easing on balance sheets

ECB	
A	**L**
Securities	Banknotes
Loans to FI	FI deposits
+Securities	+FI deposits

Financial Insitution	
A	**L**
Loans	Bonds
Securities	Deposits
ECB deposits	Loans from ECB
-Securities	
+ECB deposits	

As a result of purchase of a bond by the ECB, its stock of securities increases. Meanwhile, as compensation for the seller, its account with the ECB is deposited with the amount equal to the sale price. This results in simultaneous increases in the assets and liabilities of the ECB, leaving its balance sheet larger by the price of the security. This is the case only because the purchase is not sterilized.

Meanwhile, from the perspective of the FI, the sale of a bond leads to a change on the asset side of the balance sheet: the amount of securities decreases, but the amount of deposits at the ECB increases by a corresponding amount. Note that, in contrast to VLTRO, this program does not lead to increases in the accounting size of FI balance sheets.

Figure 1.23: Effect of asset purchase on excess reserves

ECB

A	L
Securities	Banknotes
Loans to FI	FI deposits
	-FI A deposits
	+FI B deposits

Financial Insitution A

A	L
Loans	Bonds
Securities	Deposits
ECB deposits	Loans from ECB
+Securities	
-ECB deposits	

Financial Insitution B

A	L
Loans	Bonds
Securities	Deposits
ECB deposits	Loans from ECB
-Securities	
+ECB deposits	

Excess reserves. Since 2014, excess reserves are being penalized by negative deposit rate (DR). The accounting technique can also shed light on why FIs as a group cannot decrease the total *collective* amount of excess reserves by buying assets from other FIs. To show this, Figure 1.23 captures the effects of purchase of bond by Financial Institution A from Financial Institution B.

The reason FIs cannot collectively decrease the total amount of excess reserves is because any transaction between FIs leads only to changes in distribution of deposits of individual FIs at the ECB, but not to change in total deposits. In the example of a purchase of a bond by one FI from another, the deposits of the FI that is buying the bond decrease, but the deposits of the FI that is selling the bond correspondingly increase. This is because transacitons between FI are settled in terms of deposit accounts at the ECB. The overall effect is only to rearrange excess reserves between individual FIs, and does not lead to their overall decrease or to shrinking of the ECB balance sheet.[28]

[28] Alternatively, the bank can purchase a financial asset from a non-bank investor. In such a case, the bank deposits the account of the non-bank investor with the proceeds from the sale, given that the seller does not have an account with the ECB. Once again, this does not alter the amount of excess reserves. This also means that the effect on excess reserves is identical whether the central bank buys from bank or non-bank entity. There is a small difference between those due to their different effects on the required reserves. In the first case, the reserve base in unaltered (deposits with central bank do not count); in the second case, the reserve base increases by the selling price, and hence excess reserves increase only by 99% of the selling price. There is also a different effect on the money aggregates.

Summary. The crucial aspect of the two the ECB policies, VLTRO and QE, is the fact that their effects were not sterilized. In other words, the ECB does not take any countervailing action to prevent increases in its balance sheet size. These increases in balance sheet size manifests themselves by increasing deposits of FIs at the ECB. In the world of negative DR, each bank wants to decrease its overall deposits at the ECB, but collectively they are unable to do so.

There is one way commercial banks can collectively decrease their excess reserves at the ECB: by forcing their clients to transfer their deposits into cash. This could of course be achieved by imposing (large) negative interest rates their on depositors. Presumably, the reason they do not do so is because they *collectively* benefit more from having clients, with these benefits outweighing the costs of holding excess reserves.

1.B Additional estimation results

This appendix presents additional estimation results complementing those in the main text. Appendix 1.B.1 presents results for all available maturities of IIRs. Appendix 1.B.2 presents results for the Eonia rate when alternative frequency is used instead of monthly frequency. Next, Appendix 1.B.3 shows results for 3-m Euribor when I use different proxy variables for the spread component of the series. Finally, Appendix 1.B.4 complements the results for the full sample estimation sample by showing the results for linear models with 2 regimes. Throughout the appendix, the values reported in the main text are highlighted in bold fonts in the tables to facilitate comparison of the results.

1.B.1 Additional maturities

The main text presented results for the IIRs with 1-day, 1-week, and 3-month maturity. However, there are IIRs with additional maturities, ranging from 1 week to 12 months. This appendix presents results for selected additional maturities, specifically 2-week, 1-month, 2-month, 6-month and 12-month maturities. The appendix focuses only on the robust sample, which is restricted to end in 2019M05 to avoid issues with negative spreads, most likely caused by expectations of forthcoming decrease in the DR in September 2019. This yields 53 observations. All models are estimated using the FMOLS method.

Table 1.7 shows estimation results for the linear model (equation (1.6)) for all considered maturities, while Table 1.8 shows results for the nonlinear model (equation (1.7)). The tables show that the main coefficient of interest - the coefficient on excess reserves - is negative, and is statistically and economically significant for all maturities. Table 1.7 also provides additional support for the notion that excess reserves caused compression in spreads between IIRs with different maturities, with coefficients almost uniformly larger for longer maturities, and especially large for the longest maturities. That said, the results in Table 1.8 are less uniform int this respect.

Table 1.7: Estimation results for different maturities (linear model)

Coef.	Maturity **1-day**	 **1-week**	 2-week	 1-month	 2-month	 **3-month**	 6-month	 12-month
β_0 (const.)	**0.337*** '(0.018) [18.484]**	**0.316*** '(0.024) [12.951]**	0.346*** '(0.034) [10.120]	0.352*** '(0.050) [7.101]	0.467*** '(0.039) [12.054]	**0.599*** '(0.054) [11.107]**	0.899*** '(0.083) [10.772]	1.056*** '(0.116) [9.139]
β_1 (DR_t)	**0.902*** '(0.034) [26.346]**	**1.109*** '(0.045) [24.478]**	1.101*** '(0.061) [18.029]	1.397*** '(0.092) [15.221]	1.221*** '(0.069) [17.618]	**1.111*** '(0.100) [11.106]**	0.493*** '(0.155) [3.185]	0.048*** '(0.214) [0.223]
β_2 (ER_{t-1})	**-0.045*** '(0.004) [-11.305]**	**-0.034*** '(0.005) [-6.315]**	-0.038*** '(0.007) [-5.068]	-0.022*** '(0.011) [-2.013]	-0.043*** '(0.008) [-5.093]	**-0.064*** '(0.012) [-5.413]**	-0.129*** '(0.018) [-7.021]	-0.160*** '(0.025) [-6.294]
Observations	**56**	**53**	53	53	53	**53**	53	53
Model R^2	**0.991**	**0.974**	0.973	0.975	0.981	**0.972**	0.97	0.933
MRR R^2	**0.986**	**0.96**	0.959	0.963	0.973	**0.96**	0.958	0.906
DR R^2	**0.844**	**0.627**	0.66	0.795	0.842	**0.79**	0.831	0.718

Notes: See notes under Table 1.2 for explanation of the values in the table. 1-day maturity corresponds the Eonia rate, results for which differ from those reported in the main text due to use of a shorter estimation sample.

Table 1.8: Estimation results for different maturities (nonlinear model)

Coef.	Maturity **1-day**	 **1-week**	 2-week	 1-month	 2-month	 **3-month**	 6-month	 12-month
β_0 (const.)	**0.337*** '(0.018) [18.484]**	**3.069*** '(0.633) [4.848]**	2.854*** '(0.572) [4.986]	5.088*** '(1.032) [4.933]	1.874*** '(0.351) [5.344]	**2.843*** '(0.745) [3.815]**	0.824*** '(0.477) [1.725]	0.484*** '(0.731) [0.662]
β_2 (ER_{t-1})	**-0.045*** '(0.004) [-11.305]**	**-0.926*** '(0.090) [-10.234]**	-0.876*** '(0.083) [-10.598]	-1.161*** '(0.147) [-7.876]	-0.631*** '(0.051) [-12.472]	**-0.724*** '(0.107) [-6.801]**	-0.369*** '(0.068) [-5.414]	-0.250*** '(0.105) [-2.387]
Model R^2	**0.991**	**0.698**	0.729	0.581	0.86	**0.565**	0.794	0.646
MRR R^2	**0.986**	**0.972**	0.974	0.762	0.978	**0.778**	0.943	0.88
DR R^2	**0.844**	**0.738**	0.787	-0.33	0.873	**-0.159**	0.775	0.641

Notes: See notes under Table 1.2 for explanation of the values in the table.

1.B.2 Alternative frequencies

The main text presented results when models were estimated on data with a monthly frequency. This frequency was selected in accordance with the focus of the chapter on medium frequency/horizon developments in the IIRs; see discussion in the beginning of section 1.4. This appendix contains results when an alternative frequency (daily, weekly and quarterly) is used, focusing on the Eonia rate estimated on a robust estimation sample using the FMOLS method. The main takeaway from Table 1.9 is that the results with respect to the main coefficient of interest are almost identical when different frequencies are used; if anything, the results for the nonlinear model even suggest that the use of monthly frequency understates the size of the effect of excess reserves, as the results for all other frequencies are higher, sometimes substantially so. Therefore, the only major result that varies with frequency is the DR-spread R^2, which is lower at higher frequencies. This is as expected and was one of the motivations for the use of monthly frequency in the main text: since at higher frequencies the IIRs are substantially more volatile, and since this volatility likely reflects random noise rather than any predictable movement, the higher weight on such noise leads to weaker model fit. Still, even for daily frequency the DR the spread R^2 is very high at above 0.7, providing futher endorsement of the model.

1.B.3 Results with proxy for expectations and spread

In case of IIRs with longer than daily maturity - i.e. Euribor rates - a significant amount of variation comes from the variation expectations of future policy rates and from the variation in the credit and liquidity risk components. The main text claimed that ignoring these components does not pose econometric problems from the perspective of the focus of the present chapter. This appendix first provides more detailed discussion why these omitted components are unlikely to bias the results. It then proceeds to present results from models that attempt to control for variation in these components by using several alternative proxy variables. The

Table 1.9: Estimation results for different frequencies

	Linear model				Non-linear model			
Coef.	D	W	**M**	Q	D	W	**M**	Q
β_0 (cons.)	0.308*** '(0.013) [24.536]	0.326*** '(0.017) [19.089]	**0.337*** '(0.018) [18.484]**	0.368*** '(0.007) [51.016]	0.511*** '(0.143) [3.586]	0.734*** '(0.305) [2.405]	**0.390*** '(0.256) [1.520]**	0.813*** '(0.204) [3.988]
β_1 (DR_t)	0.928*** '(0.022) [42.881]	0.915*** '(0.031) [29.276]	**0.902*** '(0.034) [26.346]**	0.930*** '(0.015) [60.175]	-	-	**-**	-
β_2 (ER_{t-1})	-0.040*** '(0.003) [-14.984]	-0.043*** '(0.004) [-11.625]	**-0.045*** '(0.004) [-11.305]**	-0.048*** '(0.002) [-28.259]	-0.503*** '(0.020) [-24.875]	-0.532*** '(0.043) [-12.296]	**-0.487*** '(0.036) [-13.337]**	-0.546*** '(0.029) [-18.575]
Observations	1175	236	**56**	19	1175	236	**56**	19
Model R^2	0.981	0.988	**0.991**	0.985	0.762	0.827	**0.886**	0.933
MRR R^2	0.97	0.981	**0.986**	0.978	0.972	0.985	**0.99**	0.99
DR R^2	0.715	0.81	**0.844**	0.808	0.732	0.85	**0.89**	0.91

Notes: See notes under Table 1.2 for explanation of the values in the table.

results provide empirical support to the arguments.

Potential sources of bias. The first potential source of bias in coefficient estimates is the absence of future policy rates in the equations. The absence is problematic because variations in the IIRs (likely) reflect changes in expectations of future policy rates. If these expectations are systematically correlated in my sample with excess reserves then the corresponding regression coefficients would be biased.

There are two reasons such a correlation is unlikely to be present in the sample under consideration. First, during the period of the QE program, policy rates were unchanged for most of the time and were expected to remain so. Second, during this period excess reserves were gradually increasing as result of gradual, pre-determined purchases of assets under the QE program. Statistically speaking, correlation between an (almost) deterministically trending variable and a stationary[29] variable is zero, so there is little reason to expect a correlation between excess reserves and expectations of future policy rates. Of course, the picture is

[29] Note that the expected difference between current and future policy rates is clearly stationary, mean-zero variable.

more complicated outside of the QE program period, which is another reason I primarily focus on the period of QE program.

There are two ways to address this potential source of bias. This appendix will present estimation results for model that controls for future policy interest rates by including proxy variables in the model. Alternatively, in Appendix 1.C, I eliminate the observations for which expectations about future policy decisions could affect current values of interbank interest rates under the expectational hypothesis of interest rates.[30]

The second potential source of bias is the absence of risk component in the model. This can lead to omitted variable bias in case the variation in financial market stress is correlated with other regressors, and especially with excess reserves.

In general, the situation is similar as in the case of policy rate expectations: the potential for such an issue is very low during the QE program period, but somewhat larger outside of this period. During the QE period the variation in stress in financial markets was very low, and in any case excess reserves were trending deterministically, making significant correlation theoretically impossible. This mirrors the discussion of policy rate expectations. In contrast, in the sample before the QE program there were periods of significant and time varying amounts of stress in the euro zone interbank market, especially during the global financial crisis of 2007-2009 and the Eurozone sovereign debt crisis of 2010-2012.The evolution of excess reserves was less uniform, with multiple increases and decreases. These facts combined raise the possibility that excess reserves and interbank market stress are correlated during this period. For example, the liquidity infusions due to VLTRO could possibly pose a problem with endogeniety, as it was introduced partly in reaction to stress in the financial markets, including in the interbank market, manifested by elevated the IIRs.[31] For this reason, the coefficient estimates for

[30]Specifically, for 1-week Euribor model estiamted at weekly frequency I drop observations for weeks that include a scheduled ECB meeting as well as 1 week prior these meetings.

[31]In contrast, the QE program was announced in reaction to macroeconomic, not financial, developments (Rostagno et al. 2019).

Euribor rates based on the pre-QE sample should be taken with a grain of salt. The way to address this problem is to use proxy variable for the stress component, as I do below.

Model. The results in the main text were obtained from equations (1.3) and (1.4). To understand the meaning of using this model, notice that it amounts to decomposing Euribor rate into its long-run (steady state or equilibrium) component and the residual from this component:

$$IIR_t = IIR_t^{eq} + Residual_t \tag{1.8}$$

Before discussing these two components, it is worth clarifying the meaning of both components. The key idea is that the residual component captures all *transitory* variations in Eurbor rates. Meanwhile, the equilibrium component is the level of the IIR that prevails in the absence of these transitory variations, and hence in steady state defined by absence of expected changes in policy rates or stress in financial markets causing transitory variations in credit and liquidity risk.

Given this definition, it is clear that the equilibrium component is the part of long-maturity the IIRs that behaves like the Eonia rate: the difference between Euribor rates and the Eonia rate is the presence of a time-varying risk component, which has been confined to the spread component. This means that in principle we can use equations (1.3-1.4) for the equilibrium component, simply replacing IIR_t with IIR_t^{eq}, and this is approach taken in the main text. This was justified by the fact these transitory variations were minimal and/or unlikely to be correlated with variables of interest during the main sample of interest.

Meanwhile, the residual component is composed of the two transitory components. The first component - expected future policy rates - means that expected decreases/increases in *future* policy rates lead to lower/higher *current* IIRs with longer maturity, reflecting the usual expectational hypothesis. One way to address this issue is to include in the model for Euribors also the expected policy rates over the maturity of the giver Euribor, an approach taken below.

The second component captures the *transitory* variation in the risk component of the IIR, and as such it can be associated with the presence of stress in the interbank market. In normal times when such stress is absent. However, in periods of stress in interbank markets, equations (1.3-1.4) omit an important driver in the movements of Euribor rates and, correspondingly, can lead to inconsistent coefficient estimates due to omitted variable bias, as discussed in detail in the main text. Again, the solution here is to use proxy measure for stress in interbank market, given that direct measures of such stress are not observable.

Putting everything together, the resulting model has following structure:

$$IIR_t = \beta_{20} + \beta_{21}DR_t + \beta_{22}f(ER_t) + \beta_{23}DR_t^{exp.} + \beta_{24}Stress_t + \epsilon_t \tag{1.9}$$

in case of linear model, and

$$log(IIR_t - DR_t) = \beta_{20} + \beta_{22}f(ER_t) + \beta_{23}DR_t^{exp.} + \beta_{24}Stress_t + \epsilon_t \tag{1.10}$$

in case of non-linear model. In both cases I focus only on the excess reserves regime.

Proxy variables. With respect to controlling for expectations of future changes in policy rates, the simplest approach is to rely on the actual observed future path of policy rates.[32] Using changes in these as proxy variables for expectations of changes in policy rates is akin to assuming perfect foresight on the part of financial market participants. While such an assumption is clearly too strong, the point is that error with respect to such an assumption is unlikely to be correlated with anything else in the model. As argued in the main text, expectations about future policy rate changes are very unlikely to be correlated with the vari-

[32]More standard approach is to rely on overnight index swaps, which are financial market measures of expected future Eonia rate, as measures of expected development in future policy rates. This approach is not viable here. The main text established that Eonia varies with excess reserves, which are themselves predictable in environment with the QE program. As a result, during the QE program there could/should be predictable movements in future Eonia rates that are not related to the expected changes in policy rates, but instead reflect expected changes in excess reserves. Including those in the model would then invalidate the empirical strategy.

able of interest. The point here is to see the effect of eliminating another source of variation in the IIRs with longer maturity.

Meanwhile, there are several possible proxy variables for stress in interbank markets, all of which rely on the fact that stress in one part of financial markets affects other parts of financial markets, and specifically interbank markets. For example, stress in government bond markets increases the likelihood of commercial banks failures and hence raises the credit risk component of the IIRs. This suggests that we can use the measure of stress in government bond markets as a proxy for stress in interbank markets.

Correspondingly, I focus on the use of the following proxies for stress in interbank markets:

1. *Spread between US 3-month libor and risk free rates.* This proxy relies on the notion that global interbank markets are, to a large degree, connected and hence stress in US interbank markets is indicative of stress of in the euro zone interbank markets. In other words, the stress components co-move together, allowing of the use measure of US stress as a proxy for the stress in the euro zone.

2. *Spread between 3-month Euribor and 3-month French bond yields.* This proxy relies on the notion that, after accounting for movements in (almost) risk-free French government bond yields, the remainder of the movements in 3-month Euribor reflect the risk components.

3. *Spread between the average the euro zone bond yield and risk free rates.* This proxy relies on the notion that increased riskiness of particular euro zone government's bonds translates into increased riskiness of banks that hold bonds of this sovereign. This suggests that part of the movement in risk components can be proxied by the movements in the measure of risk in the euro zone bond markets.

As indicated, construction of all three proxy measures requires elimination of risk-free interest rates. In case of the US-based proxy, I use the 3-month treasury yield. In the case of the second proxy, the risk-free rate is measured by the French government bond yield. Finally, in the case of bond yield proxy, I use the a German bond yield as measure of risk free rates.

Before proceeding, it is worth discussing an important limitation of the first two proxy variables. These proxy variables rely on directly on the IIRs and hence clearly suffer from endogeneity problems. This is best understood in the case of the second proxy. Here, the proxy variable is constructed from the dependent variable as such, which clearly means that it is correlated with the error term of such regressions, biasing its coefficient upwards. That said, this does not pose problems insofar as the proxy variable is not correlated with other regressors, as argued in the main text. Furthermore, to avoid this problem, I will also provide results for the TSLS estimation method, where I will use past values of the proxy variable as the instrument for its current value.

Results. Table 1.10 presents results for the 3-month Euribor when different proxy variables are used, estimating the linear model on a robust estimation sample, which has 53 observations.[33] The use of proxy variables almost uniformly leads to coefficients on excess reserves that are higher in absolute value and more statistically significant, clearly supporting the conclusion that the negative estimated coefficient reported in the main text is not an artifact of the omission of expectations and stress components. That said, with one exception, the coefficients for neither proxy variable for stress are estimated to be statistically significant, and the proxy based on US Libor has a negative estimated coefficient. This corresponds to the fact stressed in the main text that the period of the robust sample does not feature any stress in interbank markets, further supporting the conclusion of the irrelevance of a stress component for the results presented in main text.

[33]The focus here is not on the proxy variables or the variation in the expectations and risk components, but rather on how controlling for these components changes the results from the perspective of the main coefficient of interest, the coefficient on excess reserves. Throughout the appendix I focus on results for the 3-month Euribor estimated using the FMOLS estimation method.

Table 1.10: Estimation results with proxy variables

Coef.	**0**	1	2a	2b	3a	3b	4	5
	Model alterantive							
β_0 (const.)	**0.451*** '(0.071) [6.382]**	0.576*** '(0.051) [11.281]	0.502*** '(0.065) [7.677]	0.645*** '(0.135) [4.772]	0.452*** '(0.058) [7.850]	0.486*** '(0.076) [6.379]	0.439*** '(0.074) [5.938]	0.526*** '(0.045) [11.684]
β_1 (DR_t)	**1.225*** '(0.137) [8.946]**	-	1.057*** '(0.129) [8.220]	0.621** '(0.246) [2.527]	1.226*** '(0.119) [10.264]	1.028*** '(0.197) [5.210]	1.235*** '(0.139) [8.891]	-
β_2 (ER_{t-1})	**-0.038*** '(0.016) [-2.393]**	-0.059*** '(0.012) [-5.015]	-0.050*** '(0.014) [-3.605]	-0.086*** '(0.025) [-3.475]	-0.036** '(0.015) [-2.429]	-0.051** '(0.023) [-2.223]	-0.036** '(0.016) [-2.334]	-0.037*** '(0.013) [-2.973]
β_1 ($DR_t^{exp.}$)	**-**	1.140*** '(0.111) [10.268]	-	-	-	-	-	1.281*** '(0.108) [11.878]
β_3 ($Stress_t$)	**-**	-	-0.07 '(0.047) [-1.503]	-0.225 '(0.142) [-1.587]	0.009 '(0.021) [0.432]	0.012 '(0.024) [0.478]	0.008 '(0.031) [0.268]	0.038** '(0.018) [2.168]
Model R^2	**0.982**	0.975	0.986	0.98	0.983	0.994	0.982	0.974
MRR R^2	**0.974**	0.963	0.98	0.95	0.976	0.979	0.974	0.963
DR R^2	**0.863**	0.855	0.893	0.741	0.873	0.891	0.866	0.854

Notes: See notes under Table 1.2 for explanation of the values in the table. Model 0 reproduces the baseline estimates presented in the main text. Model 1 replaces the DR with a proxy for the expected DR. Models 2, 3, and 4 use a stress proxy based on the US Libor rate, Euribor rate, and 10-year bond yield spread, respectively. Model 5 combines a proxy based on Euribor with a proxy for the expected DR. Models 2b and 3b use a TSLS estimation with the past value of the proxy as an instrument for the current value of the stress proxy.

1.B.4 Linear models with 2-regimes

This appendix provides results for a linear model with 2-regimes, complementing the results for the non-linear model provided in the main text. Table 1.11 shows that the conclusions from the main text are supported by the results: the coefficients are broadly consistent with the coefficients from single regime models presented in Table 1.2, and with the coefficients from non-linear two-regime models presented in Table 1.11.

Table 1.11: Estimation results - 2 regime models

	Dependent variable		
Coefficient	the Eonia rate	1W Euribor	3M Euribor
β_{10} (constant) (NR)	0.051*** '(0.008) [6.702]	0.109*** '(0.008) [13.463]	0.241*** '(0.045) [5.327]
β_{20} (constant) (ERR)	0.474*** '(0.070) [6.780]	0.647*** '(0.125) [5.186]	0.734*** '(0.102) [7.202]
β_{21} (DR_t) (ERR)	1.057*** '(0.030) [35.721]	1.097*** '(0.046) [23.778]	1.414*** '(0.034) [41.514]
β_{22} (ER_{t-1}) (ERR)	-0.058*** '(0.012) [-4.967]	-0.087*** '(0.021) [-4.172]	-0.069*** '(0.016) [-4.173]
Observations	212	209	209
Level R^2	1.00	1.00	0.99
the MRR-spread R^2	0.93	0.91	0.64
DR-spread R^2	0.98	0.97	0.87
Threshold	22.85	17.38	85.38

Notes: See notes under Table 1.2 and Table 1.5 for explanation of all values.

1.C Robustness checks

This appendix presents additional estimation results that demonstrate the robustness of the results in the main text to multiple different changes. Appendix 1.C.1 shows results for alternative estimation methods, focusing on the maturities presented in the main text. Appendices 1.C.2 and Appendix 1.C.3 show results for the Eonia rate when alternative timing or a functional form for excess reserves are used, respectively. Appendix 1.C.4 shows results when a different ARMA structure of errors is used, or when these are replaced by a lagged the dependent variable. Appendix 1.C.5 shows that the results are unaffected when observations around the ECB decisions are dropped. Finally, Appendix 1.C.6 investigates the effect of using information about bond yields and the ECB bond holdings in addition to or instead of excess reserves. Unless otherwise stated all results correspond to the FMOLS estimation method. In all cases the number of observations is 56, as in the main text.

1.C.1 Alternative estimation methods

The main text presented results for several selected estimation methods. This appendix provides results for additional estimation methods. For the Eonia rate, and a robust estimation sample, these additional methods consist of alternative methods for co-integration methods: The canonical cointegration regression of Park (1992), dynamic OLS of Stock and Watson (1993) and the ARDL model with alternative selection criterion. For the 1-week and 3-month Euribor, and for the Eonia rate in the full estimation sample, the addition consists of those methods and the other methods discussed in the main text but not reported for these the IIRs or sample. In all cases, the results differ only in terms of the estimation method used; everything else is kept as in the results discussed in the main text to facilitate comparison.

Tables 1.12 and 1.13 show that using the other cointegration estimation methods for the Eonia rate leads to identical conclusions as those based on FMOLS estimation methods: the main coefficient of interest is mostly unchanged in terms of its economical and statistical significance. Meanwhile, Tables 1.14-1.17 show that this is mostly true also for 1-week and 3-month Euribors, albeit with the exception of the ARDL estimation method: using ARLD leads to a systematically lower (but still statistically significant) coefficient of interest; that said, it also leads to weaker model fit, suggesting that the lower estimated coefficient does not correspond to the data better than the coefficients from alternative methods. Finally, Table 1.18 shows that the results are also unaffected for the 2 regime model for the Eonia rate.

Table 1.12: Estimation results for the Eonia rate - linear model

	Method				
Coefficient	**FMOLS**	ARDL - AIC	CCR	DOLS - SIC	DOLS - AIC
	0.337*** **'(0.018)** **[18.484]**	0.261*** '(0.033) [7.983]	0.330*** '(0.017) [19.682]	0.352*** '(0.044) [7.910]	0.322*** '(0.037) [8.800]
β_0 (constant)					
β_1 (DR_t)	**0.902***** **'(0.034)** **[26.346]**	0.902*** '(0.050) [18.023]	0.901*** '(0.032) [27.928]	0.913*** '(0.038) [24.121]	0.969*** '(0.030) [31.812]
β_2 (ER_{t-1})	**-0.045***** **'(0.004)** **[-11.305]**	-0.035*** '(0.006) [-5.486]	-0.044*** '(0.004) [-11.761]	-0.047*** '(0.007) [-6.819]	-0.039*** '(0.006) [-6.721]
Lag struct.	**-**	(3,1,3)	-	(0,2)	(1,3)
Model R^2	**0.991**	0.984	0.991	0.995	0.996
MRR R^2	**0.986**	0.974	0.986	0.981	0.982
DR R^2	**0.844**	0.736	0.849	0.797	0.803

Notes: See notes under Table 1.2 for explanation of the values in the table.

Table 1.13: Estimation results for the Eonia rate - nonlinear model

	Method				
Coefficient	**FMOLS**	ARDL - AIC	CCR	DOLS - SIC	DOLS - AIC
	0.390* **	0.213***	0.414***	0.291***	-0.059***
β_0	**'(0.256)**	'(0.356)	'(0.238)	'(0.294)	'(0.361)
(constant)	**[1.520]**	[0.597]	[1.743]	[0.988]	[-0.162]
	-0.487* **	-0.463***	-0.490***	-0.474***	-0.428***
β_1	**'(0.036)**	'(0.050)	'(0.034)	'(0.041)	'(0.049)
(ER_{t-1})	**[-13.337]**	[-9.272]	[-14.403]	[-11.501]	[-8.714]
Lag struct.	**-**	(1,0)	-	(0,0)	(1,0)
Model R^2	**0.886**	0.987	0.886	0.896	0.9
MRR R^2	**0.99**	0.979	0.99	0.989	0.983
DR R^2	**0.89**	0.789	0.892	0.877	0.813

Notes: See notes under Table 1.2 for explanation of the values in the table.

Table 1.14: Estimation results for 1-week Euribor - linear model

	Method							
Coefficient	**FMOLS**	OLS	TSLS	ARDL(SIC)	ARDL(AIC)	CCR	DOLS(SIC)	DOLS(AIC)
	0.316* **	0.285***	0.127***	0.101***	0.128***	0.307***	0.317***	0.355***
β_0	**'(0.024)**	'(0.034)	'(0.057)	'(0.042)	'(0.039)	'(0.024)	'(0.087)	'(0.090)
(constant)	**[12.951]**	[8.342]	[2.211]	[2.409]	[3.304]	[12.945]	[3.658]	[3.961]
	1.109* **	1.018***	0.922***	1.084***	1.088***	1.082***	1.298***	1.247***
β_1	**'(0.045)**	'(0.044)	'(0.035)	'(0.049)	'(0.045)	'(0.044)	'(0.067)	'(0.076)
(DR_t)	**[24.478]**	[22.901]	[26.677]	[21.957]	[24.048]	[24.424]	[19.411]	[16.350]
	-0.034* **	-0.035***	-0.019***	-0.006***	-0.009***	-0.034***	-0.023***	-0.031***
β_2	**'(0.005)**	'(0.006)	'(0.007)	'(0.007)	'(0.007)	'(0.005)	'(0.014)	'(0.015)
(ER_{t-1})	**[-6.315]**	[-5.751]	[-2.614]	[-0.859]	[-1.412]	[-6.424]	[-1.597]	[-2.006]
Lag struct.	**-**	MA(2)	AR(1)	(2,2,2)	(2,2,0)	-	(4,3)	(4,4)
Model R^2	**0.974**	0.997	0.999	0.962	0.97	0.977	0.998	0.998
MRR R^2	**0.96**	0.974	0.916	0.94	0.953	0.964	0.902	0.889
DR R^2	**0.627**	0.758	0.229	0.452	0.566	0.667	0.094	-0.022

Notes: See notes under Table 1.2 for explanation of the values in the table.

Table 1.15: Estimation results for 1-week Euribor - nonlinear model

	Method							
Coefficient	**FMOLS**	OLS	TSLS	ARDL(SIC)	ARDL(AIC)	CCR	DOLS(SIC)	DOLS(AIC)
	3.069* **	1.474***	1.676***	-1.105***	-1.105***	2.522***	1.396***	2.158***
β_0	**'(0.633)**	'(0.479)	'(0.464)	'(1.607)	'(1.607)	'(0.533)	'(0.424)	'(0.511)
(constant)	**[4.848]**	[3.075]	[3.611]	[-0.688]	[-0.688]	[4.727]	[3.295]	[4.223]
	-0.926* **	-0.719***	-0.748***	-0.374***	-0.374***	-0.849***	-0.709***	-0.806***
β_1	**'(0.090)**	'(0.068)	'(0.067)	'(0.218)	'(0.218)	'(0.077)	'(0.061)	'(0.069)
(ER_{t-1})	**[-10.234]**	[-10.615]	[-11.176]	[-1.712]	[-1.712]	[-11.031]	[-11.646]	[-11.615]
Lag struct.	-	MA(2)	AR(1)	(1,0)	(1,0)	-	(3,4)	(4,4)
Model R^2	**0.698**	0.864	0.862	0.946	0.946	0.752	0.864	0.88
MRR R^2	**0.972**	0.986	0.989	0.917	0.917	0.985	0.985	0.992
DR R^2	**0.738**	0.873	0.895	0.234	0.234	0.861	0.861	0.923

Notes: See notes under Table 1.2 for explanation of the values in the table.

Table 1.16: Estimation results for 3-month Euribor - linear model

	Method							
Coefficient	**FMOLS**	OLS	TSLS	ARDL(SIC)	ARDL(AIC)	CCR	DOLS(SIC)	DOLS(AIC)
	0.599* **	0.528***	0.196***	-0.135***	0.177***	0.551***	0.827***	0.827***
β_0	**'(0.054)**	'(0.044)	'(0.437)	'(1.077)	'(0.572)	'(0.035)	'(0.049)	'(0.049)
(constant)	**[11.107]**	[12.015]	[0.447]	[-0.125]	[0.309]	[15.674]	[16.860]	[16.860]
	1.111* **	0.951***	0.485***	0.378***	0.693***	1.083***	1.198***	1.198***
β_1	**'(0.100)**	'(0.052)	'(0.072)	'(1.238)	'(0.562)	'(0.057)	'(0.021)	'(0.021)
(DR_t)	**[11.106]**	[18.369]	[6.765]	[0.305]	[1.232]	[19.016]	[58.164]	[58.164]
	-0.064* **	-0.063***	-0.045***	-0.007***	-0.031***	-0.059***	-0.089***	-0.089***
β_2	**'(0.012)**	'(0.008)	'(0.057)	'(0.108)	'(0.071)	'(0.007)	'(0.006)	'(0.006)
(ER_{t-1})	**[-5.413]**	[-7.822]	[-0.792]	[-0.069]	[-0.441]	[-8.277]	[-14.088]	[-14.088]
Lag struct.	**-**	MA(2)	AR(1)	(2,0,1)	(2,2,1)	-	(4,3)	(4,4)
Model R^2	**0.972**	0.993	0.995	0.148	0.746	0.98	0.999	0.999
MRR R^2	**0.96**	0.978	0.549	-0.249	0.628	0.971	0.741	0.741
DR R^2	**0.79**	0.884	-1.349	-5.508	-0.94	0.849	-0.351	-0.351

Notes: See notes under Table 1.2 for explanation of the values in the table.

Table 1.17: Estimation results for 3-month Euribor - nonlinear model

	Method							
Coefficient	**FMOLS**	OLS	TSLS	ARDL(SIC)	ARDL(AIC)	CCR	DOLS(SIC)	DOLS(AIC)
	2.843***	1.118***	1.084***	-0.106***	-0.106***	2.095***	4.118***	4.382***
β_0	**'(0.745)**	'(0.213)	'(0.202)	'(0.975)	'(0.975)	'(0.588)	'(0.580)	'(0.662)
(constant)	**[3.815]**	[5.242]	[5.358]	[-0.108]	[-0.108]	[3.563]	[7.100]	[6.616]
	-0.724***	-0.493***	-0.488***	-0.330***	-0.330***	-0.619***	-0.885***	-0.920***
β_1	**'(0.107)**	'(0.033)	'(0.032)	'(0.134)	'(0.134)	'(0.085)	'(0.075)	'(0.086)
(ER_{t-1})	**[-6.801]**	[-14.947]	[-15.467]	[-2.456]	[-2.456]	[-7.252]	[-11.829]	[-10.715]
Lag struct.	-	MA(2)	AR(1)	(1,0)	(1,0)	-	(3,4)	(4,4)
Model R^2	**0.565**	0.865	0.865	0.943	0.943	0.727	0.943	0.946
MRR R^2	**0.778**	0.981	0.981	0.916	0.916	0.911	-0.109	-0.411
DR R^2	**-0.159**	0.9	0.9	0.562	0.562	0.535	-4.778	-6.357

Notes: See notes under Table 1.2 for explanation of the values in the table.

Table 1.18: Estimation results for the Eonia rate - 2 regime model

				Method				
Coefficient	**FMOLS**	OLS	TSLS	ARDL(SIC)	ARDL(AIC)	CCR	DOLS(SIC)	DOLS(AIC)
α_{01}	**0.051*** **	0.052***	0.056***	0.056***	0.056***	0.051***	0.052***	0.052***
(const.)	**'(0.008)**	'(0.008)	'(0.007)	'(0.007)	'(0.007)	'(0.008)	'(0.008)	'(0.008)
(NR)	**[6.702]**	[6.925]	[8.304]	[8.304]	[8.304]	[6.702]	[6.925]	[6.925]
α_{02}	**1.044*** **	1.039***	1.326***	0.882***	0.882***	1.042***	1.102***	1.102***
(constant)	**'(0.365)**	'(0.219)	'(0.255)	'(0.316)	'(0.316)	'(0.358)	'(0.223)	'(0.223)
(ERR)	**[2.860]**	[4.744]	[5.199]	[2.789]	[2.789]	[2.908]	[4.945]	[4.945]
α_{12}	**-0.569*** **	-0.570***	-0.618***	-0.552***	-0.552***	-0.569***	-0.583***	-0.583***
(ER_{t-1})	**'(0.060)**	'(0.032)	'(0.037)	'(0.046)	'(0.046)	'(0.060)	'(0.033)	'(0.033)
(ERR)	**[-9.416]**	[-17.986]	[-16.853]	[-12.069]	[-12.069]	[-9.541]	[-17.826]	[-17.826]
Level R^2	**1.00**	1.00	1.00	1.00	1.00	1.00	1.00	1.00
the MRR R^2	**0.94**	0.94	0.95	0.95	0.95	0.94	0.94	0.94
the DR R^2	**0.98**	0.98	0.98	0.98	0.98	0.98	0.98	0.98
Threshold	**22.85**	17.38	11.39	11.39	11.39	22.85	17.38	17.38

Notes: See notes under tables 1.2 and 1.5 for explanation of values in table.

1.C.2 Alternative timing of excess reserves

The main text reported results for a model estimated with excess reserves lagged by one period, so that increases in excess reserves affect the IIRs with a one month delay. In the absence of any theoretical considerations, the lag on excess reserves was chosen in accordance with the goal of maximizing model fit. Table 1.19 shows the results for the Eonia rate when alternative timing is used, and highlights that this alternative timing does not influence the conclusions of the chapter.

Table 1.19: Estimation results for the Eonia rate - alternative timing of excess reserves

	Linear model			**Non-linear model**		
Coefficient	**0**	1	2	**0**	1	2
β_0 (constant)	**0.316***** **'(0.015)** **[20.384]**	0.352*** '(0.020) [17.544]	0.281*** '(0.015) [18.208]	**0.44** **'(0.294)** **[1.496]**	0.756** '(0.338) [2.238]	-0.073 '(0.315) [-0.233]
β_1 (DR_t)	**0.914***** **'(0.030)** **[30.063]**	0.898*** '(0.035) [25.740]	0.935*** '(0.033) [28.020]	**-**	-	-
β_2 (ER)	**-0.042***** **'(0.003)** **[-12.080]**	-0.047*** '(0.004) [-11.027]	-0.036*** '(0.004) [-10.052]	**-0.491***** **'(0.042)** **[-11.747]**	-0.534*** '(0.048) [-11.163]	-0.422*** '(0.045) [-9.356]
Model R^2	**0.993**	0.992	0.991	**0.895**	0.886	0.879
MRR-spread R^2	**0.988**	0.987	0.986	**0.992**	0.991	0.987
DR-spread R^2	**0.871**	0.856	0.842	**0.914**	0.906	0.853

Notes: See notes under Table 1.2 for explanation of the values in the table. Model 1 uses the current value of excess reserves, while model 2 uses the second lag of excess reserves.

1.C.3 Alternative functional forms

The main text reported results for a model estimated with a logarithm of excess reserves. While this functional form was chosen with the view of satisfying the apparent concavity in the relationship between excess reserves and the IIRs, the choice is arbitrary to some degree. Table 1.20 shows the results for the Eonia rate when alternative functional forms are used, and highlights that using these alternatives does not influence the conclusions of the chapter.

Table 1.20: Estimation results for the Eonia rate - alternative functional forms

	Linear model			**Non-linear model**		
Coefficient	**0**	1	2	**0**	1	2
β_0 (constant)	**0.316*** '(0.015) [20.384]**	0.133*** '(0.007) [18.722]	0.000*** '(0.000) [6.349]	**0.44 '(0.294) [1.496]**	-1.816*** '(0.096) [-18.969]	0.000* '(0.000) [1.972]
β_1 (DR_t)	**0.914*** '(0.030) [30.063]**	0.995*** '(0.036) [27.680]	0.809*** '(0.042) [19.495]	**-**	-	-
β_2 (ER_{t-1})	**-0.042*** '(0.003) [-12.080]**	-0.002*** '(0.000) [-8.033]	-0.00035 [54.21]***	**-0.491*** '(0.042) [-11.747]**	-0.034*** '(0.003) [-12.807]	-0.00233 [61.39]***
Model R^2	**0.993**	0.987	0.995	**0.895**	0.879	0.891
MRR-spread R^2	**0.988**	0.979	0.993	**0.992**	0.986	0.99
DR-spread R^2	**0.871**	0.772	0.92	**0.914**	0.846	0.894

Notes: See notes under Table 1.2 for explanation of the values in the table. Model 1 uses square root transformation of excess reserves, while model 2 uses the cubic polynomial of excess. For model 2, the table reports the sum of coefficients and the F-statistic for the test that all three coefficients are 0.

1.C.4 Alternative ARMA components

The main text reported results for a model estimated with OLS when a particular ARMA structure for the model errors was used. While this ARMA structure was chosen in accordance with the goal of maximizing model fit, alternative ARMA structures could be used. In addition, ARMA errors could be plausibly replaced by a lagged dependent variable. Tables 1.21 and 1.22 show the results for the Eonia rate when an alternative ARMA structure or lagged the dependent variables are used[34], and highlights that the results are robust to these modifications.

[34]Of course, with lagged the dependent variables the coefficient on excess reserves is lower, since it measures only the impact multiplier. After accounting for the delayed effects, the total multiplier is similar to specifications with ARMA errors.

Table 1.21: Estimation results for the Eonia rate - alternative ARMA components (linear model)

Coef.	Model **0**	1	2	3	4
β_0 (const.)	**0.323***** **'(0.018)** **[18.203]**	0.314*** '(0.015) [21.086]	0.377*** '(0.026) [14.604]	0.341*** '(0.019) [17.808]	0.213*** '(0.039) [5.491]
β_1 (DR_t)	**0.900***** **'(0.046)** **[19.748]**	0.907*** '(0.041) [22.070]	0.878*** '(0.045) [19.623]	0.886*** '(0.050) [17.602]	0.736*** '(0.145) [5.078]
β_2 (ER_{t-1})	**-0.044***** **'(0.005)** **[-9.688]**	-0.042*** '(0.004) [-10.650]	-0.052*** '(0.005) [-9.994]	-0.047*** '(0.005) [-9.494]	-0.027*** '(0.006) [-4.775]
ARMA/Lag structure	**MA(2)**	MA(1)	AR(1)	ARMA(1,1)	LDV
Model R^2	**1.060*****	0.996	0.996	0.996	0.994
MRR R^2	'(0.141)	0.989	0.982	0.987	0.991
DR R^2	[7.535]	0.878	0.804	0.86	0.897

Notes: See notes under Table 1.2 for explanation of the values in the table.

Table 1.22: Estimation results for the Eonia rate - alternative ARMA components (nonlinear model)

Coef.	Model **0**	1	2	3	4
β_0 (const.)	**0.482**** **'(0.196)** **[2.460]**	0.496** '(0.216) [2.298]	0.573** '(0.271) [2.116]	0.506** '(0.224) [2.256]	0.009 '(0.163) [0.053]
β_2 (ER_{t-1})	**-0.498***** **'(0.028)** **[-17.800]**	-0.500*** '(0.031) [-16.254]	-0.511*** '(0.039) [-13.029]	-0.501*** '(0.032) [-15.652]	-0.205*** '(0.066) [-3.098]
ARMA/Lag structure	**MA(1)**	MA(2)	AR(1)	ARMA(1,1)	LDV
Model R^2	**0.93**	0.931	0.927	0.931	0.928
MRR R^2	**0.992**	0.992	0.992	0.992	0.995
DR R^2	**0.915**	0.915	0.915	0.915	0.945

Notes: See notes under Table 1.2 for explanation of the values in the table.

1.C.5 Sample without observations surrounding the ECB decisions

To eliminate the possibility that omitted expectations of policy rate changes are driving the results presented in the main text, Table 1.23 shows that the results are virtually unchanged when we consider only the weeks that do not overlap with the ECB decisions. Specifically, the table shows results for 1-week Euribor estimated at weekly frequency when weeks just prior to scheduled the ECB decisions are omitted from the analysis.

Table 1.23: Estimation results for the Eonia rate - alternative functional forms

	Linear		Nonlinear	
Coefficient	With	Without	With	Without
β_0 (constant)	0.293*** '(0.018) [16.187]	0.284*** '(0.019) [14.600]	1.464*** '(0.258) [5.665]	1.449*** '(0.280) [5.176]
β_1 (DR_t)	0.995*** '(0.026) [38.841]	0.999*** '(0.029) [34.175]	-	-
β_2 (ER_{t-1})	-0.037*** '(0.004) [-9.779]	-0.036*** '(0.004) [-8.496]	-0.717*** '(0.039) [-18.494]	-0.715*** '(0.042) [-17.066]
Model R^2	0.995	0.995	0.869	0.854
MRR-spread R^2	0.971	0.973	0.983	0.987
DR-spread R^2	0.74	0.752	0.849	0.88

Notes: See notes under Table 1.2 for explanation of the values in the table. The column header indicates whether weeks with the ECB decisions are or are not included.

1.C.6 Specifications with bond yields and bond holdings

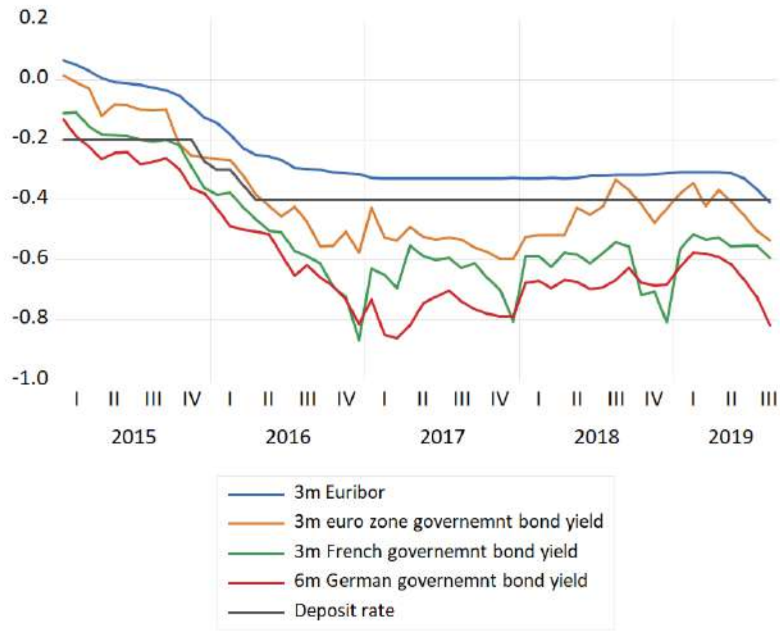
Bond yields. The chapter argued that there is a causal relationship between the amount of excess reserves and the IIRs, and that quantitative easing policy affects the IIRs through its effect on excess reserves discussed in Appendix 1.A.2. While the main text supported the presence of the link between excess reserves and the IIRs, it is a plausible that this estimated relationship is spurious. Specifically, it is plausible hypothesis that the QE does not affect the IIRs through its effect on excess reserves, but rather through its effect on bond yields, since bond yields and the IIRs are clearly closely related. This appendix argues and demonstrates that this alternative hypothesis is not supported by the data.

Before presenting empirical evidence for the hypothesis, it is worth briefly considering economic arguments against the hypothesis. The IIRs are the cost of lending/borrowing reserves between banks. Standard microeconomic theory clearly indicates that the supply of (excess) reserves should play a key role in determining the cost of lending/borrowing of these reserves between banks. Therefore, arguing that the IIRs are not related to the amount of excess reserves is clearly at odds with the standard economic model of supply and demand.

Turning to empirical evidence, there are two approaches to demonstrating that the alternative hypothesis is not valid: informal graphical and formal statistical analysis. First, Figure 1.24 shows the series of main interest - 3-month Euribor together with main 3-month bond yields in the euro zone bond markets - during the period of the QE program. While the bond yields share the same pattern as Euribor, with gradual decreases during the 2015-2016 period, this figure makes it clear that the bond yields are not driving Euribor: not only is the decrease in bond yields occurring slightly sooner than in Euribor, but more importantly, the bond yields are much more volatile throughout the sample than the Euribor. If the QE were influencing the 3-month Euribor through its effects on bond yields, then surely the Euribor would reflect the movements in bond yields throughout the sample. The notion that the QE influences both the IIRs and bond yields directly is much more plausible and explains their partial co-movement in parts of the analyzed sample.[35]

[35] An additional channel could be from Euribor to bond yields (rather than the other way around): the interest rates at which banks lend to each other should influence the interest rates

Figure 1.24: Euribor and the euro zone bond yields



As for the formal statistical analysis, Table 1.24 shows results from the alternative model that includes the euro zone bond yields[36] as an explanatory variable. The table corroborates the conclusions from the graphical analysis. The coefficient on excess reserves is broadly unaffected, with significance increasing when the restricted sample is used and decreasing when the full sample is used. In both cases, the coefficient on bond yields is positive and statistically significant, but the coefficients are relatively small in absolute size, indicating that a one percentage point change in bond yield leads only to 16-29 basis points change in Euribor.

Estimating single equation linking Euribor to bond yields is a-priori problematic from econometric point of view: it is very likely that there is simultaneity between the two series, with Euribor influencing bond yields, which poses problems for an equation linking Euribor to bond yields. For this reason, Table 1.25

at which they are willing to buy government bonds. See discussion below.

[36]The conclusions are unchanged when different bond yields are used.

Table 1.24: Estimation results for 3-m Euribor rate with bond yields

	Robust sample		**Full sample**	
Coefficient	Without	With	Without	With
β_0 (constant)	0.451*** '(0.071) [6.382]	0.426*** '(0.030) [14.285]	0.734*** '(0.102) [7.202]	0.692*** '(0.184) [3.763]
β_1 (DR_t)	1.225*** '(0.137) [8.946]	0.832*** '(0.070) [11.900]	1.414*** '(0.034) [41.514]	1.135*** '(0.106) [10.687]
β_2 (ER_{t-1})	-0.038** '(0.016) [-2.393]	-0.045*** '(0.006) [-7.272]	-0.069*** '(0.016) [-4.173]	-0.065*** '(0.029) [-2.252]
β_3 $(Yield_t)$	-	0.159*** '(0.029) [5.509]	-	0.285*** '(0.104) [2.739]
Model R^2	0.982	0.992	0.985	0.991
MRR-spread R^2	0.974	0.988	0.644	0.802
DR-spread R^2	0.863	0.939	0.867	0.92

Notes: See notes under Table 1.2 for explanation of the values in the table. Column header indicates whether bond yields are included.

shows results for the multiequation estimation method in the form of vector autoregression. The key takeaway is that excess reserves are still a statistically significant regressor in the equation for the Euribor rate.

The overall conclusion is that the observed decrease in Euribor rates during the period of the QE program is indeed due to an increase excess reserves caused by asset purchases, and not due to movements in bond yields caused by those purchases. This finding should not be surprising in the view of the full data sample. While it might be plausible that during the QE period, the link between the IIRs and excess reserves reflects the effect of the QE on bond yields, it is harder to argue that this is the case outside of a period of large-scale bond purchases by ECB. That is, however, at odds with the finding in the main text that the link between excess reserves and the IIRs is also present in the period before the QE, and more specifically with the fact that fluctuations in excess reserves are able to explain a large portion of the variation in the IIRs, even prior to the QE program.

Bond holdings. A related alternative hypothesis is that what matters for Euribor is not the amount of excess reserves, but the size of bond holdings. Table 1.26 shows that there is some support for this hypothesis when one considers the period of the QE program, but not when one considers a longer estimation sample.

In the period of the QE program, the specification which uses holdings under the asset purchase program (APP) has a substantially better model fit than the specification with excess reserves, but this is not true for specifications with holdings under the more narrow public sector purchase program. This would seem to suggest that it is bond holdings that matter, not excess reserves. However, this is at odds with the finding that excess reserves have strong explanatory/predictive power even in the period before the start of large-scale asset purchases. This is mirrored in the 3 right columns in Table 1.26: when bond holdings under PSPP or APP are used instead of excess reserves, then not only is the model fit substantially worse, but the coefficient on these alternative regressors becomes small and insignificant. This suggests that the better fit during the period of the QE program is just a reflection of the close connection between APP and excess reserves and

Table 1.25: Estimation results for 3-m Euribor rate - Vector autoregression

	Euribor	Bond yield
Euribor(-1)	1.582199	1.387832
	(0.08663)	(0.31724)
	[18.2642]	[4.37475]
Euribor(-2)	−0.666337	−2.031679
	(0.13541)	(0.49589)
	[-4.92070]	[-4.09700]
Euribor(-3)	0.010820	0.669480
	(0.06157)	(0.22547)
	[0.17573]	[2.96932]
Bond yield(-1)	0.072832	0.618684
	(0.02460)	(0.09010)
	[2.96025]	[6.86677]
Bond yield(-2)	−0.149698	−0.036041
	(0.02921)	(0.10697)
	[-5.12490]	[-0.33694]
Bond yield(-3)	0.084518	0.069883
	(0.02560)	(0.09376)
	[3.30119]	[0.74536]
C	0.096396	0.076056
	(0.03341)	(0.12236)
	[2.88493]	[0.62156]
Deposit rate	0.067856	0.499283
	(0.04300)	(0.15748)
	[1.57792]	[3.17048]
Excess reserves	−0.012979	−0.001755
	(0.00505)	(0.01850)
	[-2.56925]	[-0.09486]
R-squared	0.996485	0.937584

Notes: See notes under Table 1.2 for explanation of the values in the table. The cColumn header indicates whether bond yields are included. Sample from 2009M01 to 2019M05 potentially random noise.

Table 1.26: Estimation results for 3-m Euribor rate with bond holdings

	Robust sample			Full sample		
Coefficient	ER	PSPP	APP	ER	PSPP	APP
β_0 (constant)	0.316*** '(0.015) [20.384]	0.079*** '(0.005) [15.324]	0.242*** '(0.007) [32.284]	0.474*** '(0.070) [6.780]	0.134*** '(0.022) [6.195]	0.128*** '(0.023) [5.637]
β_1 (DR_t)	0.914*** '(0.030) [30.063]	0.926*** '(0.022) [42.337]	0.872*** '(0.024) [36.333]	1.057*** '(0.030) [35.721]	1.220*** '(0.033) [36.739]	1.227*** '(0.034) [36.039]
β_2 (ER_{t-1})	-0.042*** '(0.003) [-12.080]	-0.009*** '(0.001) [-7.305]	-0.033*** '(0.002) [-16.782]	-0.058*** '(0.012) [-4.967]	-0.001*** '(0.005) [-0.093]	0.001*** '(0.005) [0.094]
Model R^2	0.993	0.965	0.997	0.998	0.996	0.996
MRR-spread R^2	0.988	0.944	0.995	0.929	0.861	0.857
DR-spread R^2	0.871	0.387	0.943	0.977	0.954	0.953

Notes: See notes under Table 1.2 for explanation of the values in the table. The column header indicates which variable is used: ER=excess reserves, PSPP=bond holdings under the public sector purchase program, APP=asset holdings under the asset purchase program.

1.C.7 Constrained specification

In a normal regime, the spreads between the IIRs and the DR does not change with the level of policy rates: the IIRs are equally close to policy rates whether policy rates are high or low. This is a result of an active policy of ECB, which varies its liquidity provisions to ensure the IIRs are close to the desired level. Even though the ECB does not vary its liquidity provisions in the excess reserve regime, it might still be natural to expect that the spread between the IIRs and the DR does not change when the DR increases or decreases. However, this expectation is not supported by formal statistical tests. As a result, in excess reserves regime, the model includes a coefficient on the DR. Its presence allows for the possibility that the spread between the IIRs and the DR changes with changes in the level of the DR. This appendix shows results for models that force this coefficient to be equal to 1.

Table 1.27 compares the results when the coefficient on the DR is estimated and when it is forced to be equal to 1, illustrating the effect in case of the Eonia rate and for both the linear and non-linear models. The key takeaway is that the coefficient on excess reserves is still negative and both economically and statistically significant even when the coefficient on the DR is constrained to be one.

Table 1.27: Estimation results - 2 regime models

	Dependent variable		
Coefficient	the Eonia rate	1W Euribor	3M Euribor
	0.282***	0.319***	0.656***
β_0	'(0.013)	'(0.024)	'(0.069)
(const.)	[21.466]	[13.497]	[9.524]
	-0.033***	-0.040***	-0.077***
β_2	'(0.002)	'(0.003)	'(0.010)
(ER_{t-1})	[-17.404]	[-11.865]	[-7.848]
Level R^2	0.86	0.71	0.70
MRR-spread R^2	0.99	0.97	0.94
DR-spread R^2	0.86	0.71	0.70

Notes: See notes under Table 1.2 for explanation of all values.

1.D Alternative counterfactual paths for excess reserves

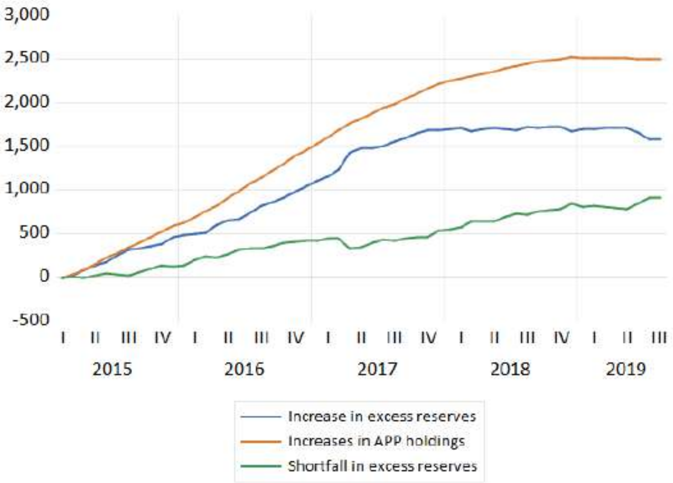
The main text highlighted the fact that constructing a counterfactual path for excess reserves in the absence of quantitative easing - and hence determining both the values of the IIRs without this program and the corresponding effect of this program on the IIRs - requires making assumptions about the behavior of multiple items of ECB's balance sheet. The main text focused on two such plausible counterfactual path and associated effects of the QE program on the IIRs. This appendix provides a more detailed discussion and presents alternative paths for counter-factual levels of excess reserves without QE.

Asset purchases under the quantitative easing policy lead to increases in the total size of ECB's assets, under the usual ceteris paribus assumption. This necessarily leads to corresponding increases in liabilities, and hence to increases in excess reserves, again ceteris paribus, as discussed in Appendix 1.A.2. However, ceteris paribus is a too strong assumption to rely on in constructing a counterfactual path for excess reserves: the other parts of ECB's balance sheet do not stay unchanged, but rather evolve over time, either for independent reasons or in connection with the asset purchases. This leads to increases in excess reserves being different from amount of assets purchased under the QE policy. As discussed in the main text and shown in Figure 1.25, this is indeed what was observed, with excess reserves increasing substantially less than the amount of total asset purchases.

In principle, the increase in excess reserves can be smaller than the increase in asset purchases for two different reasons: either the other asset categories have *decreased*, so that the overall balance sheet has changed less than one would expect solely based on asset purchases; or the other liability categories have increased, so that part of the increase in the balance sheet has been absorbed by these other liabilities. Of course, both factors can play a role; even more confusingly, there can be (and were) factors that pushed in opposite direction.

The rest of this appendix discusses the evolution of ECB's balance sheet items

Figure 1.25: Shortfall in excess reserves



Notes: Cumulative changes in different items of ECB's balance sheet since February 2015. The shortfall corresponds to the difference between cumulative change in Asset Purchase Program outstanding balance and cumulative change in excess reserves.

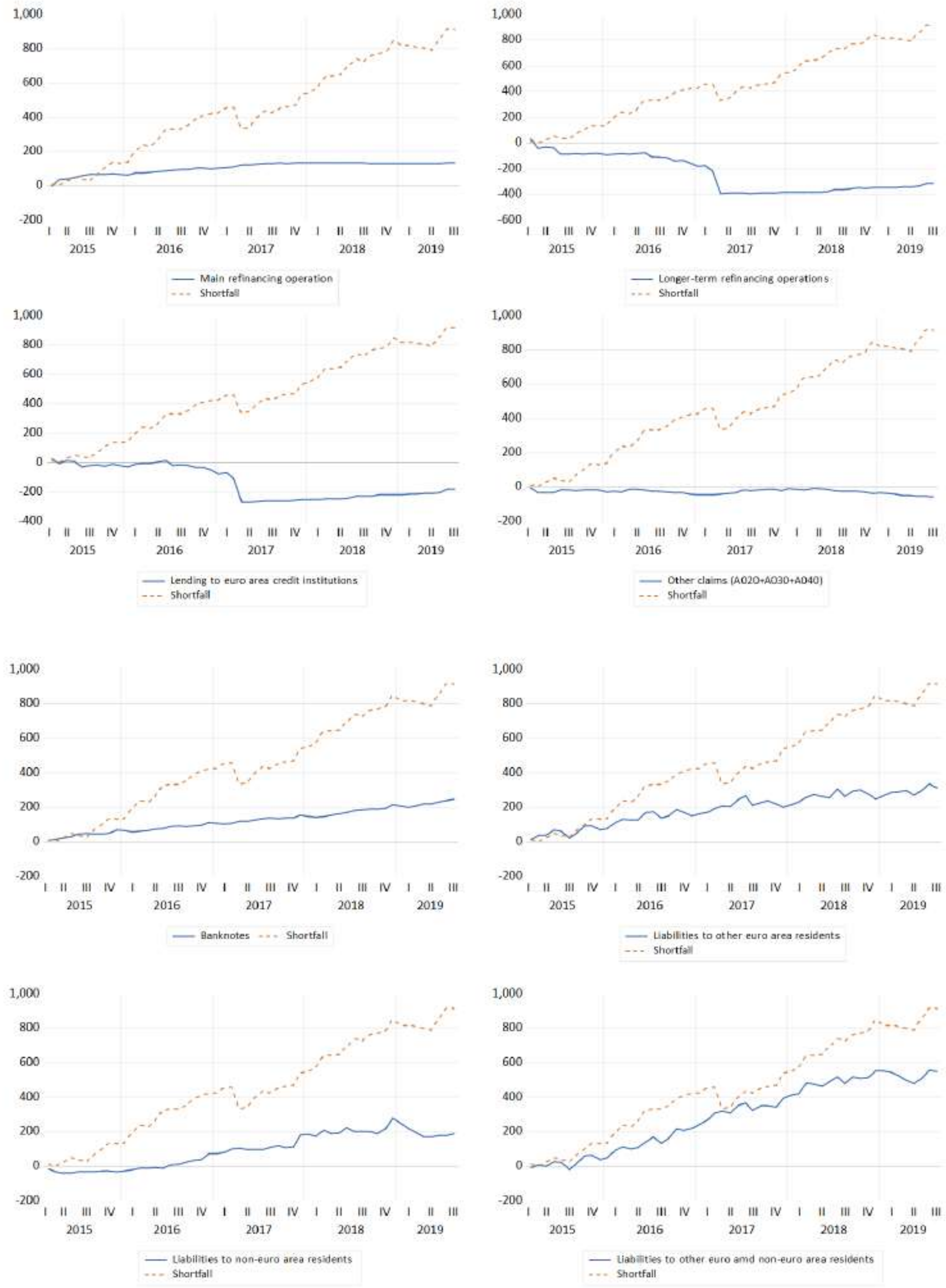
in the view of shortfall in excess reserves captured in Figure 1.25, before proceeding to constructing a counterfactual path for excess reserves using either assumptions about balance sheet items, or by estimating econometric models.

Evolution of ECB's balance sheet items. To begin the analysis, Figure 1.26 provides detailed analysis of the evolution of ECB's balance sheet items, always benchmarking them against the "shortfall" of excess reserves, i.e. the difference between increase in assets due to the QE policy and increase in excess reserves. This allows one to easily determine whether a given category contributed to the shortfall, and by how much. Looking at asset side items, for which values are reversed to facilitate interpretation, one can see that main refinancing operations contributed to the shortfall, while longer-term refinancing operations actually had the opposite effect (graphs in the first row). Since the change in the latter category was larger, the combined category of lending to credit institutions caused excess reserves to be larger than at beginning of the QE program (second row, left panel), something that was discussed in the main text and will play a role in next paragraphs. The only other significant category apart from securities are the claims on other counterparties besides credit institutions. These increased over the period of the QE program, so excess reserves are larger thanks to them, but the effect is not substantial (second row, right panel). Overall, the clear conclusion is that the asset-side items are not responsible for the shortfall, as they actually contributed to increase in excess reserves.

Turning to liability side items, all the main items contributed to the shortfall, which in other words means that all increased over the QE program period. While the increase in banknotes is not out of the ordinary and hence is unlikely to be related to the QE policy (third row, left panel), the other categories seem to be related to it, as discussed in the main text. Specifically, the liabilities to other euro area residents can explain roughly one third of the shortfall (third row, right panel), while liabilities to non-euro area residents can explain another fifth (fourth row, left panel). In total, these liabilities with other counterparties than euro area credit institutions explain more than half of the shortfall (fourth row, right panel). Importantly, the combined series tracks the shortfall very well in the medium horizon and even displays matching short-run dynamics in some periods, such as during 2019.

Counterfactual paths for excess reserves based on assumptions about balance sheet items. Turning to construction of a counterfactual path for excess reserves in the absence of the QE program, one can approach it either by accounting for developments in asset-side items, or developments in liability-side items (or both). To frame the discussion, the top row of Figure 1.27 shows the two simplest alternatives, which in effect ignore the information contained in the ECB's balance sheet. The first alternative is the most natural approach: since the most direct effect of the QE program on the ECB's balance sheet is its effect on the balance of assets under the APP, one could just assume that these would remain unchanged and hence subtract their increase from the increase in excess reserves. The figure demonstrates the point stressed in the main text and in the above paragraph: by accounting only for developments in APP balances,

Figure 1.26: Evolution of the ECB's balance sheet items



Notes: Cumulative changes in different items of ECB's balance sheet since February 2015. For asset side items the graphs display negative value of the cumulative changes. The shortfall corresponds to difference between cumulative change in the Asset Purchase Program is outstanding balance and cumulative changes in excess reserves.

one would reach the conclusion that excess reserves would become deeply negative in the absence of QE purchases.[37] Meanwhile, the second simple alternative is to simply assume that excess reserves without the QE would remain at the level prevailing before the start of QE purchases.[38] This approach assigns all the increase in excess reserves from the start of the QE program to the program itself. The problem with this approach is that it ignores all the other factors influencing the level of excess reserves. For example, this approach leads to multiple large jumps in the estimated effects of the QE on excess reserves, with the largest occurring in April 2017 (see alternative 1 in Figure 1.27). This jump is the result of the fourth round of TLTRO II loans and hence is unrelated to the QE program.

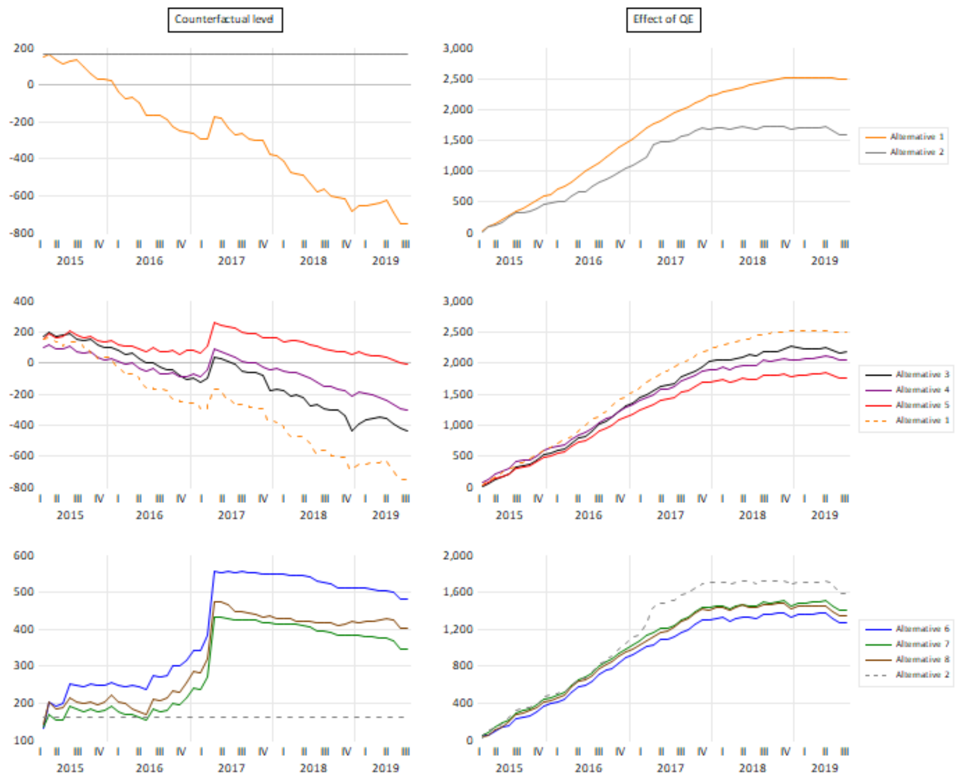
Since these simple alternatives fail to produce a plausible counterfactual path for excess reserves, one needs to enhance them by using information from the ECB's balance sheet. The second row of Figure 1.27 shows alternatives that improve upon alternative 1, while third row shows alternatives that improve upon alternative 2. Note that alternative 1 was in effect *subtracting* from the *observed* excess reserves throughout the sample, while alternative 2 was departing from the *initial* level of excess reserves and (potentially) *adding* to it. In view of their shortcomings, the way to improve alternative 1 is to subtract less from the observed excess reserves, while the way to improve alternative 2 is to add more to the initial level of excess reserves. These two closely options correspond to focusing on developments on the liability side and on the asset side of the balance sheet, respectively.

Focusing on the liability-based option, Figure 1.26 suggests a way forward. Since the shortfall is the mostly related to developments in liabilities, one should account for these developments when constructing alternative counterfactual path for excess reserves. Alternatives 3,4 and 5 do this by assuming that the liabilities to other euro area residents, liabilities to other non euro area residents, and banknotes would remain unchanged. Specifically, alternative 3 accounts only for the first

[37]The QE effect for alternative 1 is just the negative of the shortfall series in Figure 1.26.

[38]This is effectively assuming that no other balance sheet items would have changed during the QE period (or that changes on the asset and liability sides would cancel each other).

Figure 1.27: Alternative counterfactual paths for excess reserves



Notes: Level of excess reserves under alternative counterfactual assumptions (left panels), and the difference between actual and couterfactual levels of excess reserves (right panels).

balance sheet item, while alternative 4 accounts for both the first and second items and alternative 5 for all three items.

The figure clearly shows that accounting for the developments in the various liability-side items pushes the counterfactual excess reserves higher (and correspondingly, the effect of the QE lower) relative to alternative 1. This is intuitive: assuming that these liabilities wuld remain unchanged is the same as assuming that the increases in these liabilities are caused by the QE program itself; hence I subtract their increase from the amount by which observed excess reserves are decreased under the counterfactual with no the QE program, which leads to mark down in the estimate of the effect of the QE program on excess reserves. That said, even subtracting the increase in all three liability-side items - something that is problematic given that increase in banknotes was not caused by the QE program[39] - does not completely eliminate the problem: Even though alternative 5 remains positive throughout most of the sample, it still features negative excess reserves at the end of the sample.

The third row of Figure 1.27 shows the alternatives that add to the initial level of excess reserves, instead decreasing the amount subtracted from excess reserves. Again, the idea is to follow the analysis in Figure 1.26, which showed that the ECB's assets - and hence excess reserves - also increased due to other balance sheet items than the APP purchases. Specifically, alternative 6 takes the initial level of excess reserves before the start of the QE program and adds to them the increase in LTRO balances over the period of the program to account for the fact that these balances increased substantially over this period. The resulting counter-factual path for excess reserves records several step-wise increases throughout the period corresponding to successive rounds of TLTRO loans, with the largest occurring in March 2017, after which the counterfactual path of excess

[39]While banknotes would clearly also increase in the absence of the QE, this consideration potentially misses an important point: without QE-related increases in excess reserves, the increase in demand for banknotes might have lead to increase in MRO or LTRO balances and hence indirectly to increase in excess reserves. Hence accounting for changes in banknotes is not completely wrong, though it is impossible to establish what would have happened in the counterfactual world.

reserves hit its maximum of 556 billion. Corresponding to this, the effect of the QE program under this counterfactual becomes substantially smoother - corresponding to the fact that increases in LTRO balances lead to increases in excess reserves - a desirable feature given that QE purchases should have gradual effects on excess reserves.

While alternative 6 presents plausible counterfactual path for excess reserves in that it does not go negative, it likely errs in the other extreme. Specifically, the alternative does not account for change in MRO balances, decrease in which over the considered period was likely at least partially related to increase in LTRO balances, since banks treat the longer-term refinancing operations as partly substitutes for the shorter-term marginal refinancing operations. Correspondingly, alternative 7 also accounts for change in MRO balances, what results in a counterfactual path for excess reserves that is somewhat lower, and an estimated effect that is somewhat larger than under alternative 6. Finally, to complement these two alternatives, alternative 8 presents the results when one also accounts for change in other claims.

Counterfactual paths for excess reserves based on econometric models. Rather than determining the counterfactual path for excess reserves based on assumptions about developments in individual balance sheet items, one can use an alternative approach in the form of specifying and estimating an econometric model relating excess reserves to individual balance sheet items.

The starting point is realizing that all balance sheet items are related to each other by the basic accounting requirement that total assets (TA) are equal to total liabilities (TL). This means that excess reserves are related to other balance sheet items via simple identity:

$$
\begin{aligned}
TA_t &\equiv TL_t \\
TA_t &\equiv ER_t + TL_t^{OER} \\
ER_t &\equiv TA_t - TL_t^{OER}
\end{aligned}
$$

where TL_t^{OER} are total liabilities other than excess reserves. This identity is useful for understanding the meaning and validity of econometric models linking excess reserves to other balance sheet items. Consider, for example, the most straightforward approach of regressing excess reserves on APP balances:

$$ER_t = \beta_0 + \beta_1 APP_t + u_t \tag{1.11}$$

with associated coefficient estimates reported in first column of Table 1.28. Imagine one would use such an equation to predict the counterfacual path for excess reserves without the QE program. The error of the regression, u_t contains all assets other than APP balances and all liabilities other than excess reserves, which poses two problems. First, the coefficient will be unbiased only if all the balance sheet items contained in u_t are uncorrelated with APP. As discussed above and in the main text, this is clearly not the case for some of the liability series, which are positively (causally) related to APP balances and hence likely to cause negative bias in the coefficient. Second, the previous analysis also clearly highlighted that the balance sheet items contained in u_t are nonstationary, which means that the regression does not offer a useful tool for determining a counterfactual path for excess reserves, as the actual values can be arbitrarily far from the predicted values. This is illustrated in Figure 1.28, which shows predicted values for excess reserves in the absence of the APP program based on equation (1.11).[40] While the path overall is not implausible, it shares the drawback of the balance-sheet based alternatives that depart from observed excess reserves: by ignoring other, nonstationary factors

[40]Specifically, the coutnterfactual path is constructed by taking the observed excess reserves and substracting from them the *predicted* increase in excess reserves *due* to the increase in APP holdings (i.e. $\beta_1 APP_t$). Note that this is different from what the equation would predict as the value of excess reserves in the absence of APP (i.e. β_0). The difference comes from how one treats the error of the equation. The approach used includes the error in the counterfactual path, which is equivalent to assuming that other factors than APP influencing excess reserves, such as changes in LTRO balances, would occur in the absence of APP. The alternative approach would amount to assuming that APP is also responsible for these additional factors, which is an unappealing assumption. This discussion is analogical to the discussion of alternatives based on adding to the initial level of excess reserves or substracting from observed levels of excess reserves.

influencing excess reserves, it leads to implausibly low excess reserves at the end of the considered sample.

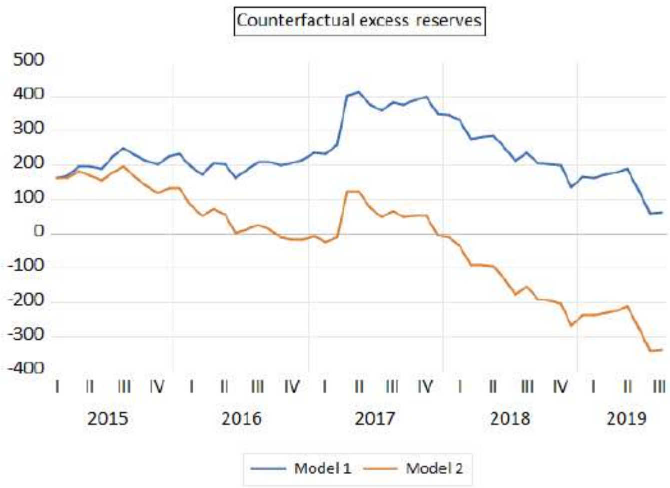
Table 1.28: Estimation results - ER models

	Model	
Coef.	**1**	2
β_0 (const.)	188*** (13.55) [13.92]	1520*** (69.56) [21.86]
β_1 (APP_t)	0.68*** (0.007) [95.25]	0.84 (0.01) [77.49]
β_2 $(LTRO_t)$	-	0.98*** (0.02) [41.39]
β_3 (MRO_t)	-	0.33** (0.15) [2.21]
β_4 (OL_t^{EU})	-	-0.76*** (0.04) [-19.97]
β_5 (OL_t^{NONEU})	-	-0.84*** (0.03) [27.05]
β_6 (OL_t^{NA})	-	-0.77*** (0.15) [-5.02]
β_7 $(Banknotes_t)$	-	-1.47*** (0.08) [-17.86]
R^2	0.97	0.99

See notes under Table 1.2 for explanation of all values.

In principle one could improve upon equation (1.11) by adding some (but not all) additional balance sheet items, so that one keeps in the error term only items

Figure 1.28: Model-based counterfactual paths for excess reserves



Notes: Level of excess reserves without the QE based on single-equation models (1.11) and (1.12).

not correlated with APP balances or those that cause nonstationarity of the error term. For example, one could estimate following regression:

$$ER_t = \beta_0 + \beta_1 APP_t + \beta_2 LTRO_t + \beta_3 MRO_t + \beta_4 OL_t^{EU} + \beta_5 OL_t^{NONEU} + \\ + \beta_6 OL_t^{NA} + \beta_7 Banknotes_t + u_t \tag{1.12}$$

where OL stands for other liabilities, with different geographical categorizations (EU, non-EU, and not allocated). The resulting coefficients in second column of Table 1.28.[41] Can this equation be used to determine the counterfactual path for excess reserves, assuming that the coefficient β_1 is econometrically valid and captures the causal effect of APP holdings on excess reserves? The answer is no, because the coefficient captures the effect of APP only ceteris paribus, i.e. holding all other regresssors constant. However, I argued in the main text and in text above that the other balance sheet items are causally related to APP holdings.

[41] Note that the coefficient on APP holdings did indeed increase, corresponding to the negative bias caused by other liabilities.

This means that the ceteris paribus assumption is invalid for the construction of the counterfactual path for excess reserves: if APP holdings were lower, then other liabilities would also be lower, while some of the asset items might be higher. This can be clearly seen in Figure 1.28, which shows the counterfactual path for excess reserves under the ceteris paribus assumption. Since the counterfactual exercise assumes that other liabilities (as well as assets) would remain unchanged, it leads to the conclusion that excess reserves would be negative. This mirrors the simple alternative based on balance-sheet analysis: excess reserves increased much less than one would expect based on APP purchases, mostly because other liabilities have increased.

The problem of the above approach clearly points towards the need to consider the evolution of balance sheet items as a system, rather than to focus on excess reserves in isolation. Specifically, if one wants to know what would happen to excess reserves in the absence of APP program, one also need to determine what would happen to other balance sheet items such as other liabilities. To do that, I specify a system of equations linking different balance sheet items together, which I estimate using the SUR method. I then use the estimated system of equations to predict values of different balance sheet items with and without the QE program.

The system of equations has 4 endogenous and 4 exogenous variables, reflecting assumptions about which balance sheet items respond to other balance sheet items, and which do not. The endogenous variables are excess reserves (ER), MRO balances, and two categories of other liabilities (EU and non-EU). Meanwhile, the exogenous variables are APP balances, LTRO balances, other liabilities not allocated, and banknotes.

The equations for the endogenous variables are:

$$\begin{aligned} ER_t =& \beta_{1,0} + \beta_{1,1} APP_t + \beta_{1,2} LTRO_t + \beta_{1,3} MRO_t + \beta_{1,4} OL_t^{EU} +; \\ &+ \beta_{1,5} OL_t^{NONEU} + \beta_{1,6} OL_t^{NA} + \beta_{1,7} Banknotes_t + u_{1,t} \qquad (1.13) \\ MRO_t =& \beta_{2,0} + \beta_{2,1} APP_t + \beta_{2,2} LTRO_t + \beta_{2,3} OL_t^{EU} + \end{aligned}$$

$$+ \beta_{2,4} OL_t^{NONEU} + \beta_{2,5} OL_t^{NA} + \beta_{1,6} ER_t + u_{2,t} \tag{1.14}$$

$$OL_t^{EU} = \beta_{3,0} + \beta_{3,1} APP_t + \beta_{3,2} LTRO_t + u_{3,t} \tag{1.15}$$

$$OL_{SA,t}^{NONEU} = \beta_{4,0} + \beta_{4,1} APP_t + \beta_{4,2} LTRO_t + u_{4,t} \tag{1.16}$$

where SA indicates that non-EU other liabilities include a seasonal adjustment factor, since the series displays end-of-quarter and end-of-year seasonality. All equation errors are assumed to follow the AR(1) process.

The first equation captures the idea that excess reserves are the residual balance sheet item, reflecting changes in both asset-side items and liability-side items. Meanwhile, the other three equations capture the bahavior of various financial institutions. MRO balances respond to LTRO balances, reflecting substitutability between the two, to excess reserves (with a lag), reflecting the decrease in demand for borrowing from the ECB when banks have excess reserves. Finally, the two categories of other liabilities that are endogenous are assumed to respond to APP and LTRO balances.

The equation estimates are in Table 1.29 and all correspond to prior expectations. Excess reserves increase with APP balances almost 1-for-1, which is a slightly stronger effect than suggested by the single-equation models in Table 1.28. They also increase 1-for-1 with LTRO balances, and with changes in MRO balances. Meanwhile, all three categories of other liabilities lead to decreases in excess reserves, with large and statistically significant effects. Turning to other endogenous variables, MRO balances, the table highlights the importance of accounting for the negative effect of higher excess reserves on MRO balances. Similarly, the equation estimates for the two categories of other liabilities show a positive relationship between them and the APP balances, and in the case of non-EU liabilities, also between them and LTRO balances.

Figure 1.29 shows the fit of the model for the 4 endogenous variables, and provides endorsement of the estimated model. The model is able to predict the evolution of all 4 the dependent variables based on knowledge of the 4 exogenous

Table 1.29: Estimation results - System of equations

Coef.	Dependent variable ER	MRO	OL^{EU}	OL^{NONEU}
β_0 (const.)	1592.93 '(85.10) [18.72]	116.27 '(5.41) [21.49]	119.76 '(9.14) [13.11]	-77.65 '(23.95) [-3.24]
β_1 (APP)	0.87 '(0.01) [61.60]	-	0.1 '(0.00) [21.36]	0.09 '(0.01) [8.66]
β_2 ($LTRO$)	0.98 '(0.03) [33.33]	-0.07 '(0.01) [-5.02]	-	0.12 '(0.03) [3.32]
β_3 (MRO)	0.62 '(0.19) [3.35]	-	-	-
β_4 (OL^{EU})	-0.81 '(0.04) [-22.69]	-	-	-
β_5 (OL^{NONEU})	-0.97 '(0.06) [-16.91]	-	-	-
β_6 (OL^{NA})	-0.58 '(0.18) [-3.19]	-	-	-
β_7 ($Banknotes$)	-1.62 '(0.10) [-15.89]	-	-	-
β_8 (ER)	-	-0.04 '(0.00) [-9.49]	-	-
R^2	0.999	0.988	0.894	0.990

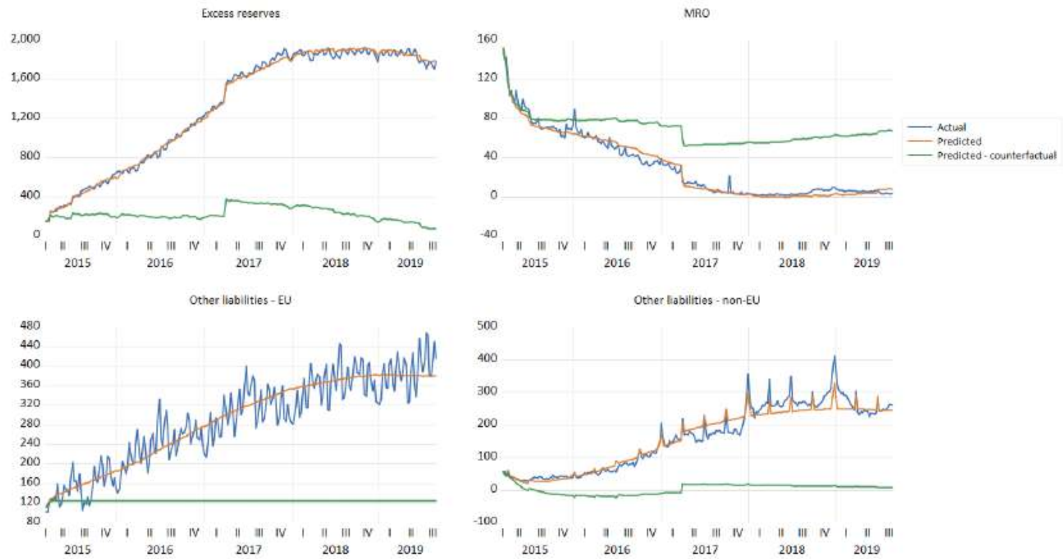
Notes: See notes under Table 1.2 for explanation of all values.

variables. While the ability to explain variation in excess reserves based on knowledge of all other relevant balance sheet items like in equation (1.12) would not be surprising, the ability to do so when 3 major drivers are treated as endogenous and hence as unknown, is a testament to the model's ability to predict excess reserves. This ability is reflected in the fact that the model is able to predict values of MRO balances and the two categories of other liabilities very well. In the case of MRO balances, the model correctly predicts that they decrease as APP and LTRO balances increase over the relevant sample. While the model is not able to explain week-to-week volatility of other liabilities to EU residents, it does explain the gradual increase in this balance sheet category over time, reflecting the fact that these liabilities increase with rises in APP balances. The variation in last endogenous variable - other liabilities to non-EU residents - is explained most weakly by the model, with prolonged periods of significant differences between actual and predicted values. However, even for this variable, the model is able to explain the secular trend, with gradual increase between 2015 and 2017, and stagnation afterwards. Again, this reflects the relationship with APP balances.

The relationship between the endogenous variables captured by the model play a crucial role in establishing the counterfactual path for excess reserves. As discussion of Figure 1.28 highlighted, one needs to know the values of other balance sheet items that respond to changes in APP balances or excess reserves to know what be the value of excess reserves in the absence of the QE program that lead to increase in APP balances. In this case, there are 3 ways in which the relationship between the 4 endogenous variables will influence the conclusions about counterfactual level of excess reserves:

- Since there is a negative relationship between MRO balances and (lagged) excess reserves, and a positive relationship between excess reserves and MRO balances, then accounting for changes in MRO balances due to changes in APP balances will lead to the conclusion that excess reserves would be higher than if the effect is ignored.

Figure 1.29: System of equations - Predicted values



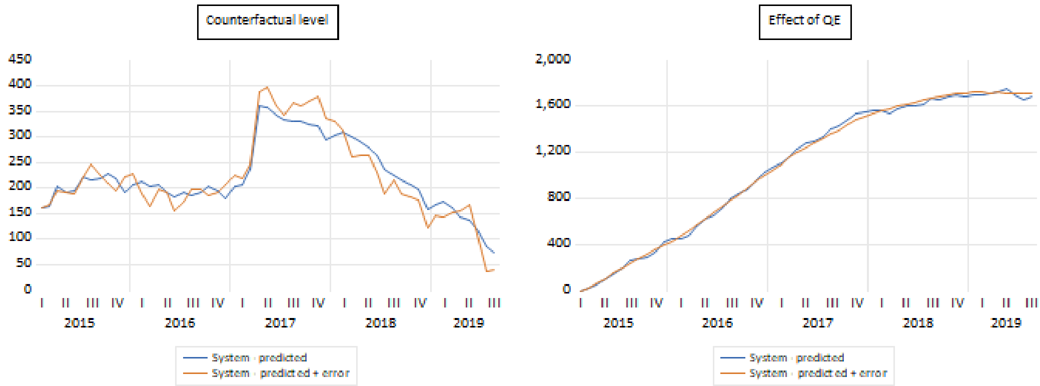
Notes: Actual and predicted values of endogenous variables based on equations (1.13-1.16), prediction sample 2015W9-2019W32. Predicted-counterfactual are values predicted in the absence of increases in APP balances).

- Since there is a positive relationship between both categories of other liabilities and APP balances, then accounting for changes in these categories due to changes in APP will lead to the conclusion that excess reserves would be higher than if the effect is ignored.

Figure 1.29 also shows the predicted values of all 4 endogenous variables in the absence of the QE program, which would mean that APP balances would remain at their pre-QE level. The main conclusion - that excess reserves would be much lower - is not surprising. The conclusion that the MRO balances would be much higher, while both categories of other liabilities would be much lower, is more surprising and important. All of these three factors imply that excess reserves would be higher then if the channels were ignored. Crucially, these three channels combined mean that the counterfactual excess reserves in the absence of the QE program are no longer predicted to be negative, unlike in Figure 1.28. Specifically, Figure 1.30 shows the counterfactual path for excess reserves and the corresponding estimated effect of the QE on excess reserves based on the system of equations, either with or without added prediction error (see footnote 40 for discussion of these two alternatives). The counterfactual path has the best features or previous paths. It avoids going negative, like the alternatives based solely on assumptions about liability items of the balance sheer (Figure 1.27) or the single-equation multiple-regression alternative (Figure 1.28). It also predicts increase in excess reserves due to increases in LTRO balances, even though it does not start from the observed excess reserves.

Conclusion. This Appendix presented several alternative counterfactual paths for excess reserves. Many of those paths are implausible, either because of their assumptions (e.g. excess reserves would remain unchanged) or because they violate the zero-lower bound on excess reserves (e.g. alternatives based on assumptions about liability-side balance sheet items). The main text used two alternatives which are most plausible a-priori, and one of which also provides a reasonable upper bound for the counterfactual level of excess reserves. The first, more plau-

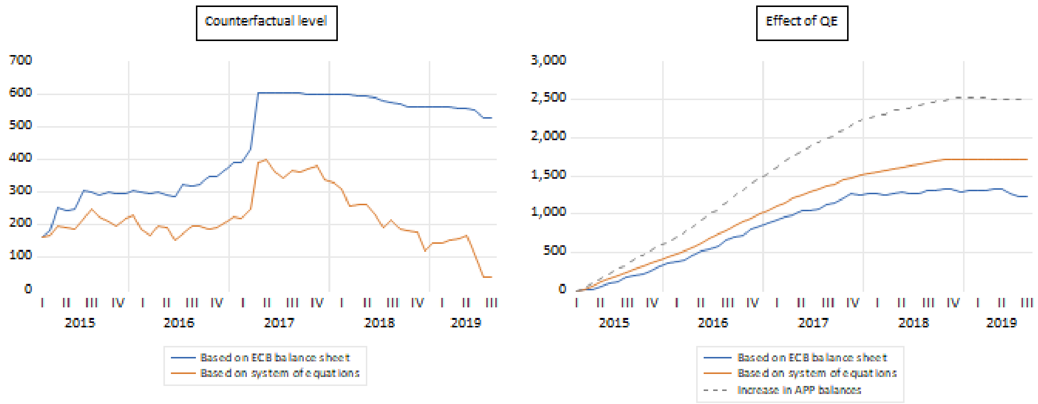
Figure 1.30: Model-based counterfactual paths for excess reserves



Notes: Level of excess reserves without the QE based on a system of equations (1.13-1.16).

sible, is based on a system of equations model. The second, more conservative, is based on the assumption that excess reserves would change from their initial level only because of changes in LTRO balances. The two alternatives are replicated in Figure 1.31.

Figure 1.31: Counterfactual paths for excess reserves



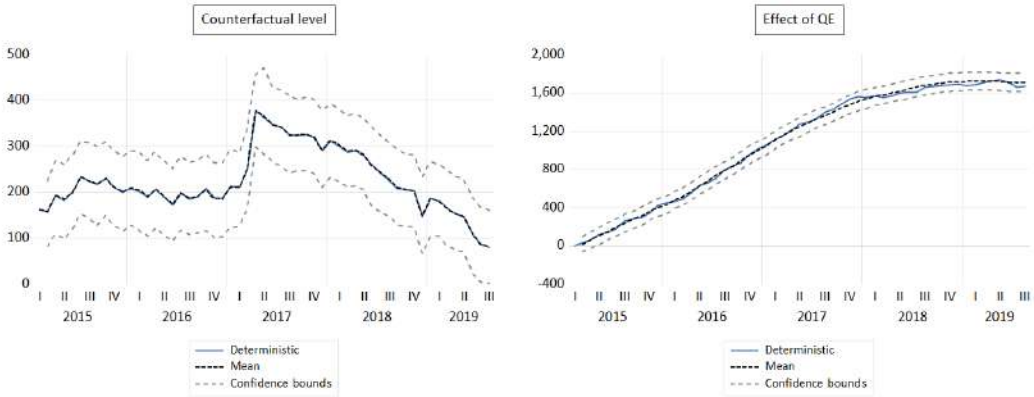
Notes: Level of excess reserves without the QE based. Alternative 6 and alternative based on system of equations.

1.E Confidence bounds for the effect of the QE program

This appendix provides confidence bounds for estimates of the effect of the QE program on the excess reserves and IIRs, and hence provides sense of statistical significance of the estimated effects presented section 1.5. All intervals are two-sided at 95% confidence and were obtained via simulation. The confidence intervals for the IIRs account for both sources of uncertainty, i.e. uncertainty originating in the effect on excess reserves and uncertainty originating in the effect on IIRs.

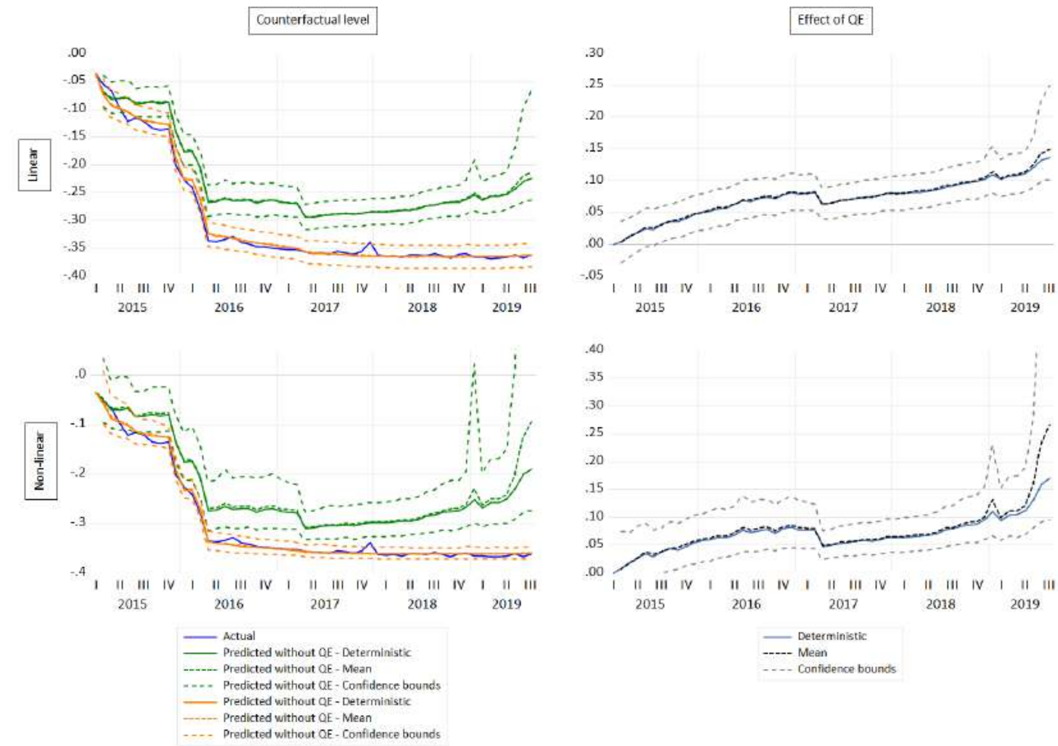
Focusing first on effect on excess reserves, left panel of Figure 1.33 displays the estimated counterfactual level of excess reserves that would prevail in absence of the QE program as estimated via the system of equations approach, reproduced from Figure 1.31. Apart from the deterministic solution outcome, the figure also includes the stochastic mean, and lower and upper confidence bounds corresponding to 5% and 95% percentile. Right panel then shows the corresponding quantities for the effect of the QE program on excess reserves. The figure shows that the confidence bounds around are slightly less than 100 billion euro.

Figure 1.32: Confidence bounds for counterfactual paths for excess reserves



Turning to effect on IIRs, the left panels of Figure reproduce results from the Figure 1.17 from the main text, this time including confidence bounds. Meanwhile, the right panels reproduce information from Figure 1.18. The Figure clearly demonstrates that the effects of the QE program on the Eonia rate are statistically significant, with confidence bounds of around 2bps compared with effect size that is 3-4 times larger.

Figure 1.33: Confidence bounds for counterfactual paths for Eonia rate



1.F Statistical test

This appendix presents the results of formal statistical tests supporting claims made in the text. The first subsection presents tests supporting the notion that there was a break in the Eonia rate the spread series. The second subsection provides evidence that the series used in the linear model are nonstationary and cointegrated, supporting the validity of the use of estimation methods presented in this chapter. The third subsection presents the results of statistical tests on various coefficient restrictions employed in the models.

1.F.1 Break test for the Eonia rate spread

The main text argued that the Eonia rate, and specifically the spread between the Eonia rate and the MRR (Eonia spread for short) underwent a change in regime. Figure 1.34 shows the spread between the Eonia rate and the main refinancing rate (MRR), which clearly indicates a break in the series sometime after the beginning of October 2008, when the ECB switched to a full allotment policy and excess reserves appeared for the first time. Table 1.30 supports this hypothesis with multiple breakpoint statistical tests. The first three rows show the results of a standard Chow breakpoint test corresponding to three alternative timings for the occurrence of breaks, while the following three rows correspond to the Chow forecast test. In all cases the tests overwhelmingly reject the null hypothesis of no break. To complement these tests with a known break date, the table also reports results from the Quandt-Andrews test, which does not focus on a concrete break date, but rather tests all possible break dates and then evaluates the test results based on maximum or average for all dates. Even allowing for unknown breaks does not change the conclusion. An additional advantage of the Quandt-Andrews test is that it allows for evaluation of the timing when the break occurred based on associated test statistics. Figure 1.35 reports these and clearly highlights that the break did occur sometime at the end of 2008 or in the beginning of 2009, further supporting the conclusions of the main text.

Figure 1.34: Eonia spread

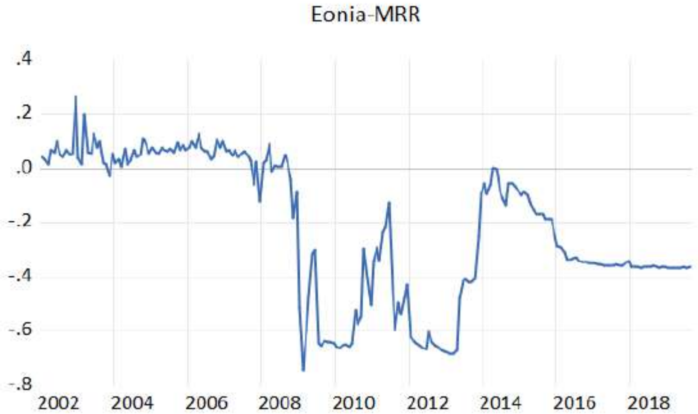


Figure 1.35: Quandt-Andrews test statistics

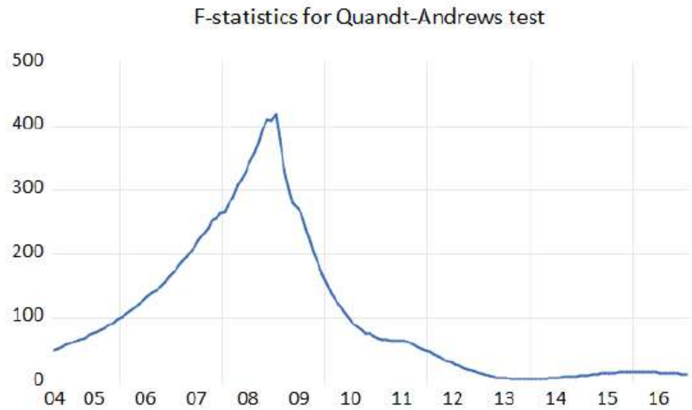


Table 1.30: Breakpoint tests for the Eonia rate spread

		Statistic		
Test type	Date	F	LR	Wald
Chow	2008M10	397.68 <0.0000	225.26 <0.0000	397.68 <0.0000
Chow	2008M11	410.65 <0.0000	229.74 <0.0000	410.65 <0.0000
Chow	2009M01	417.79 <0.0000	232.16 <0.0000	417.79 <0.0000
Chow - Forecast	2008M10	48.03 <0.0000	928.06 <0.0000	-
Chow - Forecast	2008M11	46.52 <0.0000	917.17 <0.0000	-
Chow - Forecast	2009M01	34.73 <0.0000	847.87 <0.0000	-
QA - Maximum	Uknown	-	417.79 <0.0000	-
QA - Exponential	Uknown	-	203.93 <0.0000	-
QA - Average	Uknown	-	104.26 <0.0000	-

Notes: Date indicates the start of new regime. Statistics are F-statistics (F), Log-likelihood ratios (LR) and Wald statistics (Wald). The first value refers to the value of test statistic, the second to associated p-value.

1.F.2 Stationarity and cointegration tests

The main text claimed that the series used in the modeling were nonstationary and cointergrated. This appendix supports this claim with formal statistical tests, focusing on the Eonia rate. Table 1.31 shows the results for unit root tests for all three series included in equation (1.3). It clearly shows that, according to formal statistical tests, all three variables can be considered nonstationary in the sense that they include a unit root. For the two interest rates, all the tests with a null hypothesis of unit root fail to reject this hypothesis with a significant margin, and the one test with no unit root as the null hypothesis rejects this null hypothesis. In the case of excess reserves, the conclusions are not as uniform: while the majority of tests do lead to the conclusion of the presence of a unit root, there are two tests (PP and ERS) that lead to the opposite conclusion.

Table 1.32 complements the results of unit root tests by showing that the three series together form a cointegrated set of variables, which justifies the use of the estimation methods presented in the main text, especially the estimation methods exploiting the cointegration between the series. Specifically, Table 1.32 shows the results of standard single equation tests of Engle and Granger, and Phillips and Ouliaris, which overwhelmingly reject the null hypothesis of no cointegration among the three series at multiple different considered samples.

An alternative to single equation tests for cointegration are tests based on system estimation. Note that in the current environment, single equation methods seem more suitable, given that there is clear one-way causality from the DR and excess reserves to the Eonia rate, so that there is little ambiguity about what the dependent variable in the cointegration relationship should be, which is typically considered the main drawback of single equation tests. Similarly, the presence of multiple cointegration relationships seems implausible, a second common reason for using system cointegration tests. Still, for the sake of completeness, I complement the single equation tests with tests based on system estimation, with results reported in Table 1.33. The results strongly support the presence of cointegra-

Table 1.31: Unit root tests

Test type	Eonia	DR	ER
	Variable		
ADF	-1.33 0.62	-1.64 0.46	-2.55 0.30
PP	-1.39 0.58	-1.57 0.50	-7.62 <0.0000
DFGLS	-0.17 >0.1	0.02 >0.1	-2.37 >0.1
ERS	26.15 >0.1	18.64 >0.1	-3.85 <0.01
NP1	-0.43 >0.1	-0.97 >0.1	-4.49 >0.1
NP2	-0.25 >0.1	-0.49 >0.1	-1.48 >0.1
NP3	0.58 >0.1	0.50 >0.1	0.33 >0.1
NP4	21.55 >0.1	15.88 >0.1	20.18 >0.1
KPSS	1.42 <0.01	1.24 <0.01	0.16 <0.05

Notes: Variables are the Eonia rate, the DR (DR) and logarithm of excess reserves (ER). Tests refer to Augmented Dickey-Fuller (ADF), Philips-Perron (PP), Dickey-Fuller Test with GLS Detrending (DFGLS), Elliot-Rothenberg-Stock Point Optimal (ERS), four types of Ng-Perron (NP1-NP4), and Kwiatkowski-Phillips-Schmidt-Shin (KPSS) tests. See Eviews manual for more details. In the case of the Eonia rate and the DR, the specifications includes a constant, while in the case of excess reserves, it includes a trend and a constant. Eviews default settings are used for all other settings. The first value refers to the value of test statistic, the second to the associated p-value, when available, or indicates whether the test statistic exceeds one of the 10%,5% or 1% thresholds. Sample is 2002M01-2019M08.

Table 1.32: Single equation cointegration tests

		Test type			
Sample	Statistic	EG - AIC	EG - SIC	PO - AIC	PO - SIC
2002M01	Tau	-3.41	-3.41	-11.12	-11.96
-		0.12	0.12	<0.0000	<0.0000
2019M18	Z	-25.73	-25.73	-197.10	-237.96
		0.05	0.05	0.00	0.00
2009M01	Tau	-3.67	-3.67	-4.38	-4.38
-		0.07	0.07	0.01	0.01
2019M18	Z	-29.81	-29.81	-33.69	-33.69
		0.02	0.02	0.01	0.01
2015W12	Tau	-3.62	-3.13	-9.64	-9.64
-		0.07	0.20	<0.0000	<0.0000
2019W36	Z	-31.26	-18.94	-148.19	-148.19
		0.02	0.17	<0.0000	<0.0000

Notes: Tests refer to Engle-Granger (EG) and Phillips and Ouliaris (PO) tests, with either Akaike or Schwarz information criterion used to select the order of lag. See the Eviews manual for more details. The cointegration equation has the Eonia rate as the dependent variable, and the DR and logarithm of excess reserves as indepndent variables, and includes a constant but not a trend. Eviews default settings are used for all other settings. The first value refers to the value of the test statistic, the second to the associated p-value.

tion between the three series and hence confirm the conclusions based on single equation tests.

Table 1.33: System cointegration tests

		Test type	
Trend case	Lags	Trace	Maximum
B	0	73.53	53.02
		<0.0000	<0.0000
		2	2
	1	53.52	35.39
		0.00	0.00
		1	1
	2	75.98	57.67
		<0.0000	<0.0000
		1	1
C	0	52.78	32.44
		<0.0000	0.00
		2	2
	1	46.36	28.64
		0.00	0.00
		3	1
	2	67.96	49.71
		0.00	0.00
		3	1

Notes: Tests refer to Johansen cointegration tests based on the rank of the estimated matrix of coefficients, using either trace or maximum eigenvalue. See the Eviews manual for more details. Trend case B does not include a deterministic trends in the level of the series while Trend case C does include a deterministic trend in the level of the series. In either case, the cointegration equation includes a constant. Eviews default settings are used for all other settings. The first value refers to the value of the test statistic for hypothesis of no cointegration among the series, second to the associated p-value and third to the overall number of cointegrating vectors suggested by the test. Sample 2015M03-2019M08.

1.F.3 Coefficient restrictions

The postulated models assume several coefficient restrictions relating to coefficients on different policy rates and on excess reserves. In this appendix, I provide empirical support for these restrictions, first focusing on restrictions relating to models estimated on the robust sample, and then on models estimated on the full sample.

First, for the models estimated on the robust sample - see equation (1.6) - there is a single relevant restriction, since these models cover only one of the two postulated regimes. Specifically, the models assume that the level of the other policy rate, the main refinancing rate, is irrelevant for the level of the IIRs. The model written without the restriction is:

$$IIR_t = \beta_0 + \beta_1 DR_t + \beta_2 log(ER_{t-1}) + \beta_3 MRR_t + \epsilon_t \tag{1.17}$$

Therefore, the restrictions amount to settings $\beta_3 = 0$. The results for this hypothesis are reported in Table 1.34. The results provide mixed support for the hypothesis, depending on what estimation method one uses. For basic estimation methods, the null hypothesis of a zero coefficient on the MRR cannot be rejected at any reasonable level of significance. However, for the three cointegration estimation methods, the hypothesis results are not so clear cut: for FMOLS the null hypothesis can be rejected at levels of significance just above the usual threshold, while for ARDL it is just below this threshold and for CCR it is not too far from this threshold. This would seem to suggest that the coefficient restriction is not completely supported by the data. However, for all the cointegration methods, the inclusion of the MRR leads to nonsensical coefficient estimates, with the coefficient on the MRR being negative, indicating that the IIRs increase when the MRR decreases. This is clearly at odds with theory and likely reflects the fact that the MRR nearly did not vary at all during the relevant sample, so that its relevance is hard to establish on this sample. This notion is further supported by the fact that the equivalent hypothesis is not rejected when the same estimation method is used on the full sample; see next paragraph. Overall, I conclude that the weight of evidence supports the coefficient restriction.

For the models estimated on the full sample - see equation (1.3) - multiple

Table 1.34: Coefficient restrictions - hypothesis tests (single regime model)

	Statistic	
Method	F	Chi-square
OLS	0.00 0.96	0.00 0.95
TSLS	0.69 0.41	0.69 0.41
ARDL	2.95 0.06	5.91 0.05
FMOLS	4.88 0.03	4.88 0.03
CCR	2.27 0.14	2.27 0.13

Notes: Statistics correspond to a standard F-test and a Chi-square test; see the Eviews manual for more details. The first value refers to the value of the test statistic, the second to the associated p-value. Estimation sample 2015M03-2019M08.

coefficient restrictions have been applied, corresponding to multiple regimes. In the normal regime, the model assumes that the IIRs depend only on the MRR, and not on the DR and excess reserves. Moreover, the postulated model assumes that the coefficient on the MRR was 1, which is equivalent to assuming that the spread between the IIRs and the MRR does not change with the level of the MRR. Without these restrictions, the full equation can be written as follows:

$$IIR_t = \beta_{10} + \beta_{11} MRR_t + \beta_{12} DR_t + \beta_{13} g(ER) \tag{1.18}$$

The restrictions are then following: $\beta_{11} = 1$, $\beta_{12} = 0$ and $\beta_{13} = 0$. These can be tested together or individually. With respect to the excess reserve regime, the only relevant restriction applies to the MRR, as discussed above. Without the restriction, the regime equation is:

$$IIR_t = \beta_{20} + \beta_{21} DR_t + \beta_{22} g(ER_t) + \beta_{23} MRR_t + \epsilon_t \tag{1.19}$$

and the restriction amounts to setting $\beta_{23} = 0$.

Table 1.35 presents the results for testing these hypothesis. The results overwhelmingly support the coefficient restrictions imposed in the main text, with the exception of one combination of hypotheses relating to the normal regime; moreover, the restrictions imposed in the main text are supported as a set.

Table 1.35: Coefficient restrictions - hypothesis tests (two regime model)

Hypothesis		**Statistic**	
Tested	Maintained	F	Chi-square
$\beta_{11} = 1$	$\beta_{12} = \beta_{13} = 0$	0.85 0.36	0.85 0.36
$\beta_{12} = 0$	$\beta_{11} = 1$	0.25 0.62	0.25 0.62
$\beta_{13} = 0$	$\beta_{11} = 1$	2.10 0.15	2.10 0.15
$\beta_{13} = 0$	$\beta_{11} = 1, \beta_{12} = 0$	3.75 0.06	3.75 0.05
$\beta_{12} = 0$	$\beta_{11} = 1, \beta_{13} = 0$	1.41 0.24	1.41 0.24
$\beta_{12} = 0, \beta_{13} = 0$	$\beta_{11} = 1$	4.66 0.01	9.31 0.01
$\beta_{11} = 1, \beta_{12} = 0, \beta_{13} = 0$	-	1.98 0.12	5.95 0.11
$\beta_{23} = 0$	-	0.51 0.48	0.51 0.48

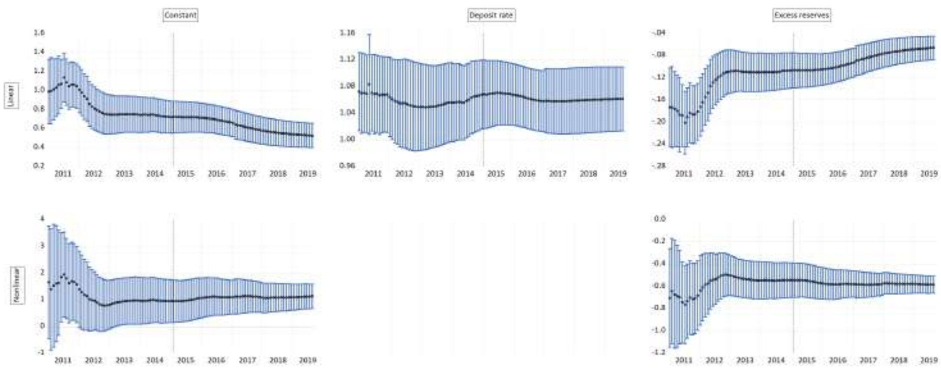
Notes: Statistics correspond to a standard F-test and a Chi-square test; see the Eviews manual for more details. The first value refers to the value of the test statistic, the second to the associated p-value. Equation estimated using FMOLS, estimation sample 2002M01-2019M08 if $D_t^{ER} = 0$.

1.F.4 Model stability

This appendix provides support for the claim that the model coefficients did not change with the advent of the QE program, so that the use of the model for constructing paths for IIRs with and without the program is valid.

Figure 1.36 shows the recursive coefficient estimates for all the model coefficients in the excess reserve regime. The top panels show the results for the linear model, while the bottom panel shows the results for non-linear model. In neither cases is there apparent break in the coefficients around the start of the QE program. While the estimated coefficients for linear model display some evolution in the later part of the sample covering the QE program, it is not significant and is likely related to evolution of excess reserves rather than to the QE program itself (see discussion in next chapter).

Figure 1.36: Recursive coefficient estimates



Notes: Recursive coefficient estimates with 2-standard deviations error bands.

Chapter 2

Forecasting euro zone interbank interest rates in the presence of excess reserves

2.1 Introduction

Forecasts for interbank interest rates (IIRs) are among the most frequently used macroeconomic forecasts, partly owing to their popularity in credit and market risk models (see e.g. Moore, Wurst, and Cramer (2019)). Moreover, small variations in forecasts for IIRs can be of great importance: the most commonly used transformation for IIRs is the spread between IIRs and a risk-free benchmark, which historically has relatively low volatility. Taken together this implies that forecasting performance of models for IIRs is of particular interest.

This chapter studies the forecasting performance of the novel structural model proposed in the first chapter. The central theme of that chapter is that IIRs underwent a change in regime due to emergence of excess reserves, which have been present in the euro zone nearly continuously since fall of 2008. The chapter shows that the structural model is important in making sense of movements in IIRs over the relevant sample, and explains these movements in terms of novel features of the model. This raises the question of whether the model is also useful in terms of out-of-sample forecasting when compared with alternative reduced form models, a possibility suggested by the nature of the novel features of the model.[1]

This chapter evaluates the forecasting performance of the structural model by comparing its performance with the universe of plausible benchmark models of the reduced form category. The chapter shows that, in ex-post pseudo out-of-sample forecasting exercises, the structural model performs substantially better than any class of reduced form models, and in most cases better than any individual reduced form model. For example, in the case of the Eonia rate and 24 month horizon the non-linear structural model has a mean average error (MAE) of 0.035, while the best reduced form model has an MAE of 0.095, and the best model class of reduced form models has an MAE of 0.14. The results for 1-week and 3-month Euribor rates are qualitatively similar, though the difference between the structural model and reduced form models is not as stark. Overall, this highlights the importance of accounting for structural factors in forecasting IIRs.

There are two caveats to these results. First, in ex-post forecasting, some reduced form models perform better than structural models at the shortest horizons of 1 and 2 months, which can be explained by the equilibrium nature of the structural model: Since it can take time for this equilibrium relationship to assert itself, and for structural forces to change significantly, models that are close to random walk can achieve better performance over short horizons.

Second and more important caveat is that the superior performance from ex-post forecasting applies only partly in ex-ante forecasting. This is an unsurprising conclusion, given that an important factor behind the difficulty of forecasting

[1]Specifically, there are two aspects of behavior of IIRs that are unlikely to be captured by usual time series models, leading to suboptimal forecasting performance. First, the presence of excess reserves changes the risk-free interest rate which anchors IIRs from the the main refinancing rate to the deposit rate. This means that models simply linking IIRs to main refinancing rate will struggle to make correct forecasts for IIRs in periods when the two policy rates either do not move in lockstep or when the difference between them is substantially different from historical averages, both of which have been true for the last several years. Second, apart from the presence of excess reserves, the IIRs also depend (negatively) on the *amount* of excess reserves, whenever excess reserves are present. This means that models that do not incorporate variations in the excess reserves will struggle to correctly forecast the variations in IIRs even when excess reserves vary over time.

market interest rates in general is having reliable out-of-sample forecasts for the conditioning monetary policy variables. In the specific case of *interbank* interest rates, it is the forecasts for excess reserves, and to a lesser degree monetary policy rates, that are substantial source of ex-ante forecast errors. That said, this chapter provides a quantitative information on the trade-off between having a more correct but more complex structural model and having a simpler but less correct model in the specific case of interbank interest rates. A key takeaway is that the structural model is better tool for forecasting interbank interest rates ex-ante as long as excess reserves can be forecast well. This is the case if the central bank is operating quantitative easing program, which is both the main driver of excess reserves and easily predictable due to the forward-looking nature of central bank announcements.

Benchmarking the structural model against reduced form models is not the primary goal of this chapter. Rather, the benchmarking provides a basis for analysis of what model features are important for forecasting IIRs in the presence of excess reserves. This analysis relies on a dual approach. First, I estimate and evaluate the whole universe of (linear) reduced form models, which allows me to draw conclusions about which features of reduced form models are correlated with good forecasting performance; second, I rely on the superior performance of the structural model to interpret why particular features are important. The answer to this question lies in the statistical nature of the series of interest. Specifically, the IIRs are both non-stationary and co-integrated with monetary policy variables, as noted in the first chapter and discussed in further here.

Correspondingly, the fourth section of this chapter shows how the suboptimal nature of forecasts from usual time series models can be explained by the statistical features of the data. First, models that ignore the nonstationarity of IIRs force the spread between IIRs and policy rates to return to historical averages. While this is a good forecasting rule for a period before the emergence of excess reserves, it is widely at odds with the behavior of IIRs over last decade. For example, due to historically high levels of excess reserves, IIRs are systematically below the main refinancing rate, something that was never observed before 2008. On the other hand, models cast in first differences, as the stationary transformation of the modeled series, ignore the cointegration relationship, and especially the role of excess reserves. This means that such models effectively assume that the spread between IIRs and policy rates will remain at current levels. With exception of period from 2017 to 2019, this assumption is at odds with the behavior of IIRs. Not only did the secular level of spreads change significantly throughout the 2010s thanks to wild fluctuations in excess reserves, but more importantly, the spreads also feature many transitory movements. Prime example of such transitory movements are jumps during periods of stress in financial markets. Models in first differences make these transitory jumps in spreads permanent, which amounts to permanent periods of stress, clearly a nonsensical forecast.

This chapter is related to literature on IIRs; see Green et al. (2016) for a recent review of this literature. However, this chapter is different from the existing literature in two important aspects. First, almost all of the papers reviewed by Green et al. (2016) focus on high frequency movements of overnight interest rates. In contrast, this chapter focuses on medium- and long-term movements in both overnight interest rates and IIRs with longer maturity. Second, it focuses on forecasting of IIRs, rather than on analysis of the relationship between the behavior of IIRs and the conduct of monetary policy. To my knowledge there is no other paper focusing primarily on the forecasting performance of alternative models of IIRs.[2] This is not completely surprising given that, prior to emergence of excess reserves, forecasting IIRs was almost equivalent to forecasting policy rates, since the latter anchored the former almost perfectly. However, with the emergence of excess reserves, this is no longer true. I attempt to fill in the gap in literature by studying the forecasting performance of alternative models and linking it to

[2]The review by Green et al. (2016) does not include a section on forecasting IIRs. Meanwhile, a google scholar search conducted in March 2021 using key words "*eonia forecasting*", "*overnight forecasting*", "*interbank forecasting*" or "*money market forecasting*" does not result in identification of any relevant research articles. The only partial exception is Marquez, Morse, and Schlusche (2013), who create several scenario forecasts for US IIRs conditional on different paths for excess reserves.

structural factors.

The rest of this chapter is divided into 4 sections. First, I present and briefly discuss alternative models used for forecasting IIRs. The forecasting performance of these models is compared in section 3. The section discusses which model features correlate with good forecasting performance and highlights the superior forecasting performance of the structural model over reduced form models. Section 4 then explains these observations by linking them to the econometric nature of the time series under analysis. Specifically, it shows how limitations of different reduced form models are linked to the nonstationarity and co-integration nature of the time series, and how these features are captured by the structural model. The last section concludes.

2.2 Alternative models

This section presents the alternative models considered in this chapter. Two goals guide my selection of models. First, given the absence of a single natural benchmark model[3] - and the absence of competitor models, as discussed in the introduction - I consider the full universe of plausible benchmark models so as not to influence my conclusions by the (un)fortunate selection of a benchmark model. Second, considering all plausible models allows me to analyze what model features are important for forecasting interbank interest rates (IIRs) in the presence of excess reserves. This will then form the basis of discussion focused on the *explanation* of the forecasting performance of various models, which constitutes the main goal of this chapter.

The section starts with a discussion of the simplest univariate models, proceeds to single-equation multivariate models, and finally turns to multiple-equation models. All the models are presented in the table at the end of this section, which also

[3]For example, macroeconomic literature and practitioners often use simple autoregressive model of second order as a benchmark model; see Edge, Kiley, and Laforte (2010) for an example of the former and further citations, and Ciccarela and Kovar (2020) for an example of the latter.

includes the equation representations of the models.

2.2.1 Univariate models

The simplest time series models are univariate models. Even within this basic class of models, one can consider multiple alternative sub-categories of models, depending on whether the dependent variable in the model is the level or first differences of the underlying variable. In order to explore the complete universe of possible models, I consider both options. The other question is which, and how many, ARMA terms one should include. I rely on the now common approach of automatically selecting ARMA terms based on information criteria, namely the Akaike Information Criterion (AIC) and the Schwarz Information Criterion (SIC).

2.2.2 Multivariate single-equation models

The second class of models I consider are multivariate single-equation models, which include information about additional variables than just the forecasted variable. The inclusion of these models is motivated by the nature of IIRs: movements in IIRs are mostly caused movements in interest rates set by central banks, and other monetary policy variables. Hence, it is important to also consider models that include information about these variables.[4]

Within this class of models, I consider two groups of models, reduced-form models and structural models. The latter group consists of variations of the structural model presented in the first chapter.

Reduced-form models. The starting point for multivariate single-equation models are models that link the level of IIRs to the level policy rate(s). The simplest model is a static regression linking the current level of IIRs to the current

[4]This is for two reasons. First, conditional on knowing the policy rates, univariate models that do not incorporate this information are very inefficient in forecasting IIRs. The second reason is conditional inconsistency: when one forecasts both policy rates and IIRs, univariate models will lead to forecasts that are potentially inconsistent with the forecast for policy rates.

level of policy rates, e.g.:

$$IIR_t = \beta_0 + \beta_1 MPR_t + a_t \tag{2.1}$$

There are two policy rates that one can consider: the main refinancing rate (MRR) and the deposit rate (DR). While prior to 2009 it was feasible to include only one of the two policy variables, since they moved in perfect lockstep, in the sample starting from 2009, both can be included.

Of course, the model in equation (2.1) is likely to suffer from strong autocorrelation in residuals, i.e. it does not capture some persistence in the dependent variable, compromising its forecasting performance. There are two alternative ways to address this problem in terms of model structure. Either one can include lag(s) of the dependent variable, creating what is commonly called an ARMAX model. Alternatively, one can include ARMA errors instead of lags of dependent variables. The difference is whether transitory movements in the independent variable will have persistent effects or not, with persistent effects being present in ARMAX models. To complement models that include persistence via inclusion of ARMA components, I consider models that include persistence via lagged values of independent variables in addition to lagged values of dependent variables, which are commonly called autoregressive-distributed-lag (ARDL) models . Whatever the form of ARMA components is included, one needs to select the order of ARMA components, and decide whether to difference the dependent variable or not. As in the case of univariate models, I consider both models in level and differences, and rely on information criteria to select ARMA components.

Finally, another class of multivariate single equation models are univariate models that use a spread between IIRs and one of the policy rates as the dependent variable. These are effectively the same as those discussed in the previous paragraph with the coefficient on the concurrent policy rate set equal to 1. While in principle such a coefficient restriction might be inefficient from the estimation perspective, it might be beneficial from a forecasting perspective, and indeed such (implicit) coefficient restriction is often used in practice. Therefore, it is valuable to include this model class among those considered.

In addition to these common reduced-form models, I also consider less common models that allow for a break in the relationship between IIRs and policy rates. These can be considered a mid-way between a pure reduced form model and the actual structural model. Specifically, I consider three types of models in levels - static regression, a model with ARMA errors, and an ARMA model in spread - and allow them to have different coefficient estimates before and after the emergence of excess reserves. To determine whether two sets of coefficients should be used, I rely on a standard Chow break test with 2008M10 as the break date. Rather than searching for an ARMA structure with the previous models - something not common in combination with breaks - I impose an ARMA(1,1) structure on an ex-ante basis.

Structural model. The structural model proposed in the first chapter also belongs to the class of multivariate single equation models. There are two differences from the models discussed above, both following from theoretical considerations supported by data analysis presented in the first chapter. First, the model imposes structure on the link between the IIRs and policy rates, which amounts to imposing multiple coefficient restrictions. Moreover, the structure relies on the presence of multiple regimes and an exogenous threshold variable. These factors combined can improve forecasting performance by not relying on potentially noisy estimation procedures to determine coefficients or breaks. Second, the model also includes another variable linked to monetary policy, the amount of excess reserves.

There are two formulations of the resulting model, one linear and one nonlinear. These are captured in equations below:

$$IIR_t = \begin{cases} \beta_{10} + MRR_t + \epsilon_t & \text{if } D_t^{ER} = 0 \\ \beta_{20} + \beta_{21} DR_t + \beta_{22} log(ER_t) + \epsilon_t & \text{if } D_t^{ER} = 1 \end{cases} \tag{2.2}$$

$$f(IIR_t) = \begin{cases} IIR_t - MRR_t = \beta_{10} + \epsilon_t & \text{if } D_t^{ER} = 0 \\ log(IIR_t - DR_t) = \beta_{20} + \beta_{22} log(ER_t) + \epsilon_t & \text{if } D_t^{ER} = 1 \end{cases} \quad (2.3)$$

where D_t^{ER} is dummy variable indicating the presence of excess reserves. In both linear and nonlinear model, the first regime amounts to assuming that the spread between IIRs and MRR is constant. Meanwhile, the second regime links the IIRs to DR, rather than MRR, and to the logarithm of the excess reserves, and correspondingly is dubbed the excess reserve regime. The difference between linear and nonlinear models is that nonlinear model uses the logarithm of spread between IIR and DR as a dependent variable, instead of using the level of IIRs. The main motivation for this formulation is to prevent the IIRs from falling below the DR, which is systematically impossible in the sample considered.

2.2.3 Multi-equation models

I also consider multiple-equation models of the usual vector-autoregression (VAR) class. These models are popular for short-term forecasting and can serve as a useful benchmark, especially for ex-ante forecasting performance. That said, the value of additional equations is likely to be low in the current environment: IIRs are unlikely to influence policy rates - with the potential exception of longer-maturity IIRs during periods of financial stress - since central banks use policy rates to control IIRs, and insofar as they are successful in this, the shocks to IIRs do not result in reactions from central banks.

There are three substantial decisions to be made when creating VAR models: (1) what variables should be included, (2) with which transformations, and (3) with how many lags. I consider VAR with MRR or DR[5], in levels, in differences and in vector error correction (VEC) form, and use either AIC or SIC to select the number of lags.

[5]I do not include VARs with both policy rates in the analysis. The issue is the high degree of colinearity - and even perfect colinearity in part of the sample - between the two policy rates. This raises issues for coefficient estimates, and indeed the VAR models which include both variables feature explosive forecasts.

2.2.4 Overview of estimated models

In total, the above discussion leads to 4 univariate models, 58 single-equation models, and 12 multiple-equation models. Table 2.1 provides a basic overview of the selected models. The models are separated into 7 groups: (1) univariate models; (2) multivariate models that use the level of IIRs as a dependent variable, (3) multivariate models that use first differences as a dependent variable, and (4) multivariate models that use the spread between IIRs and policy rates as a dependent variable; (5) multivariate models that allow for a break; (6) multiequation models; and (7) structural models.[6] Groups 1-4 together with group 6 cover the space of standard reduced-form time series models; these are complemented by models which impose more structure on the data, be it only in the form of a break, or in the form of fully specified structural models proposed in first chapter.

Table 2.1: Overview of all models estimated

#	Description	Equation
		Univariate models
1	ARMA - AIC	$IIR_t = \beta_0 + ARMA(p,q)$
2	ARMA - SIC	——"——
3	ARMA, d - AIC	$d(IIR_t) = \beta_0 + ARMA(p,q)$
4	ARMA, d - SIC	——"——
		Multivariate models - level
5	Static regerssion - MRR	$IIR_t = \beta_0 + \beta_1 MRR_t$
6	Static regerssion - DR	$IIR_t = \beta_0 + \beta_1 DR_t$
7	Static regerssion - MRR&DR	$IIR_t = \beta_0 + \beta_1 MRR_t + \beta_2 DR_t$
8	ARMAX - MRR - AIC	$IIR_t = \beta_0 + \beta_1 MRR_t + \sum_{l=1}^{p} \beta_{2,l} IIR_{t-l}$
9	ARMAX - MRR - SIC	——"——
10	ARMAX - DR - AIC	$IIR_t = \beta_0 + \beta_1 DR_t + \sum_{l=1}^{p} \beta_{2,l} IIR_{t-l}$
11	ARMAX - DR -SIC	——"——
12	ARMAX - MRR&DR - AIC	$IIR_t = \beta_0 + \beta_1 MRR_t + \beta_2 DR_t + \sum_{l=1}^{p} \beta_{3,l} IIR_{t-l}$
13	ARMAX - MRR&DR - SIC	——"——
14	ARMA errors - MRR - AIC	$IIR_t = \beta_0 + \beta_1 MRR_t + ARMA(p,q)$
15	ARMA errors - MRR - SIC	——"——
16	ARMA errors - DR - AIC	$IIR_t = \beta_0 + \beta_1 DR_t + ARMA(p,q)$
17	ARMA errors - DR -SIC	——"——
18	ARMA errors - MRR&DR - AIC	$IIR_t = \beta_0 + \beta_1 MRR_t + \beta_2 DR_t + ARMA(p,q)$

Continued on next page

[6]For brevity I will refer to the 2nd, 3rd and 4th groups as specifications in levels, specifications in differences, and specifications in spreads, while the 5th group will be referred to as models with break.

Table 2.1: Overview of all models estimated

#	Description	Equation
19	ARMA errors - MRR&DR - SIC	—"—
20	ARDL - MRR - AIC	$IIR_t = \beta_0 + \sum_{l=0}^{p} \beta_{1,l} MRR_{t-l} + \sum_{l=1}^{q} \beta_{2,l} IIR_{t-l}$
21	ARDL - MRR - SIC	—"—
22	ARDL - DR - AIC	$IIR_t = \beta_0 + \sum_{l=0}^{p} \beta_{1,l} DR_{t-l} + \sum_{l=1}^{q} \beta_{2,l} IIR_{t-l}$
23	ARDL - DR -SIC	—"—
24	ARDL - MRR&DR - AIC	$IIR_t = \beta_0 + \sum_{l=0}^{p} \beta_{1,l} MRR_{t-l} + \sum_{l=1}^{q} \beta_{2,l} DR_{t-l} + \sum_{l=1}^{r} \beta_{3,l} IIR_{t-l}$
25	ARDL - MRR&DR - SIC	—"—
	Multivariate models - differences	
26	Static regression d - MRR	$d(IIR_t) = \beta_1 d(MRR_t)$
27	Static regression d - DR	$d(IIR_t) = \beta_1 d(DR_t)$
28	Static regression d - MRR&DR	$d(IIR_t) = \beta_1 d(MRR_t) + \beta_2 d(DR_t)$
29	ARMAX, d - MRR - AIC	$d(IIR_t) = \beta_1 d(MRR_t) + \sum_{l=1}^{p} \beta_{2,l} d(IIR_{t-l})$
30	ARMAX, d - MRR - SIC	—"—
31	ARMAX, d - DR - AIC	$d(IIR_t) = \beta_1 d(DR_t) + \sum_{l=1}^{p} \beta_{2,l} d(IIR_{t-l})$
32	ARMAX, d - DR -SIC	—"—
33	ARMAX, d - MRR&DR - AIC	$d(IIR_t) = \beta_1 d(MRR_t) + \beta_2 d(DR_t) + \sum_{l=1}^{p} \beta_{3,l} d(IIR_{t-l})$
34	ARMAX, d - MRR&DR - SIC	—"—
35	ARMA errors, d - MRR - AIC	$d(IIR_t) = \beta_1 d(MRR_t) + ARMA(p,q)$
36	ARMA errors, d - MRR - SIC	—"—
37	ARMA errors, d - DR - AIC	$d(IIR_t) = \beta_1 d(DR_t) + ARMA(p,q)$
38	ARMA errors, d - DR -SIC	—"—
39	ARMA errors, d - MRR&DR - AIC	$d(IIR_t) = \beta_1 d(MRR_t) + \beta_2 d(DR_t) + ARMA(p,q)$
40	ARMA errors, d - MRR&DR - SIC	—"—
41	ARDL, d - MRR - AIC	$d(IIR_t) = \sum_{l=0}^{p} \beta_{1,l} d(MRR_{t-l}) + \sum_{l=1}^{q} \beta_{2,l} d(IIR_{t-l})$
42	ARDL, d - MRR - SIC	—"—
43	ARDL, d - DR - AIC	$d(IIR_t) = \sum_{l=0}^{p} \beta_{1,l} d(DR_{t-l}) + \sum_{l=1}^{q} \beta_{2,l} d(IIR_{t-l})$
44	ARDL, d - DR -SIC	—"—
45	ARDL, d - MRR&DR - AIC	$d(IIR_t) = \sum_{l=0}^{p} \beta_{1,l} d(MRR_{t-l}) + \sum_{l=1}^{q} \beta_{2,l} d(DR_{t-l}) + \sum_{l=1}^{r} \beta_{3,l} d(IIR_{t-l})$
46	ARDL, d - MRR&DR - SIC	—"—
	Multivariate models - spread	
47	ARMA, s - MRR - AIC	$IIR_t - MRR_t = \beta_0 + ARMA(p,q)$
48	ARMA, s - MRR - SIC	$IIR_t - MRR_t = \beta_0 + ARMA(p,q)$
49	ARMA, s - DR - AIC	$IIR_t - DR_t = \beta_0 + ARMA(p,q)$
50	ARMA, s - DR - SIC	$IIR_t - DR_t = \beta_0 + ARMA(p,q)$
51	ARMA, s, d - MRR - AIC	$d(IIR_t - MRR_t) = ARMA(p,q)$
52	ARMA, s, d - MRR - SIC	$d(IIR_t - MRR_t) = ARMA(p,q)$
53	ARMA, s, d - DR - AIC	$d(IIR_t - DR_t) = ARMA(p,q)$
54	ARMA, s, d - DR - SIC	$d(IIR_t - DR_t) = ARMA(p,q)$

Continued on next page

Table 2.1: Overview of all models estimated

#	Description	Equation
	Multivariate models - break	
55	Static regression, b - MRR	$IIR_t = \sum_{s=before,after}[\beta_{0,s} + \beta_{1,s}MRR_t]$
56	Static regression, b - DR	$IIR_t = \sum_{s=before,after}[\beta_{0,s} + \beta_{1,s}DR_t]$
57	Static regression, b - MRR&DR	$IIR_t = \sum_{s=before,after}[\beta_{0,s} + \beta_{1,s}MRR_t + \beta_{2,s}DR_t]$
58	ARMA errors, b - MRR	$IIR_t = \sum_{s=before,after}[\beta_{0,s} + \beta_{1,s}MRR_t + ARMA_s(1,1)]$
59	ARMA errors, b - DR	$IIR_t = \sum_{s=before,after}[\beta_{0,s} + \beta_{1,s}DR_t + ARMA_s(1,1)]$
60	ARMA errors, b - MRR&DR	$IIR_t = \sum_{s=before,after}[\beta_{0,s} + \beta_{1,s}MRR_t + \beta_{2,s}DR_t + ARMA_s(1,1)]$
61	ARMA, s, b - MRR	$IIR_t - MRR_t = \sum_{s=before,after}[\beta_{0,s} + ARMA_s(1,1)]$
62	ARMA, s, b - DR	$IIR_t - DR_t = \sum_{s=before,after}[\beta_{0,s} + ARMA_s(1,1)]$
	Multiequation models	
63	Levels - MRR - AIC	$\boldsymbol{y_t} = \boldsymbol{A_0} + \sum_{l=1}^{p} \boldsymbol{A_l y_{t-l}}$ where $\boldsymbol{y_t} = \{IIR_t, MRR_t\}$
64	Levels - MRR - SIC	——"——
65	Levels - DR - AIC	$\boldsymbol{y_t} = \boldsymbol{A_0} + \sum_{l=1}^{p} \boldsymbol{A_l y_{t-l}}$ where $\boldsymbol{y_t} = \{IIR_t, DR_t\}$
66	Levels - DR - SIC	——"——
67	Differences - MRR - AIC	$d(\boldsymbol{y_t}) = \sum_{l=1}^{p} \boldsymbol{A_l} d(\boldsymbol{y_{t-l}})$ where $\boldsymbol{y_t} = \{IIR_t, MRR_t\}$
68	Differences - MRR - SIC	——"——
69	Differences - DR - AIC	$d(\boldsymbol{y_t}) = \sum_{l=1}^{p} \boldsymbol{A_l} d(\boldsymbol{y_{t-l}})$ where $\boldsymbol{y_t} = \{IIR_t, DR_t\}$
70	Differences - DR - SIC	——"——
71	VECM - MRR - AIC	$d(\boldsymbol{y_t}) = \boldsymbol{A_0} + \boldsymbol{\Pi y_{t-1}} + \sum_{l=1}^{p} \boldsymbol{A_l} d(\boldsymbol{y_{t-l}})$ where $\boldsymbol{y_t} = \{IIR_t, MRR_t\}$
72	VECM - MRR - SIC	——"——
73	VECM - DR - AIC	$d(\boldsymbol{y_t}) = \boldsymbol{A_0} + \boldsymbol{\Pi y_{t-1}} + \sum_{l=1}^{p} \boldsymbol{A_l} d(\boldsymbol{y_{t-l}})$ where $\boldsymbol{y_t} = \{IIR_t, DR_t\}$
74	VECM - DR - SIC	——"——
	Structural models	
100	Linear	equation (2.2)
101	Nonlinear	equation (2.3)

Notes: ARMA is a model with autoregressive and moving average errors; ARMAX is a model with exogenous variables and lags of the dependent variable; ARMA errors is a model with exogenous variables and with autoregressive and moving average errors; ARDL is model with lags of dependent and exogenous variables.
d indicates a model in differences; s indicates a model in spread; b indicates a model with a break; MRR, DR and MRR&DR indicate models that use the main refinancing rate, the deposit rate, or both, as exogenous variables; AIC and SIC indicate a model selected by Akaike Information Criterion and Schwarz information criterion.
In equations, $ARMA(p,q)$ indicates an error process that includes p autoregressive errors and q moving average errors; $d(\cdot)$ indicates first differencing.

2.3 Forecasting performance

This section assesses the performance of alternative models in pseudo out-of-sample forecasting. Throughout the section, I highlight model features that correlate with good forecasting performance, using the superiority of the structural model as a guiding principle. This not only helps make sense of the results presented, and thus provides background for discussion in next section, but also provides additional support to the structural model.

The performance is assessed based on results from two different exercises. First, forecasts are made using actual values of monetary policy variables, in what is called ex-post or conditional forecasts. The motivation for this forecasting exercise is that the majority of variations in the level of IIRs is caused by variations in monetary policy rates that anchor IIRs. This means that when considering ex-ante forecasts from multivariate models, in which values of monetary policy variables are also forecasted, the majority of forecast errors would originate in forecasts for monetary policy rates, not in forecasts for IIRs. This hinders the evaluation of models for IIRs. To focus specifically on the model for IIRs, I start with an exercise that eliminates the forecast error originating in the monetary policy variables.[7]

The drawback of ex-post forecasts is that they do not provide a measure of how good the models would be at forecasting given a series with imperfect knowledge of future developments in independent variables. Therefore, I complement the results from the ex-post forecasting exercise with results from an ex-ante forecasting exercise. In this second exercise, the forecasts are made without knowledge of any future values of independent variables, and hence they are sometimes called unconditional.

For both exercises, I report the mean absolute error (MAE) of forecasts at several different horizons varying from 1 to 60 months. [8] I include indicator of whether the best performing specification of structural model is has better forecast performance than given benchmark model in statistically significant way.[9]

[7]Ex-post forecasting exercises have a long history in econometric literature. For a famous early example, see e.g., Meese and Rogoff (1983). For more discussion of ex-post forecasting and its informational value see for example Hyndman and Athanasopoulos (2018).

[8]The results are robust to alternative measures of forecast precision.

[9]See more details in Appendix 2.B.3. Note that the results should be taken more as illustration of the statistical significance of the magnitude of differences between forecast performance, rather than formal answer to question whether the structural model improves forecasting performance:

Additionally, I complement the numerical results by charts capturing all forecasts at the 24 month horizon, which can be found in Appendix 2.C. In all cases I focus on the sample covering period with excess reserves, January 2009 to August 2019; results for the sample also covering the period before January 2009 are in Appendix 2.B.[10]

Finally, it is worth adding a comment about the pseudo out-of-sample nature of the forecasts. While this forecasting exercise is *pseudo* out-of-sample in that the forecasts were created after the fact, and hence from models that were postulated based on knowledge about behavior of interbank interest rates, to a large degree it is close to a true out-of-sample forecasting exercise. For all reduced-form models, the model structure is not informed by developments that happened after the start of forecasts given that it is chosen dynamically - see Appendix 2.A for details of the forecasting procedures. This leaves only the structural model as potentially problematic. However, the model has been used since 2016 (Kovar 2017) and thus for part of the sample the forecasts are truly out-of-sample even for the structural model.

2.3.1 Ex-post forecasting performance

Eonia rate. Table 2.2 shows the MAE of ex-post forecasts for the Eonia rate at various horizons, considering all models discussed in the previous section.[11] The

as stressed throughout the text, the appropriate question is not whether given *single* benchmark model is better (or worse) than the structural model, but rather whether given *class* of benchmark models is better than the structural model, given that ex-ante one does not have clear guidance on which benchmark model should be chosen.

[10] I end with August 2019 because in September the ECB adopted tiering of excess reserves, and in October the definition of all considered IIRs changed. Both factors change the relationships embedded in the structural model. I leave exploration of this period for later research.

[11] To decrease the number of models, I report results only for a selected subset of models. The selection is based on two principles: similarity and performance. When models that are similar to each other and at but one has systematically better performance, I eliminate the worse performing model. In most cases, I eliminate model selected by one of the information criteria if the other information criteria led to better results; and in many cases I eliminate the model which uses one policy variable if using the alternative policy variable led to better results. Full results are reported in Appendix 2.B.

bottom of the table then displays summary statistics for each column of the table.

Table 2.2: Ex-post forecasting performance for the Eonia rate

#	Description	1	2	3	6	12	24	48	60	Avg.
		Forecast horizons (# of steps ahead)								
	Multivariate models - level									
5	Static regerssion - MRR	0.21	0.21	0.21	0.22	0.22	*0.23*	0.24	0.16	**0.21**
6	Static regerssion - DR	0.28	0.28	0.28	0.29	0.3	0.32	0.4	*0.43*	**0.32**
7	Static regerssion - MRR&DR	0.19	0.2	0.21	0.23	0.28	*0.38*	*0.63*	*0.7*	**0.35**
9	ARMAX - MRR - SIC	0.089	0.15	0.21	0.26	0.23	0.23	*0.2*	0.11	**0.18**
11	ARMAX - DR -SIC	0.076	0.14	0.22	0.39	0.44	0.44	0.54	0.57	**0.35**
13	ARMAX - MRR&DR - SIC	0.097	0.16	0.22	0.33	*0.36*	*0.48*	0.81	0.9	**0.42**
14	ARMA errors - MRR - AIC	0.058	0.082	0.099	0.13	0.22	0.25	0.24	*0.18*	**0.16**
15	ARMA errors - MRR - SIC	0.05	0.072	0.086	0.12	0.19	0.24	0.19	*0.12*	**0.13**
16	ARMA errors - DR - AIC	*0.043*	0.069	0.084	0.13	0.19	0.25	0.45	0.51	**0.22**
17	ARMA errors - DR -SIC	*0.038*	*0.057*	0.073	0.13	0.2	0.32	0.57	0.67	**0.26**
18	ARMA errors - MRR&DR - AIC	0.061	0.085	0.099	0.13	0.22	0.24	0.23	*0.17*	**0.15**
19	ARMA errors - MRR&DR - SIC	0.051	0.073	0.086	0.12	0.19	0.23	0.19	*0.15*	**0.14**
21	ARDL - MRR - SIC	0.049	0.072	0.089	0.13	0.21	0.27	0.23	0.15	**0.15**
23	ARDL - DR -SIC	*0.034*	*0.051*	*0.064*	0.11	0.18	0.23	0.46	0.55	**0.21**
25	ARDL - MRR&DR - SIC	0.063	0.09	0.11	0.17	*0.29*	*0.47*	0.95	1.22	**0.42**
Avg.		**0.09**	**0.12**	**0.15**	**0.21**	**0.27**	**0.34**	**0.57**	**0.72**	**0.31**
	Multivariate models - differences									
26	Static regerssion, d - MRR	*0.038*	0.058	0.069	0.1	0.17	0.22	0.26	0.25	**0.15**
27	Static regerssion, d - DR	*0.03*	*0.044*	*0.053*	0.077	0.12	0.11	0.17	0.18	**0.098**
28	Static regerssion, d - MRR&DR	*0.039*	0.059	0.069	0.095	0.16	0.19	0.2	0.2	**0.12**
30	ARMAX, d - MRR - SIC	*0.04*	0.06	0.073	0.1	0.17	0.23	0.29	*0.31*	**0.16**
32	ARMAX, d - DR -SIC	*0.031*	*0.044*	*0.055*	0.074	0.11	0.1	0.14	0.15	**0.088**
34	ARMAX, d - MRR&DR - SIC	*0.04*	0.059	0.071	0.094	0.15	0.19	0.23	*0.25*	**0.14**
38	ARMA errors, d - DR -SIC	*0.034*	*0.047*	*0.058*	0.078	0.12	0.11	0.16	0.17	**0.096**
42	ARDL, d - MRR - SIC	*0.043*	0.063	0.075	0.1	0.17	0.23	0.31	*0.33*	**0.17**
44	ARDL, d - DR -SIC	*0.031*	*0.044*	*0.055*	0.074	0.11	0.1	0.14	0.15	**0.087**
46	ARDL, d - MRR&DR - SIC	0.05	0.066	0.08	0.12	0.23	*0.38*	0.7	0.87	**0.31**
Avg.		**0.04**	**0.06**	**0.07**	**0.10**	**0.16**	**0.19**	**0.27**	**0.30**	**0.15**
	Multivaraite models - spread									
47	ARMA, s - MRR - AIC	0.058	0.085	0.1	0.14	0.23	0.27	0.19	*0.13*	**0.15**
48	ARMA, s - MRR - SIC	0.049	0.071	0.085	0.12	0.19	0.23	0.17	*0.1*	**0.13**
50	ARMA, s - DR - SIC	*0.038*	*0.054*	0.07	0.12	0.19	0.28	0.48	0.55	**0.22**
52	ARMA, s, d - MRR - SIC	*0.044*	0.063	0.074	0.1	0.17	0.22	0.28	0.29	**0.16**
53	ARMA, s, d - DR - AIC	*0.042*	0.064	0.074	0.092	0.12	*0.12*	0.16	0.17	**0.11**
54	ARMA, s, d - DR - SIC	*0.032*	*0.045*	*0.056*	*0.073*	0.11	0.095	0.13	0.13	**0.083**
Avg.		**0.05**	**0.07**	**0.08**	**0.11**	**0.18**	**0.21**	**0.27**	**0.28**	**0.16**
	Multivariate models - break									
55	Static regression, b - MRR	0.11	0.11	0.11	0.11	0.12	*0.1*	0.15	0.16	**0.12**
56	Static regression, b - DR	0.15	0.15	0.15	0.16	0.19	0.2	0.33	*0.36*	**0.21**
57	Static regression, b - MRR&DR	0.1	0.1	0.11	0.12	0.14	*0.13*	*0.22*	*0.22*	**0.14**
58	ARMA errors, b - MRR	0.054	0.09	0.1	0.13	0.17	*0.17*	*0.32*	0.32	**0.17**
59	ARMA errors, b - DR	0.045	0.072	0.086	0.11	0.13	0.11	*0.19*	*0.2*	**0.12**
61	ARMA, s, b - MRR	0.049	0.079	0.096	0.13	0.16	0.16	0.23	*0.25*	**0.14**
62	ARMA, s, b - DR	0.055	0.088	0.1	0.13	0.15	*0.15*	*0.22*	0.18	**0.13**
Avg.		**0.08**	**0.10**	**0.11**	**0.12**	**0.15**	**0.14**	**0.24**	**0.24**	**0.14**
	Structural models									
100a	Linear	0.066	0.07	0.075	0.084	0.089	0.098	0.085	0.12	**0.086**
100b	Linear - ARMA errors	0.048	0.061	0.066	0.075	0.081	0.088	0.075	0.088	**0.073**
101a	Nonlinear	0.044	0.047	0.046	0.046	0.044	0.035	0.035	0.024	**0.04**
101b	Nonlinear - ARMA errors	0.037	0.046	0.044	0.049	0.051	0.046	0.048	0.039	**0.045**
Avg.		**0.05**	**0.06**	**0.06**	**0.06**	**0.07**	**0.07**	**0.06**	**0.07**	**0.06**
	Summary statistics									
Mean			0.08	0.10	0.14	0.19	0.23	0.35	0.42	0.20
Std. Deviation			0.05	0.05	0.07	0.08	0.12	0.31	0.54	0.13
Minimum			0.04	0.04	0.05	0.04	0.04	0.04	0.02	0.04
1st quartile			0.06	0.07	0.10	0.13	0.12	0.19	0.16	0.12
3rd quartile			0.09	0.11	0.13	0.22	0.27	0.45	0.51	0.22
Maximum			0.28	0.28	0.39	0.44	0.61	1.78	3.64	0.81

Notes: Mean average error at given horizon. "Avg." shows average across all horizons (last column) or models (last row of each section). Green indicates low errors, yellow/orange medium errors, and red high errors; dark scale indicates more extreme values than light scale. Italics indicates that performance is not significantly worse than the best structural model.

Focusing on specifications in level and looking at all horizons together, one can easily identify types of models that perform poorly in forecasting IIRs: it is models that either link the level of IIRs only to current policy rates without any dynamic components (5-7) and hence force an immediate return to the average relationship between IIRs and policy rates; or models that incorporate dynamic components only in the form of past levels of dependent and independent variables, rather than in terms of past errors (8-13,20-25). The failure of the latter type of models, especially at short horizons, is linked to reliance on the wrong type of dynamic terms, since past values of policy rates should not influence current IIRs if monetary policy has control over them. Similar logic also explains why the regressions with ARMA errors (14-19) perform generally better than regressions with lagged values of IIRs: while both incorporate information about past values of IIRs in forecasts for future values, regressions with lagged values of IIRs suggest that future values of IIRs depend on current and past values of IIRs, which is unlikely in a situation when policy rates are changing; in contrast, regressions with ARMA errors link future values to past values only after accounting for the values of policy rates.

Turning to specifications with differences of IIRs as the dependent variable, the first conclusion is that, as a set, these models are substantially better at forecasting IIRs than models with the level of IIRs as the dependent variable: averaged across all models and horizons, the MAE is 0.15 compared to 0.31 for models in levels; see last row of the corresponding sections. This finding is somewhat unsuprising in the light of the non-stationarity of IIRs highlighted in the first chapter and discussed in greater detail in the next section; there I show how better forecasting performance is linked to the fact that equations which use the level of IIRs as a dependent variable force a return to the historically average relationship between level of IIRs and policy rates, something that is at odds with the behavior of IIRs in the presence of excess reserves.

Among the models in first differences, the variation in forecasting performance is less pronounced and less related to model structure than among the previous group of models. Rather, it is linked to which independent variable(s) are included, with models using the DR performing better than models using only (or also) the main refinancing rate. Still, there is one model that stands out in terms of forecasting performance: a static model with the deposit rate. This model links changes in IIRs to changes in the deposit rate, but does not include any additional dynamic terms. The superior performance of the model without any dynamic terms suggests that once we try to predict only changes in IIRs *from the current level*, rather then predicting the level itself, then past information is no longer useful in predicting IIRs. While this is a somewhat unsurprising realization, it is an important one with respect to the structural model that does not include information about past values of IIRs or policy rates.

The last group of standard reduced-form models use the spread between IIRs and policy rates as a dependent variable, either in levels or in differences. As a group, these models perform as well as models in differences, and hence substantially better than models in levels. Since the difference between the specifications in terms of levels and in terms of spreads is effectively in whether one imposes coefficient restrictions or not, one can conclude that such coefficient restrictions can play an important role in forecasting IIRs. Apart from this observation, the conclusions are similar to those from the previous group, with best the performing model using the DR rather than the main refinancing rate.

The following group of models with a break occupies space between pure reduced form models and models that impose some structure on the modeled series. Since one key essence of the structural model is the break in the presence of excess reserves, looking at models with a break can shed light on whether simply allowing for a break would be sufficient to obtain good forecasting performance. The dedicated section of the table shows, that while allowing for a break does improve the forecasting performance - the models as a group are the best of all groups considered so far - the improvement is not dramatic.

Finally, turning to structural models, the main takeaway is that, overall, the models are significantly better than any of the time series models: averaged across all horizons, the MAE of structural models is 0.06, compared with 0.083 for the best performing reduced-from model, and 0.14 for the best performing group of reduced form models. This better performance is true not only overall, but also for all horizons with the exception of the shortest two, at which some reduced-form models have marginally better forecasting performance. Another way to look at this is by noticing that the superiority of structural models increases with the length of the forecast horizon, and for the better performing structural model, the forecast errors are smaller by factors of two or more at the medium and long horizons.

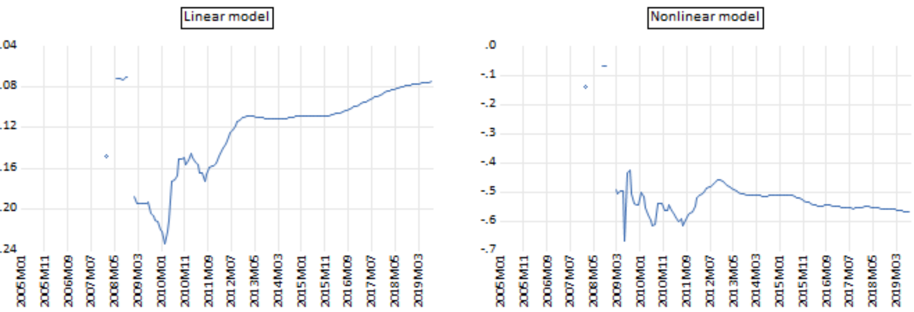
The results in Table 2.2 indicate that the differences in forecast performance between the structural model and the alternative benchmark models are statistically significant, at least for the best-performing structural model (model 101b). Moreover, more detailed results provided in Appendix 2.B.3 show that whenever a benchmark model has lower MAE, the difference is not statistically significant (with sole exception). The significantly better performance of structural models can be viewed in the context of the best reduced-form models to highlight which factors correlated with good forecasting performance - or alternatively, the variations in forecasting performance of reduced-form models can be viewed through the structural relationship embodied in the structural model. First, during a period of excess reserves, the structural model links IIRs only to the deposit rate, rather than to MRR or both policy rates. The first chapter justifies this using a formal data analysis, but the results for reduced-form models provide additional support for this result: across various types models, models that use the DR systematically outperform their analogs that rely on the MRR or both policy rates.

Second, some models which use levels of IIRs as a dependent variable are among the best performing models, despite the non-stationarity of IIRs during the relevant period. Similarly, the structural model is formulated with the level of IIRs as the dependent variable. These two observations are related through the fact that IIRs, policy rates, and excess reserves form a co-integrated relationship, justifying the use of the level of IIRs as a dependent variable. I return to this topic in greater detail in the next section.

The third way in which reduced-form models can shed light on the structural model is in terms of use of the dynamic regressors. The structural model omits all dynamic terms, either in the form of lags of IIRs or in the form of lags of policy rates. Similarly, the absence of these terms, and especially of lagged values of IIRs, is clearly an important driver of good forecasting performance among reduced form models, since static or more parsimonious reduced-form models tend to perform better than reduced form models that include many lags. This finding supports the DGP proposed by the structural model: clearly, past information is not very useful in predicting IIRs, apart from being useful as counterweights to model misspecification, as in the case of models specified in levels. Relatedly, the absence of dynamic terms in the structural model can also explain why some reduced form models perform better at very short horizons: the structural model captures cointegration relationship and hence predicts the value of IIRs in equilibrium; such equilibrium value is not necessarily a good prediction at horizons when equilibrium forces do not have time to assert themselves, an issue I return to at the end of this section.

Looking at the structural models individually, there are a few important take-aways to note. First, the nonlinear version of the structural model has substantially better forecasting performance than the linear version of the model, which clearly suggests that the nonlinear model is closer to the true DGP. The reason for this better performance is simple. As argued in the first chapter, the linear model is globally mis-specified in that it allows IIRs to fall below the deposit rate, something that was impossible in the sample considered. In contrast, the nonlinear model avoids this mis-specification by making the effect of additional excess reserves smaller when the spread between IIRs and DR is smaller. The importance of this factor can be seen in Figure 2.1. The figure shows how the coefficient estimates for excess reserves regressor evolve as the estimation sample is enlarged. It is clear that, as the sample increases, the coefficient in the linear model becomes smaller. Since the excess reserves were increasing throughout the sample, this suggests that the model is mis-specified with respect to the effect of excess reserves: as new data points with higher excess reserves are added the model reacts by decreasing the coefficient on excess reserves. The way this manifests in terms of forecasts is with systematic negative bias in the period of the largest increase in excess reserves during the quantitative easing program: Figure 2.2 shows how the linear model

Figure 2.1: Recursive coefficient estimates for excess reserves (Eonia rate)



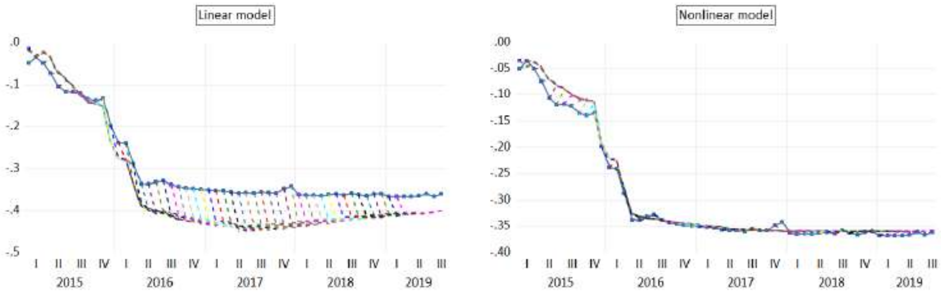
Notes: For each month the figure shows the coefficient estimate for the coefficient on the excess reserves regressor estimated on a data sample ending in the previous month. The results correspond to linear and nonlinear models estimated via OLS and with formal break tests.

would always predict IIRs lower than what was later observed during this period.

These conclusions for the linear model can be contrasted with the results for the nonlinear model. In terms of recursive coefficients, there is little trending over the sample, and the recursive coefficient estimates are stable especially during the period of massive increases in excess reserves; see right hand panel of Figure 2.1. For example, all the recursive coefficients after 2015M01 are contained in the band of one standard deviation of the coefficient estimated based on the sample before 2015M01. This is remarkable give that excess reserves increased more than 10-fold during this period. In terms of forecasts, there is basically no bias, with the forecast almost perfectly copying the actual values. Correspondingly, I will focus on the non-linear model going forward.

The second feature of the results for structural models is the relationship between forecast quality and the length of forecast horizon. The table makes it clear that for the (nonlinear) structural model, there is no positive relationship between

Figure 2.2: Recursive forecasts from structural models for Eonia rate



Notes: The figure includes actual values of the Eonia rate (blue line with filled squares symbol) and individual recursive forecasts (dashed lines without symbol).

the MAE and the forecast horizon considered, the relationship is actually negative: the MAE at the 5-year horizon is half that of at the 1-month horizon. This is in contrast to reduced-form models. This difference is in some sense expected. If the structural model correctly captures the DGP of a given variable, then there is little reason to expect that forecast errors should increase with the horizon *in a ex-post forecasting exercise*; indeed, in so far as it takes time for a structural relationship to re-assert itself after shocks, there is reason to expect that forecasts at short horizons should be *less* precise than forecasts at medium or long horizons.

This last point is also related to a final observation about the forecasting performance of structural models. As noted above, some reduced form models have significantly better forecasting performance than structural models at very short forecasting horizons, which is likely related to the immediacy of returns to an equilibrium relationship. However, this feature of results disappears as soon as one puts the structural model on the same footing as reduced-form models by allowing the structural model to also include a weak form of dependency on past values in the form or ARMA errors. The last row of the table shows that when structural models include ARMA(1,1) errors, they have superior forecasting performance at horizon of 3 months or longer, while almost reaching the lower bound for 1-month and 2-month horizons.

Euribor rates. Tables 2.3 and 2.4 report forecasting performance results for 1-week and 3-month Euribor. The tables show that many of the conclusions from the Eonia rate also apply to the longer-maturity IIRs: As in the case of the Eonia rate, the models in levels perform worse than models in differences or spreads; and among the latter two it is again models with the DR that perform better than when alternative policy rates are used. Most importantly, the structural models are again significantly better than any category of reduced form models, even though the superiority is slightly less pronounced.

Table 2.3: Ex-post forecasting performance for 1-week Euribor

#	Description	1	2	3	6	12	24	48	60	Avg.
		Forecast horizons (# of steps ahead)								
		Multivariate models - level								
5	Static regerssion - MRR	0.19	0.2	0.2	0.2	*0.2*	*0.19*	0.2	0.11	**0.19**
6	Static regerssion - DR	0.29	0.29	0.3	0.31	0.32	*0.34*	0.44	*0.48*	**0.35**
7	Static regerssion - MRR&DR	0.2	0.21	0.21	0.24	0.28	*0.33*	*0.53*	0.55	**0.32**
9	ARMAX - MRR - SIC	0.09	0.16	0.21	0.25	0.22	0.21	*0.18*	*0.081*	**0.17**
11	ARMAX - DR -SIC	0.085	0.17	0.25	0.39	0.41	0.42	0.53	0.57	**0.35**
13	ARMAX - MRR&DR - SIC	0.096	0.17	0.23	0.33	*0.35*	*0.43*	0.76	0.79	**0.4**
14	ARMA errors - MRR - AIC	0.057	0.088	0.1	0.15	0.22	0.22	0.21	*0.17*	**0.15**
15	ARMA errors - MRR - SIC	0.046	0.078	0.1	0.15	0.22	0.24	0.2	*0.14*	**0.15**
16	ARMA errors - DR - AIC	*0.041*	0.075	0.1	0.18	0.26	0.33	0.58	0.65	**0.28**
17	ARMA errors - DR -SIC	*0.034*	*0.058*	0.08	0.14	0.23	0.34	0.6	0.68	**0.27**
18	ARMA errors - MRR&DR - AIC	0.058	0.09	0.11	0.15	0.23	0.24	0.24	*0.21*	**0.17**
19	ARMA errors - MRR&DR - SIC	0.046	0.078	0.1	0.15	0.22	0.26	0.24	*0.21*	**0.16**
21	ARDL - MRR - SIC	0.043	0.074	0.097	0.15	0.21	0.23	*0.2*	*0.11*	**0.14**
23	ARDL - DR -SIC	*0.032*	*0.052*	*0.073*	0.13	0.21	0.26	0.51	0.6	**0.23**
25	ARDL - MRR&DR - SIC	0.093	0.13	0.19	0.21	*0.35*	*0.52*	0.96	1.28	**0.47**
Avg.		**0.09**	**0.13**	**0.17**	**0.23**	**0.29**	**0.34**	**0.49**	**0.57**	**0.29**
		Multivariate models - differences								
26	Static regerssion, d - MRR	*0.035*	*0.06*	0.077	*0.11*	0.2	0.26	0.26	0.28	**0.16**
27	Static regerssion, d - DR	*0.027*	*0.043*	*0.06*	*0.091*	*0.14*	*0.15*	0.19	0.21	**0.11**
28	Static regerssion, d - MRR&DR	*0.035*	0.061	0.077	*0.11*	0.19	0.23	0.22	0.24	**0.14**
30	ARMAX, d - MRR - SIC	*0.036*	*0.061*	0.08	*0.11*	0.2	0.26	0.29	0.32	**0.17**
32	ARMAX, d - DR -SIC	*0.027*	*0.045*	*0.062*	*0.089*	*0.14*	*0.14*	0.16	0.18	**0.11**
34	ARMAX, d - MRR&DR - SIC	*0.035*	0.06	0.077	*0.1*	0.18	0.23	0.23	0.26	**0.15**
38	ARMA errors, d - DR -SIC	*0.027*	*0.043*	*0.06*	*0.091*	*0.14*	*0.15*	0.19	0.21	**0.11**
42	ARDL, d - MRR - SIC	*0.037*	0.063	0.083	0.12	0.21	0.26	0.3	0.31	**0.17**
44	ARDL, d - DR -SIC	*0.027*	*0.045*	*0.062*	*0.089*	*0.14*	*0.14*	0.16	0.18	**0.11**
46	ARDL, d - MRR&DR - SIC	0.083	0.085	0.13	0.16	0.29	*0.45*	0.78	1	**0.37**
Avg.		**0.04**	**0.06**	**0.09**	**0.12**	**0.19**	**0.24**	**0.30**	**0.34**	**0.17**
		Multivaraite models - spread								
47	ARMA, s - MRR - AIC	0.054	0.086	0.11	0.16	0.26	0.32	0.24	*0.2*	**0.18**
48	ARMA, s - MRR - SIC	0.045	0.076	0.099	0.15	0.22	0.26	0.22	*0.19*	**0.16**
50	ARMA, s - DR - SIC	*0.035*	*0.059*	0.081	0.14	0.23	0.34	0.59	0.66	**0.27**
52	ARMA, s, d - MRR - SIC	*0.039*	0.063	0.079	0.12	0.21	0.26	0.27	0.28	**0.16**
53	ARMA, s, d - DR - AIC	*0.038*	*0.062*	0.087	0.12	0.18	0.16	0.2	0.22	**0.13**
54	ARMA, s, d - DR - SIC	*0.027*	*0.044*	*0.061*	*0.09*	*0.14*	*0.14*	0.18	0.19	**0.11**
Avg.		**0.04**	**0.07**	**0.09**	**0.14**	**0.21**	**0.26**	**0.32**	**0.33**	**0.18**
		Multivariate models - break								
55	Static regression, b - MRR	0.12	0.12	0.12	0.13	*0.14*	*0.11*	0.15	0.16	**0.13**
56	Static regression, b - DR	0.17	0.17	0.17	0.19	0.22	*0.23*	0.36	*0.41*	**0.24**
57	Static regression, b - MRR&DR	0.12	0.12	0.12	0.14	0.17	*0.15*	*0.27*	0.27	**0.17**
58	ARMA errors, b - MRR	0.051	0.09	0.11	0.15	*0.21*	*0.2*	0.39	0.37	**0.2**
59	ARMA errors, b - DR	0.045	0.075	0.097	0.13	0.16	*0.12*	*0.2*	*0.21*	**0.13**
61	ARMA, s, b - MRR	0.048	0.081	0.11	0.17	0.23	0.27	0.37	0.39	**0.21**
62	ARMA, s, b - DR	0.044	0.078	0.1	0.14	0.18	0.18	*0.15*	*0.089*	**0.12**
Avg.		**0.08**	**0.10**	**0.12**	**0.15**	**0.19**	**0.18**	**0.28**	**0.27**	**0.17**
		Structural models								
100a	Linear	0.085	0.09	0.096	0.1	0.12	0.14	0.14	0.2	**0.12**
100b	Linear - ARMA errors	0.047	0.066	0.078	0.095	0.12	0.15	0.15	0.2	**0.11**
101a	Nonlinear	0.059	0.065	0.068	0.07	0.069	0.049	0.058	0.055	**0.062**
101b	Nonlinear - ARMA errors	0.035	0.049	0.056	0.071	0.092	0.089	0.096	0.095	**0.073**
Avg.		**0.06**	**0.07**	**0.07**	**0.08**	**0.10**	**0.11**	**0.11**	**0.14**	**0.09**
		Summary statistics								
Mean			0.09	0.12	0.16	0.22	0.26	0.35	0.40	0.21
Std. Deviation			0.05	0.06	0.07	0.09	0.14	0.25	0.36	0.12
Minimum			0.04	0.06	0.07	0.07	0.05	0.06	0.06	0.06
1st quartile			0.06	0.08	0.11	0.18	0.16	0.19	0.20	0.13
3rd quartile			0.09	0.12	0.18	0.24	0.32	0.51	0.55	0.27
Maximum			0.29	0.33	0.40	0.58	0.88	1.21	1.65	0.65

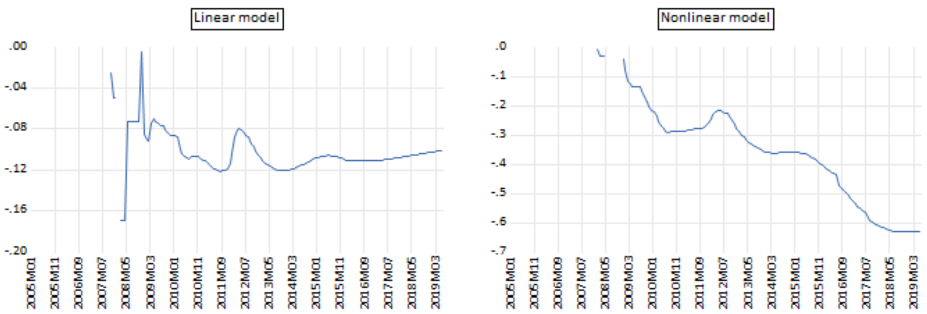
Notes: See note below Table 2.2 for explanation of values.

Table 2.4: Ex-post forecasting performance for 3-m Euribor

#	Description	1	2	3	6	12	24	48	60	Avg.
		Forecast horizons (# of steps ahead)								
		Multivariate models - level								
5	Static regerssion - MRR	0.17	0.18	0.18	0.19	*0.2*	*0.17*	*0.16*	0.1	**0.17**
6	Static regerssion - DR	0.3	0.3	0.31	0.33	0.36	*0.4*	*0.57*	*0.61*	**0.4**
7	Static regerssion - MRR&DR	0.18	0.18	0.19	0.21	*0.23*	0.19	*0.21*	*0.21*	**0.2**
9	ARMAX - MRR - SIC	0.05	0.11	0.15	0.22	0.21	*0.18*	0.17	0.12	**0.15**
11	ARMAX - DR -SIC	0.058	0.14	0.22	0.38	0.43	0.47	0.65	*0.7*	**0.38**
13	ARMAX - MRR&DR - SIC	0.057	0.13	0.21	0.34	0.44	0.54	*0.75*	*1.21*	**0.46**
14	ARMA errors - MRR - AIC	0.059	0.1	0.13	0.19	0.27	*0.27*	*0.33*	0.32	**0.21**
15	ARMA errors - MRR - SIC	0.05	0.087	0.12	0.18	0.28	0.35	*0.48*	0.47	**0.25**
16	ARMA errors - DR - AIC	0.051	0.09	0.13	0.22	0.44	0.76	0.94	0.97	**0.45**
17	ARMA errors - DR -SIC	*0.043*	0.074	0.1	0.18	0.31	0.5	0.89	0.98	**0.38**
18	ARMA errors - MRR&DR - AIC	0.057	0.1	0.14	0.24	0.44	0.67	0.63	*0.64*	**0.36**
19	ARMA errors - MRR&DR - SIC	0.045	0.08	0.11	0.18	0.29	0.44	0.7	0.74	**0.32**
21	ARDL - MRR - SIC	*0.043*	0.08	0.11	0.16	0.22	*0.24*	*0.26*	*0.25*	**0.17**
23	ARDL - DR -SIC	0.048	0.09	0.13	0.22	0.37	0.57	0.83	0.9	**0.39**
25	ARDL - MRR&DR - SIC	0.057	0.12	0.18	0.31	0.56	*0.95*	*1.92*	*3.09*	**0.9**
Avg.		**0.09**	**0.14**	**0.19**	**0.26**	**0.40**	**0.62**	**1.91**	**4.07**	**0.94**
		Multivariate models - differences								
26	Static regerssion, d - MRR	*0.037*	*0.066*	0.089	0.14	0.25	0.28	*0.3*	0.28	**0.18**
27	Static regerssion, d - DR	*0.032*	*0.053*	*0.073*	*0.12*	0.2	*0.24*	0.41	*0.49*	**0.2**
28	Static regerssion, d - MRR&DR	*0.034*	*0.06*	*0.08*	*0.13*	0.22	0.24	*0.34*	*0.37*	**0.18**
30	ARMAX, d - MRR - SIC	*0.031*	*0.059*	*0.082*	*0.13*	0.23	0.29	*0.25*	0.26	**0.17**
32	ARMAX, d - DR -SIC	*0.031*	*0.052*	*0.071*	*0.11*	*0.19*	*0.21*	0.36	*0.43*	**0.18**
34	ARMAX, d - MRR&DR - SIC	*0.03*	*0.057*	*0.08*	*0.13*	0.22	0.26	*0.26*	0.27	**0.16**
38	ARMA errors, d - DR -SIC	*0.032*	*0.053*	*0.073*	*0.12*	0.2	*0.24*	0.41	*0.49*	**0.2**
42	ARDL, d - MRR - SIC	*0.037*	*0.066*	*0.087*	*0.13*	0.22	0.28	*0.25*	0.26	**0.17**
44	ARDL, d - DR -SIC	*0.035*	*0.063*	*0.087*	*0.14*	*0.2*	*0.23*	*0.37*	*0.43*	**0.19**
46	ARDL, d - MRR&DR - SIC	0.061	0.093	0.13	0.18	0.3	0.36	0.38	0.38	**0.24**
Avg.		**0.04**	**0.07**	**0.10**	**0.15**	**0.24**	**0.30**	**0.41**	**0.46**	**0.22**
		Multivaraite models - spread								
47	ARMA, s - MRR - AIC	0.055	0.093	0.12	0.18	0.28	0.34	*0.38*	*0.36*	**0.23**
48	ARMA, s - MRR - SIC	0.051	0.088	0.11	0.17	0.25	*0.28*	*0.35*	*0.33*	**0.2**
50	ARMA, s - DR - SIC	0.047	0.081	0.11	0.19	0.32	0.5	0.82	0.88	**0.37**
52	ARMA, s, d - MRR - SIC	*0.039*	*0.068*	0.089	*0.14*	0.24	0.27	*0.27*	0.24	**0.17**
53	ARMA, s, d - DR - AIC	0.045	0.076	0.099	0.15	0.23	0.28	0.41	*0.46*	**0.22**
54	ARMA, s, d - DR - SIC	*0.033*	*0.054*	*0.072*	*0.11*	*0.19*	*0.21*	0.35	*0.41*	**0.18**
Avg.		**0.05**	**0.08**	**0.10**	**0.17**	**0.28**	**0.36**	**0.48**	**0.49**	**0.25**
		Multivariate models - break								
55	Static regression, b - MRR	0.16	0.17	0.17	0.19	*0.21*	*0.21*	*0.34*	0.37	**0.23**
56	Static regression, b - DR	0.18	0.18	0.18	0.18	0.19	*0.18*	*0.14*	*0.14*	**0.17**
57	Static regression, b - MRR&DR	0.15	0.16	0.16	0.18	*0.2*	0.16	*0.19*	*0.2*	**0.18**
58	ARMA errors, b - MRR	0.048	0.088	0.12	0.19	0.3	0.41	*0.58*	*0.56*	**0.29**
59	ARMA errors, b - DR	0.048	0.082	0.11	0.17	0.25	0.31	0.48	0.49	**0.24**
61	ARMA, s, b - MRR	*0.035*	*0.064*	0.088	0.15	0.25	0.4	0.76	0.85	**0.33**
62	ARMA, s, b - DR	*0.038*	0.073	0.098	0.16	0.25	0.32	*0.45*	*0.45*	**0.23**
Avg.		**0.09**	**0.11**	**0.13**	**0.17**	**0.24**	**0.29**	**0.45**	**0.47**	**0.24**
		Structural models								
100a	Linear	0.12	0.12	0.13	0.13	0.12	0.12	0.15	0.12	**0.13**
100b	Linear - ARMA errors	0.044	0.068	0.08	0.11	0.13	0.12	0.19	0.12	**0.11**
101a	Nonlinear	0.13	0.14	0.14	0.16	0.17	0.15	0.26	0.23	**0.17**
101b	Nonlinear - ARMA errors	0.038	0.06	0.072	0.1	0.15	0.18	0.32	0.34	**0.16**
Avg.		**0.08**	**0.10**	**0.11**	**0.13**	**0.14**	**0.14**	**0.23**	**0.20**	**0.14**
		Summary statistics								
Mean			0.10	0.13	0.19	0.29	0.41	0.92	1.67	0.47
Std. Deviation			0.06	0.08	0.09	0.19	0.53	3.51	8.96	1.62
Minimum			0.05	0.07	0.10	0.12	0.12	0.14	0.10	0.11
1st quartile			0.07	0.09	0.14	0.21	0.23	0.30	0.28	0.18
3rd quartile			0.12	0.14	0.21	0.31	0.41	0.62	0.70	0.33
Maximum			0.44	0.66	0.58	1.52	4.30	28.00	71.00	13.00

Notes: See note below Table 2.2 for explanation of values.

Figure 2.3: Recursive coefficient estimates for excess reserves (3-month Euribor)



Notes: See Figure 2.1 for explanation.

Meanwhile, results for structural models provide the only important change from the Eonia rate. While for 1-week Euribor it is still true that the nonlinear model is better than the linear model, the difference is smaller than in the case of the Eonia rate; and for 3-month Euribor the linear model produces better forecasts than the nonlinear model. This finding corresponds to the finding in the first chapter, which shows that the nonlinear model has better in-sample fit than the linear model for the Eonia rate, but not for the Euribor rate. The reason for worse forecast performance is the reversion of the finding presented in Figure 2.1: in the case of Euribor, the linear model has more stable recursive coefficients; see Figure 2.3.

2.3.2 Ex-ante forecasting performance

The only difference between ex-post and ex-ante forecasting is the treatment of the independent variables. In ex-post forecasting one uses the actual historical values for the independent variables, while in ex-ante forecasting one uses values which are al;so forecast. This means that one does not use any information about the future when creating a forecast, which makes it a true forecast. The ex-ante forecast performance answers the key question whether the structural model provides better forecasts than the benchmark models even after accounting for the difficulty of forecasting independent variables: a simpler, less correct model might prove better in forecasting than more complex, albeit more correct model, if the more complex model relies on variables that are hard to forecast.

This need to forecast independent variables raises the question as to what models should be used to forecast these variables. Here I need to determine how policy rates should be forecasted, and for the structural model, also how the excess reserves should be forecasted. After checking several alternative univariate models, I selected a simple random walk without drift for both policy rates; see Appendix 2.A for details. Apart from forecasting performance, the random walk model is also appealing for two reasons: first, the random walk model is a firmly established (benchmark) model in forecasting financial variables such as interest rates; second, in terms of economics, it amounts to expecting that policy rates will remain unchanged at current levels, which, in the sample and forecast horizons considered, is a reasonable assumption.[12]

As for excess reserves, there is a qualitative difference in terms of the information available to forecasters between the period before and after the start of the QE program. Before the start of the QE program, an economist would expect either excess reserves to remain unchanged at current levels - resulting in a ran-

[12]Note that from an economic perspective, this is likely not a good assumption for very long term horizons - it seems plausible that policy rates will eventually increase from their current record low levels. Indeed, even during the sample considered, private and institutional forecasters expected interest rates to increase *eventually*.

dom walk model - or for excess reserves to gradually return to zero - resulting in some simple ARMA model without a constant. Surprisingly, it turns out that the somewhat less plausible random walk model actually produces better forecasts for excess reserves at medium and long horizons. Meanwhile, after the start of the QE, an economist would likely rely on information about planned asset purchases: these are not only causally related to the amount of excess reserves (see the first chapter for details), but their pace is announced by the ECB for prolonged periods.

To reflect on the fact that there are significant differences in forecasting excess reserves in different parts of sample - before and after the start of the QE program- I report two sets of results. To study the effect of different models for excess reserves before the start of the QE program, I report results from two ex-ante forecasting exercises in the sample before the the start of the QE program: one which uses a random-walk model before start of the QE program, and one that uses ARMA model in that period; both models use information about planned ECB asset purchases for the period after the start of the QE program. See Appendix 2.A for more details about this and the ex-ante forecasting procedure in general.

Table 2.5 captures forecasting performance for the Eonia rate in a sample starting from 2009 and lasting until 2015M01, while Table 2.6 does so for a sample from 2015M02 to 2019M08. Before analyzing the results, it is important to stress that interpreting forecast performance is more complicated in ex-ante forecasting since the performance is no longer related only to forecast errors originating in the considered model, but also to forecast errors originating in the model for the independent variables. The key question is whether after accounting for the forecast errors in independent (conditioning) variables a particular model is still better than an alternative that relies on different set of conditioning variables.

More specifically, one can wonder whether the complex structural model that relies on large set of conditioning variables is still better in ex-ante forecasting than a simpler reduced form models, and especially reduced from models that rely on few or no conditioning variables.

Table 2.5: Unonditional forecasting performance for the Eonia rate (2009-2015)

#	Description	Forecast horizons (# of steps ahead) 1	2	3	6	12	24	48	Avg.
	Univariate models								
2	ARMA - SIC	0.081	0.14	0.2	0.44	1.04	2.00	2.71	**1.18**
4	ARMA, d - SIC	0.067	0.12	0.14	0.22	0.37	0.54	0.44	**0.3**
Avg.		**0.08**	**0.14**	**0.19**	**0.36**	**0.76**	**1.34**	**1.64**	**0.78**
	Univariate models								
7	Static regerssion - MRR&DR	0.32	0.35	0.38	0.44	0.46	0.48	0.77	**0.53**
9	ARMAX - MRR - SIC	0.17	0.29	0.41	0.52	0.49	0.62	1.01	**0.6**
13	ARMAX - MRR&DR - SIC	0.16	0.27	0.37	0.54	0.48	0.47	0.76	**0.51**
14	ARMA errors - MRR - AIC	0.11	0.16	0.21	0.31	0.46	0.54	0.82	**0.46**
21	ARDL - MRR - SIC	0.09	0.14	0.19	0.31	0.44	0.55	0.95	**0.48**
25	ARDL - MRR&DR - SIC	0.12	0.19	0.25	0.39	0.51	0.48	0.59	**0.42**
Avg.		**0.15**	**0.22**	**0.29**	**0.42**	**0.51**	**0.61**	**1.16**	**0.67**
	Multivariate models - differences								
26	Static regerssion, d - MRR	0.37	0.4	0.43	0.48	0.59	0.74	1.02	**0.66**
30	ARMAX, d - MRR - SIC	0.075	0.12	0.15	0.26	0.38	0.43	0.56	**0.34**
32	ARMAX, d - DR -SIC	0.075	0.12	0.16	0.26	0.38	0.44	0.57	**0.34**
38	ARMA errors, d - DR -SIC	0.08	0.13	0.16	0.27	0.39	0.43	0.57	**0.34**
42	ARDL, d - MRR - SIC	0.078	0.12	0.16	0.27	0.38	0.43	0.57	**0.34**
44	ARDL, d - DR -SIC	0.075	0.12	0.16	0.26	0.38	0.44	0.57	**0.34**
Avg.		**0.13**	**0.17**	**0.21**	**0.30**	**0.42**	**0.48**	**0.64**	**0.40**
	Multivariate models - spread								
47	ARMA, s - MRR - AIC	0.11	0.16	0.21	0.3	0.46	0.58	0.93	**0.49**
48	ARMA, s - MRR - SIC	0.09	0.14	0.18	0.3	0.42	0.55	0.94	**0.47**
50	ARMA, s - DR - SIC	0.093	0.15	0.2	0.32	0.47	0.61	1.06	**0.53**
52	ARMA, s, d - MRR - SIC	0.083	0.13	0.17	0.28	0.39	0.44	0.58	**0.35**
53	ARMA, s, d - DR - AIC	0.094	0.15	0.19	0.28	0.39	0.46	0.58	**0.36**
54	ARMA, s, d - DR - SIC	0.075	0.12	0.16	0.26	0.38	0.43	0.57	**0.34**
Avg.		**0.09**	**0.15**	**0.19**	**0.29**	**0.42**	**0.52**	**0.78**	**0.43**
	Multivariate models - break								
55	Static regression, b - MRR	0.18	0.21	0.25	0.33	0.41	0.47	0.68	**0.44**
56	Static regression, b - DR	0.22	0.26	0.29	0.35	0.4	0.46	0.6	**0.43**
58	ARMA errors, b - MRR	0.11	0.19	0.24	0.33	0.4	0.42	0.66	**0.41**
61	ARMA, s, b - MRR	0.1	0.17	0.21	0.33	0.42	0.48	0.78	**0.45**
62	ARMA, s, b - DR	0.11	0.19	0.23	0.33	0.39	0.46	0.73	**0.43**
Avg.		**0.14**	**0.19**	**0.24**	**0.33**	**0.40**	**0.45**	**0.69**	**0.43**
	Multiequation models								
64	Levels - MRR - SIC	0.075	0.13	0.17	0.31	0.57	0.9	1.56	**0.68**
66	Levels - DR - SIC	0.071	0.13	0.18	0.35	0.65	0.83	1.67	**0.77**
68	Differences - MRR - SIC	0.067	0.11	0.14	0.24	0.41	0.64	0.77	**0.46**
70	Differences - DR - SIC	0.066	0.11	0.14	0.25	0.4	0.54	0.57	**0.38**
72	VECM - MRR - SIC	0.075	0.13	0.17	0.31	0.57	0.9	1.56	**0.68**
74	VECM - DR - SIC	0.071	0.13	0.18	0.35	0.65	0.83	1.67	**0.77**
Avg.		**0.07**	**0.13**	**0.18**	**0.34**	**0.63**	**0.91**	**1.49**	**0.71**
	Structural models								
101a	Nonlinear - RW ERs	0.093	0.14	0.17	0.27	0.37	0.41	0.62	**0.36**
101b	Nonlinear - RW ERs - ARMA errors	0.084	0.14	0.17	0.27	0.36	0.41	0.61	**0.36**
101c	Nonlinear - ARMA ERs	0.093	0.16	0.2	0.31	0.5	0.62	1.14	**0.54**
101d	Nonlinear - ARMA ERs - ARMA errors	0.084	0.16	0.19	0.31	0.47	0.59	1.13	**0.54**
Avg.		**0.09**	**0.15**	**0.18**	**0.29**	**0.43**	**0.51**	**0.88**	**0.45**
	Summary statistics								
Mean			0.18	0.22	0.35	0.49	0.63	1.00	0.55
Std. Deviation			0.08	0.09	0.10	0.15	0.31	0.63	0.30
Minimum			0.11	0.14	0.22	0.36	0.41	0.44	0.30
1st quartile			0.13	0.17	0.28	0.39	0.44	0.59	0.36
3rd quartile			0.17	0.23	0.38	0.55	0.74	1.13	0.66
Maximum			0.46	0.49	0.68	1.19	2.21	4.41	2.48

Notes: See note below Table 2.2 for explanation of values. "RW ERs" and "ARMA ERs" refers to excess reserves being forecast by a random walk model and an ARMA model, respectively. See Appendix 2.A for details.

Table 2.6: Unonditional forecasting performance for the Eonia rate (2015-2019)

#	Description	1	2	3	6	12	24	48	Avg.
		Forecast horizons (# of steps ahead)							
		Univariate models							
2	ARMA - SIC	0.02	0.044	0.077	0.2	0.49	1.02	1.85	**0.53**
4	ARMA, d - SIC	*0.0092*	*0.015*	*0.018*	*0.032*	0.061	0.092	0.25	**0.068**
Avg.		**0.02**	**0.03**	**0.05**	**0.12**	**0.30**	**0.58**	**1.08**	**0.31**
		Multivariate models - level							
7	Static regerssion - MRR&DR	0.049	0.045	0.042	0.034	0.02	0.023	0.038	**0.036**
9	ARMAX - MRR - SIC	*0.015*	*0.027*	*0.035*	*0.044*	*0.016*	*0.015*	*0.031*	**0.026**
13	ARMAX - MRR&DR - SIC	0.022	0.043	0.064	*0.1*	*0.056*	*0.052*	0.11	**0.063**
14	ARMA errors - MRR - AIC	0.014	0.023	0.031	0.054	0.11	0.18	0.27	**0.098**
21	ARDL - MRR - SIC	*0.011*	0.02	0.029	0.055	0.1	0.17	0.32	**0.1**
25	ARDL - MRR&DR - SIC	*0.0082*	*0.012*	*0.015*	*0.021*	*0.034*	*0.024*	0.051	**0.024**
Avg.		**0.03**	**0.04**	**0.05**	**0.07**	**0.11**	**0.16**	**0.31**	**0.11**
		Multivariate models - differences							
26	Static regerssion, d - MRR	0.057	0.064	0.071	0.088	0.12	*0.17*	0.37	**0.14**
30	ARMAX, d - MRR - SIC	*0.0084*	*0.014*	*0.02*	*0.037*	*0.071*	*0.11*	0.28	**0.078**
32	ARMAX, d - DR -SIC	*0.0084*	*0.014*	*0.02*	*0.037*	*0.071*	*0.11*	0.28	**0.078**
38	ARMA errors, d - DR -SIC	*0.0088*	*0.015*	*0.021*	*0.037*	*0.071*	*0.11*	0.28	**0.078**
42	ARDL, d - MRR - SIC	*0.0091*	*0.015*	*0.021*	0.038	*0.072*	*0.11*	0.28	**0.079**
44	ARDL, d - DR -SIC	*0.0084*	*0.014*	*0.02*	*0.037*	*0.071*	*0.11*	0.28	**0.078**
Avg.		**0.02**	**0.02**	**0.03**	**0.04**	**0.08**	**0.12**	**0.29**	**0.09**
		Multivariate models - spread							
47	ARMA, s - MRR - AIC	*0.013*	0.022	0.03	0.055	0.11	0.2	0.27	**0.1**
48	ARMA, s - MRR - SIC	*0.013*	0.023	0.033	0.061	0.11	0.17	0.27	**0.098**
50	ARMA, s - DR - SIC	0.015	0.027	0.04	0.077	0.15	0.26	0.55	**0.16**
52	ARMA, s, d - MRR - SIC	*0.01*	*0.016*	*0.022*	0.039	*0.073*	*0.12*	0.29	**0.08**
53	ARMA, s, d - DR - AIC	*0.012*	*0.018*	*0.022*	*0.035*	*0.066*	*0.11*	0.27	**0.076**
54	ARMA, s, d - DR - SIC	*0.0088*	*0.015*	*0.021*	*0.037*	*0.071*	*0.11*	0.28	**0.078**
Avg.		**0.01**	**0.02**	**0.03**	**0.05**	**0.10**	**0.17**	**0.35**	**0.11**
		Multivariate models - break							
55	Static regression, b - MRR	0.06	0.066	0.072	*0.089*	*0.12*	*0.17*	*0.32*	**0.13**
56	Static regression, b - DR	0.066	0.064	0.063	0.061	0.057	0.069	0.079	**0.066**
58	ARMA errors, b - MRR	*0.0082*	*0.013*	*0.016*	*0.024*	*0.027*	*0.035*	0.061	**0.026**
61	ARMA, s, b - MRR	0.025	0.05	0.072	0.12	0.19	0.24	0.38	**0.15**
62	ARMA, s, b - DR	*0.0084*	*0.013*	*0.016*	*0.022*	*0.02*	*0.022*	0.039	**0.02**
Avg.		**0.03**	**0.04**	**0.05**	**0.07**	**0.09**	**0.12**	**0.22**	**0.09**
		Multiequation models							
64	Levels - MRR - SIC	*0.0084*	*0.014*	*0.019*	*0.037*	*0.085*	0.17	0.57	**0.13**
66	Levels - DR - SIC	0.019	0.044	0.076	0.18	0.43	0.91	2.13	**0.54**
68	Differences - MRR - SIC	*0.011*	*0.019*	*0.027*	*0.057*	*0.11*	*0.24*	0.44	**0.13**
70	Differences - DR - SIC	*0.0093*	*0.017*	*0.025*	*0.052*	*0.11*	*0.18*	0.31	**0.1**
72	VECM - MRR - SIC	*0.0084*	*0.014*	*0.019*	*0.037*	*0.085*	0.17	0.57	**0.13**
74	VECM - DR - SIC	0.019	0.044	0.076	0.18	0.43	0.91	2.13	**0.54**
Avg.		**0.01**	**0.03**	**0.05**	**0.11**	**0.27**	**0.57**	**1.20**	**0.32**
		Structural models							
101a	Nonlinear - RW ERs	0.012	0.015	0.019	0.029	0.057	0.089	0.22	**0.063**
101b	Nonlinear - RW ERs - ARMA errors	0.01	0.017	0.024	0.04	0.072	0.11	0.24	**0.074**
101c	Nonlinear - ARMA ERs	0.012	0.015	0.019	0.029	0.057	0.089	0.22	**0.063**
101d	Nonlinear - ARMA ERs - ARMA errors	0.01	0.017	0.023	0.038	0.068	0.1	0.24	**0.071**
Avg.		**0.01**	**0.02**	**0.02**	**0.03**	**0.06**	**0.10**	**0.23**	**0.07**
		Summary statistics							
Mean			0.03	0.04	0.07	0.13	0.23	0.47	0.14
Std. Deviation			0.02	0.02	0.05	0.13	0.28	0.55	0.15
Minimum			0.01	0.02	0.02	0.02	0.02	0.01	0.02
1st quartile			0.02	0.02	0.04	0.07	0.11	0.27	0.08
3rd quartile			0.04	0.06	0.08	0.13	0.23	0.44	0.14
Maximum			0.15	0.15	0.25	0.65	1.37	2.72	0.74

Notes: See note below Tables 2.2 and 2.5 for explanation of values.

There are three takeaways from the tables. First, when compared against alternative groups of models, the structural model is still significantly better than any group of reduced form models in forecasting Eonia rate.[13] That said, the improvement of forecasting performance is much smaller than in the case of ex-post performance, and there are individual reduced form models that are better at individual forecasting horizons or even overall. Second, among the reduced form models, it is again models that use first differences of IIRs (or spread from policy rates) that perform best. Third, multivariate single equation models provide only marginal improvement, if any, over the univariate models that are specified in terms of first differences. While this seems to suggest that such univariate models offer a better way forward in terms of forecasting IIRs - they perform well and do not require forecasting of other variables - one should approach this conclusion with caution: the result might be an artifact of the nature of the sample, since the model generally predicts that IIRs will remain at current levels for the whole forecast period irrespective of evolution in policy rates, which could prove to be a bad forecasting rule if and when policy rates start rising.

It is clear that structural models are significantly better than the alternatives only in the period of the QE program. Before the start of the QE program, uncertainty about the future path of excess reserves results in forecasts being only slightly better than forecasts from reduced form models. Moreover, many of the reduced form models that use first differences as the dependent variable are actually slightly better, even though as a group they are slightly worse. In contrast, during the period of the QE program, when the evolution of excess reserves can

[13]Note that this is true only when one uses a random walk model for excess reserves. I will return to this observation later in this section.

be quite precisely forecast thanks to the predictable nature of asset purchases (see Appendix 2.A), the structural model is better than alternative groups and even alternative individual models.

The previous discussion ignored an important aspect of the results which underscores this point: in the sample before the QE program, the better forecasting performance of the structural model is true only if one uses the a-priori less plausible random walk model for excess reserves. If one instead uses the ARMA model, which ensures that excess reserves return to zero values, then the forecasting performance is substantially worse. This reflects that the expectation that excess reserves would return to zero values was never borne out in the given sample; correspondingly, over the period from the beginning of 2009 to the beginning of 2015, expecting excess reserves to remain unchanged turned out to be a better forecast; see Appendix 2.A for details. Of course, this may be just a random feature of the sample, and in any case is not direct criticism of the model for IIRs.

This then points towards the key takeaway with respect to the structural model: the model provides significant value even in ex-ante forecasting *as long as one can forecast excess reserves reliably* and specifically in periods when QE program is operational, since during these periods the ECB is buying assets with pre-defined amounts. In contrast, when one is unable to make such reliable forecasts, then more simple, reduced form models using first differences of IIRs as the dependent variable might be better for forecasting purposes. This is thanks to their simpler, less correct structure, that does not include excess reserves as independent variable.

As for 3-month Euribor[14], Tables 2.7 and 2.8 show that the conclusions from the Eonia rate broadly hold even here. That said, the better performance of

[14]I relegate the results for 1-week Euribor to Appendix 2.B.

structural models is even weaker and the univariate reduced form model in first differences looks even better than for Eonia. This suggests that after accounting for uncertainty in forecasting excess reserves the structural model probably holds only limited value for forecasting IIRs with longer maturity.

2.4 Nonstationarity, cointegration and forecasting performance

The previous section discussed which model features correlate with good forecasting performance and demonstrated the superior forecasting performance of the structural model compared to standard reduced-form models. This raises the question whether one can explain which features are important and why the structural model is better than the reduced form models. This section does that by linking both to the statistical nature of the considered time series. Specifically, the IIRs over the relevant sample are both non-stationary and cointegrated with monetary policy variables, which makes it difficult for reduced-form models to predict them.

This section is separated into two parts, each corresponding to one statistical aspect of the time series. The first part shows how nonstationarity of the IIRs creates issues for models that use the level of IIRs as the dependent variable, which ignore the evolution in the relationship between IIRs and policy rates. The second part then discusses how ignoring cointegration leads to suboptimal forecasts when models with differences of IIRs as the dependent variable are used.

Table 2.7: Unonditional forecasting performance for 3-month Euribor (2009-2015)

#	Description	1	2	3	6	12	24	48	Avg.
		Forecast horizons (# of steps ahead)							
	Univariate models								
2	ARMA - SIC	*0.054*	*0.12*	*0.2*	*0.47*	0.96	1.70	2.70	**1.14**
4	ARMA, d - SIC	*0.039*	*0.078*	*0.11*	*0.24*	*0.45*	*0.62*	0.71	**0.37**
Avg.		**0.06**	**0.14**	**0.23**	**0.51**	**1.01**	**1.77**	**1.99**	**0.99**
	Multivariate models - level								
5	Static regerssion - MRR	0.24	0.28	0.32	*0.41*	*0.53*	0.69	1.02	**0.59**
9	ARMAX - MRR - SIC	*0.073*	*0.16*	*0.26*	*0.45*	*0.54*	0.7	1.07	**0.57**
13	ARMAX - MRR&DR - SIC	*0.08*	*0.18*	*0.29*	0.52	*0.63*	*0.74*	*0.59*	**0.47**
19	ARMA errors - MRR&DR - SIC	*0.092*	*0.17*	*0.23*	*0.39*	*0.59*	*0.82*	1.30	**0.63**
21	ARDL - MRR - SIC	*0.075*	*0.15*	*0.22*	*0.37*	*0.54*	0.77	1.16	**0.58**
25	ARDL - MRR&DR - SIC	*0.1*	*0.2*	*0.29*	0.54	*0.89*	*1.37*	2.78	**1.46**
Avg.		**0.13**	**0.23**	**0.32**	**0.48**	**0.74**	**1.23**	**5.01**	**3.12**
	Multivariate models - differences								
26	Static regerssion, d - MRR	*0.077*	*0.14*	*0.19*	*0.34*	*0.53*	*0.68*	0.9	**0.5**
30	ARMAX, d - MRR - SIC	*0.06*	*0.12*	*0.17*	*0.32*	*0.51*	*0.66*	*0.87*	**0.47**
38	ARMA errors, d - DR -SIC	*0.077*	*0.14*	*0.19*	*0.34*	*0.53*	*0.68*	0.9	**0.5**
44	ARDL, d - DR -SIC	*0.085*	*0.17*	*0.23*	*0.38*	*0.55*	*0.69*	*0.95*	**0.54**
Avg.		**0.09**	**0.16**	**0.22**	**0.37**	**0.56**	**0.72**	**0.96**	**0.54**
	Multivariate models - spread								
48	ARMA, s - MRR - SIC	*0.092*	*0.16*	*0.22*	*0.38*	*0.56*	*0.77*	*1.16*	**0.59**
52	ARMA, s, d - MRR - SIC	*0.08*	*0.14*	*0.2*	*0.34*	*0.53*	*0.68*	0.91	**0.5**
Avg.		**0.09**	**0.17**	**0.23**	**0.39**	**0.60**	**0.80**	**1.12**	**0.59**
	Multivariate models - break								
57	Static regression, b - MRR&DR	0.26	0.3	0.34	*0.44*	*0.56*	*0.69*	1.10	**0.63**
59	ARMA errors, b - DR	0.11	0.19	0.26	0.43	0.61	0.67	1.10	**0.6**
62	ARMA, s, b - DR	0.082	0.15	0.21	*0.36*	*0.55*	*0.8*	1.24	**0.59**
Avg.		**0.15**	**0.21**	**0.27**	**0.40**	**0.57**	**0.74**	**1.17**	**0.61**
	Multiequation models								
66	Levels - DR - SIC	*0.059*	*0.12*	*0.19*	*0.46*	0.98	1.71	2.64	**1.12**
68	Differences - MRR - SIC	*0.04*	*0.08*	*0.11*	*0.23*	*0.46*	*0.68*	*0.34*	**0.29**
74	VECM - DR - SIC	*0.059*	*0.12*	*0.19*	*0.46*	0.98	1.71	2.64	**1.12**
Avg.		**0.06**	**0.11**	**0.18**	**0.42**	**0.88**	**1.52**	**1.97**	**0.90**
	Structural models								
100a	Linear - RW ERs	0.17	0.2	0.24	0.36	0.5	0.62	0.91	**0.52**
100b	Linear - RW ERs - ARMA errors								
100c	Linear - ARMA ERs	0.17	0.21	0.26	0.39	0.56	0.69	1.16	**0.6**
100d	Linear - ARMA ERs - ARMA errors								
Avg.		**0.09**	**0.10**	**0.13**	**0.19**	**0.27**	**0.33**	**0.52**	**0.28**
	Summary statistics								
Mean			0.18	0.25	0.42	0.69	1.05	2.33	1.35
Std. Deviation			0.09	0.12	0.12	0.32	0.99	9.05	5.89
Minimum			0.08	0.11	0.22	0.44	0.46	0.27	0.26
1st quartile			0.14	0.20	0.36	0.53	0.68	0.91	0.50
3rd quartile			0.19	0.27	0.46	0.70	0.91	1.36	0.73
Maximum			0.77	1.13	1.05	2.68	8.39	80.00	52.00

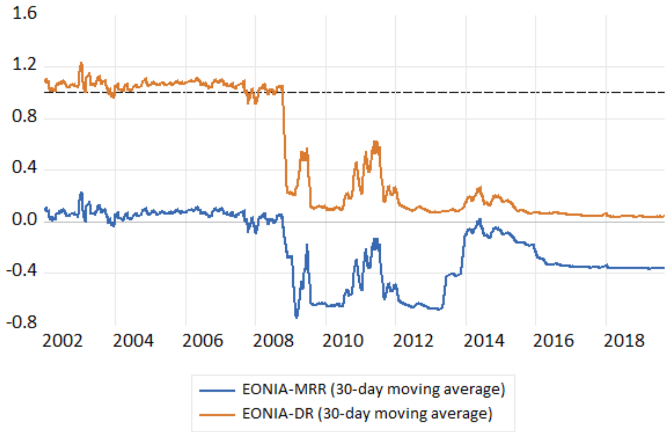
Notes: See note below Tables 2.2 and 2.5 for explanation of values.

Table 2.8: Unonditional forecasting performance for 3-month Euribor (2015-2019)

#	Description	Forecast horizons (# of steps ahead) 1	2	3	6	12	24	48	Avg.
	Univariate models								
2	ARMA - SIC	*0.019*	*0.05*	0.088	0.23	0.52	1.03	1.85	**0.54**
4	ARMA, d - SIC	*0.0047*	*0.01*	*0.016*	*0.034*	*0.079*	*0.14*	0.31	**0.084**
Avg.		**0.02**	**0.03**	**0.06**	**0.14**	**0.34**	**0.67**	**1.14**	**0.34**
	Multivariate models - level								
5	Static regerssion - MRR	0.11	0.11	0.11	*0.12*	*0.14*	*0.17*	0.18	**0.13**
9	ARMAX - MRR - SIC	*0.028*	*0.065*	0.1	0.16	*0.17*	*0.19*	*0.2*	**0.13**
13	ARMAX - MRR&DR - SIC	*0.027*	*0.063*	0.097	0.16	*0.15*	*0.13*	*0.051*	**0.096**
19	ARMA errors - MRR&DR - SIC	*0.022*	*0.044*	*0.065*	0.13	0.25	0.45	0.77	**0.25**
21	ARDL - MRR - SIC	*0.019*	*0.041*	*0.062*	*0.11*	0.19	0.27	0.33	**0.15**
25	ARDL - MRR&DR - SIC	*0.016*	*0.032*	*0.047*	*0.078*	*0.094*	*0.098*	0.04	**0.058**
Avg.		**0.05**	**0.07**	**0.10**	**0.16**	**0.24**	**0.34**	**0.44**	**0.20**
	Multivariate models - differences								
26	Static regerssion, d - MRR	*0.0082*	*0.016*	*0.024*	*0.047*	*0.095*	*0.16*	0.34	**0.097**
30	ARMAX, d - MRR - SIC	*0.0059*	*0.013*	*0.02*	*0.043*	*0.091*	*0.15*	0.33	**0.093**
38	ARMA errors, d - DR -SIC	*0.0082*	*0.016*	*0.024*	*0.047*	*0.095*	*0.16*	0.34	**0.097**
44	ARDL, d - DR -SIC	*0.0066*	*0.014*	*0.021*	*0.041*	*0.085*	*0.14*	0.31	**0.089**
Avg.		**0.01**	**0.02**	**0.02**	**0.05**	**0.09**	**0.16**	**0.33**	**0.10**
	Multivariate models - spread								
48	ARMA, s - MRR - SIC	*0.027*	*0.052*	*0.076*	0.14	0.25	0.36	0.44	**0.19**
52	ARMA, s, d - MRR - SIC	*0.0079*	*0.016*	*0.024*	*0.047*	*0.095*	*0.16*	0.34	**0.097**
Avg.		**0.02**	**0.04**	**0.06**	**0.10**	**0.21**	**0.34**	**0.52**	**0.18**
	Multivariate models - break								
57	Static regression, b - MRR&DR	0.058	0.063	0.068	0.087	0.13	0.17	0.31	**0.13**
59	ARMA errors, b - DR	0.015	0.034	0.051	0.1	0.2	0.36	0.69	**0.21**
62	ARMA, s, b - DR	0.011	0.028	0.043	0.088	0.17	0.31	0.56	**0.17**
Avg.		**0.04**	**0.06**	**0.07**	**0.11**	**0.18**	**0.28**	**0.49**	**0.17**
	Multiequation models								
66	Levels - DR - SIC	*0.02*	*0.047*	*0.08*	0.19	0.42	0.77	1.29	**0.4**
68	Differences - MRR - SIC	*0.0075*	*0.016*	*0.023*	*0.052*	*0.12*	*0.21*	0.33	**0.11**
74	VECM - DR - SIC	*0.02*	*0.047*	*0.08*	0.19	0.42	0.77	1.29	**0.4**
Avg.		**0.01**	**0.03**	**0.05**	**0.13**	**0.31**	**0.62**	**1.17**	**0.33**
	Structural models								
100a	Linear - RW ERs	0.052	0.055	0.059	0.074	0.11	0.15	0.26	**0.11**
100b	Linear - RW ERs - ARMA errors								
100c	Linear - ARMA ERs	0.052	0.055	0.059	0.074	0.11	0.15	0.26	**0.11**
100d	Linear - ARMA ERs - ARMA errors								
Avg.		**0.03**	**0.03**	**0.03**	**0.04**	**0.06**	**0.08**	**0.13**	**0.06**
	Summary statistics								
Mean			0.04	0.06	0.11	0.20	0.34	0.57	0.19
Std. Deviation			0.04	0.04	0.07	0.13	0.26	0.48	0.13
Minimum			0.01	0.02	0.03	0.08	0.10	0.04	0.06
1st quartile			0.02	0.02	0.05	0.10	0.16	0.31	0.10
3rd quartile			0.05	0.08	0.15	0.28	0.46	0.70	0.25
Maximum			0.29	0.30	0.33	0.65	1.37	2.09	0.65

Notes: See note below Tables 2.2 and 2.5 for explanation of values.

Figure 2.4: Eonia rate spreads

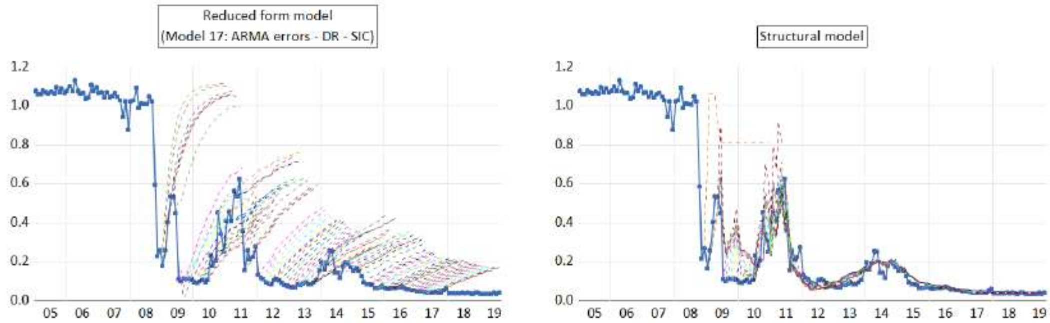


2.4.1 Consequences of nonstationarity for forecasting performance

To appreciate the nonstationary nature of IIRs, one can consult Figure 2.4, which shows the spread between the Eonia rate and the two policy rates. It is clear that while the spread has a stable mean in the period before the fall of 2008, afterwards the mean is not only different but also highly variable. The former suggests a potentially stationary series with a break, with different mean values before and after the break, something that reduced form models could deal with by allowing for a break in the relationship between IIRs and policy rates. However, as I show later, simply accounting for the break on its own does solve all the problems, since even after the break, the spread is not stationary. I illustrate these points by showing ex-post recursive forecasts from representative reduced-form models and comparing them with the forecasts from the structural model.

First, consider Figure 2.5, the left panel of which shows forecasts from one

Figure 2.5: Effect of nonstationarity on forecast performance I



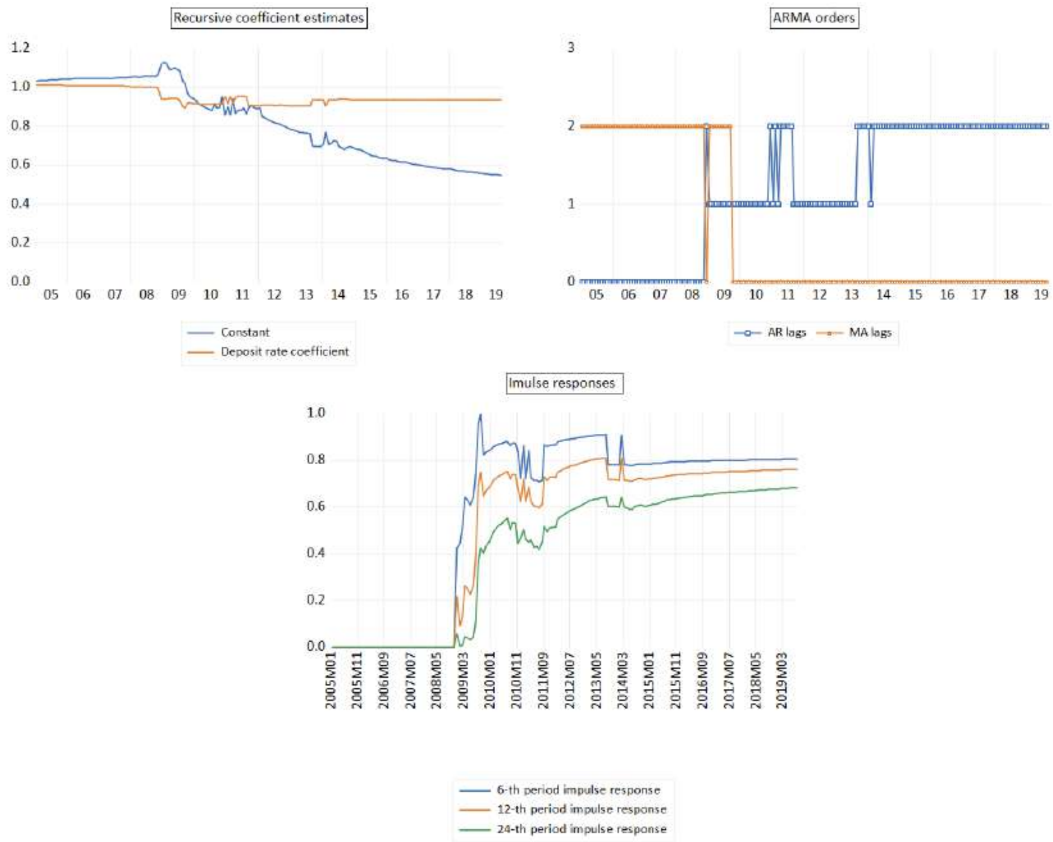
Notes: The figure shows the spread between the Eonia rate and the deposit rate. For more details see comments below Figure 2.2

of the best models using the level of IIRs as the dependent variable, the model with the DR and ARMA errors. The figure clearly shows a key obstacle facing reduced form models in levels: the models are unaware of the large permanent decrease in the average value of the spread between the Eonia and the DR around 2008M10. Correspondingly, the models consistently predict an increase in the Eonia rate as they expect a return to the historical average relationship between IIRs and policy rates. This is at odds with the observed behavior, where the Eonia rate spread remained at values close to 0 except for three periods of mild increase in 2009, 2010-2012, and 2013-2014. This poor forecast quality is in contrast to the structural model, which also uses (transformation of) the level of IIRs as the dependent variable: with the exception of the very first forecast in 2009, the model correctly predicts that the Eonia rate will remain at much lower values than prior to the break.

It is instructive to see how the nonstationarity of the Eonia spread translates into structure and coefficient estimates of the model. Figure 2.6 provides three different ways to view these effects. The top-left panel shows the recursive coefficient estimates for a constant and for a coefficient on the DR. This captures the effect of the break, with the constant gradually decreasing as the sample includes more post-break observations. That said, the decline is not perfectly monotonous and even at the end of the sample, the constant term is not stable. The right-top panel shows the selected ARMA structure. At around the break date the SIC stops selecting models with 2 moving average terms and instead starts selecting models with one, and later, two autoregressive terms. Of course, the difference between those is that models with autoregressive terms are much more likely to feature strong persistence. This is confirmed by the bottom panel, which shows the size of the impulse response after 6, 12 and 24 periods corresponding to models estimated on samples with increasing length. While before the break the models feature zero responses to shocks at these medium-to-long horizons, after the break the models rapidly start to feature strong and long persistence. This is of course in contrast to the structural model, which does not feature any persistence at all. The overall conclusion from this figure is straightforward: as the sample increases, the model starts to look more and more like a random walk model.

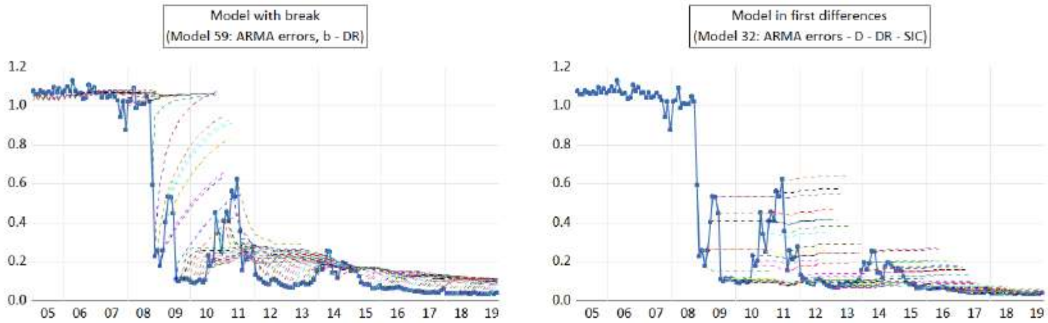
Figure 2.5 raises a question: could the problematic forecasting performance be addressed by a reduced-form model which allows for a break in the relationship between the IIRs and policy rates? To answer this question, the left panel of Figure 2.7 shows results for similar model that allows for a break in the model. Specifically, the figure shows results for model 59, which allows for a different relationship between IIRs and DR after 2008M10 if this is supported by formal a Chow break test (see Appendix 2.A for details). While the model provides marked improvement over the simple original reduced form model, it still features the

Figure 2.6: Recursive model structure and estimates



Notes: Coefficient estimates, selected ARMA orders, and impulse responses corresponding to estimation samples of different length. The date indicates the last month included in the estimation sample.

Figure 2.7: Effect of nonstationarity on forecast performance II



same problematic tendency for the forecasts to trend away from the current level: for most of the sample the forecasts are trending upwards, but in contrast to the previous model there are also periods when the forecasts are trending downwards. This reflects the other aspect of the behavior in the Eonia spread stressed above, that while the spread is lower in the period after 2008, it is no longer stable as before 2008. Correspondingly, even after allowing for the break, the model has a tendency to return to average values, reflecting its use of the level of IIRs as dependent variable; the only difference is that these values are lower than before 2008M10. Since constant (average) value is a poor prediction for the Eonia spread even if one uses only the sample after the break, the forecasting performance is still sub-optimal.

The above discussion makes clear that simply allowing for a break in the relationship between IIRs and policy rates does not completely alleviate the problem caused by the change in regime. The only other alternative in terms of reduced-foirm models is to change the transformation of the dependent variable from the level of IIRs to first-differences of IIRs. The right panel of Figure 2.7 shows results for one of the best performing first-difference models. The good news is that using first-differences instead of levels addresses the main concern raised in this section: the Eonia spread no longer has the tendency to return to its pre-208M10 values. Instead, the Eonia spread is predicted to remain unchanged at current values, potentially with some small fluctuations in near term. While this constitutes improvement, there are periods when such a forecast is problematic. Specifically, the constant-spread prediction turned out to be wrong in the three periods when the Eonia rate spread increased due to a decreases in excess reserves, and in the sample covering the QE program, when the Eonia rate kept decreasing as excess reserves were increasing. While the former forecast misses were likely obvious only with hindsight, because the future path of excess reserves was uncertain at that time, the misses during the QE period could have been (and were) anticipated.

2.4.2 Consequences of cointegration for forecasting performance

The message of the previous subsection would seem to be that reduced-form models with the level of IIRs as a dependent variable are doomed to produce problematic forecasts for IIRs, and that models using first differences should be used instead. However, as this subsection will show, first-difference models face different, equally serious obstacles in producing forecasts for IIRs.

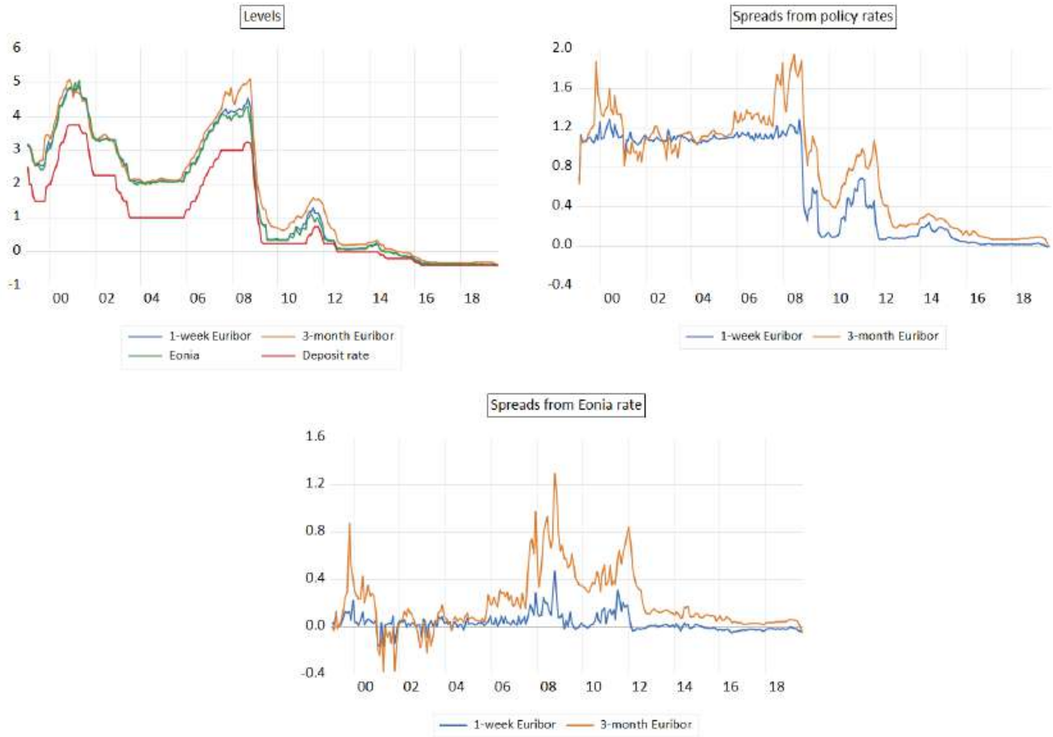
Indeed, Figure 2.7 at the end of the previous section already illustrates that changing the dependent variable from the level of IIRs to the first differences of IIRs does not lead to fully satisfactory forecasting results. However, this figure provides a somewhat flattering view of the problems faced by models in first differences, since it shows results only for Eonia rate. The picture becomes substantially worse when one considers IIRs with longer maturity - or alternatively when one would focuses on higher frequency, at which the Eonia rate features more transitory shocks.

Figure 2.8 shows the 1-week and 3-month Euribor rates together with the Eonia and deposit rate; the top left panel shows the levels, while the remaining two show spreads from the DR and the Eonia rate, respectively. The motivation for this particular series and visualization follows. The main difference between the Eonia and Euribor rates is the presence of significant risk components. The defining feature of these is that they are stationary: while risk components can fluctuate significantly, they can be expected to always return to equilibrium (near zero) value. This can be seen in bottom panel Figure 2.8, which clearly shows that deviations of Euribor from the Eonia rate are transitory.

Since reduced-form models treat all movements in the series the same way, Figure 2.8 suggests that using models in first differences will lead to suboptimal forecasting performance, since these models treat all shocks to Euribors as permanent. This is clearly not the case for movements caused by fluctuations in the risk component, such as during periods of financial stress; the bottom panel shows several such examples, most importantly in the period of the global financial crisis in 2007-2009, and the period of the euro zone sovereign debt crisis in 2011-2012. Importantly, these movements can easily swamp the other sources of movements in IIRs relative to policy rates, especially for the longer-maturity Euribor. For example, the spread between 3-month Euribor and the Eonia rates reached 1.2% during its peak month in 2008, and 0.8% during its peak month in 2011, but decreased back to near-zero values afterwards.

Therefore, models using first differences are likely to feature large forecast errors

Figure 2.8: Euribors rate spreads



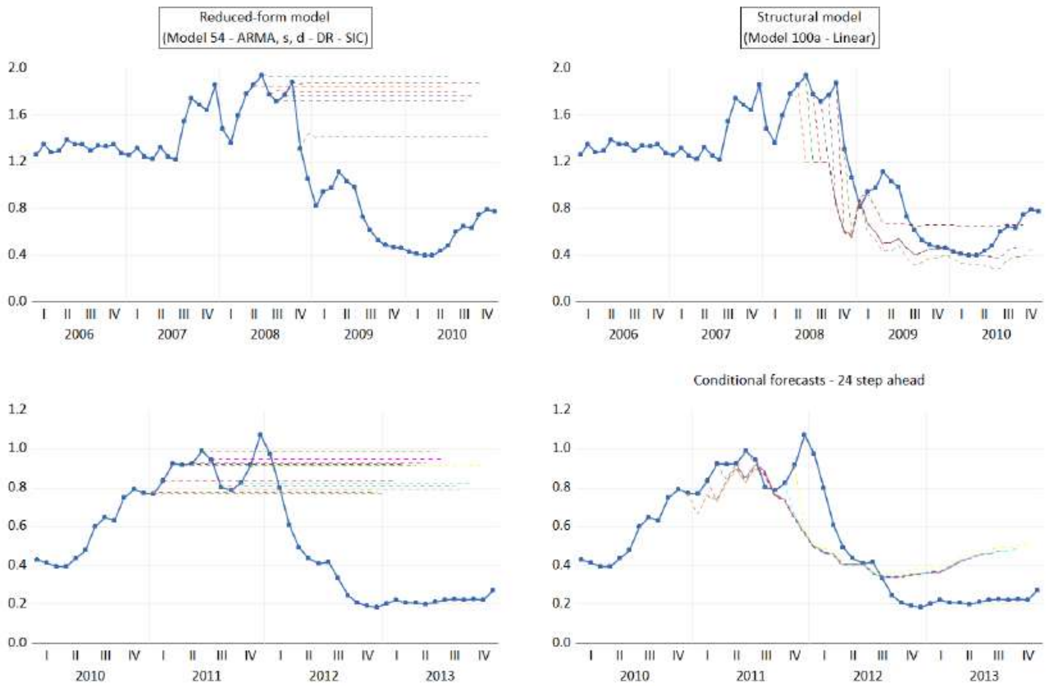
in periods following stress in financial markets. This is indeed on display in Figure 2.9, the left panels of which show the forecasts for the 3-month Euribor from one of the best performing models using first differences, the model with changes in the spread from the DR selected by SIC (model 54). The top left panel shows forecasts during the peak of the global financial crisis, while the bottom right panel shows the forecasts during the peak of the euro zone sovereign debt crisis; the right panels show corresponding forecasts from the structural model. As outlined above, the forecasts have rather poor quality, being characterized by almost constant spread between the Euribor and the DRs at the levels last observed. This is despite the fact that the spreads are elevated relative to historical values, and despite the fact that these values vary from forecast to forecast. Both of these facts are direct consequences of using first differences as the dependent variable, since such models completely ignore the level of IIRs when creating the forecasts.

This ex-ante predictable poor forecasting performance is in contrast with the forecasts from the structural model in the right panels of Figure 2.9. Since the model is specified in terms of levels of IIRs, it treats all shocks as temporary, and, correspondingly, whenever the IIR spreads are elevated due to stress in financial markets, the model predicts an (immediate[15]) decrease in the spreads.

Translating this into econometric concepts, the key realization is that, while IIRs are nonstationary, they are clearly cointegrated with policy rates, as the top-left panel of Figure 2.8 shows. In the period before the emergence of excess reserves,

[15] The forecasts are from the structural model that does not include any ARMA errors. While an econometrician would presumable include those in the face of large and persistent movements in risk components - or an economist would adjust the forecast manually - here I focus on the core of the model and ignore these aspects. Note that the structural model is also useful in understanding those movements since it eliminates the effect of movements in the equilibrium component of longer-maturity IIRs due to fluctuations in excess reserves. See Kovar (2020) for an example of such a use of the model.

Figure 2.9: Effects of cointegration on forecast performance



this cointegration means that all fluctuations in IIRs relative to policy rates can expected to correct themselves eventually, something I link to the economic concept of transitory fluctuations in risk components of IIRs. While this cointegration relationship is further complicated by the role played by excess reserves which vary over time, it is not altered completely: the only difference is that the IIRs are not cointegrated only with policy rates, but instead IIRs, policy rates and excess reserves form cointegrating relationship together. Therefore, the poor forecasting performance of models using first differences is a reflection of model mispecification which occurs when one models a cointegrated series in their first differences rather than in levels or as an error-correction model.

2.4.3 Summary

This section explained the superior forecasting performance of the structural model by linking it to the statistical nature of IIRs. Models which specify the level of IIRs as a dependent variable lead to problematic forecasts because they ignore the fact that the historical average relationship between IIRs and policy rates is no longer relevant after the emergence of excess reserves, which cause IIRs to become nonstationary. On the other hand, models with first differences of IIRs as a dependent variable do not account for variations in excess reserves that drive the variation in IIRs; more importantly, they also ignore that longer-maturity IIRs contain potentially large shocks that are transitory, stemming from the variations in their risk component. As a result, they produce misleading forecasts during periods of financial stress, arguably the most important periods for forecasting IIRs.

The key issue is that the IIRs with longer maturity have two components of different statistical nature. The equilibrium component, which can be proxied by the Eonia rate, features a break in the relationship with policy rates, and fluctuates with excess reserves. This makes the equilibrium component clearly nonstationary. However, IIRs with longer maturity also feature a risk component, that is clearly stationary since shocks to this component are eventually reversed. Reduced-form time series models effectively ignore one or the other component: models with levels of IIRs as the dependent variable ignore the fluctuations in the nonstationary component, while models with first differences as the dependent variable ignore the stationary component. This explains the superiority of the structural model: since it accounts for the only source of nonstationarity, it can be cast in term of levels of IIRs without facing the same problems as reduced form models using levels or IIRs. This means that any other fluctuations which are (primary) transitory fluctuations in the risk component are indeed treated as transitory.

2.5 Concluding remarks

This chapter shows that the structural factors highlighted in the first chapter play an important part in forecasting euro zone IIRs in period since the emergence of excess reserves; and correspondingly, that the structural model proposed there is useful not only for making sense of the movements in IIRs over last decade, but also in (pseudo) out of sample forecasting. This is true for both the Eonia rate and the longer maturity Euribor rates, but the superior forecasting quality hinges on knowledge of future movements in excess reserves with some degree of certainty.

Nevertheless, focusing on summary measures of forecasting performance masks important heterogeneity with respect to different periods and environments. Specifically, the best performing reduced form models are completely unsuitable for forecasting during periods of stress in interbank markets. Presumably, forecasts during such periods are of specific value. Moreover, even if the future path of excess reserves is not known, superior ex-post forecasting performance of the structural model means that it is much more useful in scenario analyses - forecasting the future path of IIRs under alternative assumptions about monetary policy - than reduced form models.

Establishing the appropriateness of the proposed structural model for forecasting purposes means that the model can in principle also shed light on the question of the future path of IIRs. The key question is whether and when IIRs will return to their normal values relative to policy rates. The model clearly suggests that this will only happen if and when the excess reserves in the euro zone interbank market disappear, which is intrinsically linked to the ECB's balance sheet returning to its normal size. Therefore, a return to normal is not in the cards in the next several years, or likely at any point during the next decade. That said, the prediction of future path of IIRs is complicated by changes the euro zone IIRs underwent in the fall of 2019: their definition changed, which also coincided with the introduction of tiering of excess reserves by the ECB. It is not immediately theoretically clear what these changes mean exactly, and too few observations have been accumulated thus far to provide definitive answer. I leave this question for future research.

2.A Description of forecasting procedures

This appendix describes the forecasting procedures used to create the forecasts reported in the main text. First, the specifics of the procedure for ex-post forecasting is described, before turning to specifics of ex-ante forecasting.

Ex-post forecasting. For simple reduced form models the pseudo out-of-sample ex-post forecasts are obtained by following a simple recursion. For example, consider the creation of the first forecast in the beginning of 2009M01. In ex-post pseudo out-of-sample forecasts the model coefficient estimates are based on data for dependent and independent variables prior to this date, i.e. the estimation sample runs from 1999M01 until 2008M12. The same also applies to model structure, if it is data-dependent, such as when the order of ARMA terms is determined. The resulting model is then used to create a forecast for the dependent variable, with actual observed history values for independent variables used in place of the future values.

For structural models, the pseudo out-of-sample exercise is further complicated by the regime-switching nature of the model and by the use of an exogenous threshold. In order to avoid artificially improving the out-of-sample forecasting performance by relying on information from the whole sample, I replicate the estimation procedure one would rely on at any given time by using the following decision rules for determining the presence of two regimes. First, for the given estimation sample, the two regime structure is used only if it has a better fit than the single regime structure as measured by the residual sum of squares. Moreover, the threshold value determining the prevalence of each regime is always chosen based on the available sample, using the estimation procedure proposed by Chan (1993); see the first chapter for more discussion. In addition to these rules motivated by fear of contamination of results, I also impose two additional rules motivated by reasonable econometric considerations. First, two regimes are allowed only if there are at least 6 observations for each regime. Second, the two regimes are used only if the coefficient estimates correspond to theory, which specifically means that the coefficient estimates on excess reserves is negative. Presumably, the econometrician making forecasts at given time in history would discard models that do not correspond to theoretical expectations.

While the previous decision rules should ensure that the historical equations are estimated as an econometrician would estimate them at a given time, they cannot deal with additional complications. Specifically, for structural models, the pseudo out-of-sample exercises are further complicated by the fact that the model imposes a significant amount of structure on the relationship between IIRs and monetary policy variables. This structure is clearly informed by developments observed only after the emergence of excess reserves, and presumably would not be obvious before. While one would typically ignore such considerations in a pseudo out-of-sample forecasting exercise, I aim to approximate the true out-of-sample forecasting by adding one more decision rule. The econometrician could have been alerted to the possibility of a change in the data-generating process if he/she would have performed formal statistical tests on his econometric model. In the present case, the formal statistical test is a test for a break in the relationship between IIRs and policy rates.[16]

Finally, the presence of two regimes also raises the issue of discontinuity in forecasts, and the possibility of inconsistency in the regime forecasts: in the pres-

[16]The same decision rule is also used for models that include a break.

ence of excess reserves, the IIRs should logically be lower than in the absence of excess reserves, which means that the forecast from a normal regime should act as an upper bound for the IIRs. Since this aspect cannot be imposed on the model in estimation stage, I capture it in the forecasting stage and impose a final decision rule, under which, at any point, the lower value of the two regime forecasts is used.

The model structure and the econometric nature of the structural models also raises the question of what estimation method one should use. The first chapter shows results for multiple estimation methods varying from simple OLS and TSLS to various cointegration estimation methods and shows that the results are reasonably robust to the use of different estimation methods. Here, I focus on simple OLS as the estimation method, since it is much more likely that it would be used rather than some alternative, more sophisticated method. The results using alternative estimation methods are qualitatively and quantitatively similar.

Ex-ante forecasting. As discussed in the main text and above, the only difference between ex-post and ex-ante forecasting is the treatment of the independent variables, and specifically whether one uses actual historical values or forecasts. In ex-ante forecasts one, also forecasts values for independent variables. Apart from this need to forecast independent variables, the rest of the recursion is the same: since I focus on the case where the independent variables are indeed treated as independent - there is no feedback from the main variable (IIR) to the independent variables (policy rates) - the only thing that changes is that I apply the recursion to more than one equation with suitable ordering of the forecasting. The only question to settle is what model one uses to forecast independent variables. This section delves deeper into the answer to this question, first focusing on policy rates and later on excess reserves.

Table 2.9: Forecasting performance for the main refinancing rate

Model	1	2	3	6	12	24	48	60	Avg.
	Forecast horizons (# of steps ahead)								
ARMA - AIC	0.038	0.088	0.14	0.34	0.87	1.81	2.46	2.67	**1.05**
ARMA - SIC	0.037	0.086	0.14	0.34	0.86	1.80	2.48	2.68	**1.05**
ARIMA - AIC	0.024	0.05	0.072	0.13	0.2	0.32	0.6	0.73	**0.27**
ARIMA - SIC	0.021	0.047	0.073	0.13	0.21	0.33	0.62	0.76	**0.27**
Random walk	0.029	0.055	0.08	0.13	0.2	0.33	0.67	0.89	**0.3**

Notes: See note below Table 2.2 for explanation of values.

For the two monetary policy rates - the main refinancing rate and the deposit rate - I consider either reduced form models of the ARMA class, or simple random walk. For ARMA model class, I consider both models in levels and in differences, and models selected by AIC and SIC. Tables 2.9 and 2.10 include the MAE for MRR and DR, respectively. In both cases, the results are for a sample that corresponds to the sample reported in main text, 2009M01-2019M08. Not surprisingly, the models that use the level of policy rates as a dependent variable perform much worse then models that use first differences, corresponding to the fact that expecting policy rates to return to their historically normal values proved to be a disappointing endeavor. Meanwhile, there are only slight differences between the estimated models in first differences and the random walk. Given its simplicity and easy interpretability, and greater robustness due to the lack of estimation, I use the random walk model for forecasting policy rates. Figure 2.10 shows the forecasts up to a 24 month horizon for both policy rates.

In case of structural models, ex-ante forecasting brings additional complication since one also needs to forecast the value of excess reserves. Here I take three different approaches, corresponding to three different sample periods. In each

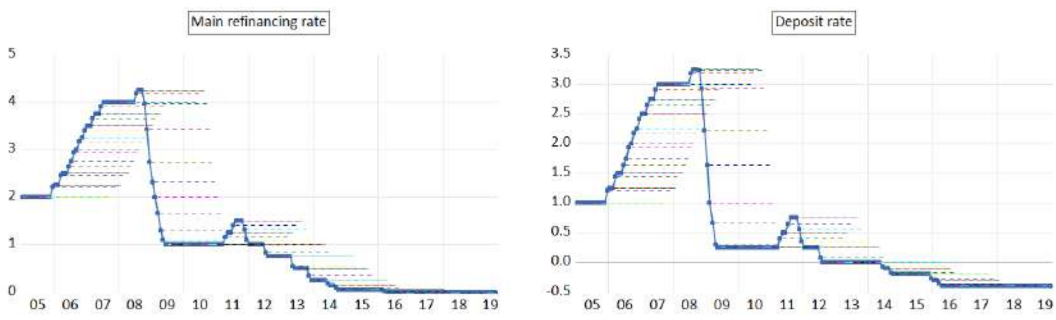
Table 2.10: Forecasting performance for the deposit rate

Model	1	2	3	6	12	24	48	60	Avg.
ARMA - AIC	0.038	0.086	0.14	0.32	0.69	1.31	1.89	2.04	**0.81**
ARMA - SIC	0.04	0.091	0.15	0.35	0.81	1.57	1.99	2.09	**0.89**
ARIMA - AIC	0.029	0.049	0.071	0.13	0.19	0.27	0.43	0.55	**0.21**
ARIMA - SIC	0.02	0.042	0.063	0.12	0.2	0.31	0.49	0.61	**0.23**
Random walk	0.028	0.052	0.072	0.11	0.16	0.25	0.45	0.56	**0.21**

Column header: **Forecast horizons (# of steps ahead)**

Notes: See note below Table 2.2 for explanation of values.

Figure 2.10: Forecasts for policy rates - random walk



case the model is meant to approximate what an econometrician would plausibly propose as a model for excess reserves, or how an economist would forecast excess reserves.

- In the period before the change in the ECB regime from a fixed allotment to a full allotment (1999M01-2008M09), the excess reserves are forecast to be equal to zero, irrespective of their previous value.

- In the period between 2008M10 and 2014M12 the excess reserves are forecast to follow either an ARMA(2,1) model without a constant, or a simple random walk without drift.

- In the period of quantitative easing (2015M01-2019M08), the excess reserves are forecast to increase in line with announced plans for asset purchases; after planned end date for purchases, the purchases are forecast to decrease to zero over a period of 6 months; finally, after end of purchases, the excess reserves are expected to stay unchanged.

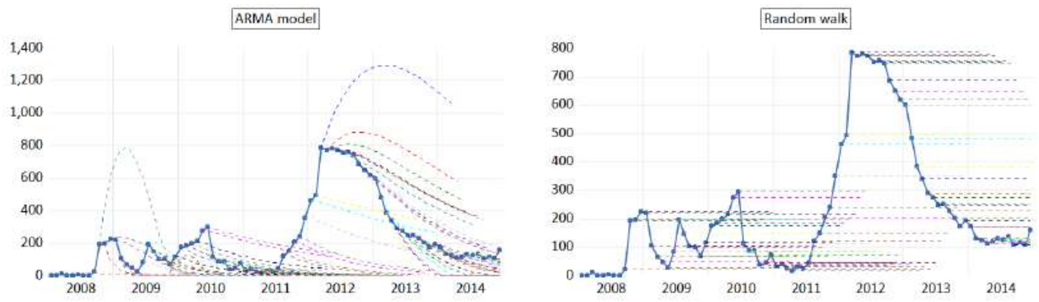
The model for the first period is very natural: during this period, banks had incentives to economize on excess reserves, naturally leading to expectations that excess reserves would always return to zero. This contrast with the period after the emergence of excess reserves. During this period, the persistence of excess reserves was almost immediately visible. Whether an econometrician/economist would consider the presence of excess reserves to be a (semi-)permanent feature, or whether he/she would expect excess reserves to return gradually to zero, is an open question. While the latter seems more reasonable as an ex-ante expectation, the expectation of things remaining as they are is not unreasonable. I consider

Table 2.11: Forecasting performance for excess reserves

	Forecast horizons (# of steps ahead)								
Model	**1**	**2**	**3**	**6**	**12**	**24**	**48**	**60**	**Avg.**
Random walk	38	67	93	158	283	354	146	42	**148**
ARMA	38	70	94	157	258	321	228	133	162

Notes: See note below Table 2.2 for explanation of values.

Figure 2.11: Forecasts for excess reserves - period before the quantitative easing program



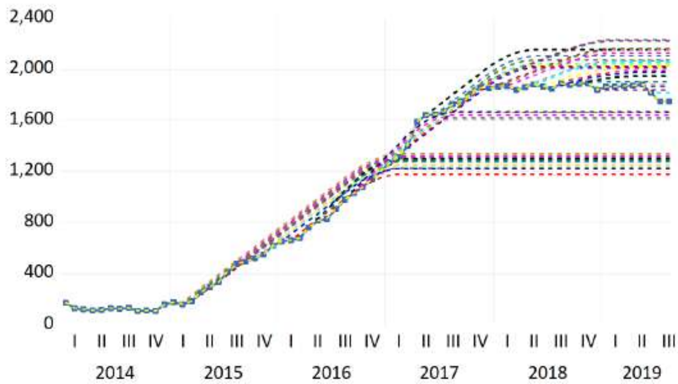
both possibilities in the main text. Of course, ex-post, we know that expecting excess reserves to become a permanent feature of the interbank market was more correct. This is documented in Table 2.11 and Figure 2.11.

Finally, once the QE program started, it was clear that excess reserves were not going to decrease to zero any time soon, but rather that they will continue to increase due to continuous asset purchases by the ECB. Since the ECB announced its asset purchase plans for long periods ahead - see Table 2.12 - the economist could easily follow these announcements and make corresponding forecasts for excess reserves. Figure 2.12 shows the resulting forecast for excess reserves.

Table 2.12: Overview of the ECB announcements about asset purchase plans

Announcement date	Amount (in bil.)	Start date	Plnanned end date
1/22/2015	60	2015M03	2016M09
3/10/2016	80	2016M04	
7/21/2016	80	-	2017M03
12/8/2016	80/60	-	2017M03/2017M12
10/26/2017	60/30	-	2017M12/2018M09
6/14/2018	30/15	-	2018M09/2018M12

Figure 2.12: Forecasts for excess reserves - period of the QE program



2.B Additional results

This Appendix contains additional results. Section 2.B.1 contains results for ex-post forecasting performance reported in the main text, but includes all models. Section 2.B.2 provides results for the full sample from 2005 until 2019.

2.B.1 All model results

This appendix contains analogs of tables 2.2, 2.3 and 2.4, with results for all models included, rather than just the subset of models with better performance. See tables 2.13,2.14 and 2.15. Results for ex-ante forecasting performance are available upon request.

Table 2.13: Ex-post forecasting performance for Eonia - All models

#	Description	1	2	3	6	12	24	48	60	Avg.
		Forecast horizons (# of steps ahead)								
	Multivariate models - level									
5	Static regerssion - MRR	0.21	0.21	0.21	0.22	0.22	0.23	0.24	0.16	**0.21**
6	Static regerssion - DR	0.28	0.28	0.28	0.29	0.3	0.32	0.4	0.43	**0.32**
7	Static regerssion - MRR&DR	0.19	0.2	0.21	0.23	0.28	0.38	0.63	0.7	**0.35**
8	ARMAX - MRR - AIC	0.088	0.15	0.21	0.26	0.23	0.24	0.21	0.11	**0.19**
9	ARMAX - MRR - SIC	0.089	0.15	0.21	0.26	0.23	0.23	0.2	0.11	**0.18**
10	ARMAX - DR - AIC	0.075	0.14	0.21	0.37	0.43	0.43	0.54	0.58	**0.35**
11	ARMAX - DR -SIC	0.076	0.14	0.22	0.39	0.44	0.44	0.54	0.57	**0.35**
12	ARMAX - MRR&DR - AIC	0.093	0.16	0.21	0.33	0.37	0.47	0.83	0.94	**0.43**
13	ARMAX - MRR&DR - SIC	0.097	0.16	0.22	0.33	0.36	0.48	0.81	0.9	**0.42**
14	ARMA errors - MRR - AIC	0.058	0.082	0.099	0.13	0.22	0.25	0.24	0.18	**0.16**
15	ARMA errors - MRR - SIC	0.05	0.072	0.086	0.12	0.19	0.24	0.19	0.12	**0.13**
16	ARMA errors - DR - AIC	0.043	0.069	0.084	0.13	0.19	0.25	0.45	0.51	**0.22**
17	ARMA errors - DR -SIC	0.038	0.057	0.073	0.13	0.2	0.32	0.57	0.67	**0.26**
18	ARMA errors - MRR&DR - AIC	0.061	0.085	0.099	0.13	0.22	0.24	0.23	0.17	**0.15**
19	ARMA errors - MRR&DR - SIC	0.051	0.073	0.086	0.12	0.19	0.23	0.19	0.15	**0.14**
20	ARDL - MRR - AIC	0.053	0.084	0.11	0.16	0.29	0.45	1.11	1.87	**0.51**
21	ARDL - MRR - SIC	0.049	0.072	0.089	0.13	0.21	0.27	0.23	0.15	**0.15**
22	ARDL - DR - AIC	0.044	0.067	0.087	0.13	0.24	0.46	1.78	3.64	**0.81**
23	ARDL - DR -SIC	0.034	0.051	0.064	0.11	0.18	0.23	0.46	0.55	**0.21**
24	ARDL - MRR&DR - AIC	0.068	0.094	0.11	0.17	0.32	0.61	1.12	1.43	**0.49**
25	ARDL - MRR&DR - SIC	0.063	0.09	0.11	0.17	0.29	0.47	0.95	1.22	**0.42**
Avg.		**0.09**	**0.12**	**0.15**	**0.21**	**0.27**	**0.34**	**0.57**	**0.72**	**0.31**
	Multivariate models - differences									
26	Static regerssion, d - MRR	0.038	0.058	0.069	0.1	0.17	0.22	0.26	0.25	**0.15**
27	Static regerssion, d - DR	0.03	0.044	0.053	0.077	0.12	0.11	0.17	0.18	**0.098**
28	Static regerssion, d - MRR&DR	0.039	0.059	0.069	0.095	0.16	0.19	0.2	0.2	**0.12**
29	ARMAX, d - MRR - AIC	0.04	0.059	0.076	0.11	0.18	0.23	0.31	0.3	**0.16**
30	ARMAX, d - MRR - SIC	0.04	0.06	0.073	0.1	0.17	0.23	0.29	0.31	**0.16**
31	ARMAX, d - DR - AIC	0.032	0.045	0.055	0.077	0.11	0.1	0.15	0.16	**0.091**
32	ARMAX, d - DR -SIC	0.031	0.044	0.055	0.074	0.11	0.1	0.14	0.15	**0.088**
33	ARMAX, d - MRR&DR - AIC	0.041	0.059	0.074	0.1	0.16	0.19	0.24	0.24	**0.14**
34	ARMAX, d - MRR&DR - SIC	0.04	0.059	0.071	0.094	0.15	0.19	0.23	0.25	**0.14**
35	ARMA errors, d - MRR - AIC	0.054	0.077	0.092	0.11	0.19	0.23	0.29	0.31	**0.17**
36	ARMA errors, d - MRR - SIC	0.044	0.064	0.075	0.1	0.17	0.22	0.29	0.31	**0.16**
37	ARMA errors, d - DR - AIC	0.042	0.065	0.074	0.093	0.13	0.12	0.16	0.17	**0.11**
38	ARMA errors, d - DR -SIC	0.034	0.047	0.058	0.078	0.12	0.11	0.16	0.17	**0.096**
39	ARMA errors, d - MRR&DR - AIC	0.045	0.063	0.073	0.095	0.16	0.2	0.24	0.24	**0.14**
40	ARMA errors, d - MRR&DR - SIC	0.045	0.063	0.073	0.095	0.16	0.2	0.24	0.24	**0.14**
41	ARDL, d - MRR - AIC	0.049	0.075	0.091	0.11	0.18	0.22	0.32	0.32	**0.17**
42	ARDL, d - MRR - SIC	0.043	0.063	0.075	0.1	0.17	0.23	0.31	0.33	**0.17**
43	ARDL, d - DR - AIC	0.036	0.054	0.065	0.079	0.1	0.089	0.13	0.13	**0.086**
44	ARDL, d - DR -SIC	0.031	0.044	0.055	0.074	0.11	0.1	0.14	0.15	**0.087**
45	ARDL, d - MRR&DR - AIC	0.069	0.091	0.12	0.13	0.23	0.35	0.74	0.93	**0.33**
46	ARDL, d - MRR&DR - SIC	0.05	0.066	0.08	0.12	0.23	0.38	0.7	0.87	**0.31**
Avg.		**0.04**	**0.06**	**0.07**	**0.10**	**0.16**	**0.19**	**0.27**	**0.30**	**0.15**
	Multivaraite models - spread									
47	ARMA, s - MRR - AIC	0.058	0.085	0.1	0.14	0.23	0.27	0.19	0.13	**0.15**
48	ARMA, s - MRR - SIC	0.049	0.071	0.085	0.12	0.19	0.23	0.17	0.1	**0.13**
49	ARMA, s - DR - AIC	0.043	0.066	0.08	0.14	0.2	0.27	0.47	0.54	**0.23**
50	ARMA, s - DR - SIC	0.038	0.054	0.07	0.12	0.19	0.28	0.48	0.55	**0.22**
51	ARMA, s, d - MRR - AIC	0.056	0.083	0.096	0.12	0.19	0.23	0.28	0.29	**0.17**
52	ARMA, s, d - MRR - SIC	0.044	0.063	0.074	0.1	0.17	0.22	0.28	0.29	**0.16**
53	ARMA, s, d - DR - AIC	0.042	0.064	0.074	0.092	0.12	0.12	0.16	0.17	**0.11**
54	ARMA, s, d - DR - SIC	0.032	0.045	0.056	0.073	0.11	0.095	0.13	0.13	**0.083**
Avg.		**0.05**	**0.07**	**0.08**	**0.11**	**0.18**	**0.21**	**0.27**	**0.28**	**0.16**
	Multivariate models - break									
55	Static regression, b - MRR	0.11	0.11	0.11	0.11	0.12	0.1	0.15	0.16	**0.12**
56	Static regression, b - DR	0.15	0.15	0.15	0.16	0.19	0.2	0.33	0.36	**0.21**
57	Static regression, b - MRR&DR	0.1	0.1	0.11	0.12	0.14	0.13	0.22	0.22	**0.14**
58	ARMA errors, b - MRR	0.054	0.09	0.1	0.13	0.17	0.17	0.32	0.32	**0.17**
59	ARMA errors, b - DR	0.045	0.072	0.086	0.11	0.13	0.11	0.19	0.2	**0.12**
60	ARMA errors, b - MRR&DR	0.052	0.08	0.091	0.1	0.12	0.12	0.22	0.2	**0.12**
61	ARMA, s, b - MRR	0.049	0.079	0.096	0.13	0.16	0.16	0.23	0.25	**0.14**
62	ARMA, s, b - DR	0.055	0.088	0.1	0.13	0.15	0.15	0.22	0.18	**0.13**
Avg.		**0.08**	**0.10**	**0.11**	**0.12**	**0.15**	**0.14**	**0.24**	**0.24**	**0.14**
	Structural models									
100a	Linear	0.066	0.07	0.075	0.084	0.089	0.098	0.085	0.12	**0.086**
100b	Linear - ARMA errors	0.048	0.061	0.066	0.075	0.081	0.088	0.075	0.088	**0.073**
101a	Nonlinear	0.044	0.047	0.046	0.046	0.044	0.035	0.035	0.024	**0.04**
101b	Nonlinear - ARMA errors	0.037	0.046	0.044	0.049	0.051	0.046	0.048	0.039	**0.045**
Avg.		**0.05**	**0.06**	**0.06**	**0.06**	**0.07**	**0.07**	**0.06**	**0.07**	**0.06**
	Summary statistics									
Mean		0.08	0.10	0.14	0.19	0.23	0.35	0.42	0.20	
Std. Deviation		0.05	0.05	0.07	0.08	0.12	0.31	0.54	0.13	
Minimum		0.04	0.04	0.05	0.04	0.04	0.04	0.02	0.04	
1st quartile		0.06	0.07	0.10	0.13	0.12	0.19	0.16	0.12	
3rd quartile		0.09	0.11	0.13	0.22	0.27	0.45	0.51	0.22	
Maximum		0.28	0.28	0.39	0.44	0.61	1.78	3.64	0.81	

Notes: See note below Table 2.2 for explanation of values.

Table 2.14: Ex-post forecasting performance for 1-week Euribor - All models

#	Description	Forecast horizons (# of steps ahead) 1	2	3	6	12	24	48	60	Avg.
		Multivariate models - level								
5	Static regerssion - MRR	0.19	0.2	0.2	0.2	0.2	0.19	0.2	0.11	**0.19**
6	Static regerssion - DR	0.29	0.29	0.3	0.31	0.32	0.34	0.44	0.48	**0.35**
7	Static regerssion - MRR&DR	0.2	0.21	0.21	0.24	0.28	0.33	0.53	0.55	**0.32**
8	ARMAX - MRR - AIC	0.094	0.17	0.23	0.26	0.23	0.21	0.19	0.073	**0.18**
9	ARMAX - MRR - SIC	0.09	0.16	0.21	0.25	0.22	0.21	0.18	0.081	**0.17**
10	ARMAX - DR - AIC	0.089	0.18	0.26	0.4	0.45	0.44	0.57	0.6	**0.37**
11	ARMAX - DR -SIC	0.085	0.17	0.25	0.39	0.41	0.42	0.53	0.57	**0.35**
12	ARMAX - MRR&DR - AIC	0.097	0.18	0.24	0.35	0.36	0.43	0.77	0.8	**0.4**
13	ARMAX - MRR&DR - SIC	0.096	0.17	0.23	0.33	0.35	0.43	0.76	0.79	**0.4**
14	ARMA errors - MRR - AIC	0.057	0.088	0.1	0.15	0.22	0.22	0.21	0.17	**0.15**
15	ARMA errors - MRR - SIC	0.046	0.078	0.1	0.15	0.22	0.24	0.2	0.14	**0.15**
16	ARMA errors - DR - AIC	0.041	0.075	0.1	0.18	0.26	0.33	0.58	0.65	**0.28**
17	ARMA errors - DR -SIC	0.034	0.058	0.08	0.14	0.23	0.34	0.6	0.68	**0.27**
18	ARMA errors - MRR&DR - AIC	0.058	0.09	0.11	0.15	0.23	0.24	0.24	0.21	**0.17**
19	ARMA errors - MRR&DR - SIC	0.046	0.078	0.1	0.15	0.22	0.26	0.24	0.21	**0.16**
20	ARDL - MRR - AIC	0.056	0.094	0.12	0.19	0.26	0.32	0.69	1.65	**0.42**
21	ARDL - MRR - SIC	0.043	0.074	0.097	0.15	0.21	0.23	0.2	0.11	**0.14**
22	ARDL - DR - AIC	0.036	0.063	0.087	0.15	0.24	0.32	0.63	0.76	**0.28**
23	ARDL - DR -SIC	0.032	0.052	0.073	0.13	0.21	0.26	0.51	0.6	**0.23**
24	ARDL - MRR&DR - AIC	0.18	0.22	0.33	0.27	0.58	0.88	1.14	1.56	**0.65**
25	ARDL - MRR&DR - SIC	0.093	0.13	0.19	0.21	0.35	0.52	0.96	1.28	**0.47**
Avg.		**0.09**	**0.13**	**0.17**	**0.23**	**0.29**	**0.34**	**0.49**	**0.57**	**0.29**
		Multivariate models - differences								
26	Static regerssion, d - MRR	0.035	0.06	0.077	0.11	0.2	0.26	0.26	0.28	**0.16**
27	Static regerssion, d - DR	0.027	0.043	0.06	0.091	0.14	0.15	0.19	0.21	**0.11**
28	Static regerssion, d - MRR&DR	0.035	0.061	0.077	0.11	0.19	0.23	0.22	0.24	**0.14**
29	ARMAX, d - MRR - AIC	0.041	0.073	0.095	0.13	0.21	0.25	0.3	0.29	**0.18**
30	ARMAX, d - MRR - SIC	0.036	0.061	0.08	0.11	0.2	0.26	0.29	0.32	**0.17**
31	ARMAX, d - DR - AIC	0.032	0.05	0.065	0.097	0.14	0.13	0.18	0.2	**0.11**
32	ARMAX, d - DR -SIC	0.027	0.045	0.062	0.089	0.14	0.14	0.16	0.18	**0.11**
33	ARMAX, d - MRR&DR - AIC	0.041	0.074	0.095	0.13	0.2	0.23	0.27	0.28	**0.17**
34	ARMAX, d - MRR&DR - SIC	0.035	0.06	0.077	0.1	0.18	0.23	0.23	0.26	**0.15**
35	ARMA errors, d - MRR - AIC	0.056	0.091	0.11	0.13	0.2	0.22	0.25	0.23	**0.16**
36	ARMA errors, d - MRR - SIC	0.037	0.062	0.079	0.11	0.2	0.26	0.26	0.28	**0.16**
37	ARMA errors, d - DR - AIC	0.035	0.058	0.077	0.11	0.14	0.14	0.17	0.19	**0.11**
38	ARMA errors, d - DR -SIC	0.027	0.043	0.06	0.091	0.14	0.15	0.19	0.21	**0.11**
39	ARMA errors, d - MRR&DR - AIC	0.038	0.064	0.08	0.11	0.19	0.23	0.22	0.24	**0.15**
40	ARMA errors, d - MRR&DR - SIC	0.038	0.064	0.08	0.11	0.19	0.23	0.22	0.24	**0.15**
41	ARDL, d - MRR - AIC	0.051	0.09	0.12	0.15	0.24	0.29	0.37	0.36	**0.21**
42	ARDL, d - MRR - SIC	0.037	0.063	0.083	0.12	0.21	0.26	0.3	0.31	**0.17**
43	ARDL, d - DR - AIC	0.03	0.05	0.065	0.096	0.15	0.14	0.17	0.2	**0.11**
44	ARDL, d - DR -SIC	0.027	0.045	0.062	0.089	0.14	0.14	0.16	0.18	**0.11**
45	ARDL, d - MRR&DR - AIC	0.12	0.12	0.17	0.2	0.39	0.67	1.21	1.47	**0.54**
46	ARDL, d - MRR&DR - SIC	0.083	0.085	0.13	0.16	0.29	0.45	0.78	1	**0.37**
Avg.		**0.04**	**0.06**	**0.09**	**0.12**	**0.19**	**0.24**	**0.30**	**0.34**	**0.17**
		Multivaraite models - spread								
47	ARMA, s - MRR - AIC	0.054	0.086	0.11	0.16	0.26	0.32	0.24	0.2	**0.18**
48	ARMA, s - MRR - SIC	0.045	0.076	0.099	0.15	0.22	0.26	0.22	0.19	**0.16**
49	ARMA, s - DR - AIC	0.04	0.071	0.1	0.19	0.27	0.34	0.61	0.7	**0.29**
50	ARMA, s - DR - SIC	0.035	0.059	0.081	0.14	0.23	0.34	0.59	0.66	**0.27**
51	ARMA, s, d - MRR - AIC	0.052	0.083	0.1	0.13	0.21	0.22	0.22	0.23	**0.16**
52	ARMA, s, d - MRR - SIC	0.039	0.063	0.079	0.12	0.21	0.26	0.27	0.28	**0.16**
53	ARMA, s, d - DR - AIC	0.038	0.062	0.087	0.12	0.18	0.16	0.2	0.22	**0.13**
54	ARMA, s, d - DR - SIC	0.027	0.044	0.061	0.09	0.14	0.14	0.18	0.19	**0.11**
Avg.		**0.04**	**0.07**	**0.09**	**0.14**	**0.21**	**0.26**	**0.32**	**0.33**	**0.18**
		Multivariate models - break								
55	Static regression, b - MRR	0.12	0.12	0.12	0.13	0.14	0.11	0.15	0.16	**0.13**
56	Static regression, b - DR	0.17	0.17	0.17	0.19	0.22	0.23	0.36	0.41	**0.24**
57	Static regression, b - MRR&DR	0.12	0.12	0.12	0.14	0.17	0.15	0.27	0.27	**0.17**
58	ARMA errors, b - MRR	0.051	0.09	0.11	0.15	0.21	0.2	0.39	0.37	**0.2**
59	ARMA errors, b - DR	0.045	0.075	0.097	0.13	0.16	0.12	0.2	0.21	**0.13**
60	ARMA errors, b - MRR&DR	0.049	0.085	0.11	0.13	0.18	0.16	0.34	0.3	**0.17**
61	ARMA, s, b - MRR	0.048	0.081	0.11	0.17	0.23	0.27	0.37	0.39	**0.21**
62	ARMA, s, b - DR	0.044	0.078	0.1	0.14	0.18	0.18	0.15	0.089	**0.12**
Avg.		**0.08**	**0.10**	**0.12**	**0.15**	**0.19**	**0.18**	**0.28**	**0.27**	**0.17**
		Structural models								
100a	Linear	0.085	0.09	0.096	0.1	0.12	0.14	0.14	0.2	**0.12**
100b	Linear - ARMA errors	0.047	0.066	0.078	0.095	0.12	0.15	0.15	0.2	**0.11**
101a	Nonlinear	0.059	0.065	0.068	0.07	0.069	0.049	0.058	0.055	**0.062**
101b	Nonlinear - ARMA errors	0.035	0.049	0.056	0.071	0.092	0.089	0.096	0.095	**0.073**
Avg.		**0.06**	**0.07**	**0.07**	**0.08**	**0.10**	**0.11**	**0.11**	**0.14**	**0.09**
		Summary statistics								
Mean		0.09	0.12	0.16	0.22	0.26	0.35	0.40	0.21	
Std. Deviation		0.05	0.06	0.07	0.09	0.14	0.25	0.36	0.12	
Minimum		0.04	0.06	0.07	0.07	0.05	0.06	0.06	0.06	
1st quartile		0.06	0.08	0.11	0.18	0.16	0.19	0.20	0.13	
3rd quartile		0.09	0.12	0.18	0.24	0.32	0.51	0.55	0.27	
Maximum		0.29	0.33	0.40	0.58	0.88	1.21	1.65	0.65	

Notes: See note below Table 2.2 for explanation of values.

Table 2.15: Ex-post forecasting performance for 3-month Euribor - All models

#	Description	1	2	3	6	12	24	48	60	Avg.
		Forecast horizons (# of steps ahead)								
	Multivariate models - level									
5	Static regerssion - MRR	0.17	0.18	0.18	0.19	0.2	0.17	0.16	0.1	**0.17**
6	Static regerssion - DR	0.3	0.3	0.31	0.33	0.36	0.4	0.57	0.61	**0.4**
7	Static regerssion - MRR&DR	0.18	0.18	0.19	0.21	0.23	0.19	0.21	0.21	**0.2**
8	ARMAX - MRR - AIC	0.057	0.12	0.17	0.24	0.24	0.22	0.24	0.23	**0.19**
9	ARMAX - MRR - SIC	0.05	0.11	0.15	0.22	0.21	0.18	0.17	0.12	**0.15**
10	ARMAX - DR - AIC	0.068	0.16	0.25	0.44	0.53	0.54	0.76	0.82	**0.45**
11	ARMAX - DR -SIC	0.058	0.14	0.22	0.38	0.43	0.47	0.65	0.7	**0.38**
12	ARMAX - MRR&DR - AIC	0.06	0.13	0.21	0.38	0.48	0.47	0.62	1.03	**0.42**
13	ARMAX - MRR&DR - SIC	0.057	0.13	0.21	0.34	0.44	0.54	0.75	1.21	**0.46**
14	ARMA errors - MRR - AIC	0.059	0.1	0.13	0.19	0.27	0.27	0.33	0.32	**0.21**
15	ARMA errors - MRR - SIC	0.05	0.087	0.12	0.18	0.28	0.35	0.48	0.47	**0.25**
16	ARMA errors - DR - AIC	0.051	0.09	0.13	0.22	0.44	0.76	0.94	0.97	**0.45**
17	ARMA errors - DR -SIC	0.043	0.074	0.1	0.18	0.31	0.5	0.89	0.98	**0.38**
18	ARMA errors - MRR&DR - AIC	0.057	0.1	0.14	0.24	0.44	0.67	0.63	0.64	**0.36**
19	ARMA errors - MRR&DR - SIC	0.045	0.08	0.11	0.18	0.29	0.44	0.7	0.74	**0.32**
20	ARDL - MRR - AIC	0.041	0.078	0.1	0.13	0.21	0.22	0.22	0.22	**0.15**
21	ARDL - MRR - SIC	0.043	0.08	0.11	0.16	0.22	0.24	0.26	0.25	**0.17**
22	ARDL - DR - AIC	0.05	0.097	0.14	0.22	0.41	0.63	0.74	0.9	**0.4**
23	ARDL - DR -SIC	0.048	0.09	0.13	0.22	0.37	0.57	0.83	0.9	**0.39**
24	ARDL - MRR&DR - AIC	0.26	0.44	0.66	0.58	1.52	4.30	28	71	**13**
25	ARDL - MRR&DR - SIC	0.057	0.12	0.18	0.31	0.56	0.95	1.92	3.09	**0.9**
Avg.		**0.09**	**0.14**	**0.19**	**0.26**	**0.40**	**0.62**	**1.91**	**4.07**	**0.94**
	Multivariate models - differences									
26	Static regerssion, d - MRR	0.037	0.066	0.089	0.14	0.25	0.28	0.3	0.28	**0.18**
27	Static regerssion, d - DR	0.032	0.053	0.073	0.12	0.2	0.24	0.41	0.49	**0.2**
28	Static regerssion, d - MRR&DR	0.034	0.06	0.08	0.13	0.22	0.24	0.34	0.37	**0.18**
29	ARMAX, d - MRR - AIC	0.036	0.07	0.098	0.17	0.27	0.31	0.31	0.31	**0.2**
30	ARMAX, d - MRR - SIC	0.031	0.059	0.082	0.13	0.23	0.29	0.25	0.26	**0.17**
31	ARMAX, d - DR - AIC	0.036	0.065	0.088	0.14	0.2	0.23	0.37	0.43	**0.19**
32	ARMAX, d - DR -SIC	0.031	0.052	0.071	0.11	0.19	0.21	0.36	0.43	**0.18**
33	ARMAX, d - MRR&DR - AIC	0.037	0.07	0.098	0.17	0.25	0.29	0.3	0.31	**0.19**
34	ARMAX, d - MRR&DR - SIC	0.03	0.057	0.08	0.13	0.22	0.26	0.26	0.27	**0.16**
35	ARMA errors, d - MRR - AIC	0.053	0.091	0.12	0.15	0.26	0.28	0.35	0.35	**0.21**
36	ARMA errors, d - MRR - SIC	0.037	0.066	0.089	0.14	0.25	0.28	0.3	0.28	**0.18**
37	ARMA errors, d - DR - AIC	0.043	0.069	0.09	0.14	0.21	0.25	0.43	0.49	**0.22**
38	ARMA errors, d - DR -SIC	0.032	0.053	0.073	0.12	0.2	0.24	0.41	0.49	**0.2**
39	ARMA errors, d - MRR&DR - AIC	0.035	0.062	0.081	0.13	0.22	0.24	0.34	0.37	**0.18**
40	ARMA errors, d - MRR&DR - SIC	0.035	0.062	0.081	0.13	0.22	0.24	0.34	0.37	**0.18**
41	ARDL, d - MRR - AIC	0.044	0.091	0.13	0.19	0.31	0.38	0.43	0.45	**0.25**
42	ARDL, d - MRR - SIC	0.037	0.066	0.087	0.13	0.22	0.28	0.25	0.26	**0.17**
43	ARDL, d - DR - AIC	0.039	0.072	0.095	0.15	0.2	0.23	0.37	0.43	**0.2**
44	ARDL, d - DR -SIC	0.035	0.063	0.087	0.14	0.2	0.23	0.37	0.43	**0.19**
45	ARDL, d - MRR&DR - AIC	0.17	0.17	0.24	0.27	0.52	0.95	1.72	2.28	**0.79**
46	ARDL, d - MRR&DR - SIC	0.061	0.093	0.13	0.18	0.3	0.36	0.38	0.38	**0.24**
Avg.		**0.04**	**0.07**	**0.10**	**0.15**	**0.24**	**0.30**	**0.41**	**0.46**	**0.22**
	Multivaraite models - spread									
47	ARMA, s - MRR - AIC	0.055	0.093	0.12	0.18	0.28	0.34	0.38	0.36	**0.23**
48	ARMA, s - MRR - SIC	0.051	0.088	0.11	0.17	0.25	0.28	0.35	0.33	**0.2**
49	ARMA, s - DR - AIC	0.052	0.092	0.13	0.24	0.44	0.73	0.9	0.9	**0.44**
50	ARMA, s - DR - SIC	0.047	0.081	0.11	0.19	0.32	0.5	0.82	0.88	**0.37**
51	ARMA, s, d - MRR - AIC	0.051	0.084	0.11	0.15	0.25	0.26	0.32	0.32	**0.19**
52	ARMA, s, d - MRR - SIC	0.039	0.068	0.089	0.14	0.24	0.27	0.27	0.24	**0.17**
53	ARMA, s, d - DR - AIC	0.045	0.076	0.099	0.15	0.23	0.28	0.41	0.46	**0.22**
54	ARMA, s, d - DR - SIC	0.033	0.054	0.072	0.11	0.19	0.21	0.35	0.41	**0.18**
Avg.		**0.05**	**0.08**	**0.10**	**0.17**	**0.28**	**0.36**	**0.48**	**0.49**	**0.25**
	Multivariate models - break									
55	Static regression, b - MRR	0.16	0.17	0.17	0.19	0.21	0.21	0.34	0.37	**0.23**
56	Static regression, b - DR	0.18	0.18	0.18	0.18	0.19	0.18	0.14	0.14	**0.17**
57	Static regression, b - MRR&DR	0.15	0.16	0.16	0.18	0.2	0.16	0.19	0.2	**0.18**
58	ARMA errors, b - MRR	0.048	0.088	0.12	0.19	0.3	0.41	0.58	0.56	**0.29**
59	ARMA errors, b - DR	0.048	0.082	0.11	0.17	0.25	0.31	0.48	0.49	**0.24**
60	ARMA errors, b - MRR&DR	0.04	0.07	0.093	0.15	0.23	0.33	0.62	0.73	**0.28**
61	ARMA, s, b - MRR	0.035	0.064	0.088	0.15	0.25	0.4	0.76	0.85	**0.33**
62	ARMA, s, b - DR	0.038	0.073	0.098	0.16	0.25	0.32	0.45	0.45	**0.23**
Avg.		**0.09**	**0.11**	**0.13**	**0.17**	**0.24**	**0.29**	**0.45**	**0.47**	**0.24**
	Structural models									
100a	Linear	0.12	0.12	0.13	0.13	0.12	0.12	0.15	0.12	**0.13**
100b	Linear - ARMA errors	0.044	0.068	0.08	0.11	0.13	0.12	0.19	0.12	**0.11**
101a	Nonlinear	0.13	0.14	0.14	0.16	0.17	0.15	0.26	0.23	**0.17**
101b	Nonlinear - ARMA errors	0.038	0.06	0.072	0.1	0.15	0.18	0.32	0.34	**0.16**
Avg.		**0.08**	**0.10**	**0.11**	**0.13**	**0.14**	**0.14**	**0.23**	**0.20**	**0.14**
	Summary statistics									
Mean		0.10	0.13	0.19	0.29	0.41	0.92	1.67	0.47	
Std. Deviation		0.06	0.08	0.09	0.19	0.53	3.51	8.96	1.62	
Minimum		0.05	0.07	0.10	0.12	0.12	0.14	0.10	0.11	
1st quartile		0.07	0.09	0.14	0.21	0.23	0.30	0.28	0.18	
3rd quartile		0.12	0.14	0.21	0.31	0.41	0.62	0.70	0.33	
Maximum		0.44	0.66	0.58	1.52	4.30	28.00	71.00	13.00	

Notes: See note below Table 2.2 for explanation of values.

2.B.2 Full sample results

This appendix contains measures of forecasting performance for the sample covering months between January 2005 and August 2019, rather than just the period of excess reserves. I report results only for ex-post forecasts and the Eonia rate; additional results are available upon request.

Table 2.16: Ex-post forecasting performance for Eonia - Full sample

#	Description	Forecast horizons (# of steps ahead) 1	2	3	6	12	24	48	60	Avg.
	Multivariate models - level									
5	Static regerssion - MRR	0.16	0.17	0.17	0.19	0.21	0.26	0.36	0.32	**0.23**
6	Static regerssion - DR	0.23	0.23	0.24	0.25	0.29	0.37	0.55	0.61	**0.35**
7	Static regerssion - MRR&DR	0.15	0.16	0.17	0.2	0.25	0.36	0.61	0.64	**0.32**
9	ARMAX - MRR - SIC	0.079	0.13	0.17	0.22	0.22	0.27	0.34	0.3	**0.21**
11	ARMAX - DR -SIC	0.075	0.13	0.19	0.32	0.38	0.45	0.64	0.69	**0.36**
13	ARMAX - MRR&DR - SIC	0.082	0.13	0.18	0.27	0.31	0.43	0.72	0.76	**0.36**
15	ARMA errors - MRR - SIC	0.047	0.065	0.083	0.12	0.19	0.27	0.34	0.31	**0.18**
16	ARMA errors - DR - AIC	0.049	0.073	0.093	0.14	0.21	0.32	0.59	0.65	**0.27**
17	ARMA errors - DR -SIC	0.043	0.06	0.08	0.13	0.22	0.36	0.67	0.75	**0.29**
19	ARMA errors - MRR&DR - SIC	0.046	0.066	0.082	0.12	0.18	0.26	0.33	0.32	**0.18**
21	ARDL - MRR - SIC	0.049	0.069	0.086	0.13	0.2	0.29	0.36	0.32	**0.19**
23	ARDL - DR -SIC	0.042	0.06	0.075	0.12	0.2	0.31	0.59	0.68	**0.26**
25	ARDL - MRR&DR - SIC	0.057	0.08	0.1	0.15	0.26	0.43	0.81	0.95	**0.36**
Avg.		**0.08**	**0.10**	**0.13**	**0.18**	**0.25**	**0.36**	**0.60**	**0.70**	**0.30**
	Multivariate models - differences									
26	Static regerssion, d - MRR	0.039	0.059	0.076	0.12	0.22	0.36	0.51	0.55	**0.24**
27	Static regerssion, d - DR	0.036	0.053	0.069	0.11	0.2	0.31	0.52	0.59	**0.24**
28	Static regerssion, d - MRR&DR	0.039	0.06	0.079	0.12	0.21	0.34	0.48	0.51	**0.23**
30	ARMAX, d - MRR - SIC	0.04	0.059	0.075	0.11	0.19	0.3	0.46	0.49	**0.22**
32	ARMAX, d - DR -SIC	0.036	0.051	0.066	0.097	0.16	0.25	0.44	0.49	**0.2**
34	ARMAX, d - MRR&DR - SIC	0.039	0.058	0.076	0.11	0.18	0.28	0.42	0.45	**0.2**
38	ARMA errors, d - DR -SIC	0.039	0.054	0.067	0.094	0.16	0.23	0.41	0.47	**0.19**
42	ARDL, d - MRR - SIC	0.041	0.059	0.074	0.11	0.18	0.28	0.44	0.46	**0.2**
44	ARDL, d - DR -SIC	0.035	0.049	0.064	0.091	0.15	0.22	0.4	0.45	**0.18**
46	ARDL, d - MRR&DR - SIC	0.046	0.064	0.083	0.12	0.22	0.39	0.71	0.82	**0.31**
Avg.		**0.04**	**0.06**	**0.08**	**0.11**	**0.18**	**0.28**	**0.46**	**0.50**	**0.21**
	Multivaraite models - spread									
47	ARMA, s - MRR - AIC	0.054	0.075	0.093	0.13	0.21	0.29	0.34	0.32	**0.19**
48	ARMA, s - MRR - SIC	0.046	0.064	0.081	0.12	0.19	0.27	0.33	0.3	**0.17**
50	ARMA, s - DR - SIC	0.042	0.058	0.077	0.13	0.21	0.33	0.61	0.68	**0.27**
52	ARMA, s, d - MRR - SIC	0.041	0.058	0.07	0.1	0.17	0.26	0.39	0.41	**0.19**
53	ARMA, s, d - DR - AIC	0.045	0.065	0.079	0.1	0.15	0.22	0.4	0.45	**0.19**
54	ARMA, s, d - DR - SIC	0.037	0.049	0.064	0.086	0.14	0.2	0.38	0.43	**0.17**
Avg.		**0.05**	**0.06**	**0.08**	**0.11**	**0.18**	**0.27**	**0.43**	**0.46**	**0.21**
	Multivariate models - break									
55	Static regression, b - MRR	0.097	0.1	0.11	0.12	0.16	0.22	0.4	0.44	**0.21**
56	Static regression, b - DR	0.12	0.13	0.13	0.15	0.19	0.25	0.42	0.45	**0.23**
57	Static regression, b - MRR&DR	0.084	0.091	0.098	0.12	0.15	0.19	0.35	0.36	**0.18**
58	ARMA errors, b - MRR	0.051	0.079	0.095	0.12	0.17	0.23	0.42	0.42	**0.2**
59	ARMA errors, b - DR	0.048	0.073	0.089	0.12	0.16	0.22	0.42	0.47	**0.2**
61	ARMA, s, b - MRR	0.049	0.077	0.094	0.13	0.18	0.25	0.45	0.5	**0.22**
62	ARMA, s, b - DR	0.05	0.077	0.092	0.12	0.16	0.21	0.36	0.35	**0.18**
Avg.		**0.07**	**0.09**	**0.10**	**0.12**	**0.16**	**0.22**	**0.40**	**0.42**	**0.20**
	Structural models									
100a	Linear	0.059	0.064	0.073	0.088	0.12	0.17	0.27	0.31	**0.14**
100b	Linear - ARMA errors	0.046	0.059	0.067	0.082	0.11	0.17	0.27	0.29	**0.14**
101a	Nonlinear	0.044	0.048	0.052	0.058	0.084	0.13	0.24	0.26	**0.11**
101b	Nonlinear - ARMA errors	0.037	0.047	0.051	0.062	0.089	0.14	0.25	0.27	**0.12**
Avg.		**0.05**	**0.05**	**0.06**	**0.07**	**0.10**	**0.15**	**0.26**	**0.28**	**0.13**
	Summary statistics									
Mean			0.08	0.10	0.13	0.20	0.29	0.48	0.54	0.23
Std. Deviation			0.03	0.04	0.05	0.06	0.08	0.20	0.33	0.09
Minimum			0.05	0.05	0.06	0.08	0.13	0.24	0.26	0.11
1st quartile			0.06	0.07	0.10	0.16	0.23	0.36	0.35	0.19
3rd quartile			0.08	0.10	0.13	0.22	0.33	0.59	0.64	0.27
Maximum			0.23	0.24	0.32	0.38	0.52	1.43	2.51	0.62

Notes: See note below Table 2.2 for explanation of values.

2.B.3 Diebold-Mariano test results

This appendix provides more details on the Diebold-Mariano tests and their results.

The test of comparison of forecast performance is the standard Diebold-Mariano test Diebold and Mariano (2002) with small sample correction suggested by Harvey, Leybourne, and Newbold (1997). The null hypothesis is specified as one-sided, i.e. the null hypothesis is that the structural model has worse forecast accuracy than the benchmark model. The cut-off for significance used in the text is p-value less than 0.05. Tables below show the actual p-values. Color-coding is provided to facilitate interpretation by the reader: p-values of less than 0.01 are dark green, less than 0.05 are light green and less than 0.1 are yellow. Values above 0.1 are either orange, if the p-value of reverse hypothesis is not less than 0.05, or red, if the p-value is less than 0.05 (indicating that the benchmark model has forecasting performance that is significantly better than the structural model). Note that at higher horizons the test statistic often cannot be computed due to negative long-run variance estimates.

Table 2.17: Statistical test of predictive accuracy equivalency - Ex-post forecasts for Eonia rate

#	Description	Forecast horizons (# of steps ahead) 1	2	3	6	12	24	48	60
	Multivariate models - level								
5	Static regerssion - MRR	0.000	0.000	0.000	0.003	0.024	0.066	0.035	NA
6	Static regerssion - DR	0.000	0.000	0.000	0.000	0.001	0.000	0.023	0.174
7	Static regerssion - MRR&DR	0.000	0.000	0.000	0.000	0.003	0.064	0.051	0.112
9	ARMAX - MRR - SIC	0.000	0.000	0.000	0.000	0.004	0.029	0.079	0.000
11	ARMAX - DR -SIC	0.000	0.000	0.000	0.000	0.000	0.002	0.023	0.023
13	ARMAX - MRR&DR - SIC	0.000	0.000	0.000	0.008	0.053	0.096	0.001	NA
14	ARMA errors - MRR - AIC	0.002	0.000	0.000	0.000	0.000	0.001	NA	0.239
15	ARMA errors - MRR - SIC	0.029	0.000	0.000	0.000	0.000	0.000	NA	0.184
16	ARMA errors - DR - AIC	0.124	0.001	0.001	0.005	0.001	0.000	0.000	0.000
17	ARMA errors - DR -SIC	0.463	0.068	0.008	0.003	0.001	0.000	0.000	0.001
18	ARMA errors - MRR&DR - AIC	0.003	0.000	0.000	0.000	0.000	0.000	NA	0.165
19	ARMA errors - MRR&DR - SIC	0.041	0.000	0.000	0.000	0.000	0.000	NA	0.269
21	ARDL - MRR - SIC	0.024	0.001	0.000	0.000	0.000	0.000	NA	NA
23	ARDL - DR -SIC	0.831	0.264	0.057	0.010	0.008	0.000	0.005	0.013
25	ARDL - MRR&DR - SIC	0.001	0.001	0.004	0.047	0.082	0.116	0.006	0.000
	Multivariate models - differences								
26	Static regerssion, d - MRR	0.434	0.038	0.005	0.011	0.003	0.018	NA	NA
27	Static regerssion, d - DR	0.963	0.596	0.137	0.035	0.023	0.012	0.000	NA
28	Static regerssion, d - MRR&DR	0.403	0.017	0.003	0.007	0.002	0.014	NA	0.000
30	ARMAX, d - MRR - SIC	0.286	0.035	0.003	0.011	0.003	0.020	NA	0.085
32	ARMAX, d - DR -SIC	0.934	0.571	0.077	0.047	0.032	0.027	0.000	0.000
34	ARMAX, d - MRR&DR - SIC	0.311	0.022	0.004	0.008	0.002	0.016	NA	0.142
38	ARMA errors, d - DR -SIC	0.802	0.409	0.057	0.032	0.026	0.020	0.000	0.000
42	ARDL, d - MRR - SIC	0.176	0.035	0.004	0.010	0.005	0.031	NA	0.109
44	ARDL, d - DR -SIC	0.932	0.569	0.075	0.048	0.033	0.028	0.000	0.000
46	ARDL, d - MRR&DR - SIC	0.019	0.016	0.007	0.039	0.045	0.076	0.000	0.038
	Multivaraite models - spread								
47	ARMA, s - MRR - AIC	0.002	0.000	0.000	0.000	0.000	0.000	NA	0.335
48	ARMA, s - MRR - SIC	0.038	0.001	0.000	0.000	0.000	0.000	NA	0.359
50	ARMA, s - DR - SIC	0.484	0.153	0.022	0.002	0.001	0.000	0.000	0.000
52	ARMA, s, d - MRR - SIC	0.139	0.009	0.003	0.014	0.005	0.031	NA	0.034
53	ARMA, s, d - DR - AIC	0.209	0.004	0.001	0.023	0.037	0.062	0.000	NA
54	ARMA, s, d - DR - SIC	0.861	0.545	0.075	0.053	0.036	0.042	0.000	0.001
	Multivariate models - break								
55	Static regression, b - MRR	0.000	0.000	0.000	0.003	0.024	0.066	0.035	NA
56	Static regression, b - DR	0.000	0.000	0.000	0.000	0.001	0.000	0.023	0.174
57	Static regression, b - MRR&DR	0.000	0.000	0.000	0.000	0.003	0.064	0.051	0.112
58	ARMA errors, b - MRR	0.003	0.000	0.000	0.014	0.045	0.138	0.051	NA
59	ARMA errors, b - DR	0.038	0.000	0.000	0.000	0.001	0.000	0.081	0.206
61	ARMA, s, b - MRR	0.007	0.000	0.000	0.000	0.000	0.000	0.000	0.063
62	ARMA, s, b - DR	0.002	0.000	0.000	0.005	0.034	0.113	0.094	NA

Table 2.18: Statistical test of predictive accuracy equivalency - Ex-post forecasts for 1-w Euribor

#	Description	Forecast horizons (# of steps ahead) 1	2	3	6	12	24	48	60
	Multivariate models - level								
5	Static regerssion - MRR	0.000	0.000	0.000	0.008	0.057	0.092	0.048	NA
6	Static regerssion - DR	0.000	0.000	0.000	0.000	0.001	0.076	0.008	0.250
7	Static regerssion - MRR&DR	0.000	0.000	0.000	0.001	0.035	0.124	0.098	NA
9	ARMAX - MRR - SIC	0.000	0.000	0.000	0.000	0.006	0.038	0.228	0.629
11	ARMAX - DR -SIC	0.000	0.000	0.000	0.000	0.000	0.000	0.009	0.018
13	ARMAX - MRR&DR - SIC	0.000	0.000	0.000	0.011	0.083	0.134	0.003	NA
14	ARMA errors - MRR - AIC	0.000	0.000	0.000	0.007	0.005	NA	NA	0.410
15	ARMA errors - MRR - SIC	0.002	0.000	0.000	0.000	0.000	NA	0.000	0.384
16	ARMA errors - DR - AIC	0.102	0.006	0.002	0.007	0.001	0.000	0.000	0.000
17	ARMA errors - DR -SIC	0.545	0.187	0.013	0.000	0.000	0.000	0.000	0.000
18	ARMA errors - MRR&DR - AIC	0.000	0.000	0.000	0.003	0.000	0.000	NA	0.296
19	ARMA errors - MRR&DR - SIC	0.001	0.000	0.000	0.000	0.000	0.000	0.001	0.264
21	ARDL - MRR - SIC	0.019	0.002	0.001	0.001	0.002	0.002	0.092	0.389
23	ARDL - DR -SIC	0.709	0.391	0.070	0.006	0.008	0.000	0.001	0.002
25	ARDL - MRR&DR - SIC	0.003	0.008	0.010	0.025	0.060	0.068	0.008	NA
	Multivariate models - differences								
26	Static regerssion, d - MRR	0.449	0.091	0.009	0.066	0.010	0.015	NA	NA
27	Static regerssion, d - DR	0.927	0.738	0.311	0.124	0.058	0.068	NA	NA
28	Static regerssion, d - MRR&DR	0.425	0.016	0.001	0.058	0.008	0.016	NA	NA
30	ARMAX, d - MRR - SIC	0.387	0.086	0.004	0.062	0.017	0.023	NA	0.001
32	ARMAX, d - DR -SIC	0.909	0.674	0.164	0.139	0.085	0.079	NA	NA
34	ARMAX, d - MRR&DR - SIC	0.432	0.032	0.001	0.090	0.017	0.023	NA	0.031
38	ARMA errors, d - DR -SIC	0.927	0.738	0.311	0.124	0.058	0.068	NA	NA
42	ARDL, d - MRR - SIC	0.293	0.035	0.001	0.033	0.011	0.022	NA	0.000
44	ARDL, d - DR -SIC	0.910	0.679	0.175	0.139	0.083	0.079	NA	NA
46	ARDL, d - MRR&DR - SIC	0.001	0.001	0.005	0.021	0.049	0.083	0.000	NA
	Multivaraite models - spread								
47	ARMA, s - MRR - AIC	0.000	0.000	0.001	0.010	0.001	0.000	NA	0.321
48	ARMA, s - MRR - SIC	0.004	0.000	0.000	0.000	0.000	0.000	NA	0.377
50	ARMA, s - DR - SIC	0.500	0.162	0.011	0.000	0.000	0.000	0.000	0.000
52	ARMA, s, d - MRR - SIC	0.120	0.025	0.004	0.044	0.009	0.017	NA	NA
53	ARMA, s, d - DR - AIC	0.252	0.060	0.000	0.000	0.000	0.019	0.000	0.017
54	ARMA, s, d - DR - SIC	0.911	0.713	0.276	0.130	0.067	0.083	NA	NA
	Multivariate models - break								
55	Static regression, b - MRR	0.000	0.000	0.000	0.008	0.057	0.092	0.048	NA
56	Static regression, b - DR	0.000	0.000	0.000	0.000	0.001	0.076	0.008	0.250
57	Static regression, b - MRR&DR	0.000	0.000	0.000	0.001	0.035	0.124	0.098	NA
58	ARMA errors, b - MRR	0.000	0.000	0.000	0.012	0.057	0.173	0.041	NA
59	ARMA errors, b - DR	0.027	0.004	0.001	0.001	0.022	0.064	0.157	0.273
61	ARMA, s, b - MRR	0.007	0.002	0.000	0.000	0.000	0.000	0.000	0.003
62	ARMA, s, b - DR	0.009	0.000	0.000	0.000	0.001	0.000	0.157	0.559

Table 2.19: Statistical test of predictive accuracy equivalency - Ex-post forecasts for 3-m Euribor

#	Description	Forecast horizons (# of steps ahead) 1	2	3	6	12	24	48	60
	Multivariate models - level								
5	Static regerssion - MRR	0.000	0.000	0.000	0.002	0.236	0.507	1.000	NA
6	Static regerssion - DR	0.000	0.000	0.000	0.000	0.019	0.145	0.388	0.450
7	Static regerssion - MRR&DR	0.000	0.000	0.000	0.001	0.091	NA	0.933	0.712
9	ARMAX - MRR - SIC	0.023	0.001	0.000	0.000	0.049	0.478	NA	NA
11	ARMAX - DR -SIC	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.051
13	ARMAX - MRR&DR - SIC	0.002	0.000	0.000	0.001	0.022	0.017	0.183	0.327
14	ARMA errors - MRR - AIC	0.000	0.000	0.000	0.000	0.001	0.068	0.318	NA
15	ARMA errors - MRR - SIC	0.000	0.000	0.000	0.000	0.000	0.028	0.149	NA
16	ARMA errors - DR - AIC	0.004	0.003	0.000	0.000	0.000	0.000	0.000	NA
17	ARMA errors - DR -SIC	0.075	0.034	0.000	0.000	0.000	0.000	NA	0.006
18	ARMA errors - MRR&DR - AIC	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.059
19	ARMA errors - MRR&DR - SIC	0.002	0.000	0.000	0.000	0.000	0.000	0.000	NA
21	ARDL - MRR - SIC	0.127	0.012	0.001	0.001	0.024	0.096	0.756	1.000
23	ARDL - DR -SIC	0.003	0.000	0.000	0.000	0.000	0.000	0.000	0.004
25	ARDL - MRR&DR - SIC	0.000	0.000	0.000	0.002	0.033	0.055	0.233	0.314
	Multivariate models - differences								
26	Static regerssion, d - MRR	0.751	0.101	0.023	0.034	0.001	0.000	0.658	NA
27	Static regerssion, d - DR	0.995	0.899	0.466	0.127	0.045	0.067	0.000	0.299
28	Static regerssion, d - MRR&DR	0.996	0.597	0.113	0.054	0.004	0.016	0.401	0.316
30	ARMAX, d - MRR - SIC	0.998	0.582	0.162	0.163	0.028	0.001	0.803	NA
32	ARMAX, d - DR -SIC	0.997	0.930	0.612	0.334	0.120	0.166	NA	0.402
34	ARMAX, d - MRR&DR - SIC	1.000	0.787	0.172	0.121	0.008	0.006	0.931	NA
38	ARMA errors, d - DR -SIC	0.995	0.899	0.466	0.127	0.045	0.067	0.000	0.299
42	ARDL, d - MRR - SIC	0.774	0.168	0.072	0.155	0.049	0.006	0.844	NA
44	ARDL, d - DR -SIC	0.866	0.362	0.145	0.122	0.069	0.175	0.341	0.391
46	ARDL, d - MRR&DR - SIC	0.048	0.034	0.043	0.029	0.013	0.002	NA	NA
	Multivaraite models - spread								
47	ARMA, s - MRR - AIC	0.000	0.001	0.002	0.000	0.003	0.037	0.316	0.395
48	ARMA, s - MRR - SIC	0.000	0.000	0.000	0.000	0.003	0.135	0.415	0.544
50	ARMA, s - DR - SIC	0.005	0.005	0.000	0.000	0.000	0.000	0.000	0.000
52	ARMA, s, d - MRR - SIC	0.263	0.065	0.029	0.081	0.009	0.003	0.876	NA
53	ARMA, s, d - DR - AIC	0.049	0.030	0.011	0.027	0.003	0.006	NA	0.375
54	ARMA, s, d - DR - SIC	0.972	0.868	0.552	0.306	0.119	0.199	NA	0.417
	Multivariate models - break								
55	Static regression, b - MRR	0.000	0.000	0.000	0.002	0.236	0.507	1.000	NA
56	Static regression, b - DR	0.000	0.000	0.000	0.000	0.019	0.145	0.388	0.450
57	Static regression, b - MRR&DR	0.000	0.000	0.000	0.001	0.091	NA	0.933	0.712
58	ARMA errors, b - MRR	0.000	0.000	0.000	0.000	0.000	0.007	0.111	0.054
59	ARMA errors, b - DR	0.015	0.008	0.000	0.000	0.002	0.023	NA	NA
61	ARMA, s, b - MRR	0.906	0.314	0.019	0.000	0.000	0.000	0.000	0.001
62	ARMA, s, b - DR	0.602	0.002	0.000	0.000	0.000	0.001	0.205	0.254

Table 2.20: Statistical test of predictive accuracy equivalency - Ex-ante forecasts for Eonia rate

#	Description	Forecast horizons (# of steps ahead) 1	2	3	6	12	24	48
	Univariate models							
2	ARMA - SIC	0.000	0.000	0.000	0.000	0.000	0.000	NA
4	ARMA, d - SIC	0.947	0.534	0.571	0.300	NA	NA	NA
	Multivariate models - level							
5	Static regerssion - MRR	0.000	0.000	0.002	0.066	0.500	0.582	1.000
6	Static regerssion - DR	0.000	0.000	0.000	0.000	0.000	0.000	0.000
7	Static regerssion - MRR&DR	0.000	0.000	0.000	0.000	0.000	0.000	NA
9	ARMAX - MRR - SIC	0.095	0.082	0.168	0.323	0.821	0.776	1.000
11	ARMAX - DR -SIC	0.000	0.000	0.000	0.000	0.000	0.000	NA
13	ARMAX - MRR&DR - SIC	0.005	0.013	0.032	0.080	0.514	0.785	NA
14	ARMA errors - MRR - AIC	0.044	0.003	0.009	0.001	0.000	0.001	NA
15	ARMA errors - MRR - SIC	0.049	0.000	0.000	0.000	0.000	0.000	NA
16	ARMA errors - DR - AIC	0.017	0.000	0.000	0.000	0.000	0.001	NA
17	ARMA errors - DR -SIC	0.034	0.000	0.000	0.000	0.000	0.000	0.000
18	ARMA errors - MRR&DR - AIC	0.000	0.001	0.007	0.007	0.000	NA	NA
19	ARMA errors - MRR&DR - SIC	0.006	0.000	0.000	0.000	0.000	0.000	NA
21	ARDL - MRR - SIC	0.670	0.004	0.000	0.000	0.000	0.000	0.000
23	ARDL - DR -SIC	0.938	0.015	0.001	0.000	0.001	0.022	NA
25	ARDL - MRR&DR - SIC	0.998	0.934	0.890	0.843	0.814	0.797	NA
	Multivariate models - differences							
26	Static regerssion, d - MRR	0.000	0.000	0.000	0.000	0.017	0.109	NA
27	Static regerssion, d - DR	0.000	0.000	0.000	0.000	0.017	0.109	NA
28	Static regerssion, d - MRR&DR	0.000	0.000	0.000	0.000	0.017	0.109	NA
30	ARMAX, d - MRR - SIC	0.999	0.661	0.258	0.057	0.090	0.232	0.000
32	ARMAX, d - DR -SIC	0.999	0.661	0.258	0.057	0.090	0.232	0.000
34	ARMAX, d - MRR&DR - SIC	0.999	0.661	0.258	0.057	0.090	0.232	0.000
38	ARMA errors, d - DR -SIC	0.998	0.597	0.214	0.055	0.091	0.232	NA
42	ARDL, d - MRR - SIC	0.994	0.532	0.173	0.047	0.093	0.229	0.000
44	ARDL, d - DR -SIC	0.999	0.661	0.258	0.057	0.090	0.232	0.000
46	ARDL, d - MRR&DR - SIC	0.990	0.513	0.163	0.045	0.092	0.228	NA
	Multivariate models - spread							
47	ARMA, s - MRR - AIC	0.118	0.002	0.005	0.002	0.002	0.011	NA
48	ARMA, s - MRR - SIC	0.117	0.000	0.000	0.000	0.000	0.000	0.000
50	ARMA, s - DR - SIC	0.005	0.000	0.000	0.000	0.000	0.000	NA
52	ARMA, s, d - MRR - SIC	0.962	0.269	0.097	0.038	0.091	0.222	NA
53	ARMA, s, d - DR - AIC	0.422	0.190	0.176	0.174	0.051	0.271	NA
54	ARMA, s, d - DR - SIC	0.998	0.595	0.214	0.056	0.091	0.232	NA
	Multivariate models - break							
55	Static regression, b - MRR	0.000	0.000	0.002	0.066	0.500	0.582	1.000
56	Static regression, b - DR	0.000	0.000	0.000	0.000	0.000	0.000	0.000
57	Static regression, b - MRR&DR	0.000	0.000	0.000	0.000	0.000	0.000	NA
58	ARMA errors, b - MRR	0.991	0.761	0.712	0.675	0.779	0.749	NA
59	ARMA errors, b - DR	0.000	0.000	0.000	0.000	0.000	0.000	NA
61	ARMA, s, b - MRR	0.000	0.000	0.000	0.000	0.000	0.000	NA
62	ARMA, s, b - DR	0.985	0.727	0.700	0.718	0.813	0.778	NA
	Multiequation models							
63	Levels - MRR - AIC	0.410	0.008	0.000	0.000	0.000	0.000	NA
64	Levels - MRR - SIC	0.982	0.621	0.490	0.300	0.177	NA	NA
65	Levels - DR - AIC	0.000	0.000	0.000	0.000	0.000	0.018	NA
66	Levels - DR - SIC	0.000	0.000	0.000	0.001	0.008	0.040	NA
67	Differences - MRR - AIC	0.764	0.175	0.166	0.175	0.301	0.235	0.000
68	Differences - MRR - SIC	0.751	0.116	0.131	0.073	0.233	0.241	0.000
69	Differences - DR - AIC	0.843	0.279	0.121	0.085	0.233	0.302	NA
70	Differences - DR - SIC	0.962	0.238	0.112	0.115	0.247	0.334	0.000
71	VECM - MRR - AIC	0.410	0.008	0.000	0.000	0.000	0.000	NA
72	VECM - MRR - SIC	0.982	0.621	0.490	0.300	0.177	NA	NA
73	VECM - DR - AIC	0.000	0.000	0.000	0.000	0.000	0.018	NA
74	VECM - DR - SIC	0.000	0.000	0.000	0.001	0.008	0.040	NA

Table 2.21: Statistical test of predictive accuracy equivalency - Ex-ante forecasts for 3-m Euribor

#	Description	Forecast horizons (# of steps ahead) 1	2	3	6	12	24	48
	Univariate models							
2	ARMA - SIC	1.000	0.737	0.001	0.000	0.000	0.000	NA
4	ARMA, d - SIC	1.000	1.000	1.000	1.000	0.970	0.581	NA
	Multivariate models - level							
5	Static regerssion - MRR	0.000	0.001	0.010	0.122	0.303	0.258	NA
6	Static regerssion - DR	0.000	0.000	0.000	0.000	0.000	0.000	NA
7	Static regerssion - MRR&DR	0.000	0.000	0.001	0.005	0.003	NA	NA
9	ARMAX - MRR - SIC	1.000	0.252	0.038	0.043	0.176	0.214	1.000
11	ARMAX - DR -SIC	1.000	0.000	0.000	0.000	0.000	0.001	NA
13	ARMAX - MRR&DR - SIC	1.000	0.240	0.017	0.031	0.345	0.540	1.000
14	ARMA errors - MRR - AIC	0.999	0.254	0.021	0.001	0.000	NA	NA
15	ARMA errors - MRR - SIC	1.000	0.776	0.144	0.001	0.000	NA	NA
16	ARMA errors - DR - AIC	1.000	0.871	0.259	0.008	0.000	0.000	NA
17	ARMA errors - DR -SIC	1.000	0.971	0.467	0.012	0.009	0.011	NA
18	ARMA errors - MRR&DR - AIC	1.000	0.501	0.079	0.000	0.000	0.000	NA
19	ARMA errors - MRR&DR - SIC	1.000	0.915	0.257	0.003	0.003	0.002	NA
21	ARDL - MRR - SIC	1.000	0.894	0.431	0.062	0.007	NA	NA
23	ARDL - DR -SIC	1.000	0.855	0.098	0.001	0.003	0.006	NA
25	ARDL - MRR&DR - SIC	1.000	0.996	0.858	0.434	0.601	0.661	NA
	Multivariate models - differences							
26	Static regerssion, d - MRR	1.000	1.000	1.000	0.978	0.689	0.450	NA
27	Static regerssion, d - DR	1.000	1.000	1.000	0.978	0.689	0.450	NA
28	Static regerssion, d - MRR&DR	1.000	1.000	1.000	0.978	0.689	0.450	NA
30	ARMAX, d - MRR - SIC	1.000	1.000	1.000	0.994	0.769	0.481	0.000
32	ARMAX, d - DR -SIC	1.000	1.000	1.000	0.990	0.736	0.468	0.000
34	ARMAX, d - MRR&DR - SIC	1.000	1.000	1.000	0.992	0.750	0.474	NA
38	ARMA errors, d - DR -SIC	1.000	1.000	1.000	0.978	0.689	0.450	NA
42	ARDL, d - MRR - SIC	1.000	1.000	1.000	0.993	0.766	0.480	0.000
44	ARDL, d - DR -SIC	1.000	1.000	1.000	0.998	0.865	0.521	NA
46	ARDL, d - MRR&DR - SIC	1.000	1.000	0.999	0.963	0.693	0.444	NA
	Multivariate models - spread							
47	ARMA, s - MRR - AIC	0.999	0.372	0.034	0.000	0.000	0.000	NA
48	ARMA, s - MRR - SIC	1.000	0.623	0.107	0.002	0.000	NA	NA
50	ARMA, s - DR - SIC	1.000	0.807	0.098	0.001	0.002	0.003	NA
52	ARMA, s, d - MRR - SIC	1.000	1.000	1.000	0.979	0.693	0.452	NA
53	ARMA, s, d - DR - AIC	1.000	1.000	0.993	0.759	0.312	0.229	NA
54	ARMA, s, d - DR - SIC	1.000	1.000	1.000	0.978	0.688	0.450	NA
	Multivariate models - break							
55	Static regression, b - MRR	0.000	0.001	0.010	0.122	0.303	0.258	NA
56	Static regression, b - DR	0.000	0.000	0.000	0.000	0.000	0.000	NA
57	Static regression, b - MRR&DR	0.000	0.000	0.001	0.005	0.003	NA	NA
61	ARMA, s, b - MRR	0.000	0.000	0.000	0.000	0.000	0.000	NA
62	ARMA, s, b - DR	0.000	0.000	0.001	0.007	0.012	0.000	NA
	Multiequation models							
63	Levels - MRR - AIC	1.000	0.991	0.476	0.000	0.000	0.000	0.000
64	Levels - MRR - SIC	1.000	0.997	0.596	0.000	0.001	0.004	NA
65	Levels - DR - AIC	1.000	0.903	0.205	0.000	0.000	0.001	NA
66	Levels - DR - SIC	1.000	0.798	0.125	0.000	0.000	NA	NA
67	Differences - MRR - AIC	1.000	1.000	0.995	0.794	0.464	0.423	NA
68	Differences - MRR - SIC	1.000	1.000	0.995	0.769	0.457	0.416	NA
69	Differences - DR - AIC	1.000	1.000	0.998	0.782	0.439	0.435	NA
70	Differences - DR - SIC	1.000	1.000	0.999	0.869	0.500	0.445	1.000
71	VECM - MRR - AIC	1.000	0.991	0.476	0.000	0.000	0.000	0.000
72	VECM - MRR - SIC	1.000	0.997	0.596	0.000	0.001	0.004	NA
73	VECM - DR - AIC	1.000	0.903	0.205	0.000	0.000	0.001	NA
74	VECM - DR - SIC	1.000	0.798	0.125	0.000	0.000	NA	NA

2.C Recursive forecast graphs

This section contains graphs of recursive forecasts with a 24 month horizon for all models for the Eonia rate, in terms of both level and spread from the deposit rate. For single equation models, I report ex-post forecasts, while for VAR models, I report ex-ante forecasts. The forecasts for other variables, and ex-ante forecasts, are available upon request.

Figure 2.13: Eonia rate forecasts - Univariate models

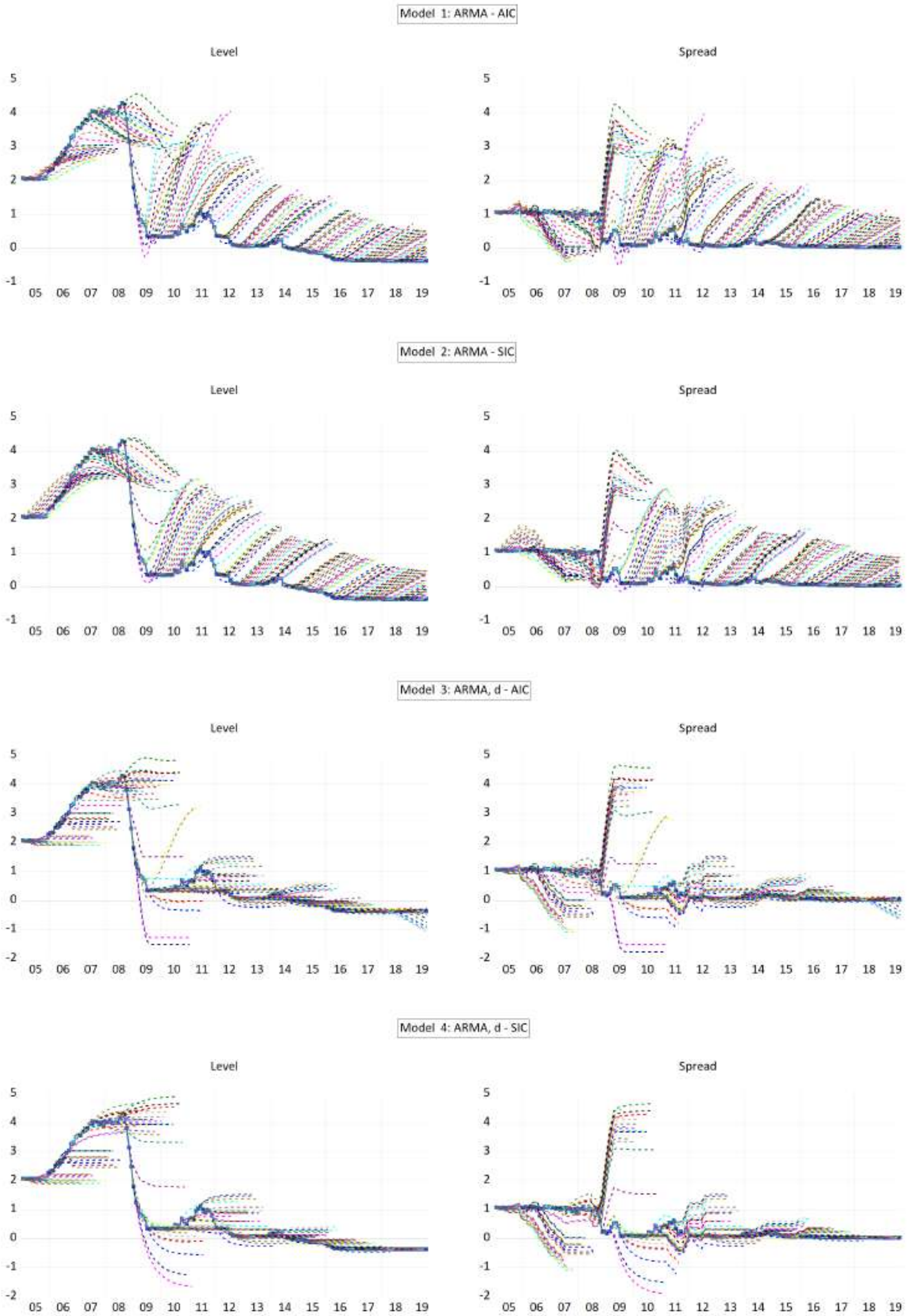


Figure 2.14: Eonia rate forecasts - Reduced-form models - Levels (part 1)

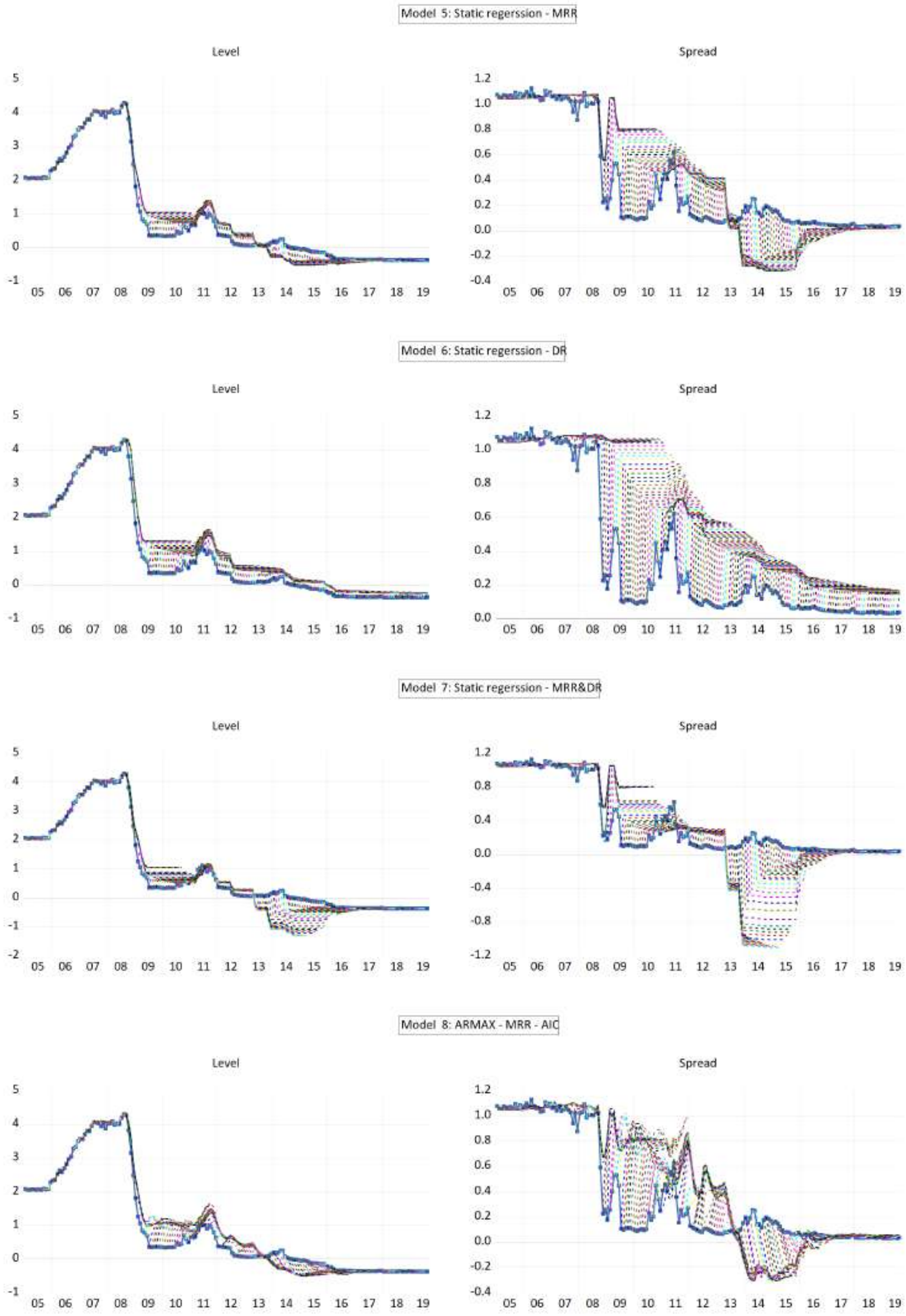


Figure 2.15: Eonia rate forecasts - Reduced-form models - Levels (part 2)

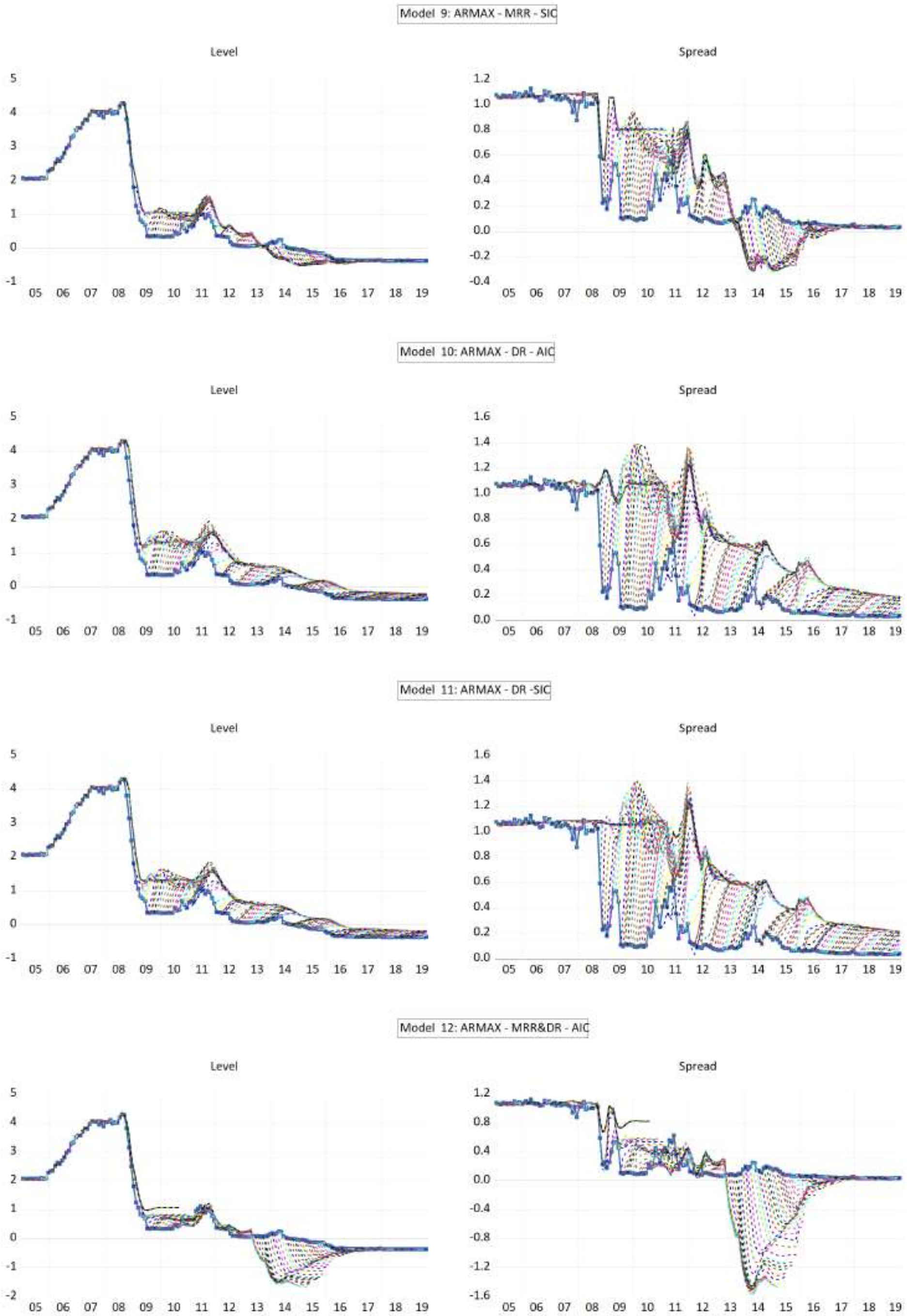


Figure 2.16: Eonia rate forecasts - Reduced-form models - Levels (part 3)

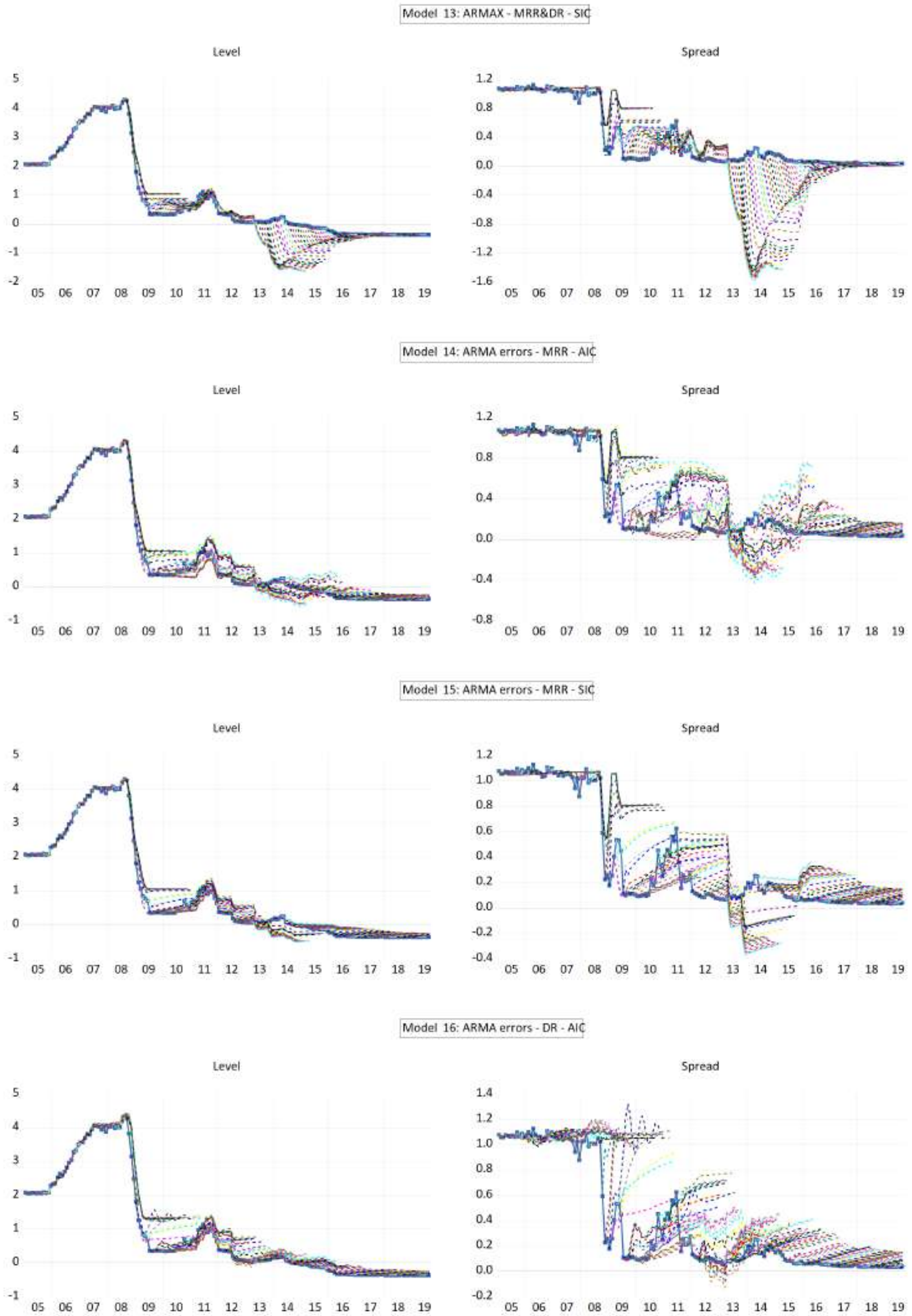


Figure 2.17: Eonia rate forecasts - Reduced-form models - Levels (part 4)

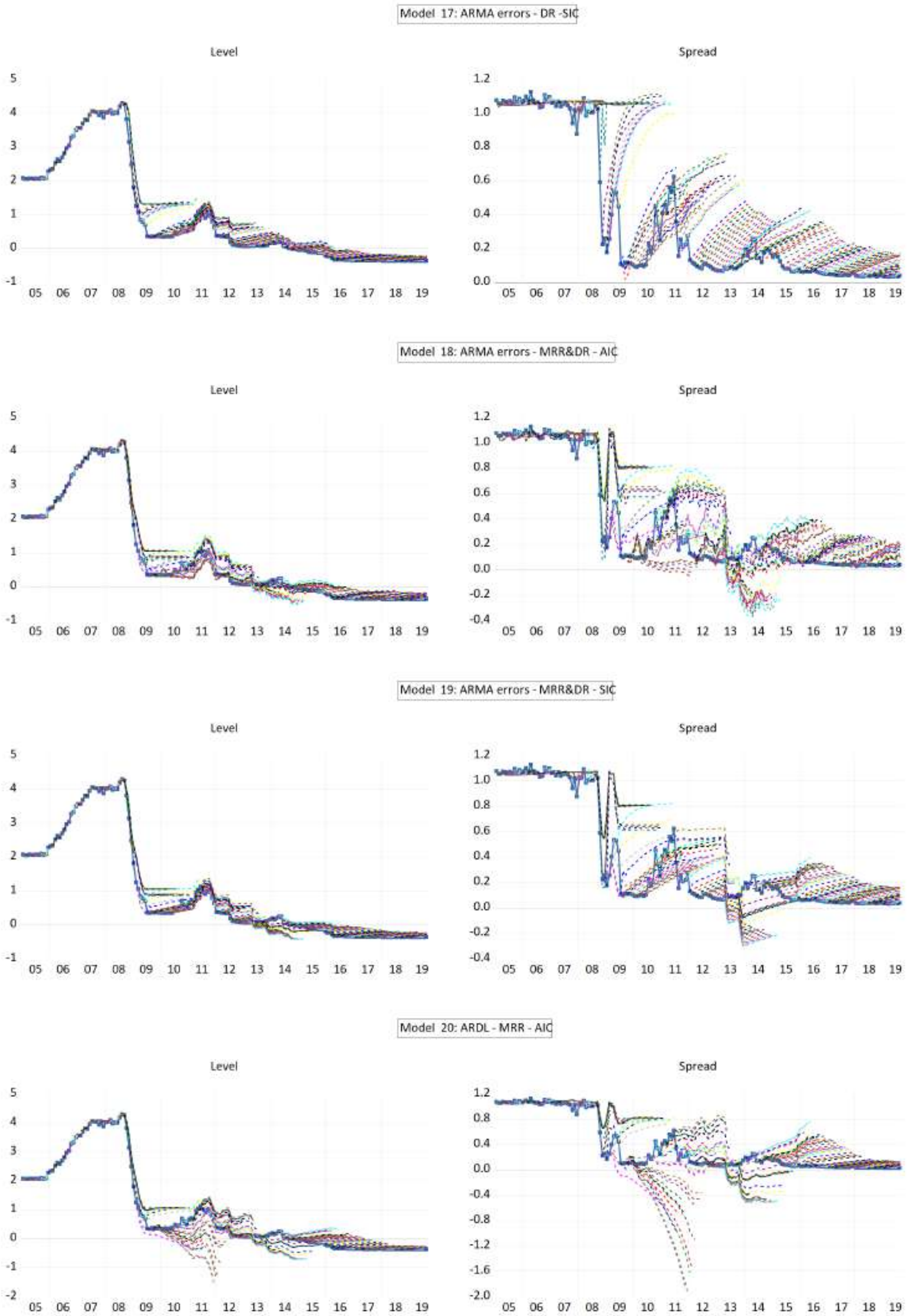


Figure 2.18: Eonia rate forecasts - Reduced-form models - Levels (part 5)

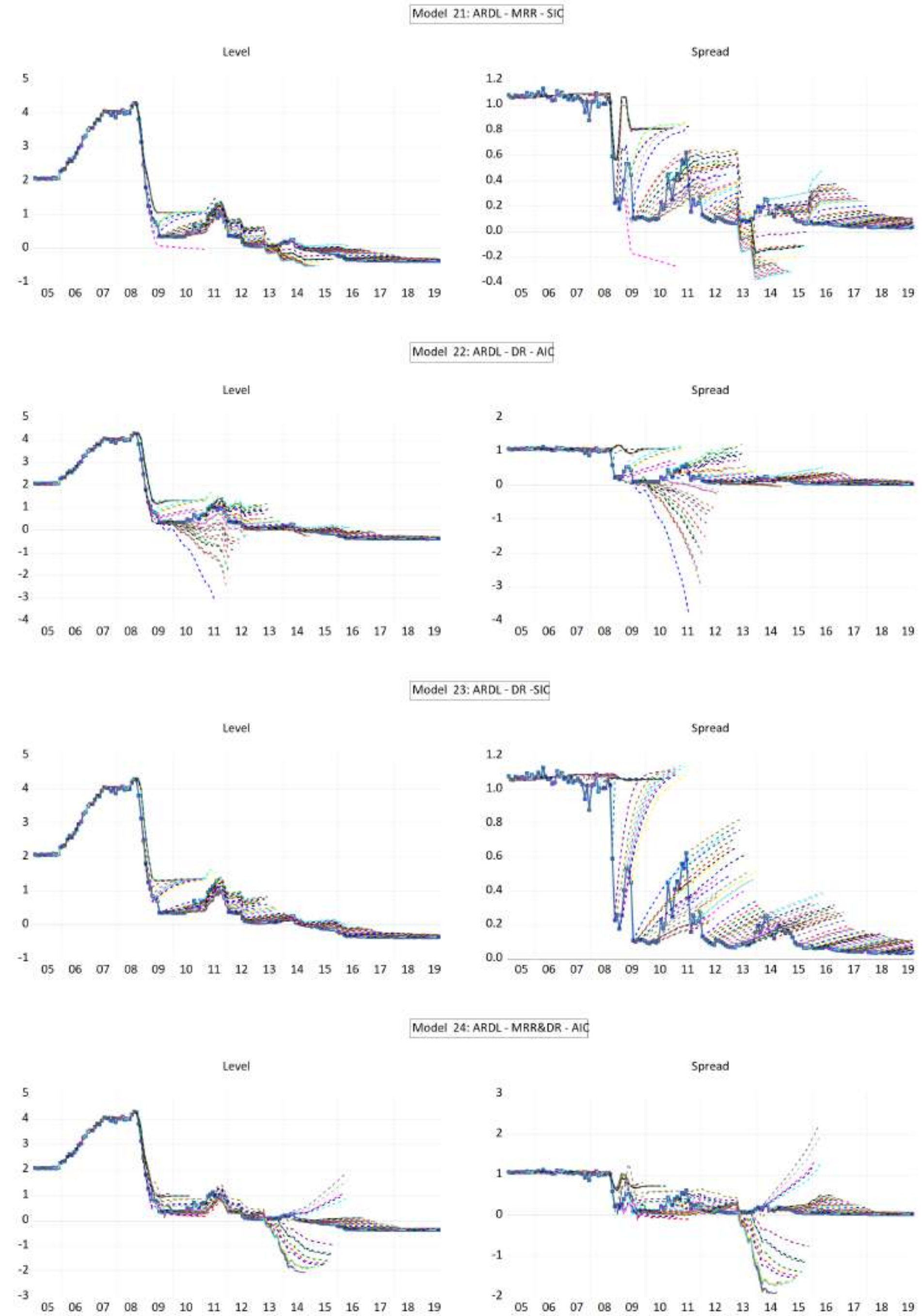


Figure 2.19: Eonia rate forecasts - Reduced-form models - Levels (part 6)

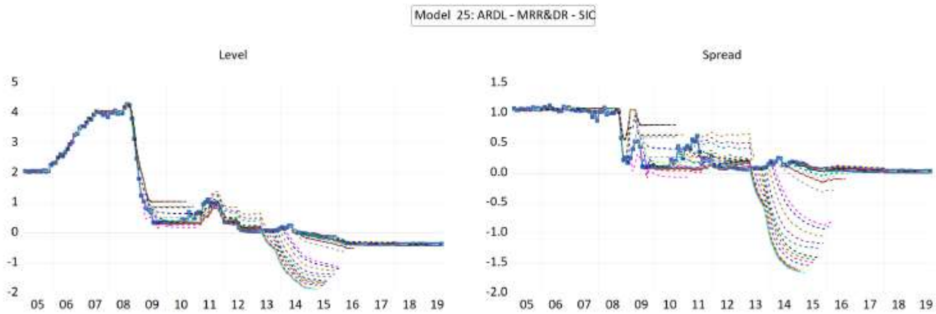


Figure 2.20: Eonia rate forecasts - Reduced-form models - Differences (part 1)

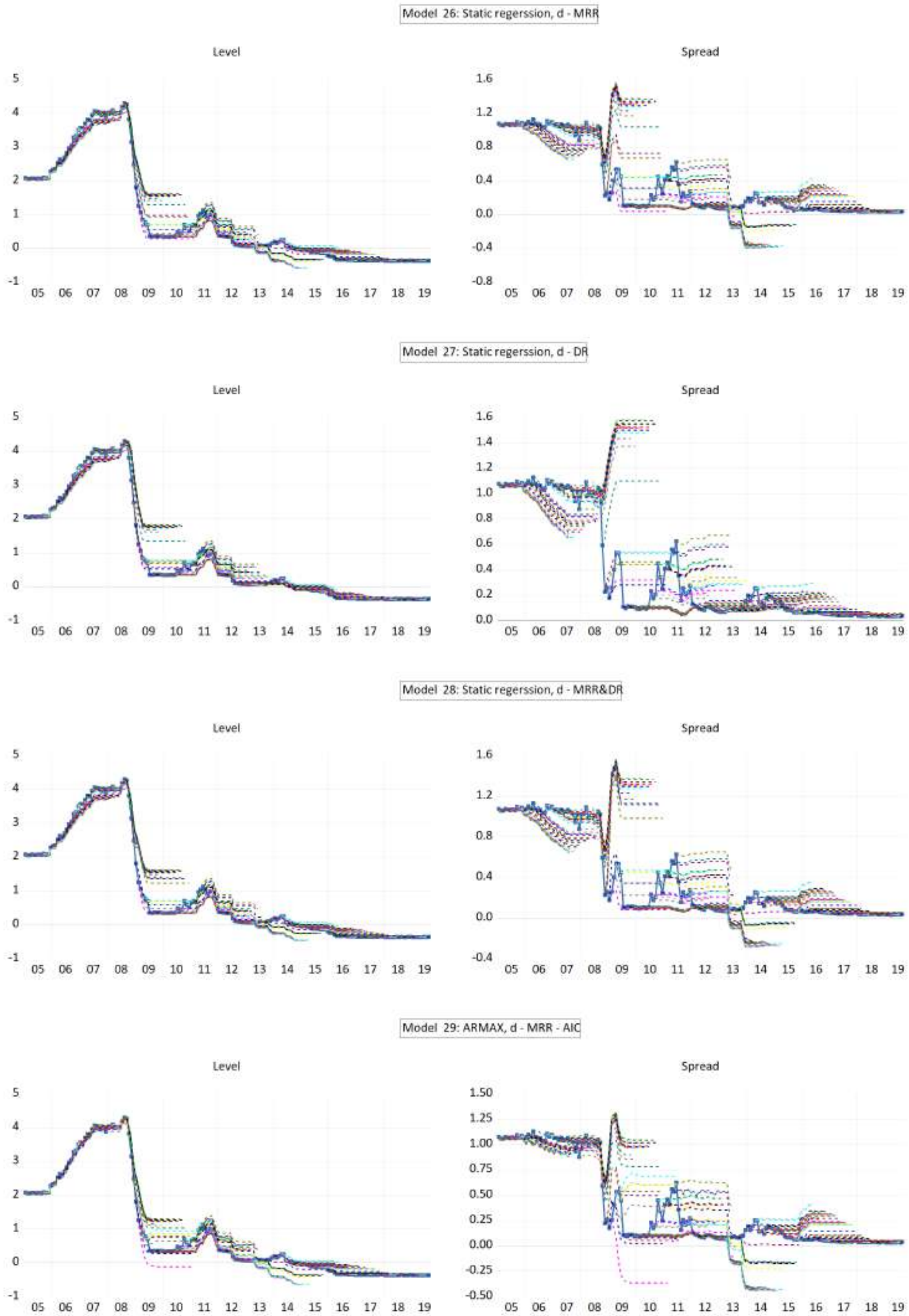


Figure 2.21: Eonia rate forecasts - Reduced-form models - Differences (part 2)

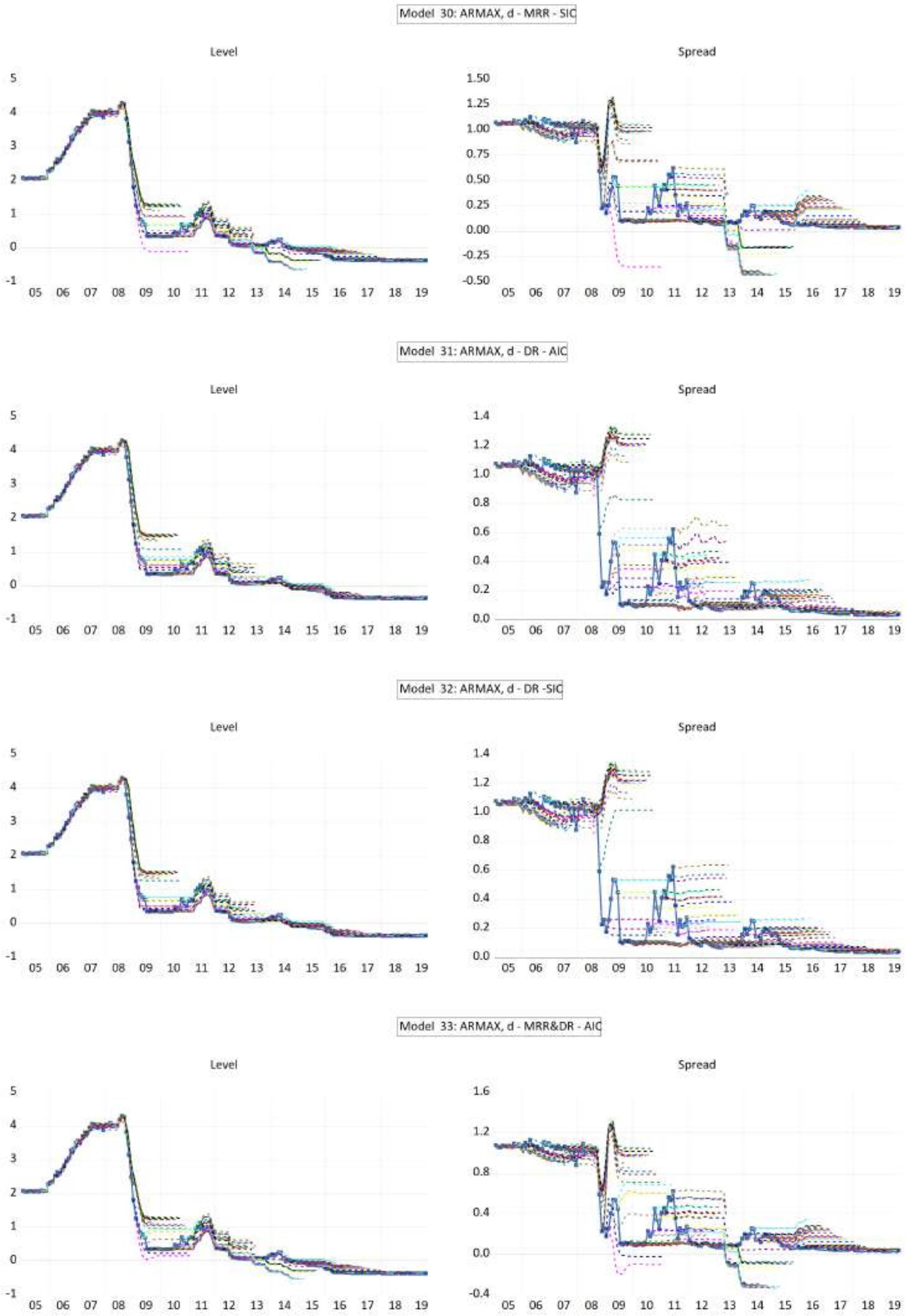


Figure 2.22: Eonia rate forecasts - Reduced-form models - Differences (part 3)

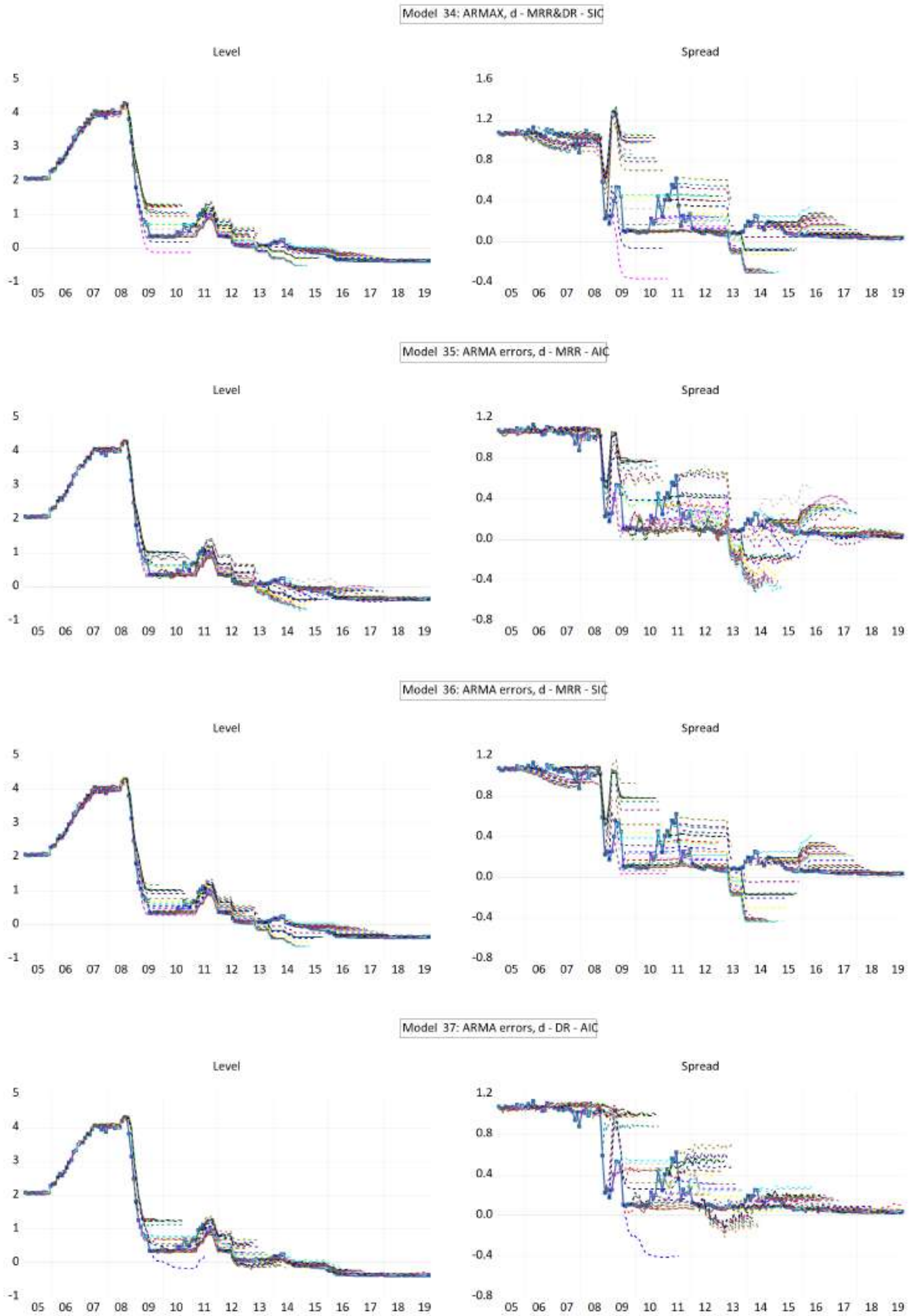


Figure 2.23: Eonia rate forecasts - Reduced-form models - Differences (part 4)

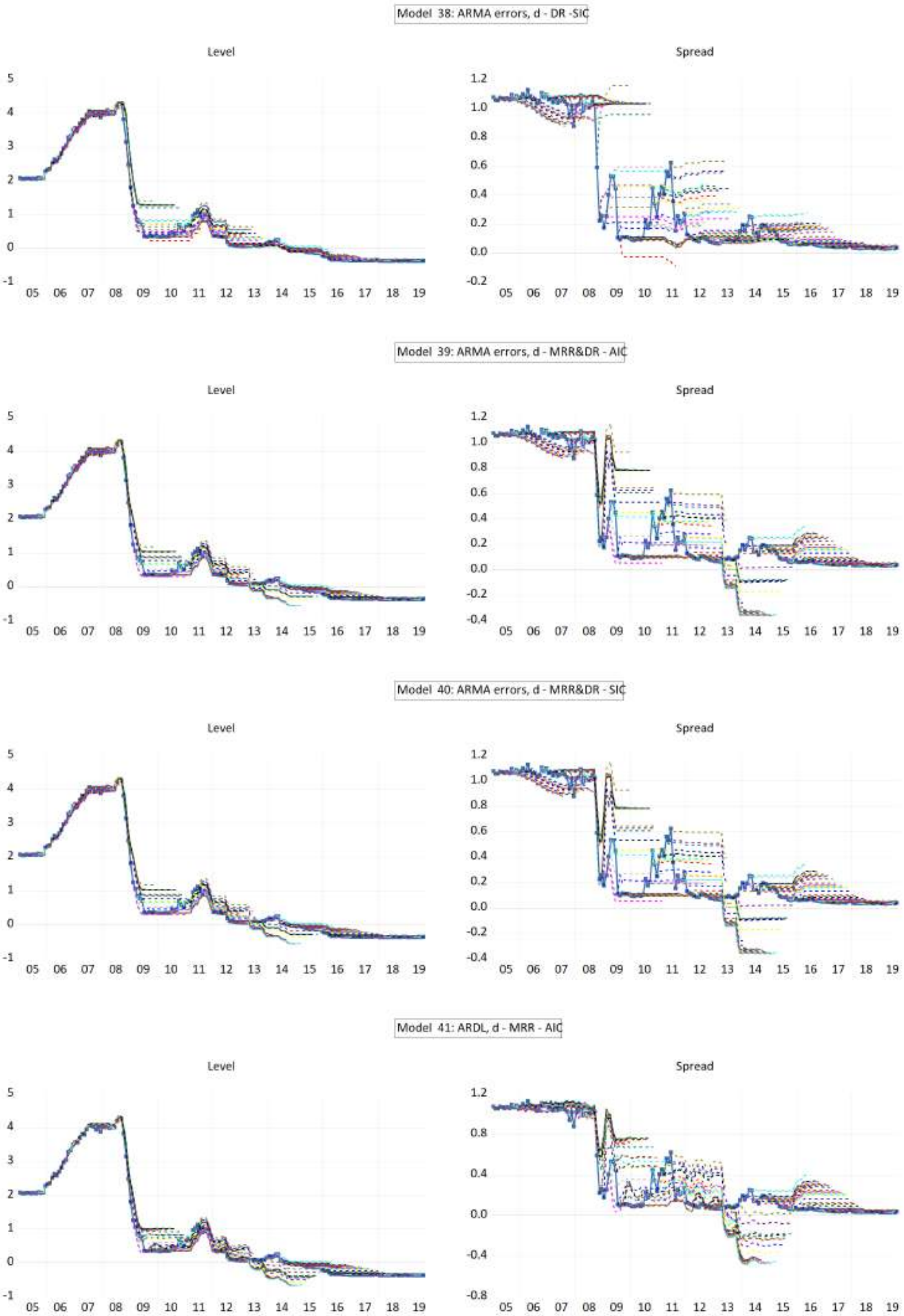


Figure 2.24: Eonia rate forecasts - Reduced-form models - Differences (part 5)

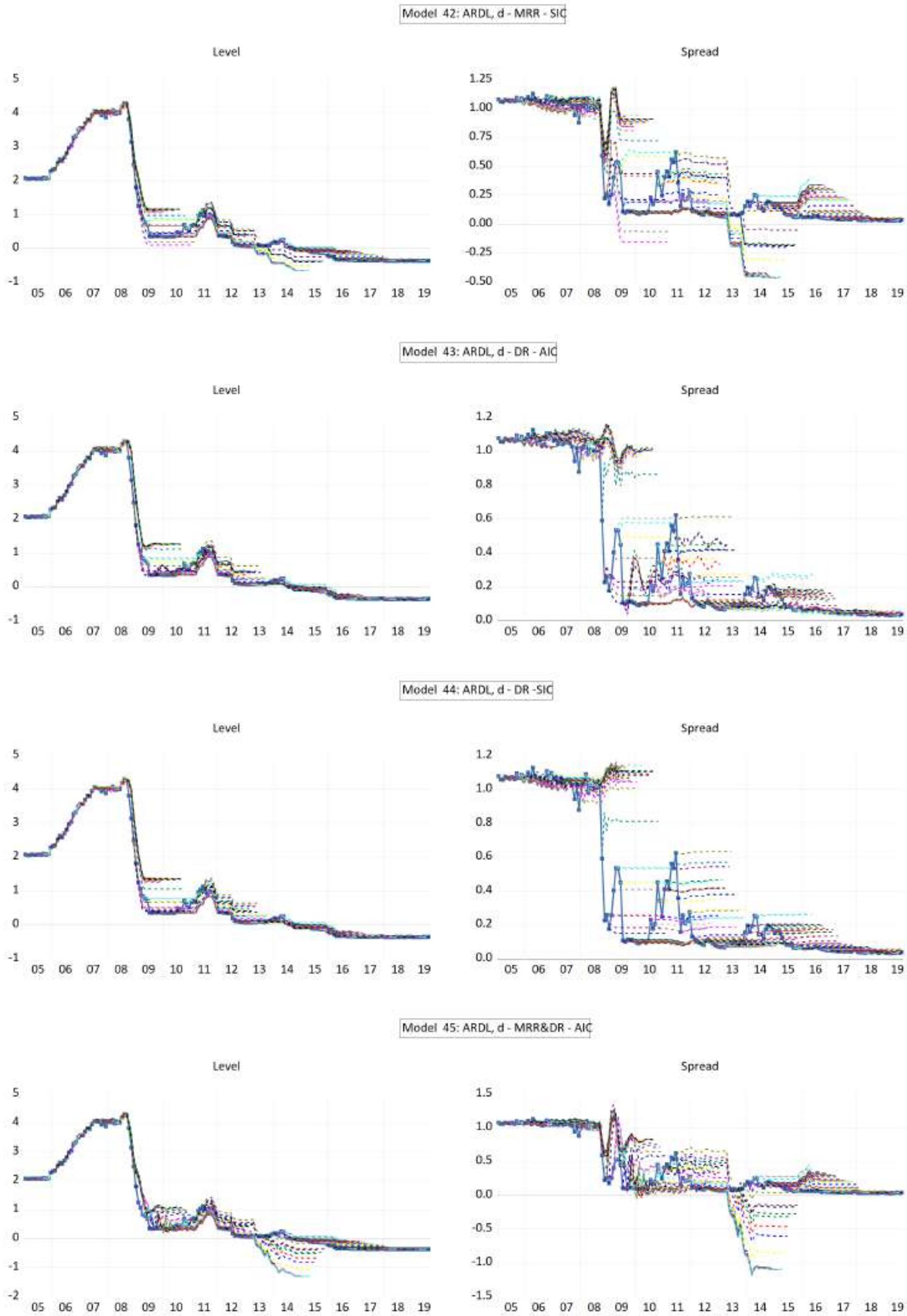


Figure 2.25: Eonia rate forecasts - Reduced-form models - Differences (part 6)

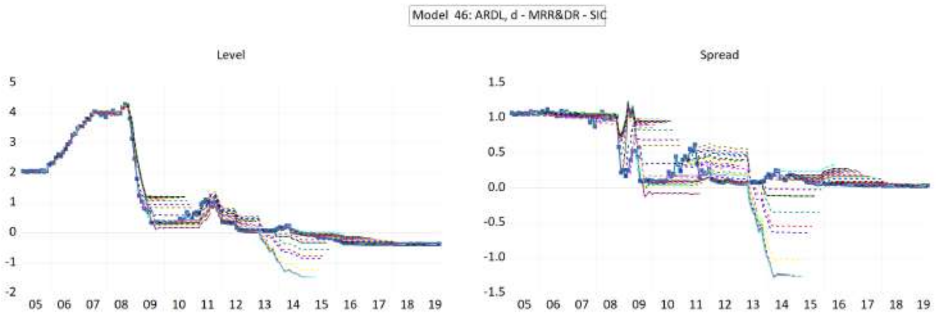


Figure 2.26: Eonia rate forecasts - Reduced-form models - Spread (part 1)

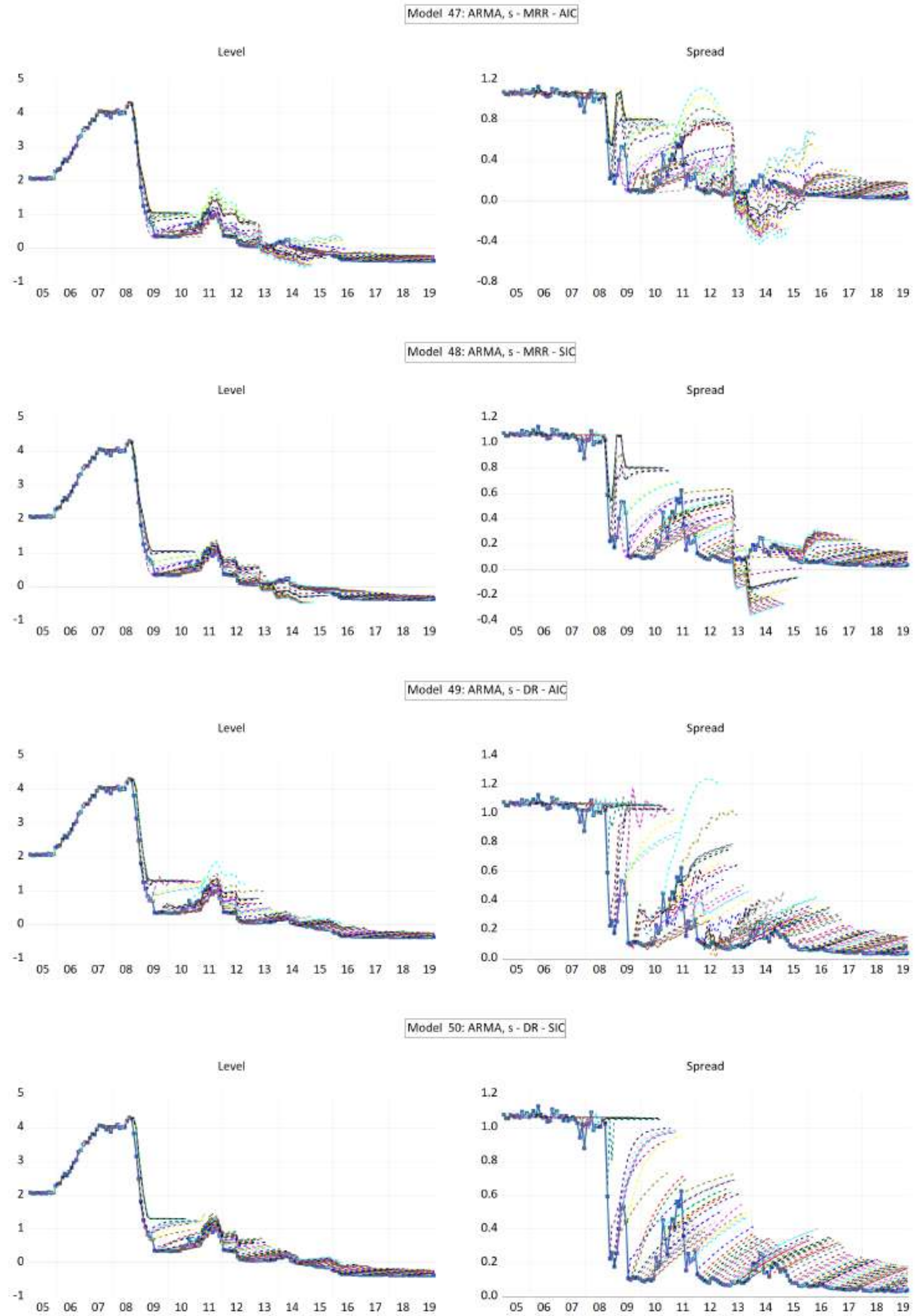


Figure 2.27: Eonia rate forecasts - Reduced-form models - Spread (part 2)

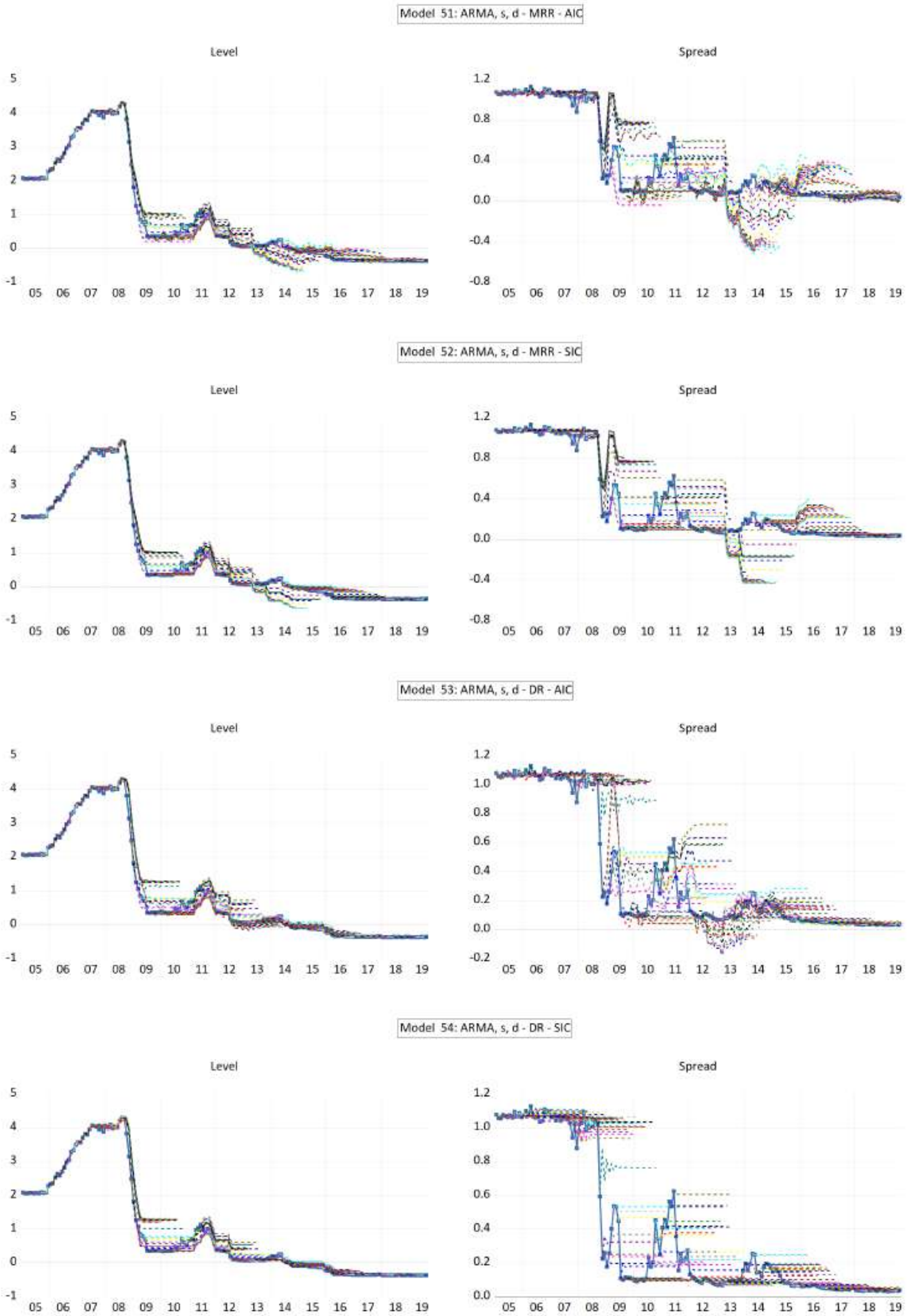


Figure 2.28: Eonia rate forecasts - Reduced-form models - Break (part 1)

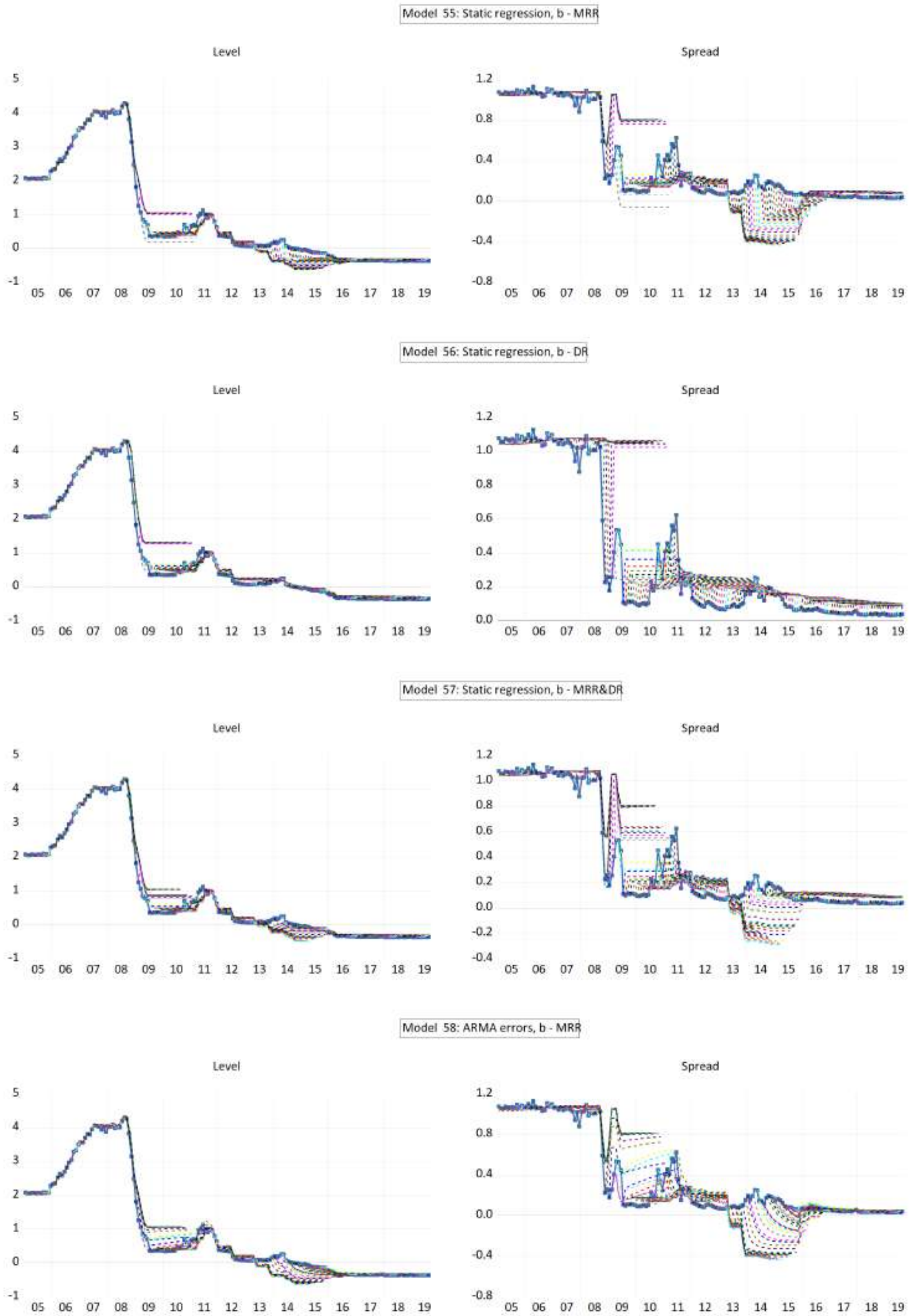


Figure 2.29: Eonia rate forecasts - Reduced-form models - Break (part 2)

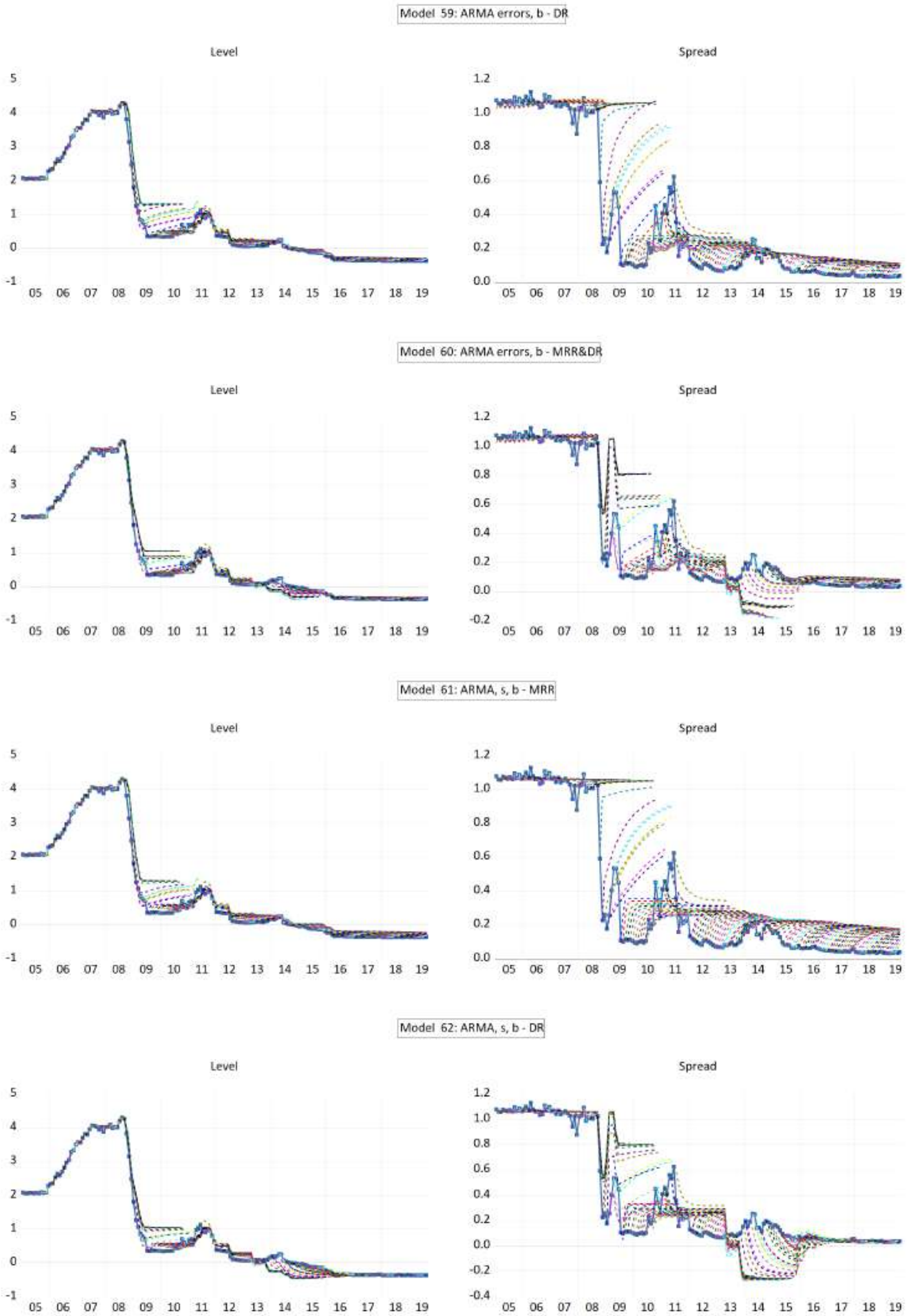


Figure 2.30: Eonia rate forecasts - VAR models (part 1)

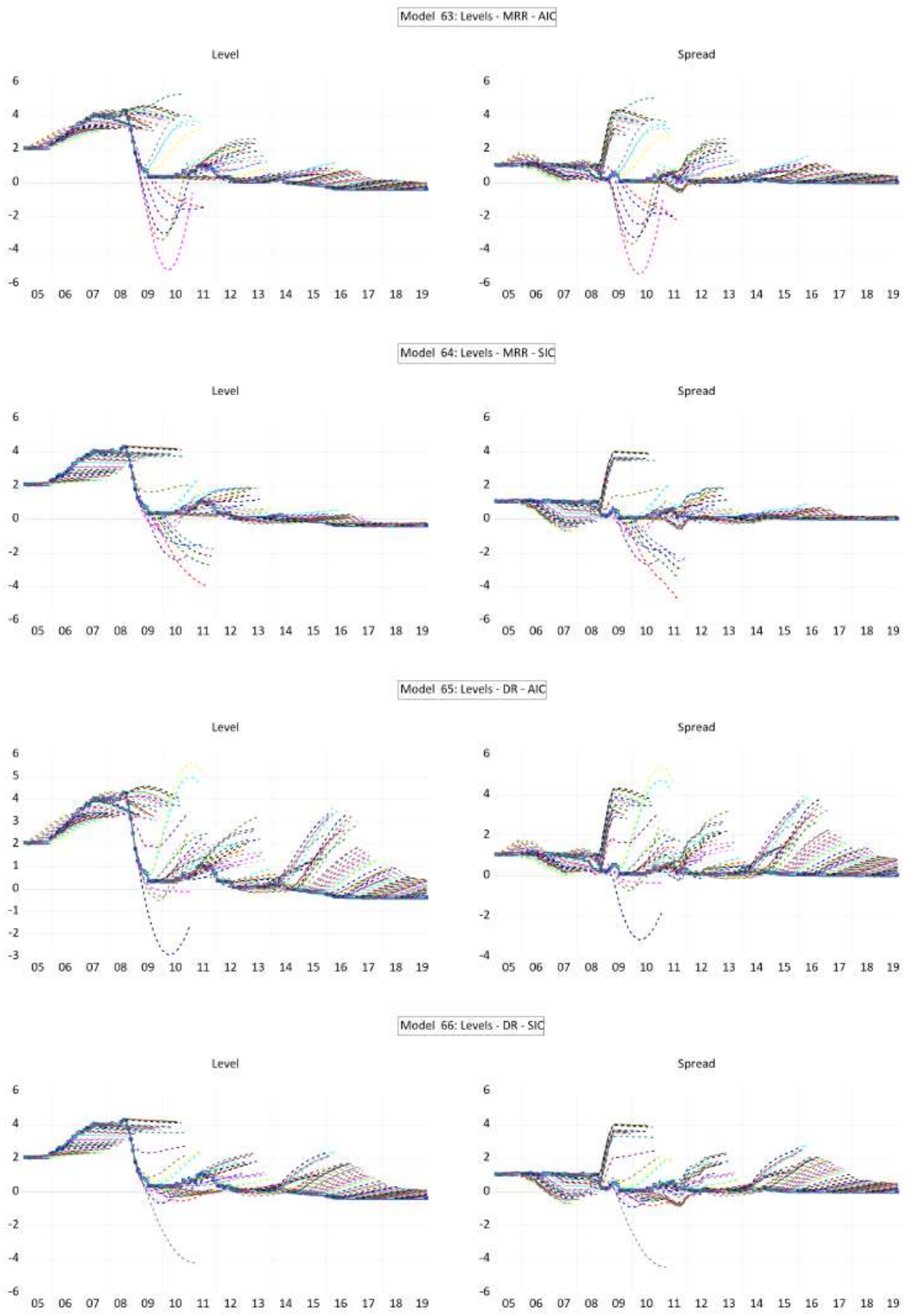


Figure 2.31: Eonia rate forecasts - VAR models (part 2)

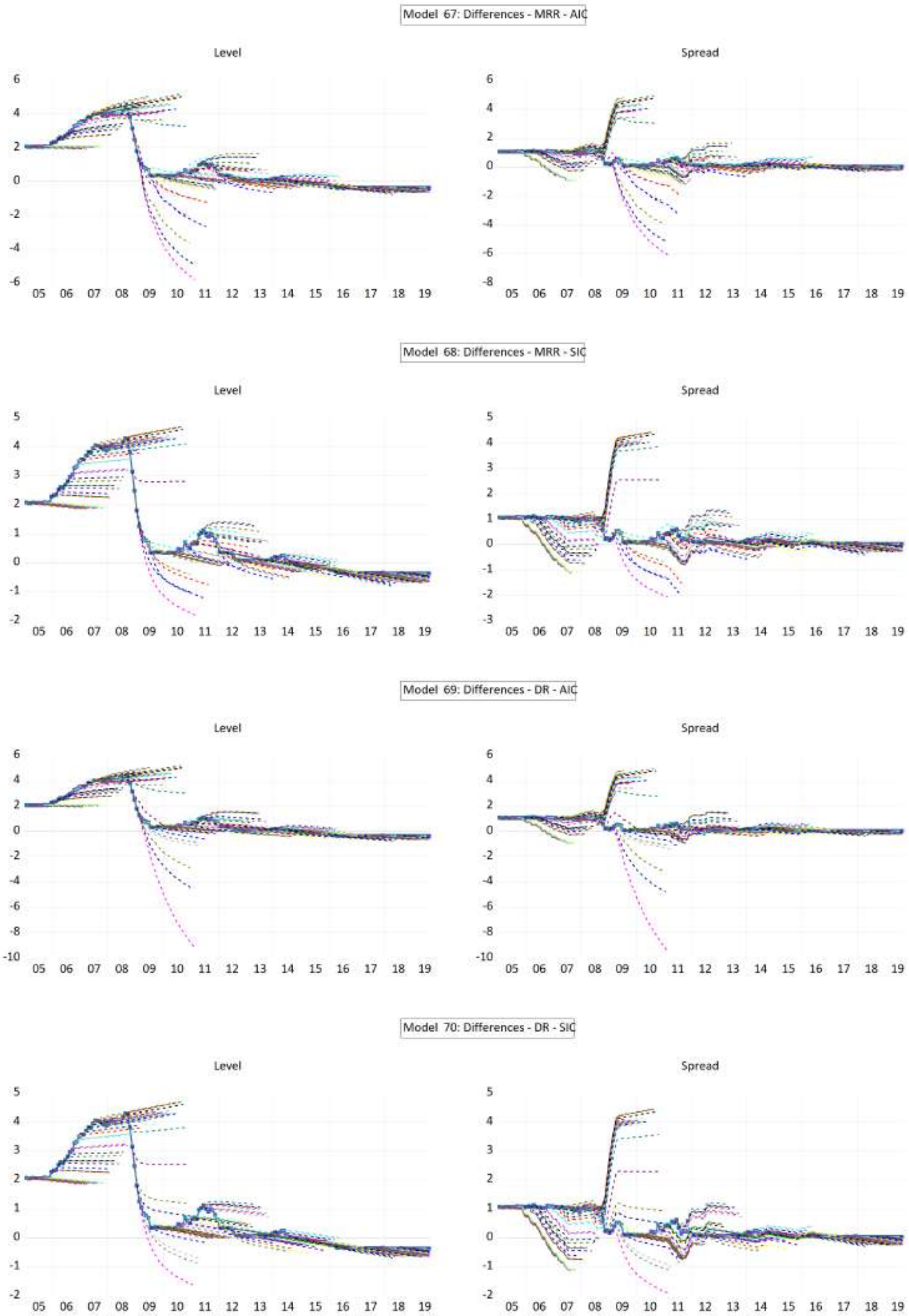


Figure 2.32: Eonia rate forecasts - VAR models(part 3)

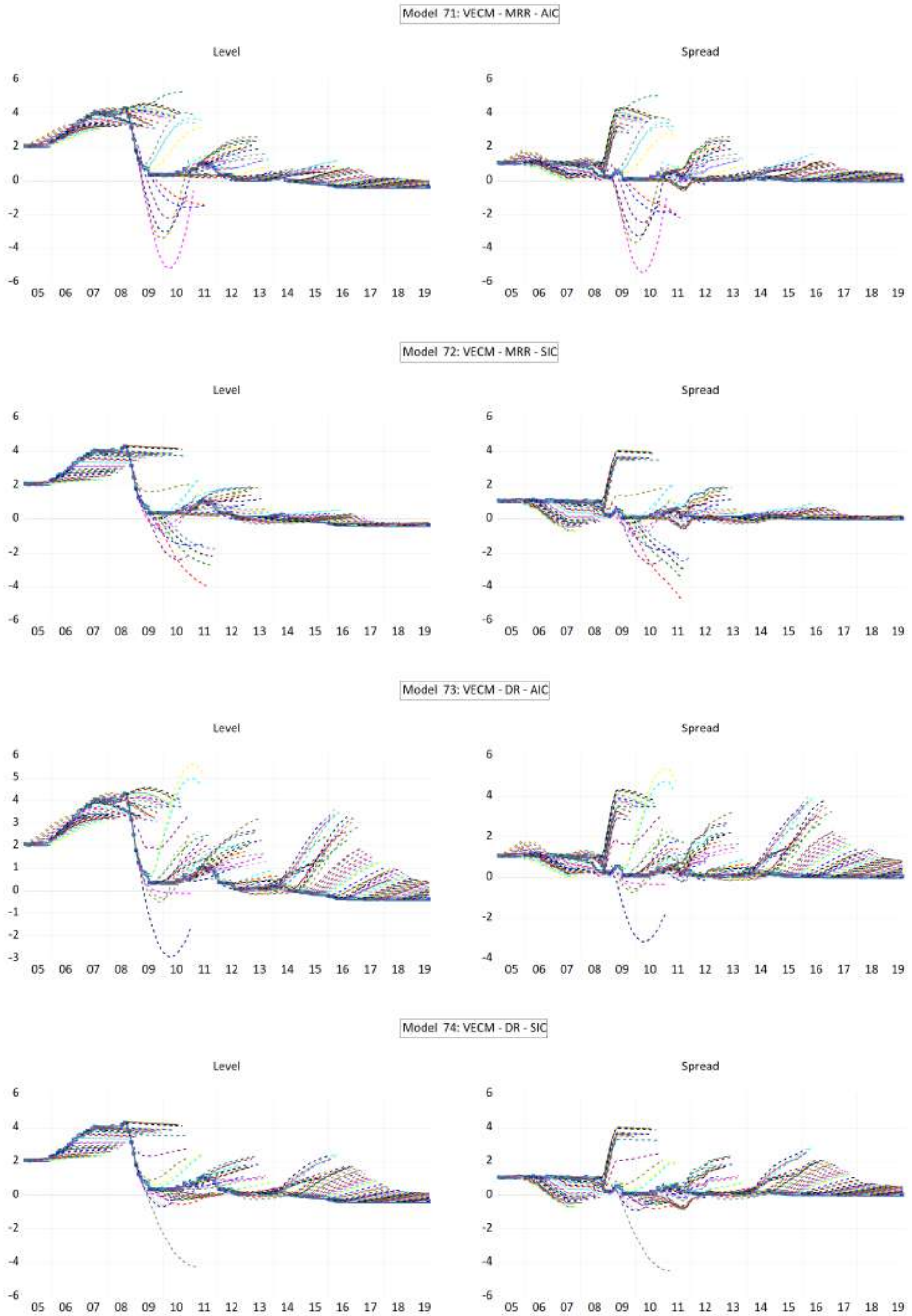
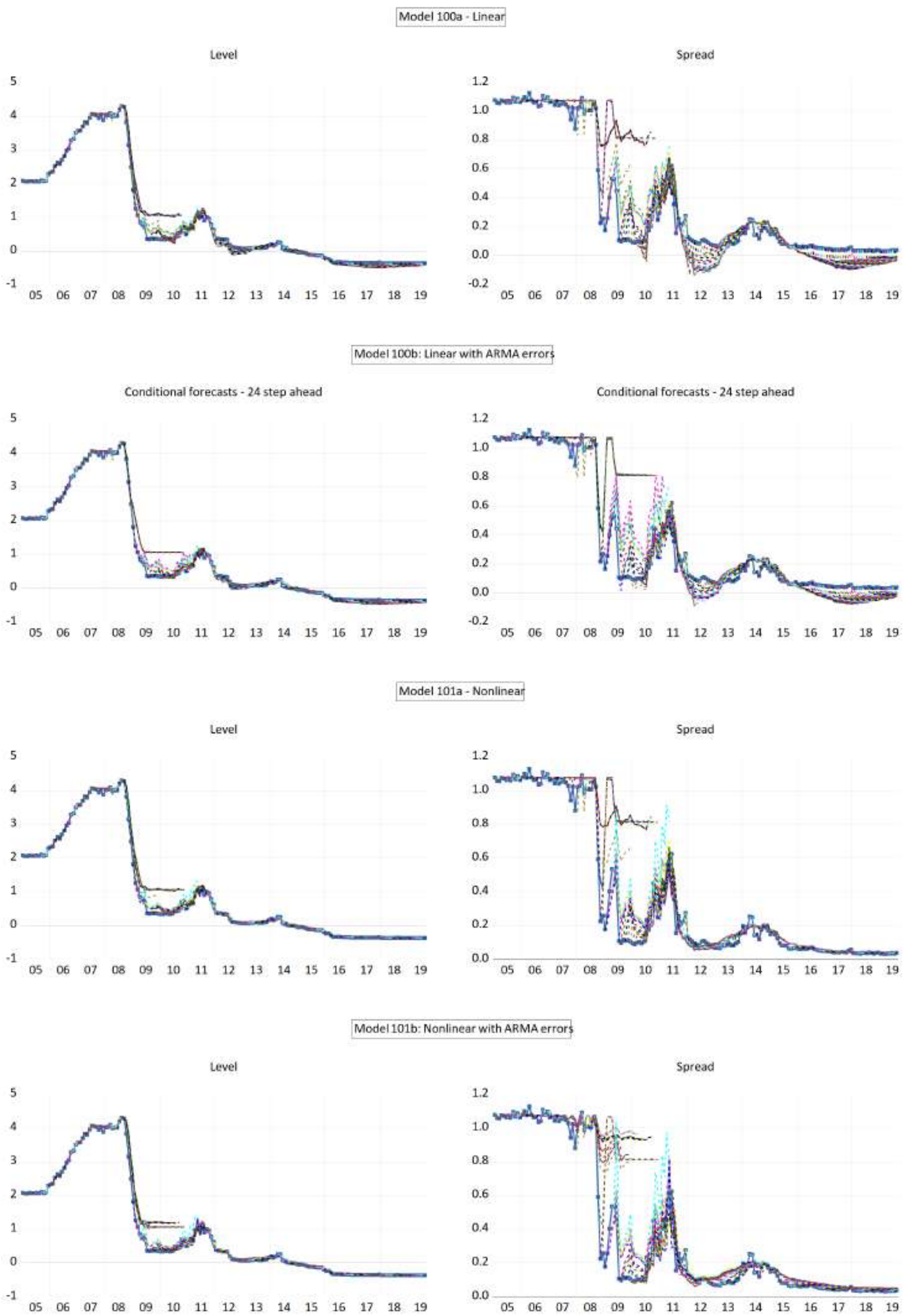


Figure 2.33: Eonia rate forecasts - Structural models



Chapter 3

Developing forecasting models using the `SpecEval` add-in for Eviews

3.1 Introduction

The proliferation of economic data and computational power over last two decades has led to an explosion in the development of forecasting models. Not only has the population of model developers increased, but even more importantly, it has diversified away from economists employed by government and international institutions and towards practitioners. This brings new challenges, since practitioners face very different constraints when developing forecasting models, especially in terms of time available for development. The objectives of practitioners are also often different: while academia and government institutions pay close attention to causality and econometric validity, thanks to their focus on policy analysis, practitioners are less likely to perform policy analysis and are more focused on capturing co-movements than causality. Similarly, while practitioners are often more focused on conditional and scenario forecasting, for institutions such as central banks, unconditional forecasting is often the primary goal.

These factors mean that many model developers need tools that will allow them to analyze model forecasting performance *quickly* and *flexibly*. The Eviews user add-in `SpecEval` belongs to the family of statistical packages aimed at facilitating fast and flexible evaluation of forecasting performance. However, in contrast to existing packages, the main goal of the add-in is much broader than simply calculating and reporting summary measures of forecasting performance. The underlying idea is that model development should be an *iterative* and *interactive* process. It should be *iterative* in that it proceeds in steps, each providing improvements over previous models. It should be *interactive* in that improvements to the model should be based on analysis of shortcomings of an initial model. Crucially, numerical summaries of forecasting performance fail on the goal of interactiveness, since a numerical summary provides little information on when and why the model fails to produce good forecasts, or how it should be improved. For this reason the `SpecEval` add-in is focused on producing graphical information about forecasting performance, mostly in form of visualizing the forecasts themselves.

Focus on visualizing forecasts is also important for a second reason: time series models are often too complex for humans to completely understand what the exact shape of forecasts produced by these models will be, and hence too complex to fully understand the model behavior. Visualizing multiple forecasts helps in this regard, since it provides illustrations of the behavior of the model under various historical or scenario conditions. The `SpecEval` also features several additional tools facilitating understanding of model forecasts: (a) ability to decompose a forecast into individual drivers, (b) ability to create and compare (conditional) scenario forecasts and (c) a capability to create shock responses for single-equation multivariate models. The key idea behind these features, which are novel in context of such packages, is that understanding of model behavior is an important source of information that can facilitate model improvements. This is especially true in a theory-heave environment such as forecasting macroeconomic time series.

In addition to its specific novel features, the add-in also facilititates model development through its flexibility and comprehensiveness, both of which are crucial for analysis of model behavior. For example, a user can choose between in-sample and out-of-sample forecasts, set starting and ending dates of backtesting, or specify multiple sub-samples over which performance should be analyzed. Moreover, the add-in allows users to control which information is treated as endogenous by either providing forecasts for exogenous series or by including additional equations/identities in the forecasting model. Taken together, the focus and functionality of the add-in make it a novel tool for building forecasting models, not only in the context of the Eviews program, but in the context of other programs used for statistical analysis such as R and Python.

This chapter is an excerpt from one of the documents provided with the add-in. It first briefly discusses basic operation of the add-in, including an illustration of the output produced by the add-in and list of all the types of tasks it can perform, and their associated objects. This provides a background for a full illustration of the add-in in context of an application, presented in section 3. The full version of this document - available as part of the documentation for the add-in - then includes additional 7 applications, brief outline of which appears in beginning of section 3.

3.2 Basic use of the `SpecEval` add-in

This section briefly describes the basic operation of the `SpecEval` add-in. The section is divided into two parts, one corresponding to operation of the add-in in case of a single specification, and the other for its operation in case of multiple specifications. The difference between these use cases is in terms of organizing and reporting results: if multiple specifications are specified, the main spool will have results organized in a way that facilitates comparison across specifications; nothing changes in terms of execution, which is still performed one specification at a time.

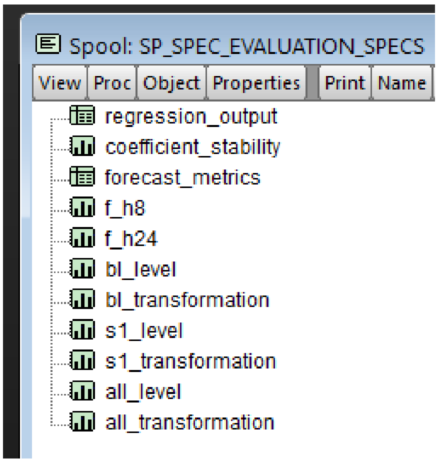
The add-in can be executed from the GUI or as an object procedure in a command line or program. In the latter case, one simply needs to specify the proc "speceval", the same way one executes other object procedures:

```
{equation}.speceval(options)
```

The options specified in the command allow a user to modify the execution of the add-in. All options are described in the documentation for the add-in. While their discussion is not inside the scope of this chapter - there are almost 40 of them - the chapter will illustrate a large share of these options as part of the application.

Single specification. The basic operation of the `SpecEval` add-in considers the case of evaluating a single specification. Executing the add-in will result in a spool called 'sp_spec_evaluation' being produced. This spool will contain multiple objects, both tables and graphs, that allow a user to perform the key goals of

Figure 3.1: Spool with output objects - single specification



the add-in: evaluation of forecast performance of a given model. See Figure 3.1 for illustration. Detailed discussion of all output objects produced by the add-in is beyond the scope for this chapter, though most will be illustrated in the application below. Table 3.1 includes overview with short description of all the outputs; detailed explanations and illustration of all the output is available in the document 'Outputs for specification evaluation add-in'. As the table makes clear, most of the outputs are focused on visualizing forecasts, either backtest forecasts (outputs 6-8), or scenario forecasts (outputs 10-13). In addition there are tables with numerical summaries of forecast performance, and graphs capturing forecast errors of the forecasts (Mincer-Zarnowitz graphs). Finally, the spool includes standard Eviews regression output with several additions and adjustments that facilitate model development, and coefficient and model stability graphs.

Though a discussion of the full range of setting options is beyond the scope of this chapter, some key options are worth highlight. First, given the large number of potential outputs the add-in allows for easy and flexible customization of the execution list. The default does not contain all the output objects, given that the

Table 3.1: List of outputs from the `SpecEval` add-in

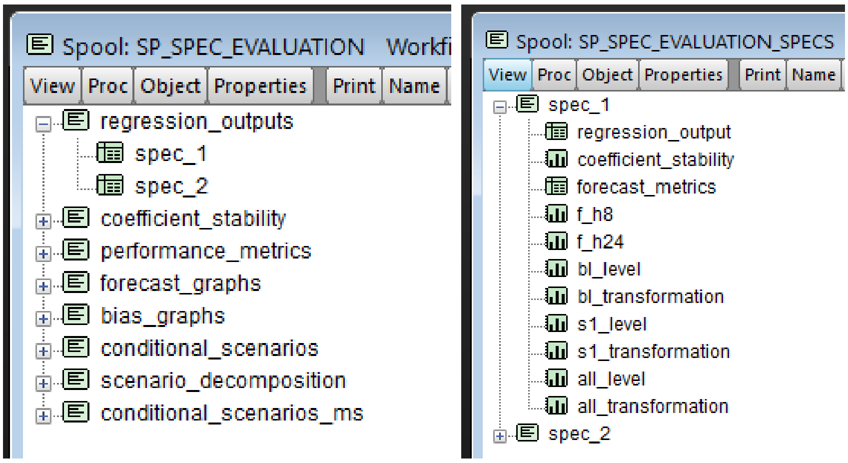
Object name	Description
Regression output	Adjusted regression output table
Coefficient stability graph	Graph with recursive equation coefficients
Model stability graph	Graph with recursive lag orders
Performance metrics table	Table with values of forecast performance metrics
Performance metrics table (multiple specifications)	Table with values of forecast performance metrics for given metric for all specifications
Forecast summary graph	Graph with all recursive forecasts with given horizons
Sub-sample forecast graph	Graph with forecast for given sub-sample
Subsample forecast decomposition graph	Graph with decomposition of sub-sample forecast
Forecast bias graph	Scatter plot of forecast and actual values for given forecast horizon (Minzer-Zarnowitz plot)
Individual conditional scenario forecast graph (level)	Graph with forecast for single scenario and specification
Individual conditional scenario forecast graph (trans.)	Graph with transformation of forecast for single scenario and specification
All conditional scenario forecast graph	Graph with forecasts for all scenarios for single speciication
Multiple specification conditional scenario forecast graph	Graph with forecasts for single scenario for multiple specifications
Shock response graphs	Graphs with response to shock to individual independent variable/regressor

full list is unlikely to be desired in most cases. The user can freely add and/or subtract from the list, using the key words corresponding to each output type, found in the documentation for the add-in. Second, the add-in allows a great deal of customization of the evaluation procedure. For example, one can change the horizons for which performance metrics will be calculated by including the settings option 'horizons_metrics'; or horizons for which forecast summary graphs will be created by including the settings option 'horizons_graphs'; and adjust the starting and/or ending date of the evaluation procedure by including the settings parameter 'tfirst_test', respectively 'tlast_test'.

Lastly, a user can customize the outputs. For example, the forecast summary graphs can be changed to several different transformations by including the settings option 'trans'. Similarly, one can adjust the sample of these graphs by including the settings parameter 'tfirst_graph', respectively 'tlast_graph', and include additional series useful for evaluation of forecast performance by specifying them in the settings option 'graph_add_backtest'. Use of these and other setting options will be the main focus of several applications below.

Multiple specifications. While the basic use of the `SpecEval` add-in is for a single specification, the great advantage of the add-in is that it is tailored so to facilitate direct comparisons across alternative specifications. This is achieved by appropriate structuring and reporting of results. Specifically, when the add-in is executed for multiple specifications, it will produce two types of spools. The main spool will be organized by the type of output object, i.e. output objects of a particular type for all specifications will be collected together in a single subspool object. The auxiliary spool will be organized by specification, i.e., for each specification it will contain subspools corresponding to single specification

Figure 3.2: Spools with output objects - multiple specifications



execution. See Figure 3.2 for an illustration.

To execute the add-in for multiple specifications, one only needs to provide a list of specifications as one of the setting options, while still executing the add-in from the single equation object. For example, if there are three specifications being considered, `EQ01`,`EQ02` and `EQ03`, one can evaluate the three specifications together by executing the following command:

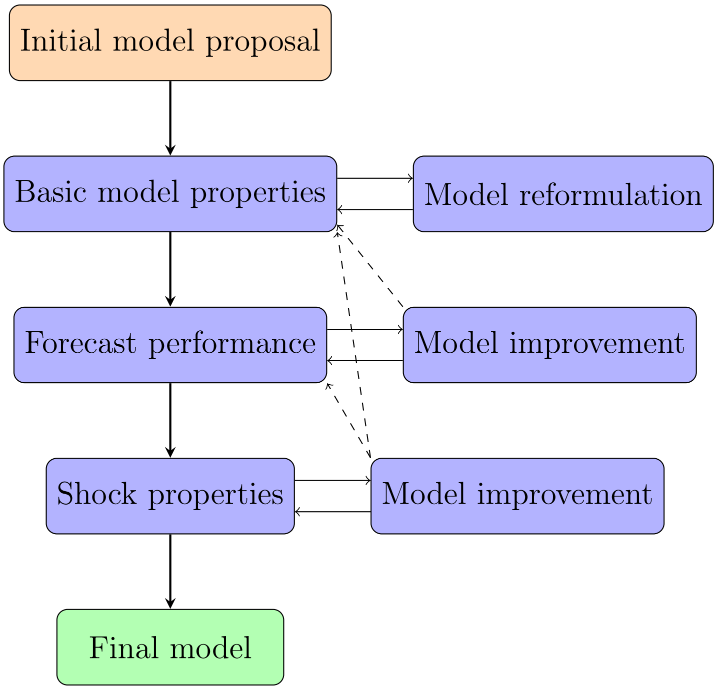
```
EQ01.speceval(spec_list="EQ01 EQ02 EQ03")
```

A few comments are in order. First, the order in the 'spec_list' matters in that the resulting outputs will follow this order. Second, in the above example, the 'spec_list' argument contains the original equation, but this is not necessary; it would be included by default as a first specification. Third, the add-in allows for use of `"*"` as a wildcard.

3.3 Applications

This section illustrates the full functionality of the add-in by focusing on the first of the 8 applications included in the full version of this document. While the primary goal is to demonstrate the options and outputs of the add-in, the actual goal in some sense is broader: the section also illustrates the interactive and iterative model building process that the add-in is meant to facilitate. The workflow of the model development process appears in Figure 3.3.

Figure 3.3: Model development workflow



The process starts with an initial proposed model, which is then evaluated using the outputs of the add-in. The initial model is best evaluated by looking at the forecast summary graphs, which allow the model developer to visualize the model forecasts and compare them with the actual values, and hence obtain some qualitative notion of forecast quality and errors: when only a single model is being evaluated, their quantitative measures of forecast precision are uninformative in most cases, since their plausible range and lower bound are typically not known.

Next, based on this evaluation of forecast quality, one can propose alternative models. Once one considers multiple model specifications, forecasting performance is more easily evaluated by using the forecast precision metric, especially if the number of models is large, when the embedded color coding plays an important role. That said, forecast summary graphs can still provide a lot of value. Finally, when potentially large numbers of forecasting models are reduced to several final candidates, one can switch focus to detailed outputs from the add-in, such as the exact profiles of sub-sample or scenario forecasts.

Plan. While this section includes only the basic application, in the full version of this document, there are 8 sub-sections, one corresponding to each application. I briefly discuss the remaining applications here. Each application is meant to demonstrate a specific functionality of the add-in, and to show how it can be leveraged as part of the iterative model building process outlined above. The first application focuses on standard trending macroeconomic variable and shows all the basic features of the add-in. The second application highlights the ability of the add-in to use different transformations of forecast variable, focusing on the growth rate transformation in the context of a variable for which growth rate transformation is commonly of interest. The section also shows the functionality of the add-in with respect to automatic model selection. The third application focuses on a different transformation - the spread between two variables. This transformation is often valuable when one variable is the source of the majority of variations in another variable, such as in example of risk-free and risky interest rates. When the latter is of interest, the movements in the risk-free interest rate obscures the evaluation of models for risky interest rate, when level transformation is used. Additional transformations are covered in the fourth application, which focuses on an exponentially growing variable, where logarithmic and/or ratio transformations are of particular value.

Table 3.2: List of applications

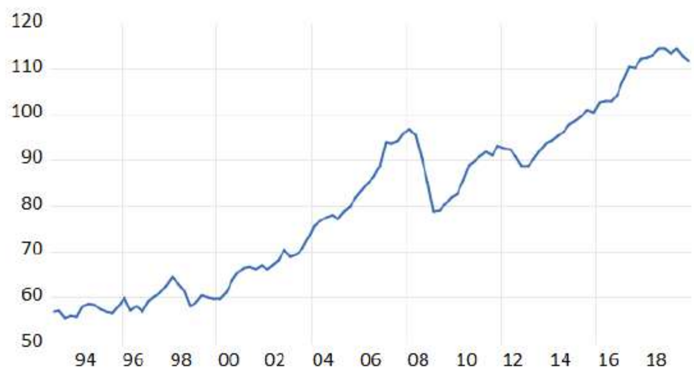
#	Primary focus	Secondary focus
1	Basic use of add-in and overview of key output objects	Iterative and interactive model development process
2	Basic use of transformations (growth)	Recursive automatic model selection
3	Advanced use of transformation (spread)	Interactive model development
4	Advanced use of transformation (log and ratio)	-
5	Unconditional forecasts I - Exogenously produced forecasts	Use for identities
6	Unconditional forecasts II - Systems of multiple individual equations	-
7	Custom re-estimation	-
8	Using intermediate objects	-

Applications 5, 6, and 7 focus on how to perform different types of forecasting exercises using the add-in. Applications 1-4 focus solely on conditional forecasting; these are complemented by applications 5 and 6, which focus on (semi-)unconditional forecasting. First, application 5 considers unconditional forecasting when forecasts for independent variables are supplied as inputs to the process. This is relevant when there is no feedback between the forecasts for the dependent and independent variables. Next, application 6 focuses on evaluating a model consisting of multiple single-equation multivariate models, with *mutual* feedback between individual endogenous variables. Finally, while the add-in allows the user to customize the forecasting process to a great extent, there are still use cases in which the user needs to perform more complex re-estimation which cannot be dealt with within the re-estimation subroutine of the add-in itself; an illustration of custom re-estimation appears in application 7. The last application shows how a user can store intermediate objects from the execution of the add-in and perform his/her own analysis on those.

3.3.1 The basic application

The section provides illustration of the most basic use of the add-in in context of an application, which will also serve to illustrate the model development process for which the add-in is meant to be used. The intermediate goal of the application will be to develop time series model for industrial production for Czechia. The series is displayed Figure 3.4.

Figure 3.4: Level of Czechia industrial production



Univariate models. As a preliminary step, consider using the Eviews automatic ARMA model selection to choose the best ARIMA model. The resulting model when one uses Schwarz Information Criterion (SIC) is ARIMA(0,1,2).[1] One

[1]Figure 3.4 makes it clear that the series is both trending and that it does not have a determin-

can apply the SpecEval add-in to the resulting equation by simply issuing the following command :

```
eq_ip_arma.speceval(noprompt)
```

This will create and display spool that includes 4 objects. Rather than discussing these, I estimate additional 2 ARIMA models and include them among the specifications: a simple ARIMA(1,1,0) model and ARIMA model selected by Akaike Information Criterion (AIC), ARIMA(4,1,4). Once those equations are estimated, I run the SpecEval add-in for all three models:

```
eq_ip_arma.speceval(spec_list="eq_ip_arma*")
```

As a result, the program will create and display a spool that, among other things, includes a table with forecast precision metrics for all three specifications; see Table 3.3.[2] The table shows that the more complex model selected by AIC has substantially worse forecasting performance at both selected horizons, but there is little difference between the two more parsimonious models. The spool also includes the regression results for each of the equations, so that one can quickly check and compare them. For example, in the present case this would reveal that the ordering of the models by RMSE is the exact opposite that one would conclude from looking at adjusted R-squared values of the three models.

In addition to the performance metrics tables, the spool also includes forecast

istic trend. At the same time, the nature of series suggests that it should be growing exponentially over long horizons, though this is not exactly clear from the figure. Correspondingly, the natural way to model the series is in log-differences. I specify logarithmic transformation and force at maximum 1 differencing as the settings in the automatic ARMA model selection.

[2]The document 'Outputs for specification evaluation add-in' includes a detailed discussion of the values and the colors. Briefly, the numbers are average RMSE for all available backtest forecasts, while colors go from green, corresponding to lowest values, to red, signaling the highest.

Table 3.3: Forecast precision metrics - RMSE - ARMA models for industrial production

	Forecast horizons		
Specification	8	24	Avg.
1	7.52	12.9	10.2
2	7.41	12.6	10.0
3	8.45	14.8	11.6

summary graphs. These are displayed in Figure 3.5. The figure makes it clear that neither model provides good forecasts during stress periods, and specifically during the Great Recession. This is the raison d'entre for the multivariate models I consider next.

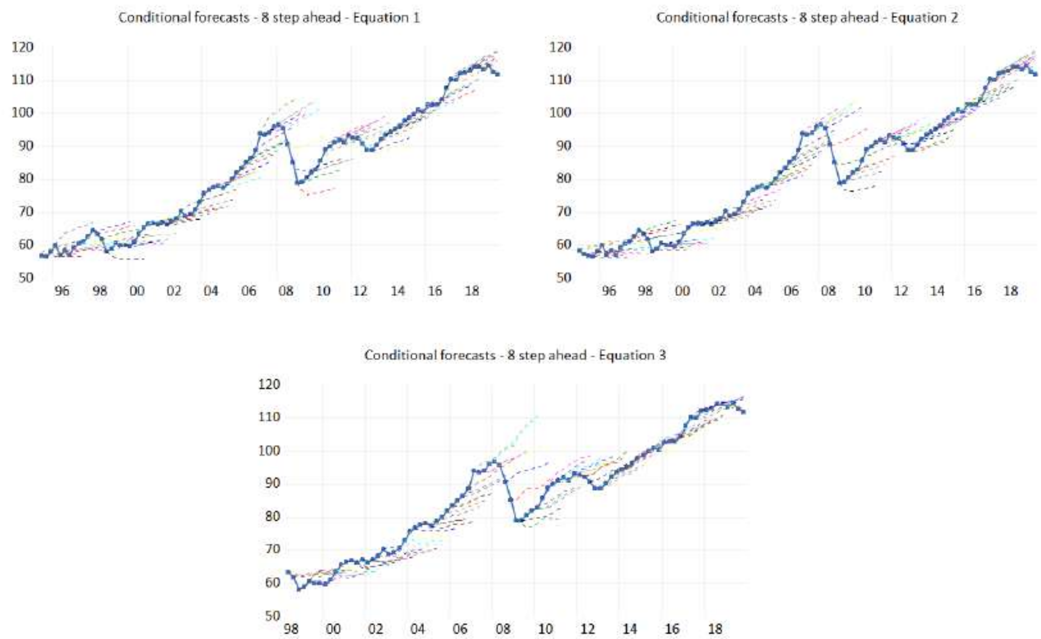
Multivariate models. If one considers using other variables for forecasting industrial production, it is natural to start with GDP: because industrial production is a part of GDP, GDP should provide a lot of information relevant for forecasting industrial production. Consider starting with simple static regression linking the log-difference of industrial production to the log-difference of GDP, as in equation (3.1):[3]

$$dlog(IP_t) = \beta_0 + \beta_1 dlog(GDP_t) \tag{3.1}$$

[3]A few warning comments: The focus here is on conditional forecasting, as opposed to true forecasting. This means that, in contrast to standard VAR models, the models below include concurrent GDP as an explanatory variable, rather than the lag of GDP. While this might seem strange from the perspective of current academic econometrics, the alternative would lead to very poor forecasts from a conditional perspective; and it would also have very little value in scenario forecasting. For example, it would lead to forecasts where industrial production would change only with a delay of at least one lag in response to changes in GDP.

Similarly, no consideration is given to the econometric validity and/or causality of the coefficients estimates. The sole focus here is on forecasting performance, from which perspective causal interpretation of coefficients is not important; simply, coefficients capture co-movement patterns.

Figure 3.5: Forecast summary graph - ARMA model for industrial production



I can evaluate the model on its own, or include one of the ARMA models as a benchmark:

```
eq_ip_static.speceval(noprompt)
```

```
eq_ip_static.speceval(spec_list="eq_arma", use_names="t",
  graph_add_backtest="gdp[r]")
```

In the second case, I specify that I want to use equation names in the output objects, as opposed to aliases (numbers), since this allows me to easily figure out which specification is which. I have also included the GDP series in the backtest graphs, assigning it to the right axis. The resulting forecast summary graphs are in Figure 3.6. The figure shows that while the static equation does a better job forecasting industrial production during the Great Recession than the univariate models, it does not improve the forecasts much. Moreover, Table 3.4 shows that the RMSE of the forecasts is actually higher than for the benchmark model. However, from Figure 3.6, one can see that this might be an artifact of the large errors in the beginning of the sample, where the coefficient estimates are based on only a few observations. To explore this issue further, one could do several different things:

- Include coefficient stability in the execution list of the program, executed by the following command:

```
eq_ip_static.speceval(exec_list="normal stability")
```

- Cut the backtesting sample to start in 2000q1 for both specifications, executed by the following command:

```
eq_ip_static.speceval(spec_list="eq_arma", use_names="t",
   graph_add_backtest="gdp[r]", tfirst_test="2000q1")
```

- Perform the backtesting for both specifications in-sample rather than out-of-sample, executed by the following command:

```
eq_ip_static.speceval(spec_list="eq_arma", use_names="t",
   graph_add_backtest="gdp[r]", oos="f")
```

Figure 3.7 shows the recursive coefficients for the static equation, and demonstrates that, indeed, the coefficients in the beginning of the sample are nonsensical, since the coefficient on GDP is negative. Presumably, such a model would not be

Figure 3.6: Forecast summary graph - static regression for industrial production

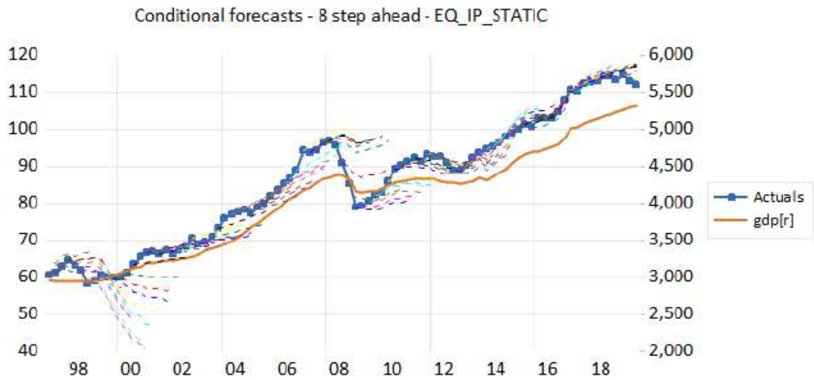


Table 3.4: RMSE - static equation for industrial production

	Forecast horizons (# of steps ahead)		
Specification	8	24	Avg.
EQ_IP_STATIC	8.22	22.2	15.2
EQ_IP_ARMA	7.52	12.9	10.2

Figure 3.7: Coefficient stability graph - static equation for industrial production

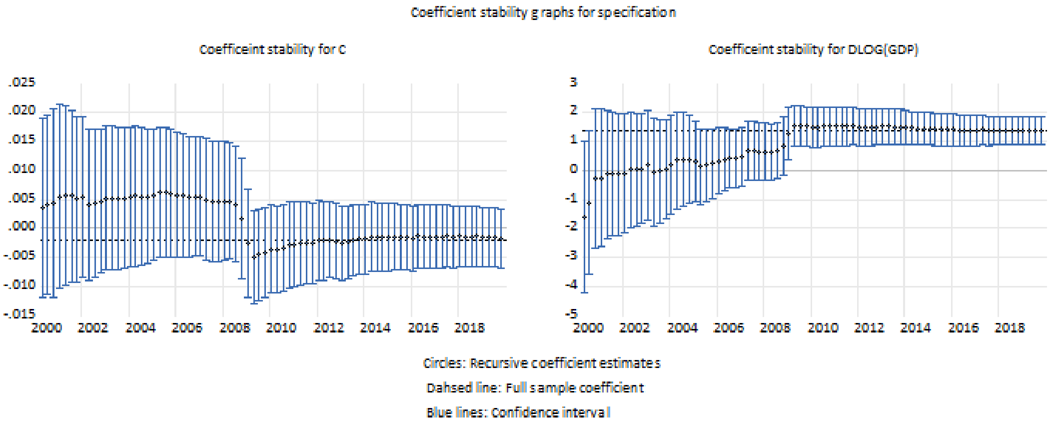
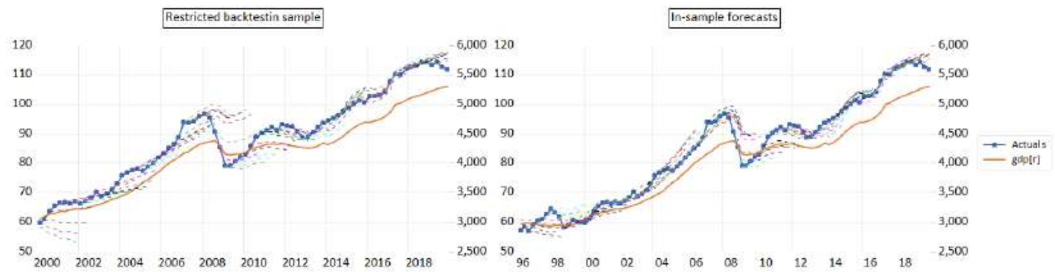


Figure 3.8: Forecast summary graph - static regression for industrial production (Adjusted)



used and these forecast should not be included in evaluation of the forecasting model. Meanwhile, Figure 3.8 shows the forecast summary graphs when the initial periods are excluded (left panel), or when the forecasts are created in-sample (right panel). In either case, the conclusions about the quality of the forecasts are improved. This is also confirmed in Tables 3.5 and 3.6, which show that when only forecasts from 2000q1 onward, or in-sample forecasts are considered, the static equation produces better forecasts than the benchmark ARMA model.

Before proceeding to including additional variables, it is interesting to consider

Table 3.5: RMSE - static regression for industrial production (in-sample)

Specification	8	24	Avg.
	Forecast horizons (# of steps ahead)		
EQ_IP_STATIC	4.36	5.55	4.96
EQ_IP_ARMA	6.70	9.25	7.98

Table 3.6: RMSE - ARMA models for industrial production (restricted sample)

Specification	8	24	Avg.
	Forecast horizons (# of steps ahead)		
EQ_IP_STATIC	6.79	14.5	10.6
EQ_IP_ARMA	8.18	13.8	11.0

a modification of the model in equation (3.1): a modeler might consider dropping the constant from the model, since the constant allows the two series to grow at different growth rates. The `SpecEval` add-in allows the modeler to quickly assess whether such a model change is a good idea. Since the answer might differ substantially with forecast horizons, it is sensible to include more horizons than the two considered above. Specifically, by issuing the following command, I include 6 different forecast horizons, while creating graphs for 3 of them:

```
eq_ip_static.speceval(spec_list="eq_ip_static*",
  horizons_forecast="1 2 4 8 16 40 80", horizons_graph="4
  8 40", alias="with without")
```

Table 3.7 shows the RMSEs for different horizons for the two specifications. Note that the command above also included the aliases for the two specifications, so that the user can easily distinguish them in the output objects. The table shows that the specification with a constant is slightly worse at forecasting at short and medium horizons, but better at very long horizons. In either case, the differences do not seem large.

Table 3.7: RMSE - static regression for industrial production (effect of constant)

Specification	Forecast horizons (# of steps ahead) 1	2	4	8	16	40	80	Avg.
with	1.56	2.77	4.84	8.22	13.7	37.7	103	24.6
without	1.51	2.59	4.38	7.44	13.5	39.9	104	24.7

The forecast summary graphs make it clear that, while including GDP does help with forecasts during recessionary periods - especially when one considers in-sample forecasts, which are more relevant for understanding the behavior of forecasts in future recessionary periods - the model is far from good enough. Moreover, the model tends to miss not only during a recession, but also during a recovery. This suggests that the forecast quality could be systematically improved by including additional information.

Figure 3.7 suggest one possible way to adjust the model. It is clear that there is a break in coefficient estimates around the Great Recession. One option to address this would be to use model with a break. An alternative that is probably more sensible is to allow the relationship between industrial production and GDP to be different during recessionary periods.[4] Consider the following equation

$$dlog(IP_t) = \beta_0 + \beta_1 dlog(GDP_t) + \beta_2 D_t^{recession} dlog(GDP_t) \tag{3.2}$$

where $D_t^{recession}$ is a recession dummy indicator. I compare this specification with equation (3.1). Since the motivation for the modification is the performance during recessionary periods, it is useful to obtain detailed information about those.

[4]The reason this is more sensible is because one does not discard the observations before the break.

I focus on two such periods, 2008q3-2009q4, and 2011q3-2012q4, and specify them as the two subsamples via the following command:

```
eq_ip_static.speceval(spec_list="eq_ip_static_dummy",
  subsamples="2008q3-2009q4,2011q3-2013q2",
  horizons_forecast="1 2 4 8",oos="f", alias="normal dummy")
```

Specifying sub-samples has two consequences. First, it creates additional tables with forecast performance metrics for given sub-samples; see tables 3.8 and 3.9. Second, it creates graphs of a single forecast that starts at beginning of the sub-sample; see Figure 3.9, the top panels show the forecast from the basic static regression, while the bottom panels show the forecast from the static regression with the recession dummy. The table and the figure make it clear that including the dummy in the regression improves the forecasting performance during the two recessionary periods. The figure suggests that this is due to making industrial production more sensitive to movements in GDP, which can be confirmed by checking the coefficients in the regression outputs (not reported): the inclusion of the interaction with the regime dummy leads to a lower coefficient on the stand-alone GDP term, but to higher combined coefficient. Investigating the table with overall RMSE (not reported) shows that this is not at the cost of worse forecasting performance overall, since RMSE for the whole sample is lower.

Table 3.8: RMSE - static regression for industrial production (2008q3-2009q4)

	Forecast horizons (# of steps ahead)				
Specification	1	2	4	8	Avg.
normal	2.65	4.39	5.40	NA	4.15
dummy	2.36	3.33	3.66	NA	3.12

While allowing industrial production to be more sensitive to GDP movements

Table 3.9: RMSE - static regression for industrial production (2011q3-2013q2)

Specification	Forecast horizons (# of steps ahead) 1	2	4	8	Avg.
normal	1.14	1.61	1.62	1.28	1.41
dummy	1.10	1.50	1.65	0.13	1.09

Figure 3.9: Sub-sample forecasts graph - static regression for industrial production

during recessionary periods improves the forecasting performance during these periods, the forecast during the Great Recession is still not satisfactory. Another approach is to consider including GDP components in addition to GDP as such. Different recessions can be characterized by different composition of declines in GDP, and some declines can be associated with larger or smaller declines in industrial production. A prime example are exports: Czechia relies heavily on exports of industrial goods, hence industrial production should be sensitive to foreign demand. As such, including exports could improve forecasting performance, especially in periods of large movements in global demand for industrial goods, such as the Great Recession. I therefore estimate following equation

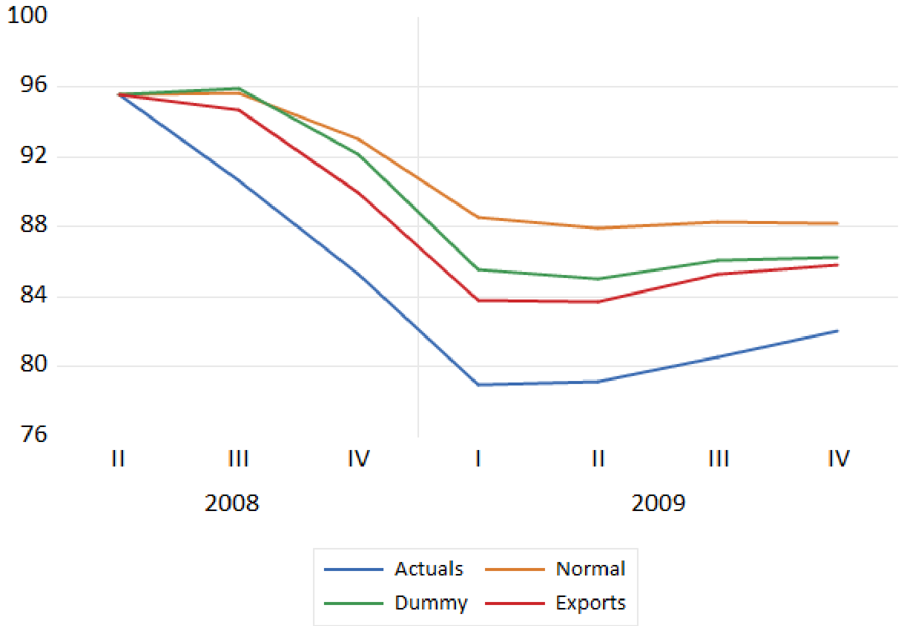
$$dlog(IP_t) = \beta_0 + \beta_1 dlog(GDP_t) + \beta_2 D_t^{recession} dlog(GDP_t) + \beta_3 dlog(Exports_t) \tag{3.3}$$

and then call the `SpecEval` as follows:

```
eq_ip_static.speceval(spec_list="eq_ip_static_dummy
  eq_ip_static_exports", subsamples="2008q3-2009q4,2011q3-2013q2",
  horizons_forecast="1 2 4 8", oos="f", alias="normal dummy
  exports", keep_forecasts="t")
```

Note that this time I have specified that I want to keep the forecasts in the workfile so that I can investigate them together, rather than one specification at a time as in Figure 3.9. I am specifically interested in the series *IP_F2008Q3_NORMAL*, *IP_F2008Q3_DUMMY* and *IP_F2008Q3_EXPORTS*, which contain forecasts starting in 2008q3 for each of the three specifications. The forecasts, together with actuals, are in Figure 3.10. It shows that including exports improves the forecast,

Figure 3.10: Sub-sample forecasts graph - static regression for industrial production (multiple specifications)



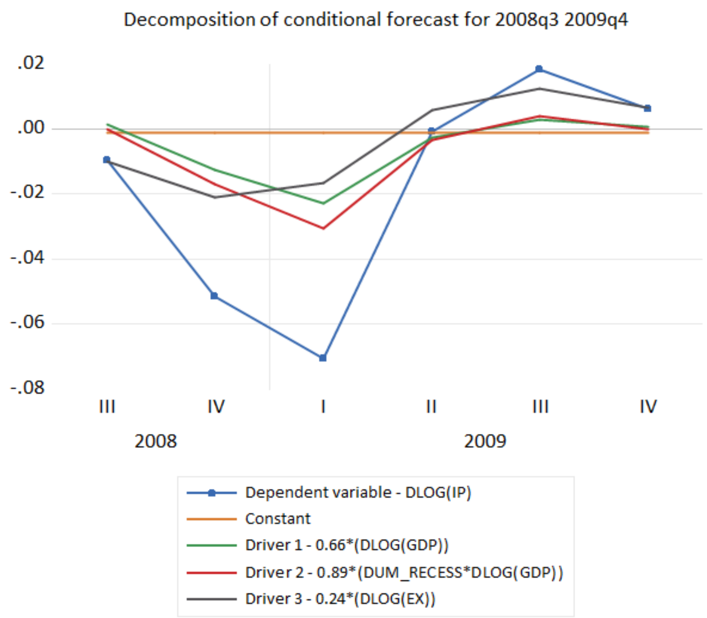
but not dramatically. This suggests that while exports might be useful for forecasting industrial production - indeed full sample and sub-sample RMSE again decreases - the decline during the Great Recession is above and beyond what one would conclude by looking at GDP and exports.

The model with exports already includes three regressors, making understanding forecasts a potentially complicated task. The `SpecEval` add-in contains functionality that helps with this task, the forecast decomposition tool. To perform forecast decomposition, just add it to the execution list as follows:

```
eq_ip_exports.speceval(exec_list="normal decomposition",
  subsamples="2008q3-2009q4,2011q3-2013q2", oos="t")
```

The spool will now include a forecast decomposition graph for each sub-sample,

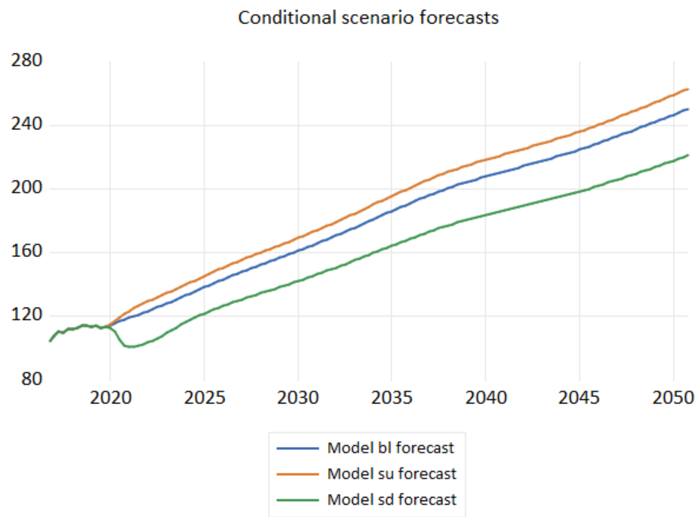
Figure 3.11: Sub-sample forecasts graph - static regression for industrial production (multiple specifications)



as in Figure 3.11. The figure shows how each of the regressors contributes to the overall forecasts in terms of the dependent variable in each of the forecast periods. Among other things, this graph can be used to figure out the source of a problematic forecast, or to understand un-intuitive forecasts. Here, it shows how exports contribute to the overall decline in 2008q4 more than the other regressors, but less in 2009q1.

So far I have focused solely on overall forecasting performance, or on forecasting performance during particular historical period(s). In either case, this amounts to analyzing the specifications in terms of their backtesting performance. The other focus of the `SpecEval` is a scenario forecast performance. While more detailed illustration is performed as part of the other applications, I introduce some basic

Figure 3.12: Scenario forecasts graph - static regression for industrial production



aspects of it here, to highlight how scenario forecasting performance can be also useful as source of information when overall forecasting performance is the focus of model building. The following command creates forecasts for 3 scenarios - baseline, SU (upside scenario) and SD (downside scenario) - based on equation (3.2), and includes a graph that contains all scenario forecasts together in the resulting spool; see Figure 3.12.

```
eq_ip_dummy.speceval(scenarios="bl su sd")
```

While containing the basic information, Figure 3.12 has some important drawbacks. First, the sample before the start of the scenarios is too short to judge the magnitude of the movements, while the sample after the start of scenarios is too long. Second, it is hard to judge movements in industrial production without knowing the size of movements in GDP. One can address these issues by adjusting the starting period of the graph, the ending period of scenario forecasts, and including GDP as a comparison variable, and then specifying that one wants to produce individual scenario graphs rather than all scenario graphs. The resulting command is:

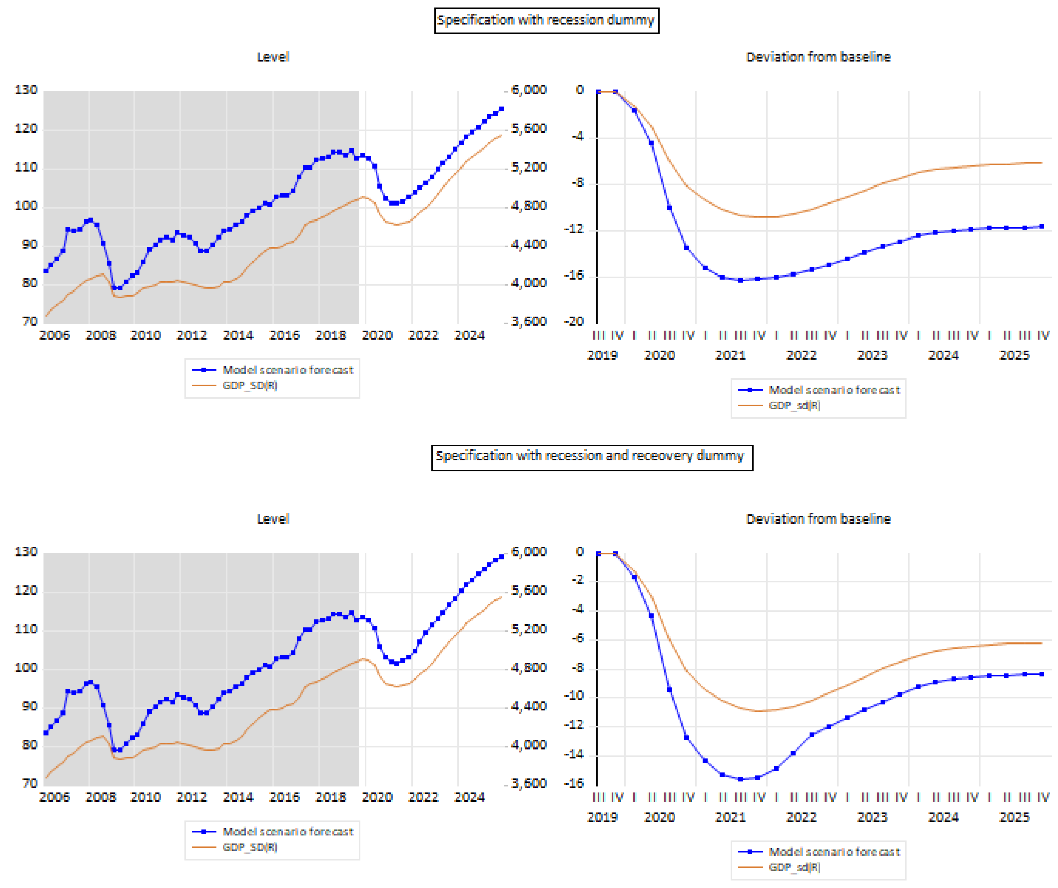
```
eq_ip_dummy.speceval(exec_list="normal
  scenarios_individual", scenarios="bl su sd",
  tfirst_sgraph="2006q1", tlast_scenarios="2025q4",
  graph_add_scenarios="gdp[r]")
```

Now the spool will include a graph for each scenario individually, in additional to a single graph with all scenarios. For example, the top left panel of Figure 3.13 shows the individual scenario forecast graph for the downside scenario. Among other things, the figure suggests that while industrial production does fall more than GDP in the beginning of the scenario, as desired based on previous analysis, the recovery seems to be too weak compared to GDP. This can be best seen if one uses a scenario transformation graph, while using deviations from the baseline as the transformation, as the following command does:

```
eq_ip_dummy.speceval(exec_list="normal
  scenarios_individual", scenarios="bl su sd",
  tfirst_sgraph="2006q1", tlast_scenarios="2025q4",
  graph_add_scenarios="gdp[r]", trans="deviation")
```

The desired chart is shown in the top right panel of Figure 3.13. The model results in industrial production falling permanently and substantially behind GDP, which is likely not reasonable. Instead, one would expect both the drop and rebound in industrial production to be larger, so that the permanent effect on industrial production is only slightly larger than for GDP. The reason the model

Figure 3.13: Scenario forecasts graph - static regression for industrial production



fails to make such a forecast is that it makes industrial production more sensitive movements in GDP only during recessions, not during recoveries. One simple way to address this is to replace the dummy indicating recession with a dummy that is equal to 1 for 4 quarters after the end of recessions. The downside scenario forecasts for the resulting specification are in the bottom panels of Figure 3.13. The graphs make it clear that adjusting the model in such a way addresses the problems raised above. Moreover, Table 3.10 shows that this modification motivated by improving scenario forecasts also leads to improvement in overall forecasting performance. This constitutes an example of how analysis scenario forecasting can be useful in improving overall forecasting performance, irrespective of whether scenario forecasting is of importance.

Table 3.10: RMSE - static regression for industrial production (different dummy variables)

	Forecast horizons (# of steps ahead)		
Specification	8	24	Avg.
short_dummy	3.94	4.46	4.20
long_dummy	3.91	4.04	3.97

3.4 Concluding remarks

This chapter presented excerpts from a document that describes how the `SpecEval` add-in can be leveraged in developing time series models used for forecasting. It offers a basic demonstration of both the variate of outputs with a focus on graphical representation of forecasts and illustrates the flexibility and comprehensiveness of the add-in. The full version of the document available on my personal website offers an additional 7 applications which show additional functions of the add-in. The previous chapter offers illustrations of how the add-in can be used in forecasting exercises which are more common in academic literature.

Chapter 4

Responding to the Inattentiveness of Others: Experimental Evidence from a Cooperative Environment

Co-authored by Jelena Plazonja and Suren Vardanyan (both CERGE-EI).

4.1 Introduction

Consider a work environment where employee is tasked to produce a report which is then passed to a manager, who reads the report and makes a decision, based on which both manager and employee are evaluated. An example is a trading company, where bonuses for both research staff and traders are linked to the profits the trader generates; or a company considering launching a new product line, where future promotions of employee and manager are linked to success of the product line. In such an environment, the success of both actors depends on two factors: (1) how much information the employee collects when producing the report, and (2) how well the manager absorbs the information in the report. Specifically, the quality of information supporting a decision is a joint product of the quality of collected information, and how well the information is absorbed by the manager. In related theoretical work (Plazonja 2018), it is shown that, if both agents find it costly to collect/absorb information due to limited attention, then the attention effort of each agent is a strategic complement for the other agent. In other words, each agent should increase his attention when the other player increases his.

This chapter presents an experiment designed to test this prediction. The experiment has following structure. We first assign participants to one of two roles, Sender and Receiver. We then successively pair each Sender with each Receiver and let them play the following game: the the Sender collects information about a randomly selected state of the world and then he communicates it to the Receiver; the the Receiver then pays attention to the information communicated by the the Sender before taking an action that influences the payoff of both players. The payoff depends on the difference between the state of the world and the Receiver's action.

To capture the idea that collecting/absorbing information is costly, we do not allow either player to observe information directly. Instead, we present both players with a task that requires them to exert attention effort. We follow other authors (Caplin, Dean, and Martin 2011; Caplin and Martin 2014; Jin, Luca, and Martin 2015), and show the participants 20 integers that add up to either the state of the world (for the Sender) or to the report provided by the the Sender (for the Receiver). Note that since the the Receiver observes a noisy version of the Sender's signal (rather than her own signal about the state), neither player can shirk on their effort in the hope of forcing the other player to compensate for his or her lesser effort. This, together with the symmetry of the payoff function, means that the game does not feature any non-cooperative component: the players agree on the desired action given the information set, and on the value of the information.

To study how subjects adjust their attention, and specifically whether they understand that their attention effort is a strategic complement, we follow the literature that studies how strategic play is affected by information about opponent's characteristics, and give both players information about how good their partner is in a given attention task. Before starting the game outlined above, we ask participants to complete an individual task that consists of 5 identical rounds in which participants complete the attention task. We then calculate the average absolute mistake (AAM) made in this individual task, and communicate the partner's AAM to both players before each round of the actual game. This allows the participants to estimate the attention costs of their partner, and hence to estimate how noisy his/her signals are likely to be.

Since we have observations on multiple plays by each subject, we can study whether subjects vary their attention in response to changes in the partner's AAM as predicted by theory. As we cannot directly observe attention effort, we use the size of mistakes as a proxy variable for attention effort, and study whether subjects make greater mistakes when they play with a partner with higher attention costs. We present results from several empirical models that all support this hypothesis, with regression coefficients being both statistically and economically significant. For example, in our main model, the average mistake made increases by 10 when the partner's AAM increases from the first to the third quartile. This is equivalent to 83% of the quartile range of mistakes observed. The effect is also significant in terms of potential earnings, leading to decreases in a subject's potential earnings of 120CZK, 30% of maximum earnings. Moreover, the average estimated effect potentially masks important heterogeneity: the effect increases substantially with the quantile considered, suggesting that the reaction is stronger when other factors make paying attention harder.

In addition to studying variations in attention, we are also interested in whether subjects reflect on (soft) information about the likely precision of their information when they take actions. Standard theory of decision-making under uncertainty suggests that actions should be closer to a prior mean when information is less precise. Even though the structure of our data does not allow us to perform a precise empirical analysis, several empirical models suggest that subjects do indeed behave according to this prediction: actions are more likely to be closer to a prior mean when the partner's attention costs are higher.

This chapter is primarily related to empirical and experimental literature on rational inattention. Existing experimental and empirical literature has focused on two questions, corresponding to two fundamental principles underlying the rational theory of attention. First, different studies have demonstrated that agents pay imperfect attention to available information, which demonstrates that attention is a scarce resource (e.g Lacetera, Pope, and Sydnor (2012)). The fact that agents pay imperfect attention to available information implies that agents' choices include mistakes even though they have available all the information they need. Second, several mostly experimental studies have shown that agents vary their attention according to the nature of the decision problem. The main experimental example is Dean and Neligh (2017), who test predictions of general and specific rational inattention models. Their main finding is that experimental subjects change their attention in systematic ways in reaction to changes in the environment, giving strong support to the theoretical notion that attention is a choice variable. Among other things, they show that subjects do adjust their attention in reaction to changes in incentives such as increasing the payoff from a correct choice or adding potentially high-payoff action. Similarly, Bartos et al. (2016) use a field experiment to show that agents adjust what they pay attention to depending on whether their choices are more or less selective: in a market where most applicants are accepted (e.g., the housing market) decision makers pay relatively more attention to negatively stereotyped candidates (i.e., minorities), while in markets where only few applicants are accepted (e.g., the labour market) decision makers pay relatively less attention to negatively stereotyped candidates. Other experimental papers are Caplin and Dean (2015),Cheremukhin, Popova, and Tutino (2015),Ambuehl, Ockenfels, and Stewart (2019) and Martin (2017a).

In summary, previous studies have shown that agents do pay imperfect attention and that they vary their attention systematically and rationally. To reach these conclusions, the experimental literature has been varying the environment a decision maker faces. Our experiment is focused on the next logical step in the empirics of rational inattention: we study whether agents react to the inattentiveness of their partners, and specifically whether and how subjects vary their attention in response to changes in the attention costs of their partner. Thus, this chapter differs from the above papers in that we focus on an interactive environment, rather than on individual decision tasks. This also makes this chapter methodologically different , because instead of exogenously changing the environment a decision maker faces, the experiment keeps the environment constant and studies whether subjects change their attention in response to changes in their partners' characteristics.

Apart from experimental and empirical literature on rational inattention, this chapter is also related to experiments studying the role of information about opponents on subjects behaviour. This literature is related to this chapter in two ways. First, these papers also focus on how subjects behave differently as a function of differing characteristics of their opponents. The main example is Gill and Prowse (2016), who show that subjects with high cognitive ability respond to the cognitive ability of their opponents. Similarly, Palacios-Huerta and Volij (2009) report that subjects are more likely to play the Nash equilibrium strategy when they play against players who are likely to be sophisticated in backward induction.[1] Meanwhile, Agranov et al. (2012),Le Coq and Sturluson (2012), and Slonim (2005) show that subjects respond to (manipulated) beliefs about the experience of their opponents. Note that due to focus on the cognitive ability and sophistication of opponents none of these papers study cooperative environment, as this chapter does.

These papers are also close to ours in terms of the experimental methodology, since this chapter also uses experiments in which subjects are informed about particular characteristics of their partners. In this respect, the closest study to this chapter is Gill and Prowse (2016), in which subjects are informed whether their partners have above or below median cognitive skills as measured by a Raven's test. The main difference is that we inform subjects about their partners' performance

[1] Specifically, students - especially chess players - are more likely to stop in the first move of a 6-node centipede game when they play against professional chess players, compared to when they play against university students. The authors connect this to the fact that chess players have extensive experience in backward induction.

in the individual version of our game, which we conceptualize in our experimental context as a proxy measure for attention costs. In contrast, Gill and Prowse (2016) inform the subjects about their partners' general characteristics (i.e. cognitive skills), as the effect of those is the main focus of their paper. Of course, our measure of attention costs is likely correlated with the more general concept of cognitive costs; indeed, attention costs can be viewed as a subset of cognitive costs. However, as our experimental task is about paying attention, we prefer to use the terminology of attention costs rather than cognitive costs.[2]

The rest of this chapter is organized as follows. The next section presents the outline of the game played by experimental subjects and details of the experimental design. The third section gives an overview of the collected experimental data. The fourth section presents our hypotheses, and empirical strategy and discussed the results. The last section concludes with a discussion of future research plans.

4.2 Experimental design

There are two roles in our experiment: the Sender (he) and the Receiver (she). The outline of the game is presented in Figure 4.1. The the Sender obtains information about the random state of the world and communicates his information to the

[2]In principle this chapter would be unchanged if we used term "cognitive costs/skills" instead of the term "attention costs". However, the information provided to subjects is not a measure of cognitive skills, rather it measures the ability to perform well in this specific task, and so referring to it as cognitive skills would be misleading. Since our task is characterized by full information, ample time and no strategic considerations, a decision by experimental subjects not to collect information seems best interpreted as a decision to exert low attention due to high attention costs. Similarly, one could claim that subjects are not reacting to the inattentiveness of other subjects, but to their low ability at a given task. We believe that such distinctions are more semantic than substantial, reflecting the fact that, in principle, one cannot distinguish between high attention costs and low cognitive ability. This chapter provides a contribution to existing economic literature irrespective of the terminology used: we are unaware of any paper that is similar outside of the literature on attention experiments.

Receiver. The the Receiver observes a noisy version of *information provided by the Sender* and then takes action.

Figure 4.1: Game outline

$$\boxed{\text{Nature}} \longrightarrow x \longrightarrow \boxed{\text{Sender's attention}} \longrightarrow s_S \longrightarrow \boxed{\text{Receiver's attention}} \longrightarrow s_R \longrightarrow \boxed{\text{Action choice}} \longrightarrow a$$

Both players have identical payoffs which depend on the distance between the Receiver's action and the actual state of the world with quadratic loss function:

$$\Pi = k - (x - a)^2 \tag{4.1}$$

where x is the value of the state and a is the action taken, while k is a constant. Importantly, the nature of information available to players together with the symmetry of the pay-off function mean that the game does not feature a non-cooperative component: the players agree on a desired action given the information set, and on the value of the information. Since the Recceiver observes a noisy version of the Sender's signal, rather than her own signal about the state, neither player can shirk on their effort in hopes of forcing the other player to compensate for his or her effort.

Calibration. The state of the world, referred to as a selected number , is randomly drawn from a set of integers between 300 and 500. Numbers do not have equal probability of being selected; instead, the probability follows (truncated) normal distribution. Both these facts are communicated to the participants in the instructions, with the distribution being communicated verbally in non-technical language and graphically (see complete instructions on my personal webpage).

The pay-off function is chosen and calibrated with the aim to make the pay-off sensitive enough to participants actions, but not too sensitive. The constant is set to 400, which translates into a pay-off of 400 experimental currency units (ECU) if the action exactly matches the state of the world. If the difference between the state and the action is more than 20, the pay-off is zero. Subjects were familiarized with the pay-off function through an example and an illustrative pay-off table specified in terms of the difference between an action and the selected number, to highlight both the quadratic nature of the function and the fact that it is the difference between action and state that is relevant for the pay-off.

The final earnings are equal to the average pay-off from two rounds randomly selected at the end of the experiment, i.e., our conversion rate between ECU and the currency of earnings (CZK) is 2-to-1. To keep the individual rounds as comparable as possible, we did not inform subjects about their payoff after each round, which should minimize competition or discouragement effects. Instead they were told their overall payoff/earnings at the end of the experiment.

The rest of this section describes the experimental design in detail, first discussing the attention task before describing the two games played by participants throughout the experiment: the individual game and the cooperative game.

4.2.1 Attention task

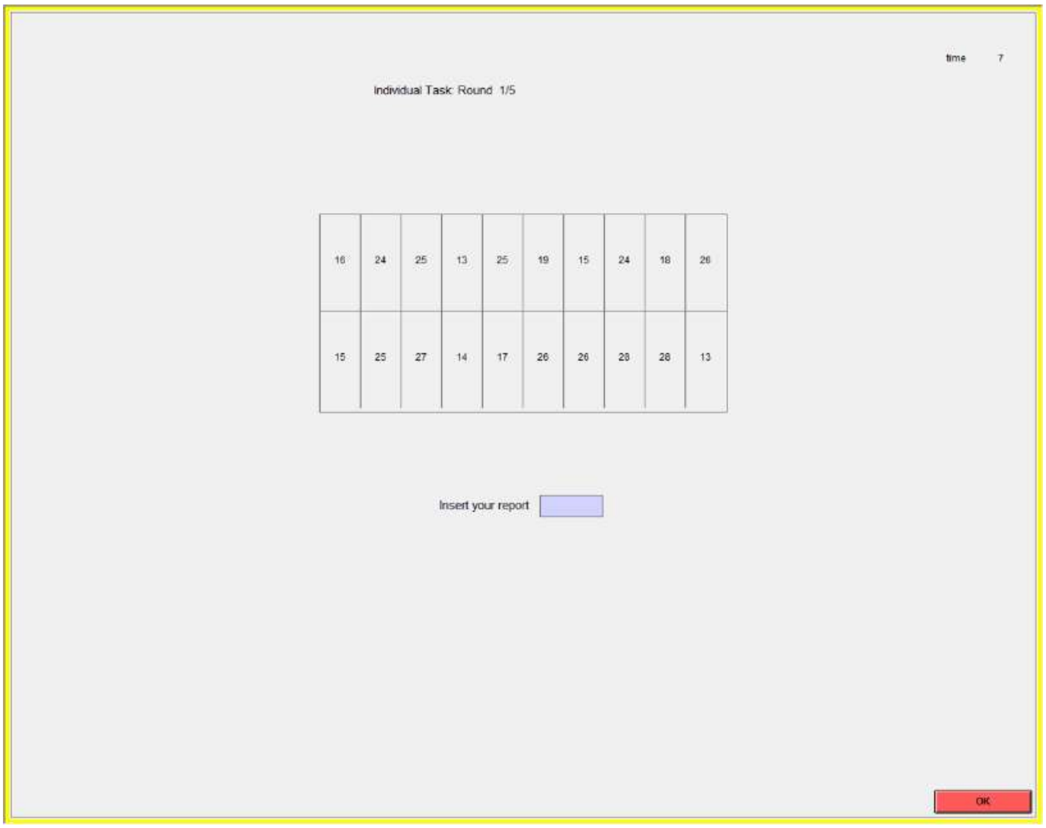
The subjects obtain information about a payoff relevant state. Our choice of experimental task is motivated by a desire to capture two features of rational inattention theory: (1) while all relevant information is available, obtaining information requires attention effort; and (2) agents choose an optimal attention strategy that minimizes attention required, rather than being restricted to a particular (set of) information strategies. We follow other authors (Caplin, Dean, and Martin 2011; Caplin and Martin 2014; Martin 2017a) in selecting the following experimental task.

Subjects are shown a set of 20 integers which are randomly drawn from between 12 and 28, excluding 20, and which add up to the value of the state. The fact that the numbers add exactly to the value of the state means that subjects have all the information about value of the state of the world that they need. It is up to them whether they decide to obtain the exact value of the state, or rather to just quickly make an imprecise estimate. At the same time, the task is hard enough so that they have to exert a substantial amount of effort to obtain a precise estimate of the state of the world.

Importantly, the task leaves subjects freedom to choose their attention strategy, since the way they obtain information about the value of the state is up to them: for example, they can add numbers one by one, or they can add the first digits of all numbers before turning to the second digits. The subjects are not allowed to use calculators or paper and pencil. From previous research we know that subjects make mistakes in this setting and that these mistakes vary systematically with the stakes and difficulty of the task. Figure 4.2 shows a print screen of the screen showed to subjects in a typical round.

The participants do not have unlimited time for the attention task. Specifically, they can inspect the numbers for 75 seconds, during which they need to provide their report about the state of the world. Throughout these 75 seconds, the

Figure 4.2: Printscreen of the attention task



subjects are informed how much time has elapsed. If they submit their report before the 75 second limit, they will not be allowed to proceed to next round before the 75 seconds elapse. If they do not submit within the 75 seconds, they have an additional 15 seconds to submit their report, but during these 15 seconds the 20 numbers are no longer displayed. If they do not submit even within these 15 seconds a random number is selected as the player's report and the experiment proceeds to the next round.

4.2.2 Individual game and role assignment

Before playing the game described above, participants play an individual game. This game consists of 5 identical rounds in which participants are presented with the attention task, including the request to provide their report. Their pay-off depends on the difference between their own report and the state as in (4.1). This means that they play a role that of a the Receiver who collects information herself.

This individual game serves several purposes. Mainly, it allows us to utilize a within-subject experimental design for the cooperative game. Specifically, to study how subjects change their attention when they face a more or less inattentive partner, we need to obtain some signal about the inattentiveness of each participant, which can then be communicated to their partners. We use the average absolute mistake (AAM) made by a given participant during the individual game: AAM measures how imprecise a players signals were during the individual game, and hence should also be indicative of how imprecise the signal will be during the cooperative game (more on this below).

We also use the results of the individual game to assign subjects to their roles in the actual experiment: we order them according to their AAM, and assign them to the roles of the Sender and the Receiver in alternating fashion, ensuring that participants encounter as diverse a group of partners as possible. The roles are fixed for the whole duration of the experiment. While this means that players experience only with one role, the fact that they initially play the individual game means that they can all familiarize themselves with the decision facing the Receiver, who chooses the pay-off relevant action. Finally, playing the individual game also ensures that they experience the attention task before the cooperative game, and so relatively little learning is happening during the cooperative game.

At the end of the individual game, we first ask subjects to estimate their own AAM and the AAM of another random subject. This information allows us to glimpse their degree of confidence in themselves and to identify subjects who are higher overconfident. After obtaining answers to these questions, we inform them of their own AAM and maximum and minimum AAM among the current set of participants.

4.2.3 Cooperative game

After the individual game the subjects proceed to the cooperative game which forms the main part of the experiment and is our main focus.[3] The game proceeds as follows. After the subjects are paired, the the Sender is presented with the attention task described in the previous subsection, with the sum of the numbers equal to the state of the world. Once he is finished, he is asked to communicate his report to the Receiver. However, the the Receiver does not observe this report

[3]Note that the cooperative game is simply referred to in the instructions as Game A or Game B, depending on the treatment, in order not to nudge participants towards particular type of behaviour.

directly. Instead, she is presented with the same attention task, where, crucially, the 20 numbers add up to the Sender's reported number (rather than to the state of the world). The game concludes with the Receiver taking action, which determines both her and Sender's pay-off according to (4.1). The experimental instructions given to participants stress two the key facts: (1) the numbers in the Receiver's attention task add up to the Sender's number rather than to the selected number, implying that the Sender cannot shirk; (2) both players' pay-offs depend on the Recever's action.

During this part of the experiment, the subjects are successively paired with different partners whose identities remain anonymous. To maximize the variation in partners of each participant, we match every Receiver with every Sender. We keep the matching order random so that we do not induce any particular behaviour.

We execute two different treatments of the cooperative game. First, the subjects play an **informed treatment (TI)**. In this treatment, once the agents are paired and before they perform their respective attention tasks, we inform subjects about the characteristics of their partner. We tell them their partner's AAM from the individual game together with the minimum and maximum AAM for all subjects. Subjects are made aware of the fact that their partners know about their characteristics. By varying the partners, and by informing each one of the attention characteristics of his/her partner, this treatment allows us to use a within-experimental design to answer our main question: Do subjects change their attention level in response to the inattentiveness of their partners? Second, participants play the **uninformed treatment (TU)**, in which they are not informed of their partners' characteristics. In this treatment we limit the number of rounds to 5.

4.2.4 Feedback and information about participants

After the cooperative game, the experiment continued with several additional small parts aimed at collecting information about the participants and their feedback. First, the participants were asked several questions relating to the cooperative game. We were interested in their self-assessed behaviour, such as whether a partner's AAM influenced them or whether they reported their exact summation.

Second, participants answered three questions and participated in two small games aimed at capturing other relevant characteristics of our participants. The three questions were the standard questions of the cognitive reflection test (Frederick 2005), which studies how are subjects inclined to override an incorrect "gut" response and engage in further reflection. Next, they played a standard dictator game, which is used to study degrees of altruistic behaviour (Hoffman, McCabe, and Smith 2008). Finally, subjects played a simple 2-move centipede game, which is suitable for studying whether participants are able to use backward induction. Knowing these characteristics allows us to better disentangle participant behaviours during the cooperative game.

At the end of the experiment, subjects were administered questionnaires, filling in their basic demographic information.

4.3 Data

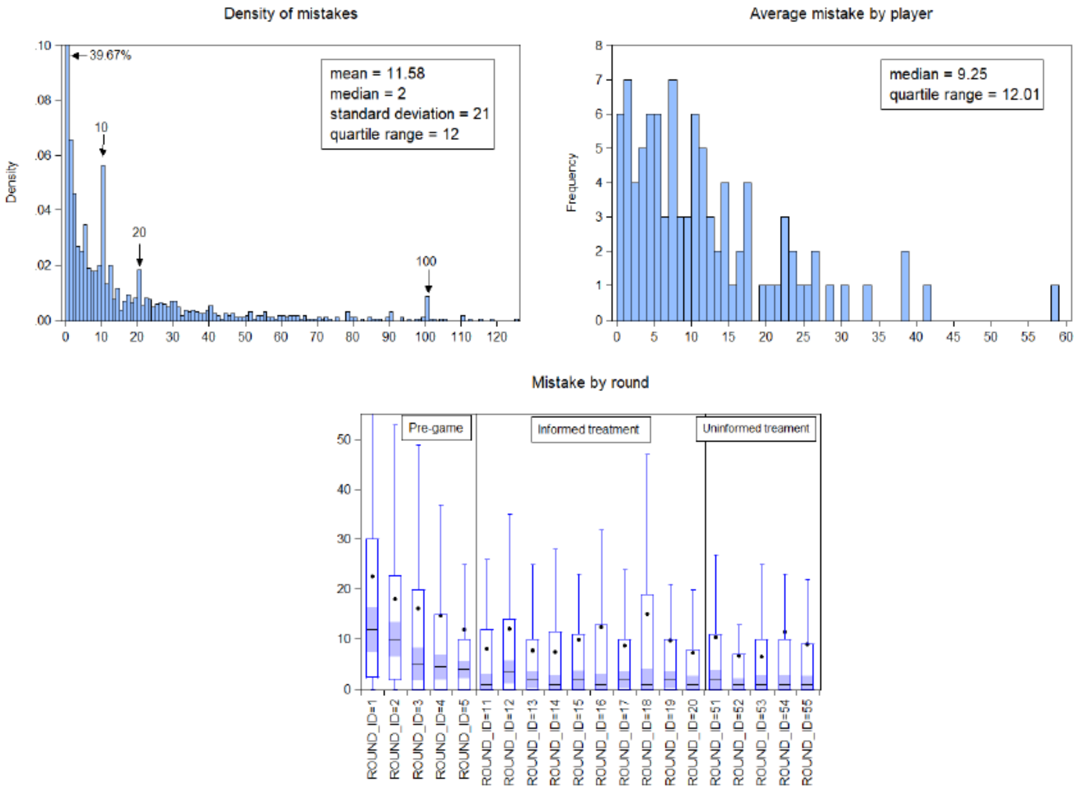
We run 5 sessions of our experiment during March and April 2018, using the Laboratory of Experimental Economics at the University of Economics in Prague. A typical experimental session took 1 hour and 45 minutes, plus 15 minutes to process payments. An average subject earned 327 CZK (14 USD), with standard deviation of 153CZK. This translates into an hourly wage of 163 CZK, above the local hourly average wage for unskilled labor. At minimum, participants received 100 CZK show up fee.

Altogether 96 subjects participated in our experiment, equally split between Senders and Receivers. All participants played the individual game and, with the exception of one session, both the informed and uninformed treatments of the cooperative game; in the other session, participants played a different version of the individual game instead of the uninformed treatment of the main game. This yields observations of 1,794 player-rounds of our attention tasks: 480 for the individual game, 924 for the informed treatment of main game, and 390 for the uninformed treatment of the main game.

As discussed below, our main variable of interest is the *size of* mistake made by a player in given round; therefore we provide a basic overview of this variable. Figure 4.3 shows the distribution of mistakes we observed: the top panel shows a histogram of mistakes in all player-rounds, together with basic summary statistics; the middle panel shows mistakes aggregated at the individual level; the bottom panel shows mistakes aggregated at the round level. A few comments are in order. First, almost 40% of the observations correspond to zero mistake. Second, while the distribution is roughly exponential if we disregard the mass at zero, there are a few points that do not fit the monotonically decreasing pattern: there is extra mass at multiples of 10 (especially 10 and 20, but also to a lesser degree at 30, 40 and 90) and at 100. These points have relatively straightforward explanation: while adding up the 20 numbers, subjects made a mistake in remembering the first digit of the cumulative sum, while correctly remembering the second digit of the cumulative sum.

Third, the distribution has a very long tail; slightly more than 10% of observations corresponds to mistakes larger than 30, and we even observe several player-rounds with mistakes above 100; the maximum mistake (not shown in the figure) is equal to 199. The average mistake made by players varies from 0 all the way to 58.5, with a median of 9.25, and a quartile range from 4.35 to 16.36. This clearly shows that there is sufficient variation between the subjects, allowing us to use within-subject estimation. Finally, looking at the mistake by round, we can see that, while the average/median of mistakes initially decreases with round, there seems to be little learning after the fourth round.

Figure 4.3: Distribution of mistakes



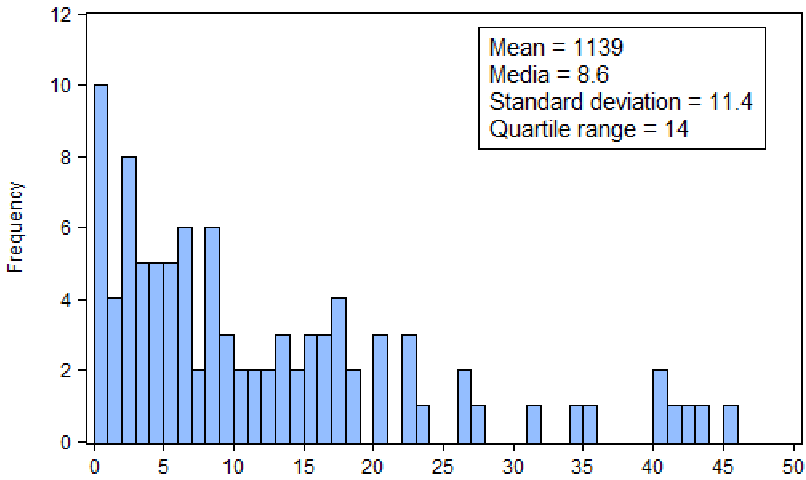
Notes: The bins in the top and middle graphs have a width of 1. The symbols in the bottom graph correspond to mean (black circle), median (black line), confidence interval for median (shaded blue area), 1st and 3rd quartile (box edges), and staples.

Since the AAM of participants in the individual game will play a crucial role in our analysis, we also present information about this quantity. Figure 4.4 shows the histogram of AAMs together with basic statistics. The individual game AAM has a mean of 9.68 and median of 6.3, and varies substantially across participants, with quartile range of 13.3. This large variation is important for us, because our hypotheses is that participants will react to variations in AAM; if the variation were too low, then our experimental design would be flawed.

Finally, in Figure 4.5, we also show the distribution of time elapsed before the participants submitted their report: the left panel shows time in all player-rounds, while the right panel shows median time by player. The aggregate distribution shows that the most typical time spent is close to our time limit. Nevertheless, almost half of reports are submitted at least 10 seconds before the time limit, with an additional quarter submitted in the last 10 seconds. Finally, in a large majority of rounds, players did spend a substantial amount of time on the task: only in 5% of the player-rounds do we observe players spending less than 30 seconds. All this suggests that the participants had a sufficient amount of time to complete the attention task, but that the amount of time was not excessive. In other words, the conditions facing participants when solving our attention task should approximate the conditions of the real world with moderate pressure to economize on attention time.

While the aggregate numbers approximate well the median time spent aggregated by players, it is important to point out that we had several participants who spent very little to no time on the attention task. These players with low median time also make very high mistakes on average: the average mistakes of the three players who spent a median time of less than 30 seconds is 5th, 6th and 11th high-

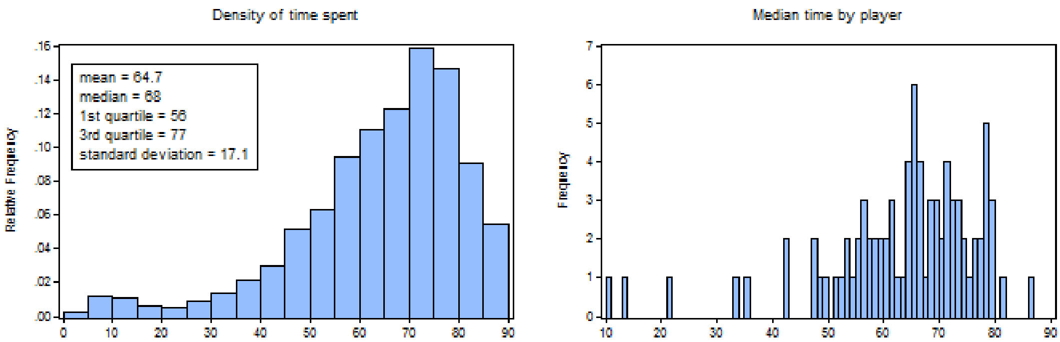
Figure 4.4: Distribution of AAM in the individual game



est. Moreover, among these players we observe clearly non-random sequences of reports: for example, one of these player reported a sequence of 415, 415, 420, 415, 415 during the individual game. This leads us to conclude that these participants are not properly motivated during the experiment.

We also collected demographic information on our participants. We asked about age, gender, nationality, knowledge of English, highest level of completed

Figure 4.5: Distribution of time spent on the attention task



education, study major and whether they had completed a course in statistics. Table 4.1 presents various summary statistics of these characteristics. The typical participant is a male economics major with a bachelor's degree and good knowledge of English, who had previously completed a course in statistics.

Table 4.1: Summary statistics for demographic information.

Characteristic	Mean	Median	Std. dev.	1st quartile	3rd quartile
Age	23.34	22	3.38	21	24
English	8.49	8	1.23	8	10
	category 0	category 1	category 2	category 3	category 4
Gender	54.47%	45.53%	-	-	-
Education	40.07%	45.7%	14.23%	-	-
Major	77.69%	6.15%	1.82%	2.33%	12.01%
Statistics	82.7%	17.3%			

Notes: "English" is self-assessed understanding of English, on a scale 0 (worst) to 10 (best). The coding of categorical variables is: gender (male=0, female=1); education (high school=0,bachelor=1,master=2)' major(economics=0, computer science=1, law=2,art=3, other=4); statistics (yes=0,no=1)

4.4 Hypotheses, empirical strategy and results

Outcomes. We are mainly interested in how subjects vary their attention. Unfortunately, we cannot measure attention directly, and hence we need to use some proxy variable for attention effort. A natural proxy is **the size of the actual mistake** made by the subject in the given round: mistake is (stochastically) related to the amount of information processed, which we refer to as effective attention; effective attention in turn depends on attention effort and attention costs of a subject. Even though a mistake made is a random variable conditional on effective attention, on average the mistake and effective attention (and hence attention effort) should be negatively related: a larger mistake should be a result of a lower attention effort.

We are also interested in how subjects reflect on variations in the expected precision of information they have when taking action: the same signal realizations should lead to different actions if are of different (expected) precision. If the expected precision is higher, according to the standard theory of decisionmaking under uncertainty, the Receiver should take an action closer to the prior mean. Therefore, we analyse whether participants' reports are closer to the prior when the (expected) precision of the signal is lower.

Independent variables. Our main interest is how subjects respond to variations in inattentiveness of their partners, i.e., changes in their partner's attention costs. We proxy attention costs by a subject's AAM in the individual game, which is our key independent variable. Individual game AAM of a subject measures the average mistake in an individualistic environment, i.e., a situation in which the subject did not have to consider the behaviour of other subjects, but was concerned only with obtaining as precise information as possible. In such an environment, the AAM should reflect only the attention costs of a given subject for given task, and hence we treat it as a proxy variable for such attention costs. Of course, AAM is a good proxy for attention costs only for our specific task, but this is not a problem for our empirical analysis. Alle we need for our empirical strategy to be valid is simple regularity of subject beliefs: we need them to believe that higher individual game AAM suggests on average less precise signals during the main game.[4] We view this as a very natural belief, as the opposite implies that subjects

[4]Note that this is indeed true in our data as discussed in the Appendix 4.B in Appendix.

expect reversals in performance between the individual and main games. This also means that our empirical strategy should be valid even in the face of fatigue or learning: we do not require attention costs to be constant throughout the game; we simply require that subjects do not expect other subjects to become fatigued or to learn faster than other subjects. Given that AAM is the only information at their disposal, such expectations seem implausible to us.

In rest of this section, we discuss the hypothesis and results that show how subjects systematically adjust their attention and choices in response to changes in their partner's attention costs. The first subsection describes the hypotheses, empirical strategy and results relating to changes in attention in response to changes in the effective attention of their partners. The next subsection focuses on how subjects' adjust their action choices in response to changes in their and their partner's attention costs. Both sections use within -subject empirical analysis.

4.4.1 Variation in attention

Hypotheses. Our expectations about subjects' reactions to variations in their partner's attention characteristics and effort are based on our related theoretical work (Plazonja 2018). There, we show that in our specific environment[5] Players' attention should vary with their partner's effective attention in an intuitive way: since in a cooperative situation both players' effective attentions are strategic complements, players should decrease their attention effort when their partners' effective attention is lower. The intuition for this follows. From the Sender's perspective, a Receiver that processes the Sender's signal with more noise decreases

[5] The specific environment is characterized by quadratic loss function and Gaussian prior and signals, as well as the entropy loss function.

the benefit the Sender derives from his attention; it is equivalent to the situation when the Sender would take action himself but would make mistakes. With lower benefits from attention and unchanged attention costs, his optimal attention effort decreases. How do Senders know how much noise a particular Receiver will add to his signal? While they cannot know how much attention effort the Receiver will exert, they can form an estimate of how much noise the Receiver will add on average to his signal simply by knowing the Receiver's AAM in the individual game, which is correlated with the underlying attention costs of Receiver. This yields our first hypothesis:

H1S: Senders pay less attention when facing a Receiver with higher attention costs.

From the Receiver's perspective, what matters is the noisiness of the Sender's signal. If the Sender's signal contains more noise, then for a given level of attention the Receiver learns less about the state of the world. Again, with lower benefits from a given level of attention and unchanged attention costs, the optimal reaction is to pay less attention. How do Receivers know how noisy the Sender's signal is? As in the case of Senders, Receivers can form beliefs about the noisiness of the signal provided by the Sender based on the Sender's individual game AAM. This yields the following hypothesis:

H1R: Receivers pay less attention when facing a Sender with higher attention costs.

Of course, we can test the hypothesis H1 for both roles together, yielding following pooled hypothesis:

H1P: Players pay less attention when facing a partner with higher attention costs.

We test these hypotheses using several different empirical models. The underlying assumed data-generating process links the observed absolute mistake of a given player in a given round (AM) to (i) the player's characteristics; (ii) the response to partner's attention costs proxied by his/her individual game AAM; (iii) round effects. The general form of the data-generating process is:

$$AM^x_{i,t} = \alpha_i + \beta^x f(AAM^p_t, AAM_i) + \delta_t + \epsilon^x_{i,t} \tag{4.2}$$

where i identifies an individual subject, t stands for the round, p stands for the partner, and x is a role (S or R, for the Sender and Receiver). Meanwhile, δ_t is a vector of controls for the experiment round. The use of general functions should stress the fact that we will also consider non-linear models.

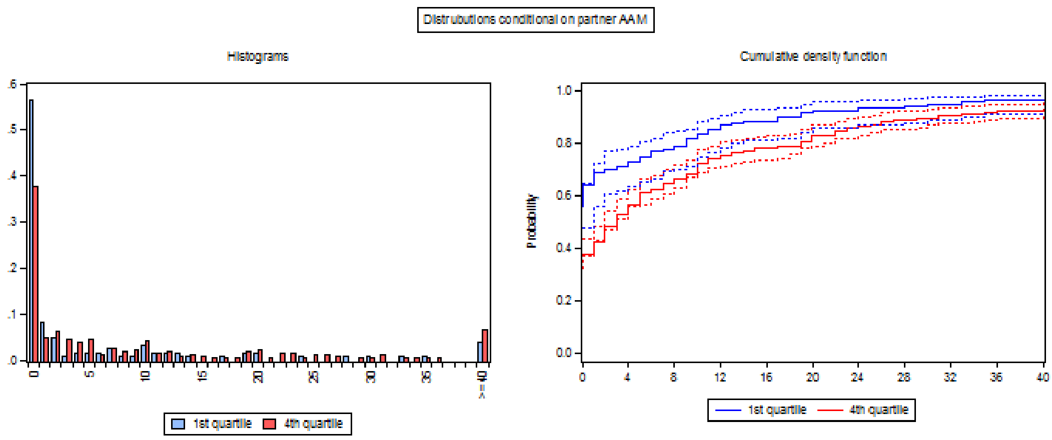
Before proceeding to the regression analysis we provide a basic statistical analysis. We divide the observations of absolute mistakes made by a subject into two groups, one corresponding to rounds where partners' AAM falls into the 1st quartile of all AAMs, and one where it falls in the 4th quartile. Roughly speaking, the two groups correspond to rounds in which players were paired with partners with very low and very high attention costs, respectively. We then compare the two groups of mistakes made, investigating whether they are statistically different. Table 4.2 shows the mean, median, and standard deviation of the groups of mistakes. The table clearly shows that the two groups are statistically different in terms of all three characteristics, with the directions of comparison always as expected. Figure 4.6 provides comparison in terms of whole distribution of mistakes. Clearly, the distribution mass shifts to the right when moving from rounds with relatively low partner AAM to rounds with relatively high partner AAM. Moreover, the difference between the distributions is clearly statistically significant. The figure also illustrates that the effect is in terms of replacing very small mistakes by medium-sized mistakes. Overall, the statistical analysis suggests a clear negative effect of a partner's AAM on subjects' attention effort, which we quantify in the following sections.

Table 4.2: Comparison of groups of mistakes

	1st quartile	**4th quartile**	**P-value**
Mean	7.2	13.3	0.0029***
Median	0.0	3.0	<0.0001***
Std. Dev.	16.9	22.2	0.0003***

Notes: Rows indicate a metric, while columns indicate whether partner AAM belonged to the 1st or 4th quartile. Last column shows p-value for test of equality between the two groups.

Figure 4.6: Comparison of conditional distribution of mistakes



Notes: Blue color is used for mistakes made when the partner AAM belong to the 1st quartile. Red color is used for mistakes made when the partner AAM belong to the 4th quartile.

Main specifications. The H1 hypotheses relate to coefficient β. Table 4.3 reports the estimated coefficients from several empirical models, together with standard errors, t-statistics. For each model, we show the results aggregated for all Senders and all Receivers separately, and results aggregated across roles. We present the individual role results to avoid cross-contaminating the estimates with possibly different behaviour in different roles. For some of the models we also present results from two samples: one set the full available sample, and one from a restricted sample.[6]

[6]We restrict the sample in several ways. First, we know that DGP does not (at least locally) hold for subjects who do not make any mistakes throughout the experiment, or who almost never spend a sufficient amount of time on the attention task. The most likely explanation for the former is that subjects are motivated beyond the experimental payoffs; the latter subjects seem not to be motivated enough. We operationalize the former concept as making any mistake in at most 2 rounds out of 20 rounds observed; there are 4 such subjects, making it clear that this operationalization is not too lenient. As for the subjects who never spend a sufficient amount of time on task, we identify them as players for whom the 3rd quartile of time spent on the task is lower than one third of the maximum allowed time (25 seconds). There three such subjects.
In addition to dropping several subjects, we also drop particular type of observations: those that correspond to mistakes of 10,20 and 100; altogether, there are 99 such observations. As discussed above, the extra mass at mistakes equal to these values suggests that (at least some of) these mistakes are unlikely to be result of lower attention effort, given that subjects kept correct count of the second digit. By including these observations, we potentially risk biasing our results. If we drop them, the only danger is lower power in testing our hypothesis.

Table 4.3: Coefficient estimates

#	Model	Player effects	Sample	Aggregation: Senders	Receivers	All
1	Mixed effects	Random	Restricted	0.35 (0.11) [3.10]***	0.14 (0.09) [1.56]	0.24 (0.07) [3.37]***
2	Fixed effects	Fixed	Restricted	0.34 (0.12) [2.94]***	0.13 (0.10) [1.41]	0.24 (0.07) [3.16]***
3	Mixed effects	Random	Full	0.26 (0.09) [2.82]***	0.09 (0.06) [1.36]	0.18 (0.06) [3.08]***
4	Fixed effects	Fixed	Full	0.26 (0.10) [2.67]***	0.09 (0.07) [1.29]	0.17 (0.06) [2.89]***
5	Mixed effects with controls	Random	Restricted	0.34 (0.11) [3.10]***	0.14 (0.09) [1.53]	0.23 (0.07) [3.20]***
6	Tobit	None	Restricted	0.82 (0.23) [3.53]***	0.46 (0.23) [1.99]**	0.70 (0.19) [3.78]***
7	Robust LS	None	Restricted	0.14 (0.02) [6.08]***	0.12 (0.02) [6.44]***	0.12 (0.01) [8.90]***
8	Quantile reg. (median)	None	Restricted	0.12 (0.04) [3.25]***	0.07 (0.03) [2.70]***	0.10 (0.02) [4.27]***
9	Quantile reg. (90th percentile)	None	Restricted	0.91 (0.32) [2.84]***	0.39 (0.21) [1.82]*	0.56 (0.16) [3.47]***
10	Ordered logit	None	Restricted	0.033 (0.009) [3.86]***	0.033 (0.009) [3.67]***	0.033 (0.006) [5.35]***

Notes: The table displays estimates for β^x in equation (4.2). It reports coefficients, standard errors (in round brackets), and the t-statistics (in square brackets). Stars indicate significance at the usual 10%,5% and 1% levels. All models include fixed round effects (except model 10). Standard errors are clustered at period level for all models except for models 6 and 7, where we use the Huber Sandwich method, and models 8 and 10, where we use the Huber/White method. All models include session dummies (except models 2,4 and 10).
We summarize other regression statistics here. The unrestricted sample has 895 observations (96 players with up to 10 rounds), while the restricted sample has 762 observations (89 players with up to 10 rounds). The R-squared is between 0.03 and 0.05 when individual fixed effects are not included, and between 0.25 and 0.4 when they are included.

The first line of the table reports the results from our most basic model: a linear model with cross-sectional (individual) random effects and round fixed effects.[7] The results strongly support hypotheses H1S, with coefficient 0.35 being both statistically (t-stat=3.2) and economically significant. The coefficient indicates that Senders mistakes were on average 0.35 higher when they faced players with AAM higher by 1. Since the quartile range of AAM is 14, then an increase in AAM from the first to third quartile translates into an average increase in absolute mistakes of 4.9. In terms of predicted payoffs, this corresponds to decrease in payoff by 98CZK (245CZK vs 343CZK). In contrast, the results for Receivers provide substantially weaker evidence for hypothesis H1R: even though the coefficient is positive as expected, it is not statistically significant (t-stat=1.5), and the size of the coefficient is less than one third of the coefficient for Senders; the coefficient is too small to be truly economically meaningful. When we pool observations for Senders and Receivers together, we obtain a coefficient that is somewhere in between the coefficients for Senders and Receivers, and which remains statistically significant at a high level. Therefore, we conclude that there is sufficient evidence to support hypothesis H1P.

[7] The use of individual random effects (as opposed to fixed effects) is warranted because individual fixed effects cannot by design be correlated with the other independent variables in our environment: the only other variables in the model at this point are the partner's attention costs and round fixed effects, variations in which clearly cannot be correlated with subject characteristics.
The standard errors are clustered at period level. The use of period-clustered (cross-section robust) standard errors is motivated by our belief that correlation across periods for a given individual is more likely than correlation across individuals within a round: the performance of the individual player in a particular round is quite likely to have effects on performance in later rounds, causing correlation across periods for the individual; correlation between the performance of different individuals within a given round (beyond the fixed effects of rounds, which are captured by the inclusion of round fixed effects) would require *regular* session-specific shocks affecting all players. For results with alternative specifications for covariance estimator; see appendix 4.A.1.

We also present results from a model with fixed effects for both cross-sectional and time dimensions (see line 2). The results further support hypothesis H1S, with coefficients almost unaffected in size or significance. Meanwhile, the coefficient for Receivers remains small and insignificant. The pooled estimate is again in between the two coefficients, and is unchanged from the original model. To further check the robustness of our results, we also estimate models which use a full set of demographic controls we collected on the subjects, instead of using individual fixed effects. These controls include participant behaviour in the additional games from section 4.2.4, so that we are potentially controlling for altruistic and other characteristics. The results are in line with those from previous models, with coefficients and their significance almost unchanged (see line 5). This is encouraging, given that we increased the number of estimated coefficients from to 11 to 29.

The results in lines 1 and 2 of Table 4.3 use the restricted sample. When we use the full sample, the coefficients generally become smaller and less significant (see lines 3 and 4 of Table 4.3), but the effect is not large and the coefficients are still significant at the usual significance levels. Given this and the problematic nature of the excluded observations, we believe that coefficient estimates from the restricted sample reflect the actual behavior of subjects more closely, and so in the rest of this chapter we will report results from models using the restricted sample.[8]

All previous specifications postulate a linear relationship between partner AAM and mistake made by player in a given round. However, such a linear relationship might not be globally valid. Specifically, some subjects might pay full attention (i.e., exert maximum attention effort) *for a range of partner attention costs*, while

[8]Appendix 4.A.2 presents results from full and restricted samples side by side for all the models from Table 4.3, showing that the conclusions we draw are broadly unaffected.

still lowering their own attention in response to increases in their partners' attention costs as long as these are high enough. In this case, we would observe a positive relationship between size of mistakes made and partner attention costs on a range of values for partner attention costs, but the observed relationship would break down below a particular threshold.[9]

We account for this possibility by estimating a standard Tobit model with left censoring at 0, corresponding to the fact that variance of signal noise cannot be negative (i.e., at best, subjects can make zero mistakes on average). Line 6 in Table 4.3 shows the results: the coefficients are greatly increased, while the statistical significance levels either remain the same or also increase. Therefore, accounting for the censoring changes our conclusions about the statistical and economic significance of the effects. First, the coefficient for Receivers also becomes (borderline) significant. Second, the implied effects are very large for Senders, and meaningful for for Receivers and for pooled observations: the effect for Senders, Receivers, and pooled effect are 189CZK, 52CZK and 120CZK.

Should GLS or Tobit results be trusted more? Almost half of our observations are potentially affected by censoring due to corner solution issues. Due to the positive relationship between our dependent and main independent variable, not accounting for corner solutions leads us to believe that the response of the dependent variable (mistake made) to change in the independent variable (partner AAM) is smaller than it actually is. For this reason, we consider the results from the Tobit model as our main results. It is worth highlighting the difference between the predicted relationship when using the Tobit model and the linear

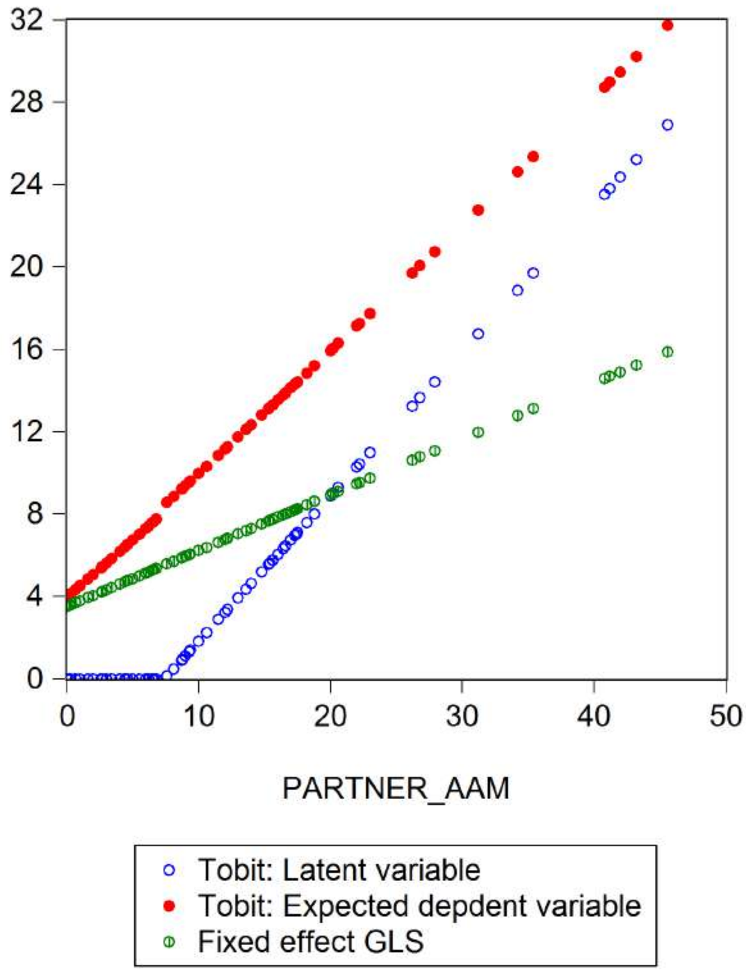
[9] This is an example of a corner solution model. Appendix 4.D provides a graphical illustration of corner solution outcomes in our situation.

fixed effect specification. Figure 4.7 shows the predicted values for both models, setting round and individual fixed effects to zero. We show both the prediction for the latent variable of the Tobit model and the expected dependent variable. The former shows the true effect of partner AAM on average mistakes, while the latter shows the effect on the conditional expected value we should observe. The figure clearly shows how the Tobit predicts a significantly steeper relationship, and also how the relationship does not hold below a particular level of partner AAM. Notwithstanding, the observed dependent variable is of course expected to be positive everywhere, and, interestingly, is almost identical to the value predicted from the fixed effect model when partner AAM is 0.

Robust specifications. The empirical models presented in this subsection on main specifications all use least squares estimation, and there is the potential for the results to be driven by outlier observations. Moreover, our data have clear potential for such issues, since the distribution of mistakes is one sided (i.e., mistakes can be only positive) and it has a long tail corresponding to relatively large mistakes, as evidenced by Figure 4.3. It is possible that observations with such large mistakes could randomly drive our results, something we investigate using alternative estimation methods.

Before discussing results from alternative estimation approaches, it is worth highlighting that the presence of large mistakes is in some sense expected. The main specifications postulated a continuous (linear) relationship between attention of a given subject and partner attention costs. However, subjects might not be able to adjust their attention in a perfectly continuous way. Instead, they might be limited to choosing between a finite number of different attention levels. An extreme example would be when they switch between paying maximum attention

Figure 4.7: Predicted values for GLS and Tobit models



and no paying any attention at all. If this is the case, we would expect to see a distribution that is somewhat binomial at the individual level, which in aggregation would naturally lead to a distribution with a relatively large number of observations that look like outliers. However, these observations would not really be outliers, but rather would correspond to rounds when subjects ***choose*** to pay no attention. In principle, this kind of behaviour would still support the hypotheses that agents change their attention in response to inattentiveness of their partners.

Our first approach is to use a robust least squares (RLS) estimation, which replaces the quadratic loss function of least squares with an alternative function that gives less weight to outliers. Line 7 shows coefficient estimates when using such estimation method. The coefficients are positive and strongly statistically significant for all three levels of aggregation. On the other hand, the coefficients are reduced when compared to the estimates we obtained from least squares estimation, so that they barely remain economically significant. Similarly to our Tobit model results, but in contrast to results from least squares estimation, the coefficient is clearly statistically significant even for Receivers.

When estimating robust least squares, one needs to make somewhat arbitrary choices of weighting (and scaling) functions, with different functions yielding potentially significantly different estimation results.[10] An alternative estimation method that uses less arbitrary weighting function and is still robust to outlying observations is quantile regression. The coefficient estimates from the median regression are presented in line 8 of Table 4.3. Encouragingly, we obtain coefficients that are also significant for all three roles, and similar in size to RLS estimation.

Based on the RLS and quantile regression results, one could conclude that while subjects clearly react to partner AAM in expected ways, the size of the reaction is relatively muted and bordering on economic insignificance. However, looking at other quantiles than median suggests that this conclusion is too simplistic. First consider the estimated effect on the 90th percentile in line 9 of Table 4.3: not only

[10]The table reports results based on the Cauchy weighting function. As Appendix 4.A.1 shows, in our case the size of the coefficients obtained from robust least squares is somewhat sensitive to this choice, with the reported estimates in line 6 somewhere in the middle of the possible coefficient estimates.

are all the coefficients still significant, they are 5 to 7 times larger than the median effects, and more than twice as large as mean effects estimated by least squares estimation. This suggests that the conditional mean and median understate the effects for some part of the distribution. To provide further evidence for this feature of the data, the top panel of Figure 4.8 shows coefficient estimates at all deciles. Clearly, the effect of partner AAM increases with the considered quantile, and is significantly positive for all quantiles at or above median and for each level of aggregation. The bottom panel of Figure 4.8 shows the reason for this heterogeneity across quantiles: our data are clearly heteroskedastic. While mistakes are concentrated close to zero (with occasional large values) when partner AAM is low, they are distributed more widely when partner AAM is high, with only rare low values. This feature of the data is expected, since lower attention leads to increases in both mean and variance of mistakes made. The figure also shows different regression lines corresponding to mean and different quantiles, demonstrating how the effect of partner AAM is increasing with the considered quantiles.

There are two possible interpretations of the heterogeneity in quantile effects. First, it might reflect genuine heterogeneity across participants, with some participants reacting more strongly to being matched with an inattentive partner, and some reacting less (or not at all). A standard regression, which estimates effects on conditional mean will estimate the effect averaged across these two behavioural types. In contrast, the conditional 90th percentile of mistakes is likely driven only by the latter behavioural type. Under this interpretation, the results strongly support our hypothesis for around a quarter of our subjects, reject the hypothesis for 40% and give limited support for the rest. However, this interpretation requires some strong regularity on the response functions to be true.[11] An alternative, more robust explanation is: if unobserved factors - e.g., how complicated an attention task is or how tired a subject is in a round - would suggest large mistake, the effect of these factors would be larger if a partner has high attention costs. In other words, an unobserved round-specific shock to individual attention costs is amplified by the partner attention costs.

[11] Appendix 4.C provides two simple simulated examples, one in which the estimated coefficients correspond to actual responses to changes in AAM, and one when they do not.

Finally, an approach similar to quantile regression is to classify mistakes relative to other mistakes made by a player. For each player, we classify mistakes into terciles according to their position in the distribution of mistakes for the particular player. By doing this we avoid any problems with our estimates being driven by outliers: it no longer matters how large a particular mistake is; it only matters whether it is larger or smaller than other mistakes made by a player. We then run ordered logit on the resulting classification, including period fixed effects. The resulting coefficient estimates answer the following questions: Are player's mistakes ordered with respect to partner AAM? Does the probability of making a larger mistake increase with partner AAM? The last line of Table 4.3 shows the coefficient estimates, which strongly support our hypothesis, with coefficients being strongly statistically significant for all levels of aggregation. Since interpretation of marginal effects in ordered models is complicated we present the predicted probability of each tercile as a function of partner AAM. While the probability of first tercile mistake is almost 70% when partner AAM is zero, it drops to less than 30% when partner AAM is at maximum. Correspondingly, the probability of the third tercile increases by 31%. This clearly highlights that the coefficient estimates assign relatively large effects to partner AAM.

Figure 4.8: Quantile regression coefficients and lines

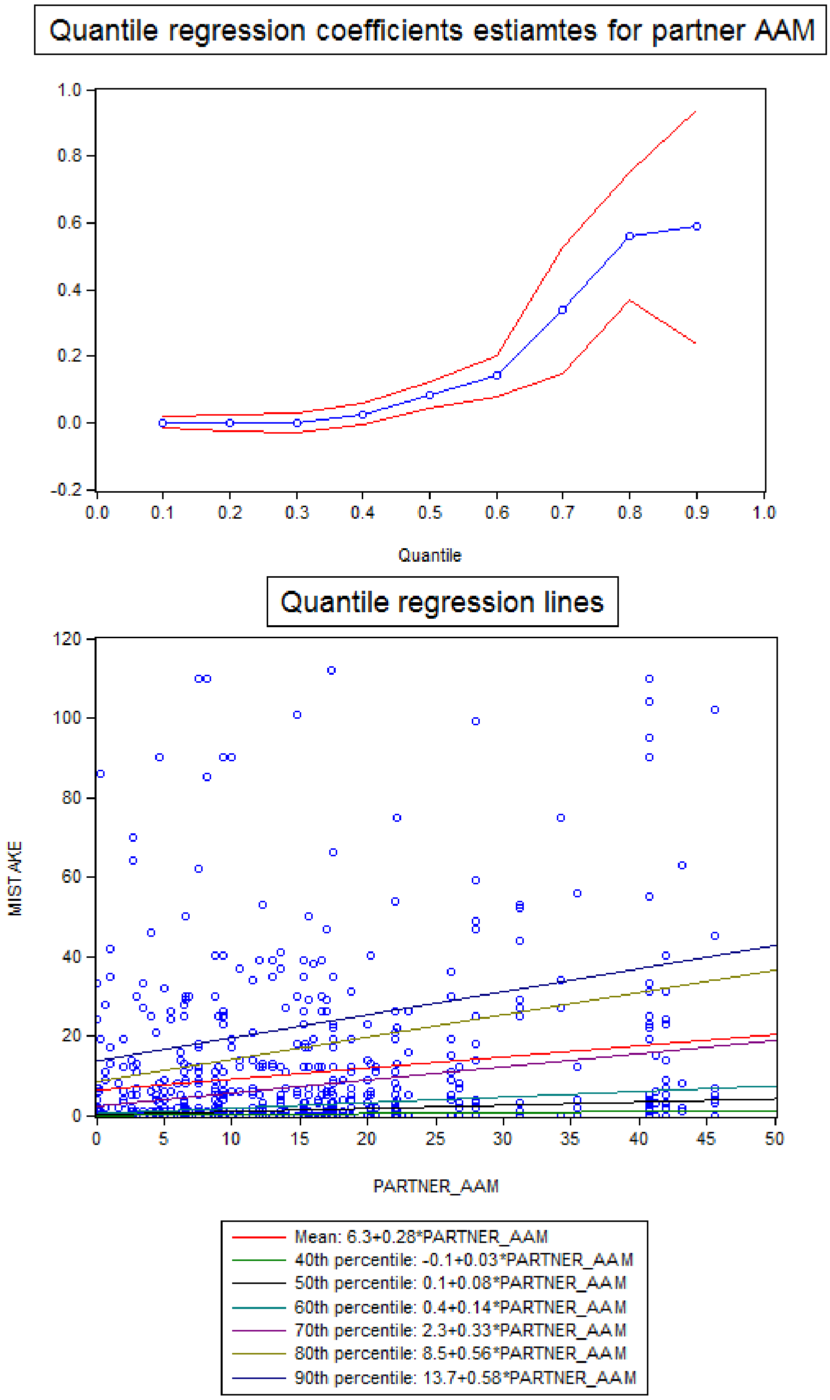
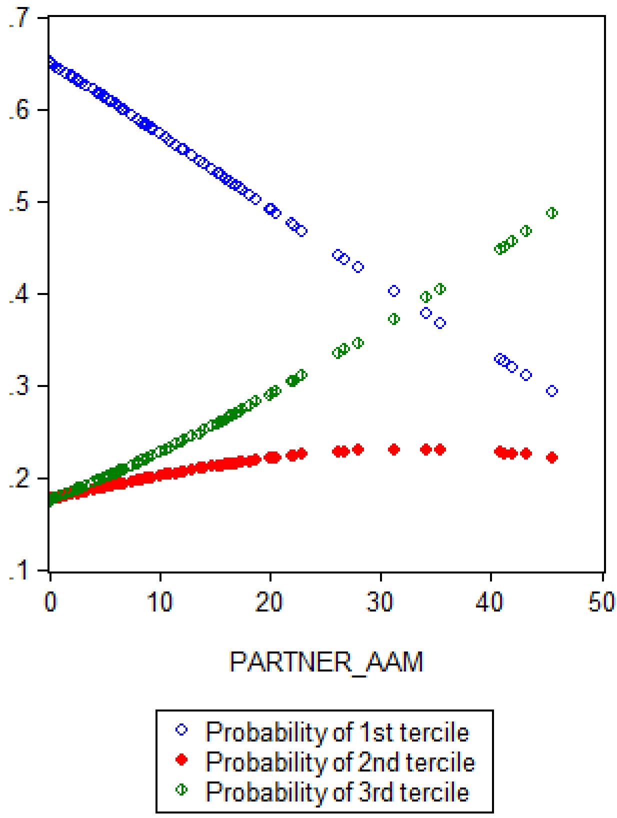


Figure 4.9: Effect of partner AAM on a probability of mistake belonging to different terciles



Individual coefficients. The structure of our data also allows us to estimate individual coefficients β. Given the number of participants, we present only a summary of the results in Table 4.4. Three fifths of individual coefficients are positive, with the majority of those being significant at the 1% level; meanwhile, coefficients with a negative sign are mostly insignificant. Together, this makes positive and significant coefficients the largest group, more than three times as large as group of negative and significant coefficients. We can conclude that the estimated positive and significant relationships from table 4.3 are results of behaviour observed in the majority of participants, giving further support to hypothesis H1.

Table 4.4: Individual coefficients

		Significance		
		< 1%	1-10%	> 10%
Sign	+	33 [37%]	4 [5%]	20 [22%]
	-	10 [11%]	5 [6%]	17 [19%]

Having individual coefficients also allows us to analyse whether they are related to either demographic characteristics or to participant behaviour in one of the three small games administered at the end of the experiment (CRT, dictator game, and centipede game). Given the relatively small sample, we start with a simple correlation analysis. Out of demographic characteristics gender, education, some majors and having taken a statistics course seem to be significantly correlated with the individual regression coefficients, so we include only those in further regression analysis. We also include behaviour in the dictator and centipede games, and answers to CRT. Table 4.5 shows results from 4 regressions: first two columns show results for models regressing the level of the individual coefficient on our set of explanatory variables, while the latter two show results for models regressing a dummy variable for positive coefficient on our set of explanatory variables; in each case we present results with and without information about education and the major of participants.

The results offer a few interesting conclusions. First, the regressor most strongly associated with the level of individual coefficients is the answer to the first CRT question, but the direction is opposite of what we would expect: the correct answer leads to a lower coefficient, so that participants who apply reflection tend not to react as strongly to partner AAM. One way to interpret this is that the results we observe reflect the behavior of subjects who are reacting based on intuition rather than on optimization, but this conclusion would require further analysis. Second, a course in statistics has a positive and significant association with both the level of the individual coefficients and the probability that they are positive, in line with what we would expect. Lastly, female participants have lower coefficients and are more likely to have negative coefficients than males, with the latter being statistically significant at the 5% significance level. It is also noteworthy that play in the dictator game is not related to behaviour in the main experiment. This suggests that altruistic considerations do affect how subjects were behaving in the main experiment. Finally, note that the demographic information included in the regressions can explain a relatively large portion of variations in individual coefficients.

Table 4.5: Regression results for individual regression coefficients

	Level (small)	Level (full)	Prob. (small)	Prob. (full)
Constant	0.2 (0.23) [0.89]	0.21 (0.22) [0.94]	0.7 (0.18) [3.91]***	0.59 (0.21) [2.85]***

(continued on next page)

Table 4.5, continued

	Level (small)	Level (full)	Prob. (small)	Prob. (full)
Amount given in dictator	0 (0) [1.50]	0 (0) [1.71]*	0 (0) [1.40]	0 (0) [1.32]
Correct CRT1	-0.46 (0.18) [-2.53]**	-0.52 (0.2) [-2.66]***	-0.08 (0.11) [-0.78]	-0.1 (0.11) [-0.90]
Correct CRT2	-0.04 (0.17) [-0.22]	-0.14 (0.17) [-0.81]	-0.02 (0.14) [-0.18]	-0.09 (0.13) [-0.68]
Correct CRT3	0.24 (0.21) [1.15]	0.5 (0.25) [2.03]**	-0.1 (0.14) [-0.71]	0.03 (0.15) [0.19]
Stop in centipede	-0.19 (0.18) [-1.06]	-0.22 (0.17) [-1.28]	-0.1 (0.12) [-0.83]	-0.12 (0.12) [-1.02]
Course in statistics	0.36 (0.17) [2.11]**	0.45 (0.16) [2.86]***	0.2 (0.12) [1.63]*	0.36 (0.14) [2.57]**
Female gender	-0.43 (0.16) [-2.63]**	-0.45 (0.16) [-2.78]***	-0.29 (0.1) [-2.84]***	-0.3 (0.11) [-2.81]***
Bachelor education		-0.26 (0.19) [-1.37]		-0.11 (0.12) [-0.94]
Master education		0.02 (0.28) [0.07]		-0.03 (0.18) [-0.18]
Computer science major		-0.04 (0.17) [-0.24]		0.15 (0.22) [0.67]
Law major		0.79 (0.24) [3.33]***		0.17 (0.21) [0.82]
Arts major		0.56 (0.26) [2.19]**		0.79 (0.19) [4.18]***
Other major		-0.52 (0.27) [-1.91]*		-0.1 (0.16) [-0.62]
R-squared	0.15	0.28	0.1	0.17

Summary of results. This section shows that subjects systematically adjust their attention effort in response to variations in partner attention costs: when partner attention costs are higher, decreasing the average precision of his/her signal, subjects react by lowering their attention, as predicted by rational inattention theory. This effect is both statistically and economically significant, and results from quantile regression suggest that it increases when other factors make paying attention harder. Overall, this section contributes to existing literature on rational inattention by showing that experimental subjects adjust their attention not only in response to changes in environment, but also in response to changes in their teammates' characteristics, demonstrating that the theory also has a bite in interactive situations.

4.4.2 Action choice adjustment

We also study how agents adjust their action choices in response to variations in their own and partners' attention costs. The expectations about behaviour come from standard theory of decision-making under uncertainty. Optimal Bayesian updating suggests that action in our game should be equal to a weighted average of the signal value and the prior mean, i.e., subjects should shade their signals towards the prior mean:

$$A = \omega \cdot \pi + (1 - \omega) \cdot S^R \tag{4.3}$$

where S^R is the Receiver's signal and π is the prior mean.

The weight put on the prior mean, which we will refer to as the degree of shading towards the prior, should depend on the noisiness of the signal, with a noisier signal leading to more shading, ceteris paribus, i.e., for given signal realization. Crucially, in our environment, the noisiness of the signal depends on both Sender's and Receiver's effective attention, so on the attention costs of both players. This means we can study whether players behave in line with predictions from standard Bayesian theory by analyzing whether Receivers vary the degree of shading towards a prior mean in response to variations in both partners' attention costs. Specifically, Bayesian theory predicts that we should expect that higher attention costs lead to higher weight on the prior mean, yielding the following set of hypothesis.

H2S: Receivers put more weight on the prior mean when Sender's attention costs are higher.
H2R: Receivers put more weight on the prior mean when their attention costs are higher.

It would seem that we could test these hypotheses by calculating the weights based on observed actions according to equation (4.3)[12] and then regressing them on our variables of interest. However, we have data only on Receivers' actions, not on Receivers' signals. We could potentially use the fact that Receiver signals should on average equal the Sender signal and proxy the former by the latter. However, this approach has the drawback that the noise introduced by using a proxy variable, combined with the division operation, yields series that are unfit for such analysis.

Instead, we focus on studying the distance between actions taken by the Receivers and signals provided by Senders. Since Receiver signals are equal to Sender signals plus noise due to imperfect attention, we can express S^R as the sum of S^S and unobserved error coming from a Receiver's inattention ϵ^R. Substituting for

[12] With knowledge of A, π and S^R we can express ω as $\frac{A-S^R}{\pi-S^R}$.

$S^R = S^S + \epsilon^R$ in (4.3) and rearranging yields following expression:

$$A - S^S = (\pi - S^S) * \omega + (1 - \omega)\epsilon^R \tag{4.4}$$

Our hypothesis is that ω depends on the Sender and Receiver attention costs, $\omega = g(AC^S, AC^R)$. If we approximate this general function by a linear function, $\omega_{i,t} = \gamma_0 + \gamma_1 AC^S + AC^R$, and proxy attention costs by AAM as before, then we can run the following regression:

$$A_{i,t} - S^S_{i,t} = \gamma_0 * (\pi - S^S_{i,t}) + \gamma_1 * (\pi - S^S_{i,t}) * AAM^S_{i,t} + \gamma_2 * (\pi - S^S_{i,t}) * AAM^R_{i,t} + \eta_{i,t} \tag{4.5}$$

where $\eta_{i,t}$ is the composite error term, reflecting both the effect of the Receiver's inattention error, $(1 - \omega)\epsilon^R$, and random noise.

Establishing that γ_1 and γ_2 are positive would support our hypotheses H2S and H2R, which state that ω is increasing in AC^S and AC^R, respectively. Unfortunately, $\eta_{i,t}$ is clearly correlated with our explanatory variables, because it also includes ω and hence $AAM^S_{i,t}$ and $AAM^R_{i,t}$, which will bias our results. Nevertheless, the sign of the bias still allows us to draw some conclusions. The correlation between the included regressors and the error term is clearly negative, because ω enters η with negative sign. Since our hypothesis is that Receivers put *higher* weight on a prior mean when attention costs are higher, then establishing that γ_1 and γ_2 are positive *despite* the negative bias provides clear-cut support for our hypotheses.

First line in Table 4.6 provides results for regression (4.5). The estimated γ_1 coefficient is marginally statistically significant, giving some support to hypothesis H2S. Nevertheless, the coefficients are relatively small, so that the estimated (downwardly biased) effect of Sender attention costs on the degree of shading towards a prior is estimated to be rather small. Moreover, if we drop observations where the Receiver choose a prior mean, likely corresponding to no attention effort on his part, the effect becomes smaller and insignificant, as line 2 shows. Meanwhile, the estimated coefficient γ_2 does not support hypothesis H2R, with the coefficient negative and insignificant.

Given the problem with the previous approach, we use an alternative, less direct approach, to study shading towards a prior mean. The above approach studies the *distance* between a Receiver action and prior mean compared to the distance between Sender signal and the prior mean, relating it to the (expected) noisiness of a Sender signal. Here we instead study the *probability* that the Receiver's action is closer to the prior than the Sender's signal. In a sense, the previous approach answered whether *degree* of shading down varies with the attention costs of both players. Here, instead, we study whether the *probability* of shading down increases with the attention costs of both players.

The Receiver's signal is equally likely to be above or below Sender's signal. However, as discussed above, if Receivers are shading down, the Receiver's *action* should be more likely to be closer to the prior than further away. More importantly, since the shading down should be stronger when attention costs of the Sender or the Receiver are higher, then the probability of being closer to a prior mean should be increasing in these attention costs. To study whether this is true, we would ideally link probability of shading down to attention costs in the following model:

$$P(shade_{i,t} = 1) = h(AC^S, AC^R) \tag{4.6}$$

Table 4.6: Coefficient estimates for shading down behaviour

#	Model	Player effects	γ_1	γ_2
1	Distance	Fixed	0.0038 (0.0019) [1.96]**	-0.0023 (0.0024) [-0.96]
2	Distance (restricted sample I)	Fixed	0.0028 (0.002) [1.40]	-0.0021 (0.0024) [-0.90]
3	Probability: LPM	Random	0.007 (0.002) [3.53]***	0.0018 (0.0026) [0.68]
4	Probability: LPM (restricted sample I)	Random	0.0056 (0.0022) [2.50]***	0.0025 (0.0026) [0.98]
5	Probability: LPM (restricted sample II)	Random	0.0089 (0.0024) [3.77]***	0.0017 (0.0026) [0.64]
6	Probablity: Logit	Random	0.031 (0.01) [3.06]***	0.0077 (0.0099) [0.77]

Notes: See notes below Table 4.3 for explanation of values in the table. All models include round and session fixed effects. Standard errors are clustered at period level for all models except model 4, where we use Huber/White robust covariances.
We do not report other regression statistics. We use the full available sample with the exception of 3 players with low median time spent (see footnote 6), yielding 453 the Receiver actions. The R-squared is around 0.05 for models with individual random effects and fixed period effects, and around 0.18 for models with individual fixed effects and fixed period effects.

where $shade_{i,t} = 1$ is defined as $|A_{i,t} - \pi| < |S^R_{i,t} - \pi|$. However, since we do not observe S^R we need to use S^S as before, so that $shade_{i,t} = 1$ will be proxied by $|A_{i,t} - \pi| < |S^S_{i,t} - \pi|$, i.e., the distance from the prior being smaller for the action than for the Sender signal. Assuming a linear relationship, we estimate the following regression model:

$$P(shade_{i,t} = 1) = \gamma_{0,i} + \gamma_1 AAM^S_{i,t} + \gamma_2 AAM^R_{i,t}\delta_t + \epsilon_{i,t} \tag{4.7}$$

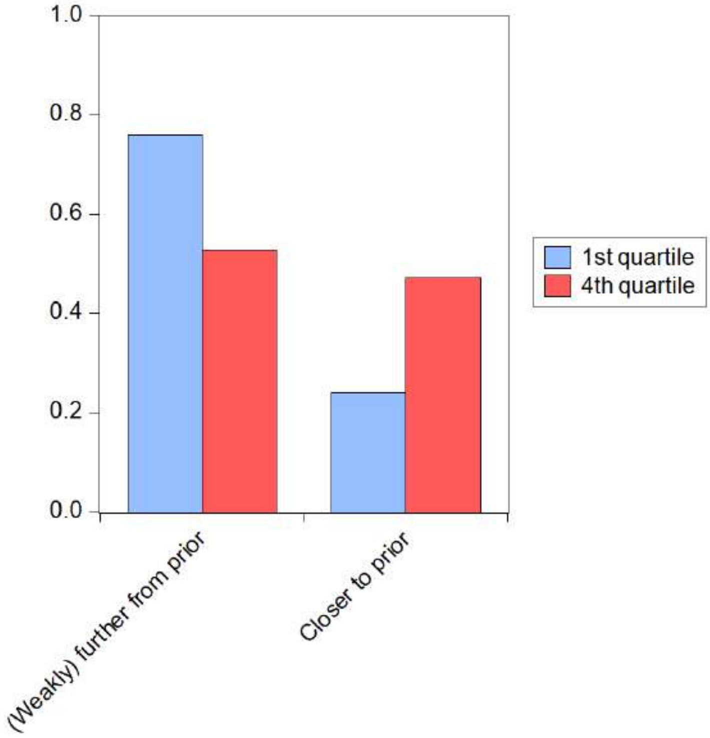
The results are presented in the bottom of Table 4.6. The results give qualified support to hypothesis H2S: the coefficient on partner AAM is positive and strongly statistically significant across different models. The coefficient in the linear probability model is 0.007. This coefficient implies that the probability that Receiver's action is closer to a prior than the Sender's signal increases by 0.7% for one unit increase in Sender AAM, translating to an increase in the probability of shading of 9.8% when partner AAM increases from the first to 3rd quartile.

In addition to regression results, we also provide classifications of actions conditional on the Sender AAM belonging to the 1st or 4th quartile. Figure 4.10 shows that the conditional classifications are very different, with the probability of being closer to the prior being higher by 28% in the latter group. Overall, we can conclude that Receivers are reacting to higher Sender AAM by shading down more or with higher probability.

Meanwhile the results are not supportive of hypothesis H2R, with coefficient being small and insignificant. However, this conclusion is complicated by the fact that individual AAM is potentially correlated with other individual unobserved characteristics affecting the tendency to shade down, and hence the coefficient is potentially biased.

We check our results by estimating several related models. First, as before, we drop observations when the Receiver takes an action equal to the prior mean, to make sure that it is not only this extreme behaviour driving our results. The

Figure 4.10: Classification of actions by quartile.



main coefficient in line 3 is somewhat smaller and less significant, but remains economically and statistically significant. Second, using the absolute value distance to the prior mean creates a non-linearity at the prior mean which might influence our results. To eliminate this potential problem, we restrict our attention to observations where the Receiver's action lies in the same direction from the prior as Sender's signal. Model 4 shows that the coefficient and its significance actually increase.[13] Finally, we use a logit rather than a linearity probability model. The γ_1 coefficient has the expected sign and is slightly more significant than in the LPM model, but implies on average a somewhat smaller effect on the predicted probability (8.8% compared to 9.8%). Conclusions about H2R are unchanged.

Summary of results. The results in this subsection show that Receivers clearly take into account the noisiness of the information they obtain about the payoff relevant variable, with their actions being closer to a prior when their information is likely to be more noisy. This comes with two caveats. First, while the effect is statistically significant, the estimated effect is relatively small, even though this could be caused by downward bias in our estimates. Second, the effect applies only to noise coming from partner inattentiveness; we were unable to find any evidence that Receivers are reflecting on their own attention costs, even though the negative result can be easily caused by multiple sources of bias.

While our positive finding is important within rational inattention literature, it also has a bearing on literature on decision making under uncertainty. The literature shows that subjects reflect only on their own signals, and not on the noisiness of a signal or information about a prior distribution, which is termed base-rate neglect (Kahneman and Tversky 1973; Lyon and Slovic 1976). In contrast, we find that Receivers take actions closer to prior mean when their signal is likely to be more noisy, as the theory suggests that they should. Importantly, in our experiment, changes in noisness of the signal do not result from changes in the environment exogenously engineered by the experimentator, but rather are the result of variation in partners from one round to another. This distinguishes our

[13]The restriction eliminates 28 out of 453 observations. Note that this sample selection is likely causing downward bias in the coefficient. We therefore take full sample estimates as the main estimates.

approach from existing studies on Bayesian updating, where the change in noisiness coincides with changes in environment.

4.5 Conclusion

This chapter presents results from our experiment testing whether people react to inattentiveness of their partners in cooperative tasks. The results strongly support the hypotheses, showing that participants did pay less attention when they were partnered with a player who had higher attention costs, realizing that the effective attention of their partner is complementary in a strategic sense to their own attention. Moreover, the subjects also internalized this soft information about the likely precision of the information, taking actions closer to a prior when information was less precise as Bayesian theory would suggest.

We view our results as a first step in the analysis. In our experiment the players could only react to average measure of inattention of their partner, i.e. attention costs, which we refer to as inattentiveness rather than inattention. In our next experiment, we plan to also communicate to players a proxy for the partners' actual attention effort in a given round, so that players can also react to intention, not only inattentiveness. In addition to studying changes in attention in cooperative behaviour, we also plan to study whether agents change their strategic behaviour in response to changes in the attention costs of their opponents, as in Martin (2017a). We plan to follow the work of Martin (2017b), but to add information about player's partners, so that we can study changes in the behaviour of subjects across their opponents.

4.A Additional estimates

This appendix presents additional regression estimates, showing that our conclusions are fairly robust to alternative estimation methods and model specifications. The appendix has 2 sections dedicated, respectively, to robustness checks and effects of restricting the estimation sample.

4.A.1 Robustness checks

This appendix provides some robustness checks for our results, focusing on alternative estimation options, including covariance estimators, distributional assumptions and the weighting function in RLS.

First, table 4.7 shows estimates from a mixed effect model estimated on a restricted sample with standard errors and t-statistics when we use other than the period-robust covariance estimator, results for which were reported in the main text, and are replicated in the first line of the table. In almost all cases, the standard errors are smaller and hence the significance of the estimated coefficients is higher, as we would expect if the errors are serially correlated.

Second, table 4.8 shows results from a Tobit model when we assume other distributions than the extreme value distribution used in the main text. While the coefficients are systematically and significantly lower, the significance is either similar or higher. As argued in the main text, we believe that the extreme value distribution is most appropriate and hence consider results based on that distribution as the main results. The table also includes alternative estimates of standard errors which allow for overdispersion: the results in the main text use Huber/White quasi-maximum likelihood robust standard errors, while the alternative

Table 4.7: Standard error estimates with alternative covariance estimators

Covariance estimator	Aggregation Senders	Receivers	All
Period robust	0.35 (0.11) [3.10]***	0.14 (0.09) [1.56]	0.24 (0.07) [3.37]***
Ordinary	(0.08) [4.25]***	(0.07) [2.04]**	(0.05) [4.59]***
Heteroskedasticity robust	(0.11) [3.11]***	(0.08) [1.66]*	(0.07) [3.49]***
Cross-section robust	(0.08) [4.10]***	(0.08) [1.74]*	(0.07) [3.69]***
PCSE heteroskedasticity robust	(0.09) [3.91]***	(0.07) [1.90]*	(0.06) [4.24]***
PCSE period robust	(0.09) [3.98]***	(0.07) [2.02]**	(0.06) [4.40]***
PCSE cross-section robust	(0.09) [3.80]***	(0.07) [2.00]**	(0.06) [3.80]***

Notes: The table displays coefficient estimates (first line) together with standard errors and t-statistics corresponding to alternative covariance estimators. The main text used period-robust estimator reported at the top of the table.

is to use generalized linear model robust standard errors that allow for conditional heteroskedasticity. In general, the standard errors are smaller and correspondingly, coefficients are significant at higher levels of significance.

Finally, we present coefficient estimates obtained when we use an alternative weighting functions in robust least squares. Table 4.9 provides estimates of coefficient β^x from equation (4.2) for 8 different weighting functions. The coefficient estimates reported in main text, which were based on the Cauchy weighting function and are reported in the first line of the table, are in the middle of the range of coefficient estimates.

Table 4.8: Tobit estimates with alternative distribution assumptions

Distribution	Aggregation Senders	Receivers	All
Extreme value	0.82	0.46	0.70
Huber	(0.23)	(0.23)	(0.19)
	[3.53]***	[1.99]**	[3.78]***
GLM	(0.19)	(0.17)	(0.13)
	[4.23]***	[2.66]***	[5.37]***
Normal	0.59	0.31	0.45
Huber	(0.17)	(0.13)	(0.11)
	[3.48]***	[2.37]**	[4.18]***
GLM	(0.15)	(0.12)	(0.09)
	[4.03]***	[2.61]***	[4.79]***
Logistic	0.39	0.26	0.31
Huber	(0.13)	(0.09)	(0.07)
	[2.97]***	[2.94]***	[4.20]***
GLM	(0.12)	(0.09)	(0.07)
	[3.30]***	[2.82]***	[4.28]***

Notes: The table displays coefficient estimates together with standard errors and t-statistics for alternative distributions assumptions and alternative covariance estimation methods, as indicated in the first column. Main text used the extreme value distribution with Huber robust standard errors reported in the first line.

4.A.2 Estimates with full sample

This section presents coefficient estimates for all estimation methods reported in the main text, with coefficients based on restricted and unrestricted samples side by side. The general conclusion from table 4.10 is that exclusion of our particular observations does not alter our conclusions: the coefficients are mostly - but not always - slightly smaller and less significant, but remain economically and statistically significant. Encouragingly, the effect is smallest on the estimation methods that are more robust to outlying observations, including robust least squares, quantile regression, and ordered logit, for which the coefficients sometimes

Table 4.9: Coefficient estimates from robust least squares

#	Weighting functon	Aggregation Senders	Receivers	All
1	Cauchy	0.137	0.117	0.122
		-0.023	-0.018	-0.014
		[6.08]***	[6.44]***	[8.90]***
2	Bisquare	0.071	0.12	0.10
		-0.037	-0.03	-0.03
		[1.92]*	[3.43]***	[3.87]***
3	Fairl	0.20	0.13	0.16
		-0.02	-0.01	-0.01
		[11.57]***	[9.53]***	[15.66]***
4	Huber	0.17	0.12	0.14
		-0.05	-0.04	-0.03
		[3.65]***	[3.23]***	[4.85]***
5	Huber-Bisquare	0.078	0.119	0.098
		-0.031	-0.03	-0.021
		[2.54]***	[3.93]***	[4.66]***
6	Logistic	0.168	0.119	0.139
		-0.045	-0.037	-0.028
		[3.74]***	[3.23]***	[4.97]***
7	Talworth	0.089	0.125	0.111
		-0.048	-0.039	-0.027
		[1.87]*	[3.19]***	[4.09]***
8	Welsch	0.156	0.128	0.132
		-0.005	-0.005	-0.003
		[28.72]***	[28.01]***	[39.08]***

Notes: The table displays regression coefficient estimates on partner AAM. All regressions use a restricted sample and include round fixed effects. Player fixed effects are not included due to the resulting large number of coefficient estimates. All estimates are based on Huber scaling and the standard errors are based on the Huber type I method.

increase. This further supports our conclusion that eliminating the observations is the correct way to proceed.

Table 4.10: Comparison of coefficient estimates with restricted and unrestricted samples

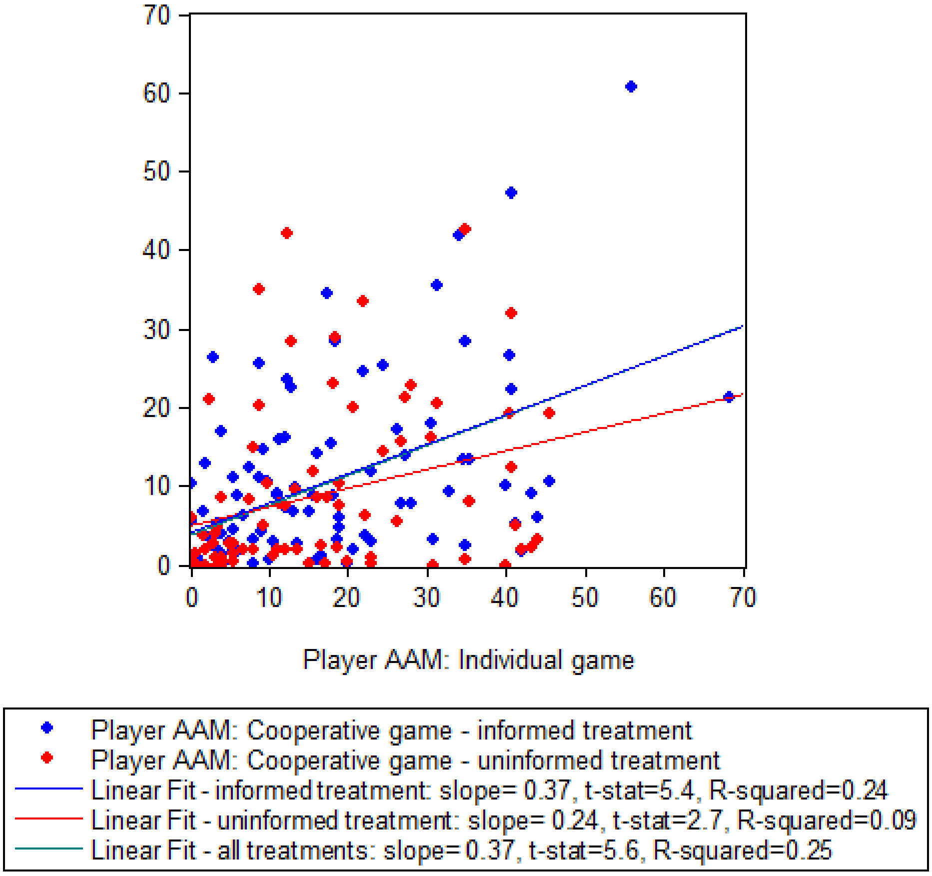
#	Model	Player effects	Restricted sample Sender	Restricted sample Receiver	Restricted sample All	Full sample Sender	Full sample Receiver	Full sample All
1	Mixed effects	Random	0.35 (0.11) [3.18]***	0.14 (0.09) [1.51]	0.25 (0.07) [3.42]***	0.26 (0.09) [2.82]***	0.09 (0.06) [1.36]	0.18 (0.06) [3.08]***
2	Fixed effects	Fixed	0.34 (0.12) [2.94]***	0.13 (0.10) [1.41]	0.24 (0.07) [3.16]***	0.26 (0.10) [2.67]***	0.09 (0.07) [1.29]	0.17 (0.06) [2.89]***
5	Mixed effects with controls	Random	0.34 (0.11) [3.10]***	0.15 (0.10) [1.52]	0.25 (0.07) [3.36]***	0.24 (0.09) [2.61]***	0.09 (0.06) [1.40]	0.17 (0.06) [2.88]***
6	Tobit	None	0.82 (0.23) [3.58]***	0.44 (0.23) [1.89]*	0.71 (0.18) [3.88]***	0.52 (0.21) [2.50]***	0.36 (0.22) [1.65]*	0.48 (0.17) [2.87]***
7	Robust LS	None	0.13 (0.02) [5.29]***	0.12 (0.02) [5.86]***	0.12 (0.02) [7.91]***	0.12 (0.02) [5.27]***	0.12 (0.02) [7.23]***	0.12 (0.01) [8.85]***
8	Quantile reg. (median)	None	0.12 (0.04) [3.25]***	0.07 (0.03) [2.70]***	0.10 (0.02) [4.27]***	0.10 (0.04) [2.53]***	0.07 (0.03) [2.52]***	0.13 (0.03) [4.46]***
9	Quantile reg. (90th percentile)	None	0.91 (0.32) [2.84]***	0.39 (0.21) [1.82]*	0.56 (0.16) [3.47]***	0.69 (0.25) [2.73]***	0.31 (0.21) [1.51]	0.41 (0.16) [2.52]***
10	Ordered logit	None	0.033 (0.009) [3.86]***	0.033 (0.009) [3.67]***	0.033 (0.006) [5.35]***	0.33 (0.14) [2.42]**	0.51 (0.13) [3.97]***	0.42 (0.09) [4.53]***

4.B Individual game performance as signal

This appendix presents data on the association between the average absolute mistake (AAM) made by players in the individual game and the AAM in the cooperative game. We show that the measures are strongly correlated in statistical significant way. This implies that subjects are correct in treating the individual game AAM as a signal of the precision of signals in the cooperative game and responding to it, as we have shown that they do.

Figure 4.11 shows a scatter plot of individual game AAM and informed and uninformed treatments of cooperative game AAM, together with linear fit and associated regression statistics. Clearly, the association between individual and cooperative game AAM is very strong, especially when we consider only the informed treatment. The slope coefficient of 0.37 indicates that cooperative game AAM increases by 0.37 when individual game AAM is higher by 1, a high enough effect to justify subjects' behavior. The t-stat is 5.4 and the R-squared of 0.24 indicates that we are able to explain 24% of the variance in cooperative game AAM by taking into account the individual game AAM. In addition to simple regression analysis, which is equivalent to correlation analysis and hence is a quantitative measure of association, we also check Kendall's Tau between the two AAMs, which relies only on observation ranks and hence is a qualitative measure of association. This also confirms a strong and statistically significant association between the two AAMs, with Kendall's Tau equal to 0.33.

Figure 4.11: Scatter plot of AAMs



4.C Simulation analysis of quantile estimates

This appendix provides a simulation analysis of a possible interpretation of quantile coefficient estimates. The simulation exercise consists of 11 subjects and 11 rounds. In each round, subjects absolute mistakes' are given by the simple relationship

$$AM_{i,t} = \alpha_i + \beta_i AAM_t^p + u_{i,t} \tag{4.8}$$

where i identifies the subject and t identifies the round, while p stands for partner. We explicitly consider heterogeneity across participants in that their average absolute mistakes and reaction to partner AAM are potentially different.

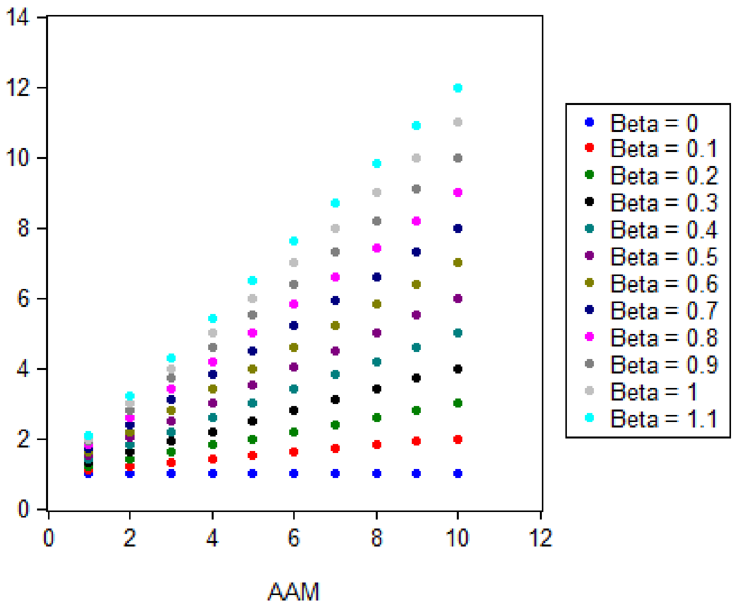
The rest of this appendix presents a graphical illustration under two possible cases: one in which the estimated quantile coefficients can be interpreted as measures of heterogeneity across participants, and one in which they cannot be interpreted in such a way.

Quantile coefficients as measures of heterogeneity. Consider a situation in which the average absolute mistake of all participants is the same, so that α_i is identical for all participants. For simplicity we set it to 1. Meanwhile, the reaction coefficients are indeed heterogeneous, with values $1, 2, 3, \ldots, 10$ and $u_{i,t}$ set to 0.[14] Figure 4.12 shows the resulting dataset of absolute mistakes, with different player mistakes coded with different colors. There are three aspects of this figure we wish to highlight. First, the figure shares a key feature with our actual dataset: the data are heteroskedastic, with variance increasing with partner AAM. Second, the mistakes for partner AAM always have the same order: the subject with the lowest

[14] Here we choose to use an (almost) deterministic "simulation" to completely highlight our conclusions. Adding noise would not materially alter them.

reaction coefficient always makes the smallest mistake (light blue dots), while the subject with the highest reaction coefficient makes the largest mistakes (light grey dots). Finally, the two observations yield the conclusion that the heteroscedasticity present in this dataset reflects the heterogeneity among participants in terms of their responses to partner AAM.

Figure 4.12: Simulated dataset I



In this dataset, estimates of quantile coefficients would correspond to the true response reaction coefficients β_i, as shown in columns 4&5 of table 4.11. Therefore, if our data were generated according to equation (4.8) with coefficients as specified, it would be correct to associate the estimated quantile coefficients with the quantiles of the actual reaction coefficients. Correspondingly, we would be able to claim that the estimated coefficients suggest that the predicted behaviour is observed among 40% of our subjects. However, as we show in the next simulation exercise, this conclusion does not hold in general.

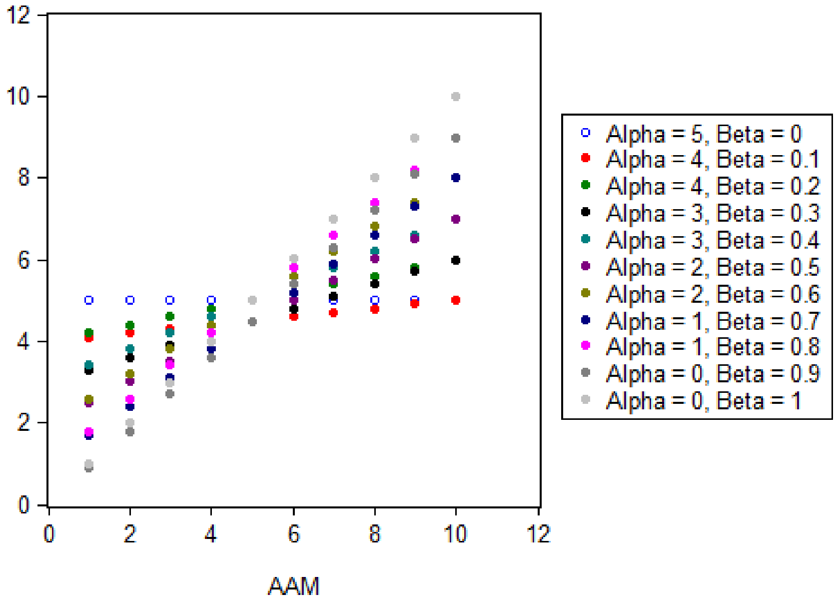
Table 4.11: Quantile coefficients estimated on simulated dataset I

Quantile	True α_i	True β_i	Estimated α_i	Estimated β_i
0.1	1	0.1	1	0.1
0.2	1	0.2	1	0.2
0.3	1	0.3	1	0.3
0.4	1	0.4	1	0.4
0.5	1	0.5	1	0.5
0.6	1	0.6	1	0.6
0.7	1	0.7	1	0.7
0.8	1	0.8	1	0.8
0.9	1	0.9	1	0.9

Quantile coefficients without relation to individual heterogeneity. One situation when the quantile coefficients cannot be interpreted as quantiles of true response coefficients is when the average mistake of each individual, α_i, is itself heterogeneous. Figure 4.13 shows the situation when the average mistake is negatively correlated with the response coefficient, while everything else is the same as in simulated dataset I. Table 4.12 shows the true coefficients α_i and β_i and quantile effect coefficient estimates. The quantile effect estimates clearly do not correspond to the true coefficients in any way. Hence, in this situation it would be incorrect to associate estimated quantile coefficients with the quantiles of actual reaction coefficients.

Note that the possibility illustrated in figure 4.13 clearly does not correspond closely to our actual data, since the heteroskedasticity is not increasing with partner AAM. A more interesting situation when the interpretation of quantile coefficients can be flawed is the situation captured in Figure 4.14. The observations in this figure were generated by equation (4.8), with $u_{i,t} = \gamma AAM_t^p \epsilon_{i,t}$ and $\epsilon_{i,t} \sim N(0, 2)$. Note that this implies an endogenous amplification mechanism:

Figure 4.13: Simulated dataset II



positive unobserved shock $\epsilon_{i,t}$ is amplified when partner AAM is high, so that the effect of the shock is increasing with partner AAM. This in turn naturally leads to heteroskedasticity as we observe in the data.

True coefficients α_i and β_i and quantile effect coefficient estimates are in table 4.13. Though the actual reaction coefficients are homogeneous across subjects, the estimated quantile coefficients increase with the considered quantile. Hence, once again, we cannot link the quantiles coefficients to the quantiles of the reaction coefficients.

Conclusion. This appendix demonstrates that the increasing quantile coefficients presented in the main text can, in principle, be reflecting two different aspects of our data generating process. They can be either reflecting the true heterogeneity in reaction coefficients across participants, or they might be reflect-

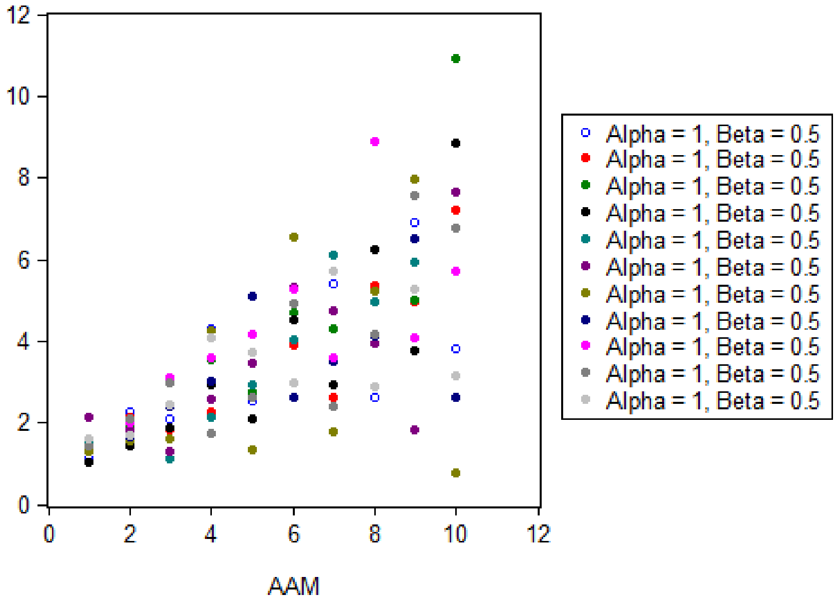
Table 4.12: Quantile coefficients estimated on simulated dataset II

Quantile	True α_i	True β_i	Estimated α_i	Estimated β_i
0.1	4.0	0.1	1.32	0.46
0.2	4.0	0.2	1.72	0.46
0.3	3.0	0.3	2.00	0.50
0.4	3.0	0.4	2.00	0.50
0.5	2.0	0.5	2.20	0.50
0.6	2.0	0.6	2.33	0.53
0.7	1.0	0.7	2.53	0.53
0.8	1.0	0.8	2.76	0.54
0.9	0.0	0.9	3.31	0.54

Table 4.13: Quantile coefficients estimated on simulated dataset III

Quantile	True α_i	True β_i	Estimated α_i	Estimated β_i
0.1	1.0	0.5	1.02	0.23
0.2	1.0	0.5	1.08	0.33
0.3	1.0	0.5	1.20	0.41
0.4	1.0	0.5	1.13	0.48
0.5	1.0	0.5	1.08	0.54
0.6	1.0	0.5	1.01	0.62
0.7	1.0	0.5	1.01	0.70
0.8	1.0	0.5	1.45	0.69
0.9	1.0	0.5	1.40	0.76

Figure 4.14: Simulated dataset III



ing the presence of an amplification mechanism, when unobserved shocks causing higher average absolute mistakes are amplified when partner AAM is high. Distinguishing between these two explanation is not possible in our context.

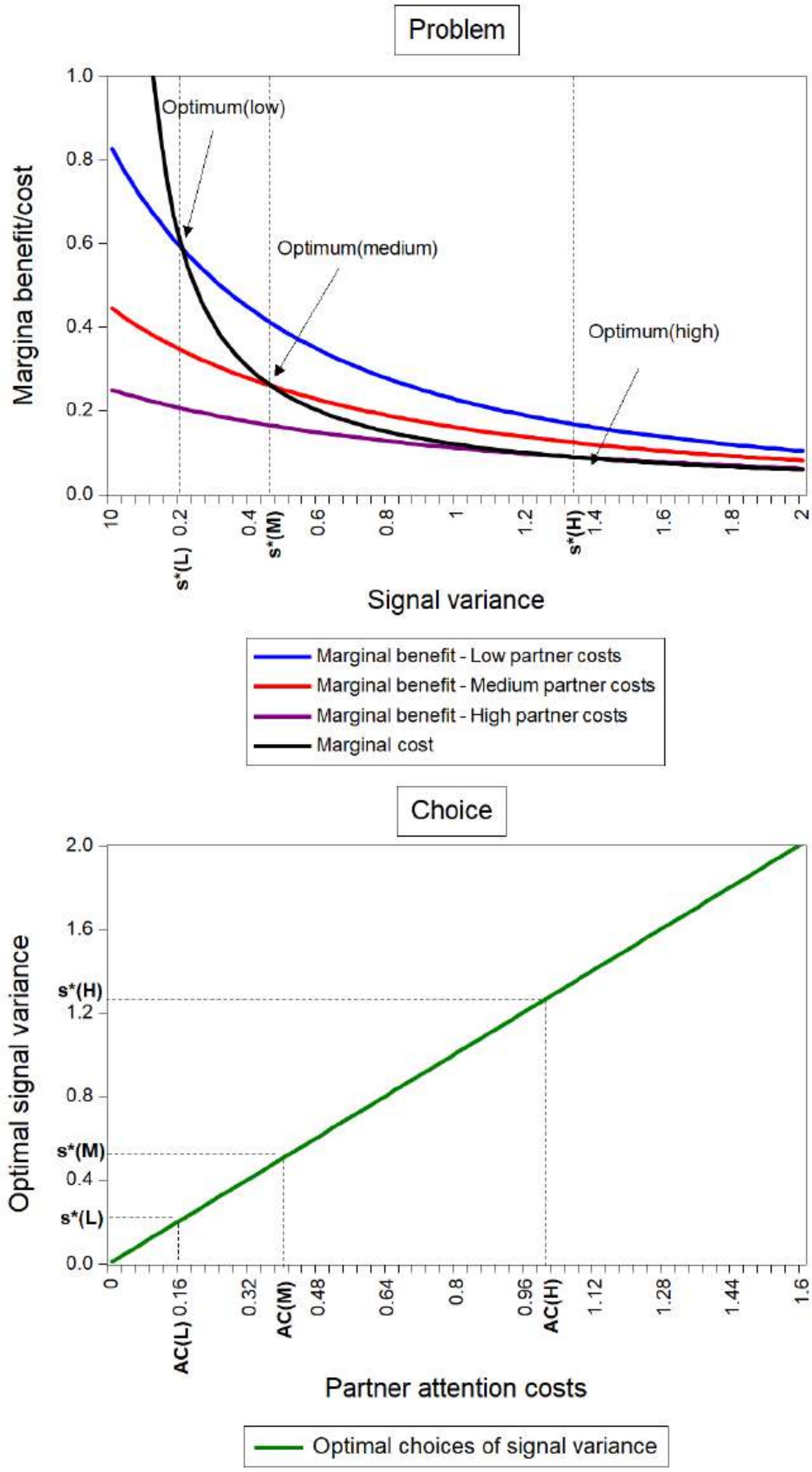
4.D Theoretical justification for Tobit models

This appendix presents theoretical justification for the use of Tobit models. It presents graphical illustrations under two possible cases: one when corner solutions would not be obtained and one when they would.

No corner solution. Figure 4.15 presents a situation in which corner solutions could not occur. It displays marginal costs (MC) as a function of the variance of noise in the signal, a choice variable for the agent in our theoretical model. In addition to the MC curve, the figure also displays three hypothetical marginal benefit (MB) lines; these three curves correspond to three different hypothetical partners with three different effective attention levels (low, medium and high). Here we associate the three different effective attentions with three different partner attention costs: the key insight here is that the expected marginal benefit is higher (i.e. the MB curve is higher) when the partner has lower attention costs, since lower attention costs should translate into more precise signal from a Sender/more precise processing of a signal by a Receiver.

Corresponding to these three levels of partner attention costs are three different optimal signal noise variances chosen by the agent. The optimum points are captured in the figure and the optimal signal noise variances are labelled as $s*(L)$,$s*(M)$ and $s*(H)$, respectively. The bottom panel of Figure 4.15 translates this into the space of optimal choice corresponding to three different values of partner attention costs (AC(L), AC(M) and AC(H)), highlighting the three outcomes captured in the top panel. As seen in the figure, higher partner attention costs translate into higher chosen signal noise variance. The key feature of the two figures is the fact that since marginal costs are infinite in the zero variance limit (i.e. $lim_{\sigma_*^2 \to 0}$MC $\to \infty$), the optimal choice of signal noise variance is positive even when the partner attention costs are zero. This corresponds to alack of corner solutions.

Figure 4.15: Problems and choices without corner solutions

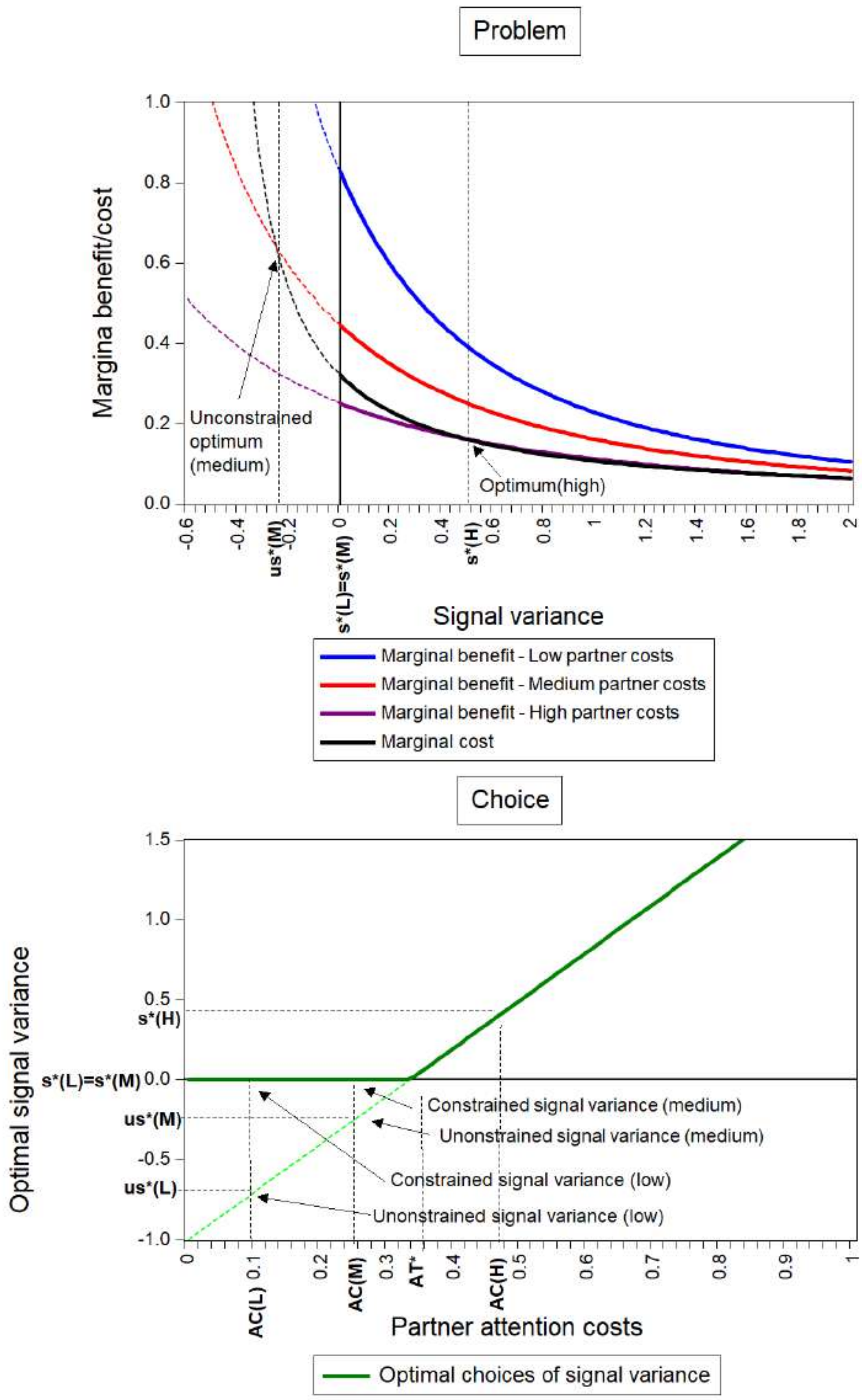


Corner solutions. Figure 4.16 presents a situation in which corner solutions can occur. As before, the figure includes a marginal cost curve and three marginal benefit curves corresponding to low, medium, and high levels of partner attention costs. The key difference from Figure 4.15 is that the marginal costs are no longer infinite in the zero signal noise variance, but are finite. While this is irrelevant when considering a situation with high partner attention costs, it translates into the same *chosen* signal noise variance with medium and low attention costs, namely zero signal noise variance. Translating this into the space of optimal signal noise variance as a function of partner attention costs (bottom panel) leads to a data generating process typical of corner solution models: while above AT^* the relationship between attention costs and optimal signal noise variance is still linear, below the threshold the two variables do not have any relationship. As the two figures highlight, if it were possible, the agent would choose variances lower than zero, but the logical/technological constraint prevents this, leading to a choice of zero signal noise variance in cases of medium and low attention costs.

As is the case in corner solution models, ignoring this non-linearity would lead us to underestimate the effect of partner attention costs on optimal signal noise variance. This can be seen by considering the effect of moving from high to medium partner attention costs, where the effect is muted by crossing the threshold, or by considering the effect of moving from medium to low attention costs, where there is no effect at all, despite the fact that the agent would optimally choose to lower his signal noise variance.

Conclusion. This appendix demonstrates that it is plausible that our data generating process is subject to corner solutions. Therefore, using a Tobit regression model is appropriate, motivating our focus on results from this model.

Figure 4.16: Problems and choices with corner solutions



Bibliography

Adrian, Tobias, Karin J Kimbrough, and Dina Marchioni. 2010. "The federal reserve's commercial paper funding facility." *FRB of New York Staff Report*, no. 423.

Agranov, Marina, Elizabeth Potamites, Andrew Schotter, and Chloe Tergiman. 2012. "Beliefs and endogenous cognitive levels: An experimental study." *Games and Economic Behavior* 75 (2): 449–463.

Altavilla, Carlo, Luca Brugnolini, Refet S Gürkaynak, Roberto Motto, and Giuseppe Ragusa. 2019. "Measuring euro area monetary policy." *Journal of Monetary Economics* 108:162–179.

Altavilla, Carlo, Lorenzo Burlon, Mariassunta Giannetti, and Sarah Holton. 2021a. "Is there a zero lower bound? The effects of negative policy rates on banks and firms." *Journal of Financial Economics.*

Altavilla, Carlo, Giacomo Carboni, and Roberto Motto. 2021. "Asset purchase programmes and financial markets: lessons from the euro area." *International Journal of Central Banking* 17 (4): 1–48.

Altavilla, Carlo, Wolfgang Lemke, Tobias Linzert, Jens Tapking, and Julian von Landesberger. 2021b. "Assessing the efficacy, efficiency and potential side effects of the ECB's monetary policy instruments since 2014." *ECB Occasional Paper*, no. 2021278.

Ambuehl, Sandro, Axel Ockenfels, and Colin Stewart. 2019. "Attention and selection effects." *Rotman School of Management Working Paper*, no. 3154197.

Andreeva, Desislava, and Miguel Garcia-Posada. 2019. "The impact of the ECB's targeted long-term refinancing operations on banks' lending policies: the role of competition."

Arrata, William, Benoît Nguyen, Imene Rahmouni-Rousseau, and Miklos Vari. 2020. "The scarcity effect of QE on repo rates: Evidence from the euro area." *Journal of Financial Economics* 137 (3): 837–856.

Bartos, Vojtech, Michal Bauer, Julie Chytilova, and Filip Matejka. 2016. "Attention discrimination: Theory and field experiments with monitoring information acquisition." *American Economic Review* 106 (6): 1437–75.

Bech, Morten, and Cyril Monnet. 2016. "A search-based model of the interbank money market and monetary policy implementation." *Journal of Economic Theory* 164:32–67.

Bech, Morten L, and Aytek Malkhozov. 2016. "How have central banks implemented negative policy rates?" *BIS Quarterly Review March.*

Beckworth, David. 2018. "The Great Divorce: The Federal Reserve's Move to a Floor System and the Implications for Bank Portfolios." *Mercatus Research, Mercatus Center at George Mason University. https://www. mercatus. org/system/files/beckworth-great-divorce-mercatus-research-v6. pdf.*

Berger, Allen N, Lamont K Black, Christa HS Bouwman, and Jennifer Dlugosz. 2015. "The federal reserve's discount window and taf programs: 'Pushing on a string?'." *Available at SSRN 2429710.*

Bernanke, Ben S. 2013. Statement by Ben S. Bernanke before the Joint Economic Committee.

Bottero, Margherita, Ms Camelia Minoiu, José-Luis Peydró, Andrea Polo, Mr Andrea F Presbitero, and Enrico Sette. 2019. *Negative monetary policy rates and portfolio rebalancing: Evidence from credit register data.* International Monetary Fund.

Boucinha, Miguel, and Lorenzo Burlon. 2020. "Negative rates and the transmission of monetary policy." *Economic Bulletin Articles*, vol. 3.

Bowman, David, Etienne Gagnon, and Mike Leahy. 2010. "Interest on Excess Reserves as a Monetary Policy Instrument: The Experience of Foreign Central Banks."

Boyson, Nicole M, Jean Helwege, and Jan Jindra. 2015. "Thawing Frozen Capital Markets and Backdoor Bailouts: Evidence from the Fed's Liquidity Programs." *Available at SSRN 2498672.*

Brunnermeier, Markus K, and Yann Koby. 2016. "The reversal interest rate: An effective lower bound on monetary policy." *unpublished paper, Princeton University.*

Caplin, Andrew, and Mark Dean. 2015. "Revealed preference, rational inattention, and costly information acquisition." *American Economic Review* 105 (7): 2183–2203.

Caplin, Andrew, Mark Dean, and Daniel Martin. 2011. "Search and satisficing." *American Economic Review* 101 (7): 2899–2922.

Caplin, Andrew, and Daniel Martin. 2014. "A testable theory of imperfect perception." *The Economic Journal* 125 (582): 184–202.

Chan, Kung-Sik. 1993. "Consistency and limiting distribution of the least squares estimator of a threshold autoregressive model." *The annals of statistics*, pp. 520–533.

Cheremukhin, Anton, Anna Popova, and Antonella Tutino. 2015. "A theory of discrete choice with information costs." *Journal of Economic Behavior & Organization* 113:34–50.

Christensen, Jens HE, and Signe Krogstrup. 2019. "Transmission of quantitative easing: The role of central bank reserves." *The Economic Journal* 129 (617): 249–272.

Christensen, Jens Henrik Eggert, and Signe Krogstrup. 2016. "A portfolio model of quantitative easing." *Peterson Institute for International Economics Working Paper*, no. 16-7.

Ciccarela, Steve, and Kamil Kovar. 2020. "Moody's Anlytics global macroeconomic model: Validation report - United States edition." Technical Report, Moody's Analytics.

Claeys, Gregorygory, Ãlvaro Leandro, and Allisan Mandra. 2015. "European Central Bank quantitative easing: the detailed manual." *Bruegel Policy Contribution*, vol. 2.

Corradin, Stefano, Jens Eisenschmidt, Marie Hoerova, Tobias Linzert, Glenn Schepens, and Jean-David Sigaux. 2020. "Money markets, central bank balance sheet and regulation." Technical Report 9289944005, ECB Working Paper.

D'Amico, Stefania, William English, David López-Salido, and Edward Nelson. 2012. "The Federal Reserve's large-scale asset purchase programmes: rationale and effects." *The Economic Journal* 122 (564): F415–F446.

Dean, Mark, and Nathaniel Neligh. 2017. "Experimental tests of rational inattention." Technical Report, Working Paper, Columbia University.

Demiralp, Selva, Jens Eisenschmidt, and Thomas Vlassopoulos. 2021. "Negative interest rates, excess liquidity and retail deposits: Banksâ reaction to unconventional monetary policy in the euro area." *European Economic Review* 136:103745.

Diebold, Francis X, and Robert S Mariano. 2002. "Comparing predictive accuracy." *Journal of Business & economic statistics* 20 (1): 134–144.

Draghi, Mario. 2014a, 5.6.2014. Introductory statement to the press conference, June 2014.

———. 2014b. "Introductory statement to the press conference, September 2014."

———. 2014c. "Monetary policy in the euro area: Opening keynote speech by Mario Draghi, President of the ECB at the Frankfurt European Banking Congress Frankfurt am Main, 21 November 2014."

Duygan-Bump, Burcu, Patrick Parkinson, Eric Rosengren, Gustavo A Suarez, and Paul Willen. 2013. "How Effective Were the Federal Reserve Emergency Liquidity Facilities? Evidence from the Asset-Backed Commercial Paper Money Market Mutual Fund Liquidity Facility." *The Journal of Finance* 68 (2): 715–737.

ECB. 2015. The impact of negative short-term rates on the money market fund industry.

Edge, Rochelle M, Michael T Kiley, and Jean-Philippe Laforte. 2010. "A comparison of forecast performance between federal reserve staff forecasts, simple reduced form models, and a DSGE model." *Journal of Applied Econometrics* 25 (4): 720–754.

Eggertsson, Gauti B, Ragnar E Juelsrud, Lawrence H Summers, and Ella Getz Wold. 2019. "Negative nominal interest rates and the bank lending channel." Technical Report, National Bureau of Economic Research.

Ennis, Huberto M, and Todd Keister. 2008. "Understanding monetary policy implementation." *FRB Richmond Economic Quarterly* 94 (3): 235–263.

Eser, Fabian, Wolfgang Lemke, Ken Nyholm, Sören Radde, and Andreea Vladu. 2019. "Tracing the impact of the ECBâs asset purchase programme on the yield curve."

Frederick, Shane. 2005. "Cognitive reflection and decision making." *Journal of Economic perspectives* 19 (4): 25–42.

Gagnon, Joseph, Matthew Raskin, Julie Remache, and Brian Sack. 2011. "The financial market effects of the Federal Reserve's large-scale asset purchases." *International Journal of Central Banking* 7 (1): 3–43.

Gill, David, and Victoria Prowse. 2016. "Cognitive ability, character skills, and learning to play equilibrium: A level-k analysis." *Journal of Political Economy* 124 (6): 1619–1676.

Green, Christopher, Ye Bai, Victor Murinde, Kethi Ngoka, Isaya Maana, and Samuel Tiriongo. 2016. "Overnight interbank markets and the determination of the interbank rate: A selective survey." *International Review of Financial Analysis* 44:149–161.

Greenlaw, David, James D Hamilton, Ethan Harris, and Kenneth D West. 2018. "A skeptical view of the impact of the fedâs balance sheet." Technical Report, National Bureau of Economic Research.

Hammermann, Felix, Kieran Leonard, Stefano Nardelli, and Julian von Landesberger. 2019. "Taking stock of the Eurosystem's asset purchase programme after the end of net asset purchases." *Economic Bulletin Articles*, vol. 2.

Hartmann, Phillip, and Frank Smets. 2015. "The transmission of the ECB's recent non-standard monetary policy measures." *Economic Bulletin* 7:32–51.

Harvey, David, Stephen Leybourne, and Paul Newbold. 1997. "Testing the equality of prediction mean squared errors." *International Journal of forecasting* 13 (2): 281–291.

Heider, Florian, Farzad Saidi, and Glenn Schepens. 2019. "Life below zero: Bank lending under negative policy rates." *The Review of Financial Studies* 32 (10): 3728–3761.

Hoffman, Elizabeth, Kevin McCabe, and Vernon Smith. 2008. "Reciprocity in ultimatum and dictator games: An introduction." *Handbook of experimental economics results* 1:411–416.

Hyndman, Rob J, and George Athanasopoulos. 2018. *Forecasting: principles and practice.* OTexts.

Jackson, Harriet. 2015. "The International Experience with Negative Policy Rates." *Bank of Canada Staff Discussion Paper*, no. 2015-13.

Jin, Ginger Zhe, Michael Luca, and Daniel Martin. 2015. "Is no news (perceived as) bad news? An experimental investigation of information disclosure." Technical Report, National Bureau of Economic Research.

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

Kahneman, Daniel, and Amos Tversky. 1973. "On the psychology of prediction." *Psychological review* 80 (4): 237.

Kovar, Kamil. 2017. "Validation of the Moody's Analytics euro zone financials model." Technical Report, Moody's Analytics.

———. 2020. "Disentangling movements in euro zone interbank interest rates during 2010-2012."

Lacetera, Nicola, Devin G Pope, and Justin R Sydnor. 2012. "Heuristic thinking and limited attention in the car market." *American Economic Review* 102 (5): 2206–36.

Le Coq, Chloe, and Jon Thor Sturluson. 2012. "Does opponents' experience matter? Experimental evidence from a quantity precommitment game." *Journal of Economic Behavior & Organization* 84 (1): 265–277.

Lyon, Don, and Paul Slovic. 1976. "Dominance of accuracy information and neglect of base rates in probability estimation." *Acta Psychologica* 40 (4): 287–298.

Marquez, Jaime, Ari Morse, and Bernd Schlusche. 2013. "The Federal Reserve's balance sheet and overnight interest rates: Empirical modeling of exit strategies." *Journal of Banking & Finance* 37 (12): 5300–5315.

Martin, Daniel. 2017a. "Strategic pricing with rational inattention to quality." *Games and Economic Behavior* 104:131–145.

———. 2017b. "Strategic pricing with rational inattention to quality." *Games and Economic Behavior* 104:131–145.

McLeay, Michael, Amar Radia, and Ryland Thomas. 2014. "Money creation in the modern economy." *Bank of England Quarterly Bulletin*, p. Q1.

Meese, Richard A, and Kenneth Rogoff. 1983. "Empirical exchange rate models of the seventies: Do they fit out of sample?" *Journal of international economics* 14 (1-2): 3–24.

Moore, Damien, Martin Wurst, and Kristopher Cramer. 2019. "Extending macroeconomic forecasts to financial varaibles: A satellite modelling approach." Technical Report, Moody's Analytics.

Ogawa, Kazuo. 2007. "Why commercial banks held excess reserves: the Japanese experience of the late 1990s." *Journal of Money, Credit and Banking* 39 (1): 241–257.

Palacios-Huerta, Ignacio, and Oscar Volij. 2009. "Field centipedes." *American Economic Review* 99 (4): 1619–35.

Park, Joon Y. 1992. "Canonical cointegrating regressions." *Econometrica: Journal of the Econometric Society*, pp. 119–143.

Pesaran, M Hashem, and Yongcheol Shin. 1998. "An autoregressive distributed-lag modelling approach to cointegration analysis." *Econometric Society Monographs* 31:371–413.

Phillips, Peter CB, and Bruce E Hansen. 1990. "Statistical inference in instrumental variables regression with I(1) processes." *The Review of Economic Studies* 57 (1): 99–125.

Plazonja, Jelena. 2018. "Persuading the Inattentive."

Poole, William. 1968. "Commercial bank reserve management in a stochastic model: implications for monetary policy." *The Journal of finance* 23 (5): 769–791.

Rognlie, Matthew. 2015. "What lower bound? monetary policy with negative interest rates." *Unpublished manuscript, Department of Economics, Harvard University (November 23).*

Rostagno, Massimo, Carlo Altavilla, Giacomo Carboni, Wolfgang Lemke, Roberto Motto, and Arthur Saint Guilhem. 2021. "Combining Negative Rates, Forward Guidance and Asset Purchases: Identification and Impacts of the ECB's Unconventional Policies."

Rostagno, Massimo, Carlo Altavilla, Giacomo Carboni, Wolfgang Lemke, Roberto Motto, Arthur Saint Guilhem, and Jonathan Yiangou. 2019. "A tale of two decades: the ECB's monetary policy at 20."

Ryan, Ellen, and Karl Whelan. 2021. "Quantitative easing and the hot potato effect: Evidence from euro area banks." *Journal of International Money and Finance* 115:102354.

Selgin, G. 2016. "Interest on Reserves and the Fed's Balance Sheet." *Testimony before the Monetary Policy and Trade Subcommittee of the House Committee on Financial Services (May 17).*

Slonim, Robert L. 2005. "Competing against experienced and inexperienced players." *Experimental Economics* 8 (1): 55–75.

Stock, James H, and Mark W Watson. 1993. "A simple estimator of cointegrating vectors in higher order integrated systems." *Econometrica: Journal of the Econometric Society*, pp. 783–820.

Valiante, Diego. 2015. "The 'Visible Hand' of the ECB's Quantitative Easing." *CEPS Working Document*, vol. 407.

Vogel, Edgar. 2016. "Forward looking behavior in ECB liquidity auctions: Evidence from the pre-crisis period." *Journal of International Money and Finance* 61:120–142.

Wu, Jing Cynthia, and Fan Dora Xia. 2016. "Measuring the macroeconomic impact of monetary policy at the zero lower bound." *Journal of Money, Credit and Banking* 48 (2-3): 253–291.

———. 2020. "Negative interest rate policy and the yield curve." *Journal of Applied Econometrics* 35 (6): 653–672.