Conditionality in Official Lending:
Compliance through strategic assessment[*]

Daragh Clancy[†]

Aitor Erce[‡]

Andreja Lenarcic[§]

Ramon Marimon[¶]

September 3, 2021

Abstract

Countries must comply with loan conditions in order to receive official-sector financial assistance. Although critical for the success of official lending, there is little evidence of what makes for an effective design of such conditions. Using a unique dataset detailing compliance with conditionality in euro area programmes, we provide such evidence. We show compliance is more likely for conditions with explicit numerical targets. We study the drivers of the official lenders' decision to assess a condition, and whether this decision affects debtors' ability to meet the condition being evaluated. We find that programme revisions, throught which official lenders agree to modify the assesment schedule, help debtor countries meet the conditions. Our results show that, from the perspective of boosting compliance, the design of official-loan conditionality should be brought closer to the state-contingent approaches employed in the theoretical literature.

Keywords: Conditionality; Official lending; Sample selection

JEL Classification: C21; E61; F33; F34

[†]Email: daragh.clancy@centralbank.ie; Central Bank of Ireland.

[‡]Email: erce.aitor@gmail.com; Navarra Public University, LUISS University.

[§]Email: Andreja.lenarcic@gmail.com; International Monetary Fund.

[¶]Email: Ramon.Marimon@eui.eu; European University Institute, UPF - Barcelona GSE, CEPR and NBER.

1 Introduction

When market access for a sovereign borrower is impaired, it often requests financial support from official crisis lenders such as the International Monetary Fund (IMF). The provision of this support is usually contingent on the implementation of a set of policies aimed at addressing the weaknesses that contributed to the crisis. Countries must comply with these conditions in order to receive disbursements of funds. Conditions also help ensure the official lender can recoup the loaned funds. Thus, the design of an adequate set of conditions regarding policy reform is critical for the success of official lending.[1]

In this paper we study the likelihood of compliance with official-loan conditionality. In contrast with the existing literature, which largely focuses on IMF conditional lending, we use a unique dataset with condition-level information on the various official lending programmes provided during the recent European sovereign debt crises.[2] The granular nature of our data allows us to decompose the compliance process into the creditor's decision to assess a condition, and the debtor's decision to comply with it. We use a selection model (Heckman, 1979) to jointly model these two decisions, and to study the extent to which the timing of the assessment has a systematic effect on compliance with conditionality. The strategic interaction between these two decisions has, to the best of our knowledge, never been analyzed together.

We show that analyzing creditors' assessment and debtors' compliance decisions jointly is crucial in understanding the drivers of compliance with conditionality. We find that compliance rests on a relatively small number of critical conditions. In particular, conditions with explicit numerical targets. We also find evidence of continued negotiation throughout the programme, with official lenders and debtor counries agreeing changes in conditions that help boost compliance.[3]

---

[1]A key finding from the theoretical literature analyzing official lending is the need to find a balance on loan conditionality that provides the right incentives to both creditors and debtors. Sachs (1989) argues for a limited role of conditionality due to weak enforcement. Similarly, (Fink and Scholl, 2016) argues that the intensity of conditionality lowers borrowers willingness to participate in financial assistance programmes. On the other hand, (Muller et al., 2019) notes that light conditionality increases moral hazard and makes creditors less likely to provide assistance, and Marchesi and Thomas (1999) point out that conditionality serves as a signal of a country's willingness to reform. Diwan et al. (1992) find that a loan must be sufficiently large to convince a constrained government to accept conditions with immediate adjustment costs and eventual benefits, while Boz (2011) shows that conditions can help make official-sector loan contracts more enforceable than those from the private sector.

[2]The European Stability Mechanism provides a publically-available database of the six euro area financial assistance programmes carried out between 2010 and 2018. A first contribution of this paper is the correction and extension of this database that renders it usable for analytical purposes. We provide an overview of euro area official-sector financial assistance process in Appendix A, and discuss an example loan condition highlighting some of the main features of the database in Appendix B. In Appendix C we describe some of the main elements of the construction of this analytical database. Our analytical database is available here.

[3]Counter to the emphasis in programme documents, compliance is more likely for conditions focused on fiscal

---

1

We extract a number of lessons from our econometric analysis that could help enhance the design of official-loans. First, the focus on conditions set ex-ante makes loan conditionality difficult to enforce.[4] As the programme evolves and lenders and debtors better understand the evolving macroeconomic and financial situation, they engage in a process of delaying the evaluation of specific conditions (often including some redrafting the same).Using a selection model, we provide evidence that conditions are not assessed until compliance is more likely. One interpretation of these results is that the focus on ex-ante forward-looking conditions makes official loans too rigid, forcing official lenders and debtor governments to find innovative ways to gain flexibility.[5] Second, compliance is greater for quantifiable loan conditions. These conditions were assessed less regularly, were subject to less delays, and were generally complied with. Third, efforts to boost the productive capacity of the economy should be less prescriptive. We find a robust a negative relationship between the total number of conditions and likelihood of compliance. [6]

Taken together, our results indicate that when designing conditionality for official loans, lenders should focus on fewer conditions, stated in quantifiable terms. They should also design in-built mechanisms for transparently adjusting these conditions as the situation evolves. Such move towards transparent ex-post adjustments of conditionality would bring the practice of official lending closer to an state-contingent approach.[7]

Literature review:

There are two broad strands to this literature. The first analyses the economic effects of conditional loans (Conway, 1994; Przeworski and Vreeland, 2000; Barro and Lee, 2005; Mody and Saravia, 2006; Erce and Riera-Crichton, 2015; Alesina et al., 2020). The evidence from this literature is mixed. Some studies find a negative effect on growth from entering a financial assistance programme, while others find a positive effect, especially when the countrys fundamentals have not yet deteriorated significantly. A key issue in this literature is sample selection,

---

and financial sector stabilisation. Compliance is less likely for structural reforms aimed at enhancing economies' productive capacity.

[4]This is different from ex-ante eligibility criteria required for access to precautionary official loans.

[5]Although not part of our econometric exercise, which focuses on the timing of the assessment, we also find that redrafted conditions get scrutinized more often but are on average less likely to be be complied with.

[6]According to Sachs (1989), given the complexity of structural adjustment and the difficulty of enforcing conditionality, he believes conditional loans are more likely to be successful if macroeconomic stabilisation receives priority over structural reforms.

[7]This would be in line with recent theoretical literature such asAbrahám et al. (2019), who develops a model of a Financial Stability Fund serving a union of sovereign countries. Abrahám et al. (2019) shows substantial welfare gains from improved risk sharing through state-contingent lending compared to the use of conventional sovereign bonds, even in the face of limited enforcement and moral hazard constraints.

with the use of political economy variables often used as instruments to overcome endogeneity concerns.[8]

Dreher (2006) uses information from the IMF's Monitoring of Fund Arrangements (MONA) database on the percentage of conditions met to analyse the effect of conditionality on economic growth. He finds that while conditional loans reduce growth, this effect is smaller when there is better compliance with conditionality. Steinwand and Stone (2008) provide a review of the literature assessing the compliance with, and effect of, conditionality. They conclude that often overlooked aspects such as the influence of major IMF shareholders, the Funds own organizational imperatives, and domestic politics within borrowing countries are crucial in understanding the conditions under which IMF lending can have negative or positive effects.

Our paper is more closely related to the second strand, which uses information on loan conditionality itself. Mourmouras et al. (2003) assess the implementation of IMF programmes using measures of programme interruptions, compliance with conditionality, and the share of committed funds disbursed. They construct a number of indicators that provide quantitative information on implementation rates by type of condition.[9] Due to their small sample size, they use the MIMIC (multiple indicators, multiple causes) model, random-effects IV and Amemiyas Tobit to evaluate the importance of political conditions, IMF effort, conditionality, initial and external conditions for programme implementation. They find that borrowers political economy (special interests, political instability and inefficient bureaucracy) weaken programme implementation. They also find no evidence that IMF staff effort and the extent and structure of conditionality materially influences programme prospects.[10] Of particular note is their finding that the imposition of large numbers of prior actions had limited success.

We structure our paper as follows. Section 2 describes our dataset, while Sections 3 and 4

---

[8]The literature presents various methods to deal with the selection problem. Most studies pursue an instrumental variables approach (Barro and Lee, 2005) or some variant of Heckman estimator (Mody and Saravia, 2006), matching (Hutchison, 2004; Essers and Ide, 2019) or re-weighting schemes (Ahokpossi, Allain, and Bua, Ahokpossi et al.). All of these approaches have benefits and drawbacks. The challenge with the instrumental variables approach is to find variables arguably affecting the probability of participation, but not economic outcomes other than through their impact on participation. Estimating the participation equation and then including the inverse Mills ratio (Heckman, 1979) depends on assumptions about the distribution of error terms. The problem of finding the correct counterfactual haunts the matching and re-weighting approaches, where matching of "treatment" and "control" groups, or the corresponding re-weighting of observations, would only result in unbiased estimates if the decision to enter a conditional loan is accounted for by the matching/re-weighting procedure. Re-weighting schemes acknowledge this limitation and focus on reducing selection on observables.

[9]Each of these indices captures an important dimension of programme implementation, including significant programme stoppages, the proportion of approved assistance actually delivered and the actual duration of the programme compared to the scheduled one.

[10]IMF staff effort is the estimated dollar cost of IMF programmes, representing the hours spent by staff on programme implementation (both preparation and supervision) and the average salaries of staff by grade.

describe our empirical methodology and results respectively. Section 6 concludes.

2 Data

The European Sovereign Debt Crisis saw several euro area member states suffer difficulties securing market financing at sustainable rates. Given the lack of crisis-lending experience in European institutions, the authorities initially leaned heavily on IMF expertise. As the crisis evolved, euro area policymakers developed their own crisis resolution framework. This placed greater emphasis on fiscal consolidation and structural reforms, compared to IMF financial assistance programmes that focus more on exchange rates and monetary policy.[11] The intergovernmental agreements establishing the temporary European Financial Stability Facility (EFSF) and the permanent ESM both explicitly state that financial assistance is subject to *strict conditionality.*

We analyse conditionality in these programmes using the EFSF \ESM programme database, which contains a rich set of cross-sectional (six programmes in five countries) and time series (2010-2018) data.[12] The database contains lending information (loan disbursements, principal repayments, past interest, and fee payments) and data on programme monitoring (loan conditionality and economic and financial indicators, including real-time nowcasts \forecasts).[13] Loan conditions aimed to tackle underlying issues that contributed to the need for financial assistance, and are grouped into sectors. Broadly speaking, the conditions targeted an improvement in the public finances ("Fiscal" and "Fiscal-Structural"), financial stability ("Financial") and the productive capacity of the economy ("Structural Labour", "Structural Product" and "Other").

---

[11]Corsetti et al. (2020) provide a detailed description of this institutional development process, and study the effect of different lending terms on market borrowing costs and fiscal performance. We discuss the specificities of the euro area official-sector financial assistance process in Appendix A

[12]The countries are Cyprus, Greece, Ireland, Portugal and Spain. Spain represents a special case, with all loan conditions exclusively focused on the Spanish financial sector (i.e. there was no macro-fiscal conditionality). The loan conditions for Spain were all continuous (i.e. scheduled for repeated assessment), they were always assessed and they remained unchanged for the entire duration of the financial assistance programme (i.e there was no redrafting). Just over 50% of the conditions were fulfilled by the first programme review, while just over 90% of them were completed by the end of the programme. The institutions deemed the remaining conditions as not relevant because a group of banks recapitalised without State aid. The highly-specific nature of the Spanish programme means we exclude it from our empirical analysis.

[13]A value added of the database is its granularity, with 2,279 unique conditions tracked throughout the programmes. Multiple assessments of individual conditions means that there are 8,152 assessment observations in total. We provide an example loan condition highlighting the main features of the database in Appendix B and a description of how we rendered the publically-available database useful for empirical analysis in Appendix C.

The sectoral distribution of conditions reflects the degree to which these problems affected each country. The Cypriot and Irish programmes were focused on financial sector policies, while the Greek and Portuguese programmes were largely fiscal (Table 1, panel a). The proportion of conditions with explicit numerical targets varies widely across countries and sector ("Numeric", panel b). Almost 20% of Portuguese conditions have an explicit numerical target, while for Ireland it is well under 10%. In principle, all conditions are subject to assessment during the lifetime of the financial assistance programme. While some are due for one-off assessment, others are scheduled for repeated assessment throughout the duration of a programme. This also differs considerably across country and sector ("Continuous", panel c), with the majority of Greece's fiscal-structural targets being subject to continuous assessment and Ireland having none in most sectors. A large proportion of conditions were redrafted throughout the duration of a financial assistance programme ("Redrafted", panel d), ranging from nearly 65% in Cyprus to just over 30% in Ireland.[14]

Although seemingly straight forward conceptually, in reality the conditionality assessment process proved quite complicated. One complication is the division of some conditions into multiple sub-conditions to track the fulfilment of specific objectives. This leads to multiple assessments of the same aggregate condition. We assign a unique assessment for conditions with multiple subcomponents (and therefore multiple assessments).[15] To do so, we follow IMF (2012) and assign a value of 0 to unmet conditions, 0.5 to partially met conditions and 1 to completely fulfilled conditions. We then calculate an average across assessments of the subconditions and assign a final (i.e. unique) compliance category for the aggregate condition by comparing the average to simple thresholds of 1/3 and 2/3.[16]

In Table 2 we demonstrate the distribution of assessments across condition categories. Although there is substantial heterogeneity across country and sector, it is clear that a substantial per-

---

[14]The programme database differentiates between significant redrafts, which materially alter the condition or its terms (e.g. a change in the deadline for completion), and minor wording changes that do not change the substance. We only consider significant redrafts in the summary statistics presented in Table 1 and our empirical analysis.

[15]This reduces the number of unique policy conditions to 2,034 and the assessments to 6,562, as we convert 2,165 assessments of multiple sub-conditions into 586 unique values. Our aggregation of multiple subcomponents of a given condition has little material effect on the aggregate structure of the data.

[16]We have only three possible values for each subcondition: 0, 0.5 and 1. When we take the average across all subconditions, we get a value between 0 and 1. Any value less than 1/3 is classified as 0, between 1/3 and 2/3 as 0.5 and above 2/3 as 1. These are, therefore, quite naive thresholds. In an alternative approach, we derive the thresholds by minimizing the distance between empirical distributions before and after the aggregation of sub-conditions. This approach delivers a lower threshold greater than 1/3 but a very similar upper threshold. Since our analysis focuses on completely fulfilled conditions, the upper threshold is of far greater importance. Therefore, our results are not affected by our choice of method for deriving thresholds.

centage of conditions were never assessed ("Total", panel a). Indeed, for Ireland and Portugal the majority of conditions were never assessed. The assessment rate of numerical conditions is larger for all countries except Portugal (panel b), while the assessment rate of continuous (panel c) and redrafted (panel d) conditions is generally no greater than for regular conditions. The programme database includes the dates for which a given condition is scheduled for assessment. Therefore, we can also identify cases where a condition was assessed as scheduled ("On time") or with a delay ("Delayed").[17] Again this varies widely across countries and sectors. Around one third of Greek conditions were assessed only after a delay, while this occured in only 1% of Irish conditions.

In Table 3 we present the distribution of compliance outcomes for assessed conditions. The database contains a range of possibilities for compliance, representing the degree to which the condition is complied with. For the descriptive statistics and our empirical analysis, we only consider conditions categorised as "Observed". This represents conditions that are fully complied with.[18] While just over 50% of Cypriot conditions were fulfilled, the compliance rate was less than 20% in Ireland and Portugal (panel a). There is substantial sectoral heterogenity in the compliance rate, with over 60% of financial sector conditions and less than 5% of structural-labour conditions fully complied with in the Irish programme, for example. The compliance rate is generally greater for numerical (panel b) and continuous (panel c) conditions, while it is about the same for the redrafted conditions as for normal conditions (panel d). Finally, conditions assessed with a delay generally have a lower compliance rate (panel e).

3   Empirical approach

Our interest lies in understanding the determinants of compliance with conditionality. But our data overview shows that programme compliance is not a straightforward process. Creditors can only deem a condition as fulfilled if they assess it first. In fact, many conditions are only assessed after a delay (and often a redraft) or are never assessed. This has an important econometric implication, as it implies that our dependent variable, the ability to comply with conditionality,

---

[17]We consider an assessment as delayed relative to the expected due date in the latest MoU. In the relatively few cases with several due dates, we consider the first one, as later dates typically reappeared in the redrafts of the same condition and are thus relevant due dates for those redrafted versions of conditions.

[18]Therefore, conditions assessed as "Largely observed", "Partial/Ongoing", "Ongoing", "Not relevant" are also included in the "Not observed" categorisation. This is a conservative choice, but greatly reduces the influence of subjectivity in the fulfillment category that is the key variable of interest in our analysis by reducing ambiguity over the extent to which a condition is complied with.

can be missing as a result of another process, the willingness of creditors to assess it. As described in Greene (2003), this selection process can bias the estimation of the determinants of compliance with conditionality if the decision to assess a condition is not random.

In order to account for this bias, we use a two-step approach to correct for sample selection (Heckman, 1979). More specifically, we want to estimate the effect of specific factors on the success ($s_j = 1$) in meeting a pre-agreed condition j. A model for whether a condition successfully passes an assessment could look like:

$$s_j = \beta x_j + \epsilon_j \tag{1}$$

where $x_{j,t}$ is a set of explanatory variables and $\epsilon_{i,t}$ is an error term.

However, creditors can decide not to assess a condition at the pre-agreed time or to even ignore it completely. A model of the decision to assess is:

$$d_j = \delta z_j + \nu_j \tag{2}$$

where z_j is a set of explanatory variables and ν_i is an error term.

In this framework, (s_j) can only be observed if $d_j = 1$. Therefore, if $E(s_j|z_j, d_j = 1) \neq 0$, the estimates are biased and require a correction to capture selection effects.

Following Heckman (1979), we first estimate $\hat{\delta}$ from a probit regression and use it to obtain an estimate of the Inverse Mill's Ratio, $\hat{\lambda}_j$. We then use $\hat{\lambda}_j$ to control for sample-induced endogeneity, i.e. the extent to which the decision to evaluate influences also the result of the evaluation.[19] Therefore, the fulfillment equation we estimate becomes:

$$s_j = \beta x_j + \gamma \hat{\lambda}_j + \epsilon_j. \tag{3}$$

In order to model the drivers of the *assessment decision*, we estimate Equation (2) where d_j is a binary variable that takes a value of one when a loan condition is assessed and zero otherwise. We start from the assumption that creditors set conditions (i.e. policy measures) and deadlines

---

[19]We use bootstrapping to compute the standard errors of the fulfillment decision regression. See Cameron and Trivedi (2005) for a discussion.

within which to implement them according to their current expectations for the current and future state of the economy. National authorities must fulfill the conditions by a set date. At the due date, the official creditors might decide to postpone the assessment. We assume they based their decision on the information set available to them at the time an assessment is due. As control variables, we include:

- a dummy for conditions with an explicit numerical target;

- a dummy for conditions set for continuous assessment;

- a dummy for conditions that were previously redrafted;

- the expected time to assessment (in quarters);

- the number of conditions imposed in the MoU a condition was set;

- the interaction of the numerical and continuous dummies;

- the institutions' nowcasts of annual real GDP growth;

- the nowcast error;

- the 10-year sovereign bond spread in the current quarter;

- the upcoming gross disbursement relative to the GDP nowcast;

- dummies for different policy sectors.

We include the nowcasts as a proxy for the institutions' *expectations* of the economic situation at the time of the programme design and in subsequent reviews. The inclusion of nowcast errors, which we define as the difference between the nowcasts and their ex-post realisation, allow us to ascertain the impact of unexpected innovations to key assumptions underpinning the conditionlaity design.[20] The 10-year sovereign spread (to the German Bund) provides a signal of market scrutiny. The upcoming disbursement, relative to the nowcast of GDP, provides information on the financial relevance of compliance. The number of conditions imposed serves as a proxy for programme complexity, while the dummy for explicit numerical targets allows us to ascertain if this formulation of conditionality increases the likelihood of assessment.

In turn, in order to model the *fulfilment decision*, we assume that compliance with conditionality

---

[20]Another interpretation of the forecast erros is that they measure the optimism of the official sector. A notable example of this alternative is the underestimation of the size of fiscal multipliers, which resulted in overly optimistic growth forecasts during the European sovereign debt crisis (Blanchard and Leigh, 2013).

largely depends on the effort of the borrowing country. In analysing the compliance decision, we set the dependent variable equal to one if the condition is fulfilled and zero for any other degree of compliance[21]. This allows us to focus only on what drives the borrowers efforts to fulfill the conditions set in a given MoU. It also allows us to focus on the information set at the time of a given MoU's drafting (i.e. at the "design" of the condition).

To assess the drivers of the fulfillment decision, we estimate Equation (3) where s_j is a binary variable that takes a value of one when a loan condition is fulfilled and zero otherwise. The statistical significance of explanatory variables in the selection equation, the inverse Mills ratio $\hat{\lambda}_j$ and the correlation between ϵ_j and ν_j give us an indication whether there is a sample selection issue or not (Certo et al., 2016).

We drop the GDP nowcast as an explanatory variable in our fulfillment equation, thereby serving as an exclusion restriction.[22] In order to be a useful instrument, a variable must affect the decision to assess, but not to decision to fulfill. Comfortingly, this variable is highly significant in the assessment regression and insignificant when included in the fulfillment regression[23]. Our rationale for why this is a good restriction is that while the creditors' expectations for economic growth may affect their decision to assess a condition, it is unlikely to influence the debtors' decision to fulfill a condition, over and beyond any unexpected surprises, which the regression also controls for.

We estimate Equations (2) and (3) treating our data as a pooled cross section. Although our dataset contains programmes from different countries that evolve over time, treating our dataset as a panel is inappropriate for a number of reasons. First, since data collection only begins at each programme's initial MoU negotiation mission, it is highly unbalanced.[24] Second, there is irregular spacing in the time series component of the dataset, as there is at least one quarter in each programme during which no review took place.[25] Finally, and most importantly, an

---

[21]Creditors only have discretion over the assessment of conditions that are not fully complied with (i.e that fall into the categories "not fulfilled", "not relevant", "ongoing" or "partially fulfilled"). In contrast, they can only categorise fulfilled conditions as "fulfilled", so there is no room for ambiguity.

[22]If the explanatory variables included in the selection equation (i.e. our assessment decision) and outcome equation (i.e. our fulfilment decision) are identical, the Heckman correction relies on the nonlinearity of the inverse Mills ratio for identification. While not always necessary (Wilde, 2000), exclusion restrictions can help reduce this reliance (Puhani, 2000).

[23]This holds whether we use the nowcast set at the MoU, as is standard for our assessment regressions, or the nowcast set at the CA, as is standard for our fulfillment regressions.

[24]Table 4 shows the structure of our database. The "M"s represent MoUs, while the "CA"s represent compliance assessments. See Appendix A for a general description of the financial assistance process and Appendix B for a specific example that illustrates the evolution of a loan condition throughout the programme.

[25]This is particularly problematic for the estimation of dynamic panels. See Millimet and McDonough (2017) and Sasaki and Xin (2017) for approaches to overcome the issues with unequally spaced panel data.

observation for a condition is only present if some action is taken (i.e. it is assessed, redrafted or fulfilled) in a given programme review. This means that many conditions are unobserved for substantial parts of the sample, with largely different conditions in each period and the majority of conditions never repeated, rendering analysis with panel data techniques ineffective.

We ensure the robustness of our results by also including year and country dummies in some regression specifications. Year dummies should capture period-specific factors that may have influenced the assessment or fulfillment decisions, over and beyond the included explanatory variables. Country dummies should capture any time-invariant within-country effects, such as national or institutional preferences and competence.

4   Results

In Tables 5 and 6 we report the results from the regressions in Equations (2) and (3) respectively.[26] The dependent variable in these equations has a value of one when creditors assess the condition, regardless of whether it does so on time or with a delay (i.e. even if this assessment occurs later than originally scheduled). In the tables we show several specifications, to understand how the results change as we consider different factors. However, we only comment on the results with sector-specific, year and country dummies. This is column 5 in each table. In the fulfilled regression output tables, we also include a column that omits the inverse mills ratio. This demonstrates the effect of not controlling for sample selection.[27] We allow for clustering in the standard errors at condition level, as a sizable portion of conditions are assessed on more than one occasion and the (seemingly) seperate decisions to assess and fulfill on each occasion may be correlated.[28]

Previous redrafts, larger upcoming disbursements, higher GDP growth nowcasts and sovereign bond spreads all increase the likelihood of assessment (Table 5). The number of conditions and whether the condition is a fiscal measure are negatively associated with the decision to assess. Debtors are more likely to fulfill conditions with explicit numerical targets, that are scheduled for continuous assessment (Table 6). While conditions in all sectors are significant, fiscal measures stand out as particularly important. A longer expected time to assessment and

---

[26]The coefficients on the dummies are omitted as our interest lies in whether our explanatory variables and not on the specific effect of a given period or country.

[27]We provide the results without the inverse mills ratio for all regression specifications in Appendix D.

[28]We also run these specifications without clustered standard errors and the results are largely unchanged.

a larger total number of conditions make it less likely a condition is fulfilled.

The inverse Mills ratio is statistically significant, providing evidence that correcting for sample selection is appropriate.[29] The positive cofficient on the inverse Mills ratio indicates that the non-random behaviour of creditors when deciding whether to assess a condition helps boosts the likelihood of compliance. We discuss this important aspect of our results in more detail in Section 5.

5 Strategic assessment of loan conditions: A discussion

Our results show a strongly significant and positive sign for the mills ratio on the second stage, and also that lenders consider the macro-financial environment when deciding whether to assess a given loan condition. This points towards an strategic timing of the assessment, which is set to influence positively the likelihood of fulfillment. Greater flexibility on the part of creditors boosts overall compliance with loan conditionality.

We have seen the importance of this strategic behaviour for compliance. But what is its underlying motivation? One possibility could be that lenders employ discretion over the timing of assessment to take advantage of any new information that arrived after they set the condition. This would be in the spirit of the official lender in Abrahám et al. (2019). Another, less favourable, interpretation could be that lenders are unsure whether the conditions they set are appropriate, and strategically time assessments to minimise external scrutiny of the compliance process.[30] This would resemble a principal-agent relation in the spirit of ?, whereby the lenders face a trade-off between providing the debtor with incentives to exert optimal effort and incentives to communicate transparently. By delaying a large portion of assessments, lenders can buy themselves time until debtors comply with the conditions.

In order to figure out the extent to which these different motivations drive delays in assessment, next provide the results from a regression analysing the drivers of the decision to delay assessment. In this case, the dependent variable has a value of one only if a condition is assessed with a delay. We find that the expected time to assessment, the total number of conditions,

---

[29]In column 6, we display the results for a regression excluding the inverse Mills ratio. Redrafted conditions and the GDP nowcast error are now statistically significant. More importantly, the statistical significance of the upcoming disbursement suggests a negative relationship between financial incentives and the decision to fulfill.

[30]This could be, for example, because lenders publically demand that all conditions are complied with, but privately believe that only a select number are truly relevant.

the GDP growth nowcast error and if it is a fiscal measure make a condition less likely to be assessed with a delay (Table 7).

The GDP growth nowcast and its error, and the size of the upcoming disbursement affect the likelihood the assessment of a condition is delayed. The positive sign on the GDP growth nowcast and negative sign on the nowcast error are supportive of incoming information affecting creditors' decisions to assess (as in Abrahám et al. (2019)). Instead, the large and positive coefficient on the upcoming disbursement and the statistical significance on the negative coefficient for the interaction term between numerical and continuous conditions are supportive of the principal-agent relation. Therefore, it is difficult to be definitive as to what was the true underlying motivation for strategic behaviour, as both elements appear to have played a role.

6 Conclusion

We find that compliance with official-loan conditionality rests on a relatively small number of critical conditions. Compliance is more likely for conditions with explicit numerical targets. We also find evidence of continued negotiation throughout the programme, with official lenders and national authorities agreeing changes in conditions that help boost compliance.

There are a number of lessons from our analysis that could improve the design of official-loan conditionality. First, the focus on ex-ante conditionality makes the programmes difficult to enforce. As the programme evolves and the institutions and national authorities better understand the underlying macroeconomic and financial situation, they engage in a process of redrafting the conditions. We find that these redrafts get scrutinised more often than other conditions, but are on average less likely to be be complied with. We also provide some evidence that conditions are not assessed until compliance is more likely. These results demonstrate that the focus on ex-ante conditions makes financial assistance programmes more opaque, as the institutions and national authorities have to find innovative ways to gain flexibility.

Second, compliance is greater for quantifiable loan conditions. These conditions were assessed less regularly, were subject to less redrafts, and were generally complied with. Third, efforts to boost the productive capacity of the economy should be less prescriptive on individual structural reforms. These made up the vast majority of conditions, and we demonstrate a negative

relationship between the total number of conditions and likelihood of compliance.

Taken together, our results indicate that when designing conditionality for official loans, lenders should focus on fewer conditions, stated in quantifiable terms, that aim first and foremost at macroeconomic and financial stabilisation. They should have an in-built mechanism for transparently adjusting these conditions as the financial assistance programme evolves. Finally, they should focus on the aggregate outcome of aims to boost economic productive capacity rather than prescribe explicit reforms to undertake.

These changes would bring the practice of financial assistance programmes closer to the optimal design emerging from theoretical work (Abrahám et al., 2019). Such models allow for an examination of the incentives facing debtors and creditors, and simulations of how changes in the intensity of conditionality affect these trade-offs. This work demonstrates the benefits of a contract that features state contingent ex-post conditionality, i.e. a contract where conditions change in a foreseeable way, depending on the outcomes and effort of the borrower.

Future work could use the results from our empirical analysis to improve the design of such state-contingent contract models, by explicitly incorporating these features. This modelling framework would enable the use of the stylised facts we have uncovered to assess how they affect optimal programme design, moral hazard considerations, and the attractiveness of the financial assistance for debtor countries. Simulating different alternatives in such a model environment provides valuable insights that could enhance future official-sector financial assistance programme design and avoid repeating past mistakes.

Our analysis also provides some lessons for the design of post-pandemic international support schemes. For example, the European Union has created a Recovery and Resilience Facility to provide loans and grants to support reforms and investment in Member States. To access these funds, Member States must prepare recovery and resilience plans, in which they specify their plans for the accessed funds. The European Commission (EC) will assess these plans at the outset and also monitor their implementation throughout the 2021-2026 period, with disbursements conditional on compliance with pre-agreed targets and milestones.[31]

Our results suggest that the EC and national authorities should give greater priority to macroe-

---

[31]This process is therefore closely related to a financial assistance programme, albeit with the very important difference that no country has lost market access. Consequently, there is not the same urgency to comply with the loan conditions to receive further disbursements. Another difference is that the national authorities are primarily responsible for the design of the plans, although technical assistance is available from the EC upon request.

conomic and financial stabilisation in the initial phase of their recovery and resiliance plans. Longer-run elements of the plans should focus on a relatively small number of critical reforms and investments, while providing sufficient flexibility to adapt the plans as necessary. This approach should help boost absorption rates, which are typically slow for EU structural funds (Darvas, 2020). Initial plans seem to contain a significant amount of detail on numerous individual reforms and investment projects. It is possible that the greater ownership of these plans, which are entirely produced by the national authorities, will boost compliance with stated conditions. If not, there is a risk of repeating past mistakes.

References

Abrahám, A., E. Cárceles Poveda, Y. Liu, and R. Marimon (2019). On the optimal design of a financial stability fund.

Ahokpossi, M. C., L. Allain, and G. Bua. A constrained choice? impact of concessionality requirements on borrowing behavior. Technical report.

Alesina, A. F., D. Furceri, J. D. Ostry, C. Papageorgiou, and D. P. Quinn (2020). Structural reforms and elections: Evidence from a world-wide new dataset. Technical report, National Bureau of Economic Research.

Barro, R. J. and J.-W. Lee (2005). Imf programs: Who is chosen and what are the effects? *Journal of Monetary Economics 52*(7), 1245–1269.

Blanchard, O. J. and D. Leigh (2013). Growth forecast errors and fiscal multipliers. *American Economic Review 103*(3), 117–20.

Boz, E. (2011). Sovereign default, private sector creditors, and the ifis. *Journal of International Economics 83*(1), 70–82.

Calmfors, L. and S. Wren-Lewis (2011). What should fiscal councils do? *Economic Policy 26*(68), 649–695.

Cameron, A. C. and P. K. Trivedi (2005). *Microeconometrics: methods and applications.* Cambridge university press.

Certo, S. T., J. R. Busenbark, H.-s. Woo, and M. Semadeni (2016). Sample selection bias and heckman models in strategic management research. *Strategic Management Journal 37*(13), 2639–2657.

Conway, P. (1994). Imf lending programs: Participation and impact. *Journal of Development Economics 45*(2), 365–391.

Corsetti, G., A. Erce, and T. Uy (2020). Official sector lending during the euro area crisis. *The Review of International Organizations 15*(3), 667–705.

Darvas, Z. (2020). Will european union countries be able to absorb and spend well the blocs recovery funding?

Diwan, I., D. Rodrik, et al. (1992). *External debt, adjustment, and burden sharing: a unified framework.* International Finance Section, Department of Economics, Princeton University.

Dreher, A. (2006). Imf and economic growth: The effects of programs, loans, and compliance with conditionality. *World Development 34*(5), 769–788.

Erce, A. and D. Riera-Crichton (2015). Catalytic imf? a gross flows approach.

Essers, D. and S. Ide (2019). The imf and precautionary lending: An empirical evaluation of the selectivity and effectiveness of the flexible credit line. *Journal of International Money and Finance 92*, 25–61.

Fink, F. and A. Scholl (2016). A quantitative model of sovereign debt, bailouts and conditionality. *Journal of International Economics 98*, 176–190.

Greene, W. H. (2003). *Econometric analysis.* Prentice hall.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica: Journal of the econometric society*, 153–161.

Hutchison, M. M. (2004). Selection bias and the output costs of imf programs. Technical report, EPRU Working Paper Series.

IMF (2012). 2011 review of conditionality overview paper.

Marchesi, S. and J. P. Thomas (1999). Imf conditionality as a screening device. *The Economic Journal 109*(454), 111–125.

Millimet, D. and I. McDonough (2017). Dynamic panel data models with irregular spacing. *Journal of Applied Econometrics 32*, 725–743.

Mody, A. and D. Saravia (2006). Catalysing private capital flows: do imf programmes work as commitment devices? *The Economic Journal 116*(513), 843–867.

Mourmouras, A., A. Ivanova, G. C. Anayotos, W. Mayer, et al. (2003). What determines the implementation of imf-supported programs? Technical report, International Monetary Fund.

Muller, A., K. Storesletten, and F. Zilibotti (2019). Sovereign debt and structural reforms. *American Economic Review 109*(12), 4220–59.

Przeworski, A. and J. R. Vreeland (2000). The effect of imf programs on economic growth. *Journal of development Economics 62*(2), 385–421.

Puhani, P. (2000). The heckman correction for sample selection and its critique. *Journal of economic surveys 14*(1), 53–68.

Sachs, J. D. (1989). Conditionality, debt relief, and the developing country debt crisis. In *Developing Country Debt and Economic Performance, Volume 1: The International Financial System*, pp. 255–296. University of Chicago Press.

Sasaki, Y. and Y. Xin (2017). Unequal spacing in dynamic panel data: Identification and estimation. *Journal of Econometrics 196*, 320–330.

Steinwand, M. C. and R. W. Stone (2008). The international monetary fund: A review of the recent evidence. *The Review of International Organizations 3*(2), 123–149.

Wilde, J. (2000). Identification of multiple equation probit models with endogenous dummy regressors. *Economics letters 69*(3), 309–312.

Table 1: **Distribution of loan conditions by sector and country (%)**

	Financial	Fiscal	Fiscal-Structural	Structural Labour	Structural Product	Other	Total Obs
a. Total							
Cyprus	0.537	0.077	0.186	0.082	0.094	0.025	1.000
Greece	0.080	0.022	0.567	0.041	0.260	0.030	1.000
Ireland	0.672	0.063	0.089	0.054	0.085	0.037	1.000
Portugal	0.108	0.151	0.415	0.050	0.276	0.000	1.000
b. Numeric							
Cyprus	0.412	0.368	0.059	0.074	0.015	0.074	0.168
Greece	0.051	0.101	0.595	0.034	0.186	0.034	0.131
Ireland	0.564	0.200	0.073	0.018	0.073	0.073	0.070
Portugal	0.068	0.531	0.372	0.017	0.011	0.000	0.199
c. Continuous							
Cyprus	0.783	0.000	0.000	0.174	0.043	0.000	0.057
Greece	0.096	0.025	0.645	0.021	0.169	0.044	0.140
Ireland	0.000	0.000	0.000	0.000	0.000	1.000	0.025
Portugal	0.350	0.007	0.552	0.000	0.091	0.000	0.081
d. Redrafted							
Cyprus	0.523	0.046	0.225	0.073	0.107	0.027	0.649
Greece	0.063	0.014	0.576	0.050	0.259	0.039	0.384
Ireland	0.685	0.052	0.080	0.056	0.084	0.044	0.318
Portugal	0.157	0.131	0.394	0.051	0.266	0.000	0.430

Notes: Data from the EFSF \ESM programme database, covering the Cypriot (2013-2016), Greek (2011-2014, 2015-2018), Irish (2010-2013) and Portuguese (2011-2014) financial assistance programmes. Obs: 6,381. Rounding may affect totals.

Table 2: **Distribution of assessed loan conditions by sector and country (%)**

	Financial	Fiscal	Fiscal-Structural	Structural Labour	Structural Product	Other	Total Obs
a. Total							
Cyprus	0.514	0.082	0.221	0.058	0.102	0.024	0.728
Greece	0.072	0.017	0.546	0.036	0.305	0.025	0.546
Ireland	0.567	0.050	0.119	0.070	0.129	0.065	0.255
Portugal	0.094	0.009	0.406	0.071	0.420	0.000	0.325
b. Numeric							
Cyprus	0.396	0.377	0.038	0.075	0.019	0.094	0.779
Greece	0.086	0.059	0.589	0.043	0.196	0.023	0.570
Ireland	0.750	0.000	0.063	0.031	0.063	0.094	0.582
Portugal	0.129	0.000	0.839	0.000	0.032	0.000	0.088
c. Continuous							
Cyprus	0.933	0.000	0.000	0.067	0.000	0.000	0.652
Greece	0.099	0.009	0.641	0.024	0.179	0.047	0.443
Ireland	0.000	0.000	0.000	0.000	0.000	1.000	0.550
Portugal	0.091	0.000	0.636	0.000	0.273	0.000	0.077
d. Redrafted							
Cyprus	0.515	0.045	0.255	0.045	0.115	0.025	0.763
Greece	0.053	0.007	0.575	0.030	0.310	0.024	0.531
Ireland	0.433	0.000	0.200	0.067	0.133	0.167	0.120
Portugal	0.124	0.010	0.390	0.095	0.381	0.000	0.413
e. Delayed							
Cyprus	0.364	0.061	0.303	0.091	0.121	0.061	0.112
Greece	0.045	0.007	0.530	0.032	0.366	0.020	0.299
Ireland	0.500	0.000	5.000	0.000	0.000	0.000	0.010
Portugal	0.053	0.000	0.309	0.085	0.553	0.000	0.163

Notes: Data from the EFSF \ESM programme database, covering the Cypriot (2013-2016), Greek (2011-2014, 2015-2018), Irish (2010-2013) and Portuguese (2011-2014) financial assistance programmes. Obs: 2,935. Rounding may affect totals. Last column represents totals of assessed in a certain category, e.g. numeric, except for panel e, where it is assessed with delay in total assessed.

Table 3: **Distribution of fulfilled conditions by sector and country (%)**

	Financial	Fiscal	Fiscal-Structural	Structural Labour	Structural Product	Other	Total Obs
				a. Total			
Cyprus	0.536	0.116	0.198	0.048	0.067	0.034	0.512
Greece	0.081	0.023	0.534	0.036	0.308	0.018	0.358
Ireland	0.605	0.061	0.102	0.048	0.109	0.075	0.186
Portugal	0.097	0.013	0.420	0.103	0.367	0.000	0.180
				b. Numeric			
Cyprus	0.364	0.455	0.023	0.045	0.000	0.114	0.647
Greece	0.096	0.072	0.536	0.042	0.241	0.012	0.371
Ireland	0.800	0.000	0.033	0.033	0.033	0.100	0.545
Portugal	0.182	0.000	0.773	0.000	0.045	0.000	0.063
				c. Continuous			
Cyprus	0.929	0.000	0.000	0.071	0.000	0.000	0.609
Greece	0.114	0.013	0.614	0.032	0.222	0.006	0.330
Ireland	0.000	0.000	0.000	0.000	0.000	1.000	0.500
Portugal	0.000	0.000	0.571	0.000	0.429	0.000	0.049
				d. Redrafted			
Cyprus	0.526	0.066	0.255	0.044	0.073	0.036	0.523
Greece	0.068	0.009	0.569	0.027	0.311	0.014	0.313
Ireland	0.429	0.000	0.071	0.000	0.143	0.357	0.056
Portugal	0.134	0.018	0.390	0.146	0.311	0.000	0.215
				e. Delayed			
Cyprus	0.421	0.105	0.211	0.105	0.053	0.105	0.092
Greece	0.039	0.011	0.529	0.025	0.386	0.011	0.297
Ireland	0.000	0.000	1.000	0.000	0.000	0.000	0.007
Portugal	0.075	0.000	0.302	0.113	0.509	0.000	0.166

Notes: Data from the EFSF \ESM programme database, covering the Cypriot (2013-2016), Greek (2011-2014, 2015-2018), Irish (2010-2013) and Portuguese (2011-2014) financial assistance programmes. Obs: 1,896. Rounding may affect totals. Last column represents totals of fulfilled in a certain category, e.g. numeric, except for panel e, where it is fulfilled with delay in total fulfilled.

Table 4: **Structure of the EFSF \ESM programme database**

	IE	PT	GR (EFSF)	CY	GR (ESM)
2010Q3					
2010Q4	M00				
2011Q1					
2011Q2	CA1, M01	M00			
2011Q3	CA3, M02	CA1, M01			
2011Q4	CA4, M03	CA2, M02	M05		
2012Q1	CA5, M04	CA3, M03	CA0, M06		
2012Q2	CA6, M05	CA4, M04			
2012Q3	CA7, M06	CA5, M05			
2012Q4	CA8, M07	CA6, M06	CA1, M07		
2013Q1	CA9, M08				
2013Q2	CA10, M09	CA7, M07	CA2, M08	M00	
2013Q3	CA11, M10	CA8, M08	CA3, M09	CA1, M01	
2013Q4	CA12	CA10, M09		CA2, M02	
2014Q1		CA11		CA3, M03	
2014Q2			CA4, M10	CA4, M04	
2014Q3			CA4b	CA5, M05	
2014Q4					
2015Q1					
2015Q2				CA6, M06	
2015Q3				CA7	M00, CA0ba, M01, CA0a
2015Q4					CA0d, M13, CA0e20
2016Q1					
2016Q2					CA1, M02
2016Q3					CA1a
2016Q4					CA1b
2017Q1					
2017Q2					CA2, M04
2017Q3					
2017Q4					CA2a
2018Q1					CA3aa, M07
2018Q2					CA4, M11
2018Q3					

Notes: Data from the EFSF \ESM programme database, covering the Cypriot (2013-2016), Greek (2011-2014, 2015-2018), Irish (2010-2013) and Portuguese (2011-2014) financial assistance programmes. The "M"s represent MoUs, while the "CA"s represent compliance assessments. See Appendix A for a description of the financial assistance process.

Table 5: **Assessment of conditions**

	(1)	(2)	(3)	(4)	(5)
Numerical	-0.348***	-0.163**	-0.309***	-0.332***	-0.111
	(0.075)	(0.081)	(0.076)	(0.074)	(0.079)
Continuous	0.017	0.029	-0.009	0.017	-0.014
	(0.075)	(0.074)	(0.071)	(0.074)	(0.068)
Redrafted	0.094**	0.084**	0.152***	0.097**	0.106***
	(0.038)	(0.038)	(0.039)	(0.038)	(0.040)
Expected time to assessment	-1.557***	-1.516***	-1.714***	-1.559***	-1.678***
	(0.098)	(0.101)	(0.111)	(0.104)	(0.114)
Number of conditions	0.004	-0.045	-0.272***	0.049	-0.841***
	(0.031)	(0.035)	(0.057)	(0.056)	(0.185)
Numerical · Continuous	0.356	0.248	0.361*	0.341	0.278
	(0.222)	(0.224)	(0.205)	(0.213)	(0.205)
GDP growth nowcast	-0.113***	-0.112***	0.023	-0.128***	0.271***
	(0.017)	(0.017)	(0.021)	(0.018)	(0.045)
Nowcast error	-0.203***	-0.207***	-0.204***	-0.173***	-0.070
	(0.029)	(0.030)	(0.048)	(0.030)	(0.055)
Spread at assessment	-0.065***	-0.064***	0.062**	-0.059***	0.063**
	(0.015)	(0.015)	(0.025)	(0.016)	(0.027)
Next disbursement-to-GDP	12.983***	12.923***	15.214***	12.229***	14.266***
	(0.730)	(0.728)	(0.957)	(0.789)	(1.028)
Financial		-0.180			-0.222
		(0.153)			(0.147)
Fiscal		-0.809***			-0.743***
		(0.190)			(0.188)
Fiscal-structural		-0.083			-0.018
		(0.145)			(0.139)
Structural-labour		-0.089			-0.090
		(0.181)			(0.181)
Structural-product		0.167			0.207
		(0.149)			(0.144)
Constant	-0.219	0.092	0.041	-0.365	2.941***
	(0.169)	(0.237)	(0.261)	(0.256)	(0.759)
N	6160	6160	6160	6160	6160
pseudo R^2	0.209	0.222	0.244	0.212	0.269
Year dummies	No	No	Yes	No	Yes
Country dummies	No	No	No	Yes	Yes

Notes: Our dependent variable is a dummy that equals one if a condition is assessed during the programme, and zero otherwise. See Section 3 for details. Standard errors, clustered at condition level, in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 6: **Fulfillment of conditions**

	(1)	(2)	(3)	(4)	(5)	(6) est6
Numerical	0.260**	0.227**	0.191**	0.104	0.195**	0.239**
	(0.122)	(0.114)	(0.094)	(0.107)	(0.099)	(0.106)
Continuous	0.676***	0.730***	0.670***	0.629***	0.713***	0.713***
	(0.134)	(0.135)	(0.116)	(0.122)	(0.130)	(0.146)
Redrafted	-0.013	-0.000	-0.109	-0.010	-0.057	-0.154*
	(0.060)	(0.058)	(0.071)	(0.059)	(0.073)	(0.081)
Expected time to assessment	-1.403***	-1.497***	-1.607***	-2.057***	-1.833***	-0.773***
	(0.187)	(0.193)	(0.273)	(0.177)	(0.223)	(0.141)
Number of conditions	-0.040	0.004	-0.300***	-0.368**	-0.373***	-0.379*
	(0.051)	(0.048)	(0.070)	(0.149)	(0.124)	(0.199)
Numerical · Continuous	0.141	0.176	0.224	0.260	0.322	0.151
	(0.347)	(0.381)	(0.381)	(0.404)	(0.450)	(0.462)
GDP nowcast error	0.152***	0.163***	0.014	0.277***	0.064	0.106**
	(0.029)	(0.028)	(0.035)	(0.030)	(0.045)	(0.046)
Avg spread btw MoU and CA	-0.100***	-0.098***	0.064***	-0.072***	-0.011	-0.024
	(0.013)	(0.010)	(0.021)	(0.013)	(0.041)	(0.045)
Next disbursement-to-GDP	6.781***	6.511***	0.645	4.317***	1.806	-4.039**
	(1.178)	(1.141)	(1.636)	(1.301)	(1.895)	(1.752)
Delayed	-0.066	-0.048	0.041	-0.157*	0.128	0.135
	(0.088)	(0.070)	(0.087)	(0.081)	(0.099)	(0.095)
Redrafted · Delayed	-0.064	-0.060	0.021	-0.035	-0.013	0.040
	(0.113)	(0.102)	(0.140)	(0.113)	(0.140)	(0.121)
Financial		0.465*			0.554**	0.693***
		(0.245)			(0.262)	(0.261)
Fiscal		1.405***			1.419***	1.833***
		(0.378)			(0.386)	(0.434)
Fiscal-structural		0.273			0.416*	0.414*
		(0.242)			(0.250)	(0.241)
Structural labour		0.333			0.536*	0.569**
		(0.268)			(0.298)	(0.268)
Structural product		0.299			0.575**	0.447*
		(0.246)			(0.248)	(0.252)
Inverse mills ratio	0.521***	0.542***	0.877***	1.168***	1.183***	
	(0.127)	(0.135)	(0.218)	(0.141)	(0.180)	
Constant	0.876***	0.320	6.343***	2.517***	5.832***	7.042***
	(0.296)	(0.350)	(0.400)	(0.617)	(0.522)	(0.805)
Observations	2730	2730	2730	2730	2730	2730
Pseudo R^2	0.066	0.078	0.119	0.108	0.144	0.133
Year dummies	No	No	Yes	No	Yes	Yes
Country dummies	No	No	No	Yes	Yes	Yes

Notes: Our dependent variable is a dummy that equals one if a condition is fulfilled during the programme, and zero otherwise. See Section 3 for details. Standard errors, clustered at condition level, in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 7: **Assessment of conditions, only after a delay**

	(1)	(2)	(3)	(4)	(5)
Numerical	-0.159*	-0.027	-0.094	-0.203**	0.114
	(0.088)	(0.092)	(0.098)	(0.093)	(0.106)
Continuous	-0.501***	-0.450**	0.025	-0.276	0.117
	(0.133)	(0.182)	(0.187)	(0.186)	(0.187)
Redrafted	-0.112**	-0.164***	0.043	0.023	-0.070
	(0.057)	(0.061)	(0.069)	(0.071)	(0.072)
Expected assessement time	-2.915***	-2.970***	-4.216***	-3.600***	-4.416***
	(0.194)	(0.175)	(0.243)	(0.249)	(0.265)
Number of conditions	0.949***	0.893***	0.712***	2.720***	-0.656*
	(0.064)	(0.101)	(0.181)	(0.323)	(0.344)
Numerical · Continuous	-0.344	-0.538	-0.505*	-0.405	-0.642*
	(0.501)	(0.502)	(0.272)	(0.419)	(0.329)
GDP growth nowcast	0.109***	0.114***	0.551***	0.229***	1.026***
	(0.022)	(0.033)	(0.084)	(0.043)	(0.226)
Nowcast error	-0.785***	-0.823***	-1.077***	-1.129***	-1.666***
	(0.051)	(0.102)	(0.257)	(0.154)	(0.370)
Spread at assessement	-0.105***	-0.113***	-0.136*	-0.180***	-0.134
	(0.020)	(0.020)	(0.069)	(0.025)	(0.157)
Next disbursement-to-GDP	25.997***	26.742***	40.656***	39.844***	42.257***
	(1.085)	(1.883)	(3.126)	(3.563)	(3.289)
Financial		-0.240			-0.331
		(0.195)			(0.246)
Fiscal		-0.676**			-0.772**
		(0.288)			(0.351)
Fiscal-structural		0.073			0.037
		(0.171)			(0.228)
Structural-labour		0.228			0.153
		(0.220)			(0.270)
Structural-product		0.493***			0.459**
		(0.174)			(0.227)
Constant	-7.043***	-6.882***	-9.980***	-15.507***	-4.888**
	(0.369)	(0.689)	(0.885)	(1.694)	(2.212)
N	6160	6160	6160	6160	6160
pseudo R^2	0.397	0.415	0.547	0.450	0.584
Year dummies	No	No	Yes	No	Yes
Country dummies	No	No	No	Yes	Yes

Notes: Our dependent variable is a dummy that equals one if a condition is assessed after its scheduled due date, and zero otherwise. See Section 3 for details. Standard errors, clustered at condition level, in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Appendix A EFSF \ESM financial assistance programmes

The EFSF and ESM (co-)financed six financial assistance programmes in the euro area between 2010 and 2018.[32] There are three broad phases to each programme. After the national authorities requested financial assistance, they negotiated a Memorandum of Understanding (MoU) with the EC, European Central Bank (ECB) and the IMF (hereafter, "the institutions").[33] The MoU details the institutions' assessment of current and future (up to three-years ahead) economic and financial situation, as well as the policy conditions to be implemented in exchange for financial assistance.[34] These conditions aimed to eliminate or reduce weaknesses in the beneficiary countrys economy and/or financial sector that contributed to their need for financial aid. This is the programme *design* phase.

Upon the successful conclusion to the MoU negotiations between the national authorities and the institutions, the EFSF \ESM proposed financial conditions for the loan, including the disbursement and replayment schedules, interest and fee payments. The EFSF \ESM Board of Governors, where all euro area countries are represented and each country has voting rights commensurate with their number of shares in the ESM capital stock, gave the final approval of the financial assistance package.[35] The database contains all the current and projected economic and financial indicators, policy conditions and loan arrangements of the approved financial assistance package at each programme review.

During the programme *implementation* phase, the institutions carried out periodic assessments of the national authorities' compliance with the agreed-upon policy conditions in the MoU (programme "reviews"). The institutions provided the EFSF \ESM Board of Directors with a "compliance assessment" report, which was the basis for disbursement decisions.[36] At each

---

[32]The ESM is a permanent institution that replaces the temporary EFSF. The EFSF provided loans to Ireland, Portugal and Greece, while the ESM provided loans to Spain and Cyprus. The EFSF cannot provide new loans and is a seperate legal entity from the ESM. However, due to the need for post-programme monitoring until full repayment of the loans, the EFSF and ESM share staff and facilities. The ESM update the database after both EFSF and ESM post-programme monitoring missions.

[33]The IMF directly participated in four of the six financial assistance programmes. Under the proposed ESM reform, the EC and ESM are jointly responsible for the design and implementation of policy conditionality in future financial assistance programmes.

[34]The institutions, and for the third Greek programme, the ESM, also produced longer-term forecasts that underpin their Debt Sustainability Analysis. These forecasts, however, are not included in the EFSF \ESM programme database.

[35]All major EFSF \ESM decisions were generally unanimous. However, there is an emergency voting procedure in the ESM Treaty, whereby financial assistance can be granted if supported by a qualified majority of 85 % of the votes. This effectively provides a veto to Germany, France and Italy.

[36]While the first phase is broadly consistent across countries and programmes, substantial differences arise during the second phase. Most notable is the period between programme reviews. While the median gap between reviews was three months, there were often six months, and, in the case of Greece, up to a year between

of these programme reviews, the institutions and national authorities also produced a revised MoU in which the required policy conditions could be different from the previously-agreed MoU. The institutions also updated the economic and financial data and forecasts available at each programme review, and the EFSF \ESM at times adjusted the financial conditions of the loans.[37]

The final phase covers the programme *exit*. This incorporated a final assessment of remaining policy conditions, as well as updates of the current and forecasted economic and financial situation. This phase also has considerable cross-country and programme differences. Only Ireland and Greece (in the third and final programme) completed the final review. Portugal, Cyprus and Spain did not complete all the necessary conditions and therefore did not draw down the final disbursement. The institutions' assessment that policy conditions were not being met meant that the first and second Greek programmes went off track.

Appendix B   EFSF \ESM programme database: An example condition

The ESM provides a publically-available database of the six euro area financial assistance programmes carried out between 2010 and 2018.[38] It includes information from around 60 different reviews across all EFSF and ESM financial assistance programmes.[39] To illustrate the features of the database, we provide an example of a loan condition and the elements we use to analyse compliance.

As a precondition for loan disbursements, the institutions required the Portuguese authorities

---

reviews. This uneven structure prohibits the use of regressions with time fixed effects. Instead, we use time dummies to capture period-specific disturbances that could effect our estimates.

[37]The terms of the loans to Ireland, Portugal, and Greece were altered during or after their programmes. These alterations included extensions to the weighted average maturity of loans and reduced fees. Corsetti et al. (2020) analyse the effect of these changes and find that they helped restore market access. They were therefore crucial in achieving a primary goal of the financial assistance programmes.

[38]The creation of such a database was one of the recommendations of the 2017 EFSF \ESM Financial Assistance Evaluation Report. Specificallym the report recommends that the "ESM should implement mandatory public reporting, including a database for the dissemination of harmonised data on country programmes". The database, made publically available in October 2019, fulfills this recommendation.

[39]The database uses EC reports of programme reviews, produced in liaison with the ECB. Data from IMF reports are not included in the database. Despite using the same institutional source, the reporting format across (and even within) programmes was not always consistent. An important aim of the programme database is to make certain key concepts consistent, to the extent possible without re-evaluating or altering historical records. The ESM applied automatic and manual matching techniques to better relate assessments to conditions and provide a consistent evolution of conditions across time. They also harmonised the degree of compliance and sectoral groupings. Programme experts from the EC and ESM reviewed and approved these aspects.

to establish a fiscal council. These advisory bodies can help strenghten fiscal discipline and complement existing procedures, such as fiscal rules (Calmfors and Wren-Lewis, 2011). We outline the assessment and compliance process for this condition in Table B1. During the programme design phase (M00, May 2011), the institutions specified a completion deadline of the third quarter of 2011. They did not assess compliance with this condition during the September 2011 (CA1) programme review. The change in deadline to the fourth quarter of 2011 in the updated MoU (M01) constituted a significant redraft. The institutions deemed the condition as complied with during the December 2011 (CA2) review, given the approval of the statutes of the Fiscal Council and it's expected operationalisation by the end of the year.[40] The change in the MoU text (M02) was only a minor revision.

However, after the failure of the Portuguese authorities to operationalise the Fiscal Council by the end of 2011, the institutions revised their assessment of compliance with the condition during the March 2012 review (CA3). They also substantially redrafted the text in the MoU (M03) by changing the deadline to the first quarter of 2012. The institutions completed the assessment process for this condition in the June 2012 review (CA4). With the Portuguese Fiscal Council operational since February 2012, the institutions deemed the condition was complied with and dropped it from subsequent MoUs.

At the start of each programme, the institutions collect and analyse quantitative macroeconomic and financial indicators in order to form an opinion on a countrys current situation and the outlook for the remainder of the programme. As the programme advances, the institutions may revise these projections. We use these data to analyse the role that they play in the assessment and compliance process. For our example condition, despite better than expected GDP growth and budget deficit in 2011, public debt increased by considerably more than projected (Table B2). This was due to larger-than-forecast bank recapitalisations. The interest rate on this debt was roughly in line with projections, as expected for a slow-moving variable.

---

[40]It is important to note that the date of the reviews match the date of publication of these documents. There was often a lag of at least one month between the completion of the mission, whereby staff from the institutions travelled to the country for discussions with the national authorities, and publication of the corresponding report. In this particular example, the mission took place in November 2011. For this reason, the institutions could not be sure of compliance with the condition, and based their assessment on their expectation of completion.

Table B1: Condition: Establish a Portuguese Fiscal Council

Review	Document	Wording
M00 (May 2011)	MoU	**3.16.** Adopt the Statutes of the Fiscal Council, based on the working group report of 6 April 2011. The Council will be operational in time for the 2012 budget. [Q3-2011]
CA1 (Sept 2011)	Compliance	**No assessment**
M01 (Sept 2011)	MoU	**3.16.** Adopt the Statutes of the Fiscal Council. [Q3-2011] The Fiscal Council will be operational by [Q4-2011].
CA2 (Dec 2011)	Compliance	**Observed:** The Statutes of the Fiscal Council were approved on 8 September and entered into force on 20 October (Law 54/2011 on 19 October). Fiscal Council will be operational by end of the year as stipulated by the MoU.
M02 (Dec 2011)	MoU	**3.14.** The Fiscal Council will be operational by [Q4-2011].
CA3 (Mar 2012)	Compliance	**Broadly observed:** The Fiscal Council was been established, its operation should start in March.
M03 (Mar 2012)	MoU	**3.9.** The Fiscal Council was been established and will be operational by [Q1-2012].
CA4 (Jun 2012)	Compliance	**Observed**

Notes: The labels for the MoUs and compliance assessments (e.g. M00, CA1) align with those found in the database. The dates for the MoU and CA documents generally accord to date of publication. However, due to publication lags, in some instances the publication date falls in the quarter after the review mission. Since we sometimes use controls in our regressions that depend on information available in a given quarter, we manually adjust these cases to align them with the timing of the review, to better reflect the information available at the time the documents were produced.

Table B2: Economic and financial indicators: Portugal

Review	$t-1$	t	$t+1$	$t+2$	$t+3$
GDP growth (%)					
M00 (May 2011)	1.3	-2.2	-1.8	1.2	2.5
M01 (Sept 2011)	1.3	-2.2	-1.8	1.2	2.5
M02 (Dec 2011)	1.4	-1.6	-3.0	0.7	2.4
M03 (Mar 2012)	-1.5	-3.3	0.3	2.1	1.9
M04 (Jun 2012)	-1.6	-3.0	0.2	2.1	1.9
Primary deficit (% of GDP)					
M00 (May 2011)	6.1	1.7	-0.3	-2.1	-2.8
M01 (Sept 2011)	6.1	1.7	-0.3	-2.0	-2.8
M02 (Dec 2011)	6.7	1.6	-0.5	-2.2	-3.0
M03 (Mar 2012)	-0.1	-0.3	-1.7	-2.6	-3.0
M04 (Jun 2012)	0.4	-0.2	-1.7	-2.6	-3.0
Public debt (% of GDP)					
M00 (May 2011)	93.0	101.7	107.4	108.6	107.6
M01 (Sept 2011)	92.9	101.1	106.2	107.3	106.4
M02 (Dec 2011)	93.3	107.2	116.2	118.1	116.0
M03 (Mar 2012)	106.8	112.4	115.3	114.4	112.7
M04 (Jun 2012)	107.8	114.4	118.6	117.7	115.7
Interest rate (%)					
M00 (May 2011)	3.7	4.4	4.7	4.8	4.9
M01 (Sept 2011)	3.7	4.4	4.7	4.8	4.9
M02 (Dec 2011)	3.7	4.5	4.6	4.5	4.6
M03 (Mar 2012)	4.3	4.4	4.3	4.4	4.4
M04 (Jun 2012)	4.1	4.2	4.2	4.3	4.3

Notes: The labels for the MoUs and compliance assessments (e.g. M00, CA1) align with those found in the database. Nowcasts and forecasts are annual, and therefore $t-1$ represents the year before, t the year of and $t+1$ the year after the production of a given MoU.

Appendix C Constructing the *analytical* EFSF \ESM programme database

The publically-available EFSF\ESM programme database provides granular lending information (loan disbursements, principal repayments, past interest, and fee payments) and data on programme monitoring (loan conditionality and economic and financial indicators, including real-time forecasts) in 6 EFSF\ESM programmes. However, it is unsuitable for empirical analysis. This appendix provides an overview of the substantial work we undertook to enable use of the database for analytical purposes.[41]

The first issue is that the data comes in five separate files: one each for disbursements, repayments, interest and fees, conditionality and economic indicators. This means that one cannot immediately use all of the available variables. Merging the data, however, is not straightforward. Different files, because they focus on distinct aspects, can have different numbers of observations. The conditionality file contains 8,282 observations, which represents the number of conditions scheduled for assessment. The economic indicators file instead contains 45,648 observations, and the disbursement file has 1,509.

To merge the individual files, we matched information on individual conditions, including their type and assessment and compliance status,[42] with information on financial assistance and real-time macroeconomic forecasts.

The next step was to add condition identifiers and transform the data into a quarterly frequency. To determine the correct quarter, we utilised cut-off dates for data used and the timing of country missions as reported in the programme documentation and assessed what was the relevant information set available to the creditors during the production of a given document. For example, a programme document published very early in a quarter was assigned to the previous quarter, as that was more representative of the information set available to the institutions when the document was produced.

---

[41] This is over and beyond the various mistakes we noticed in the publically-available database. Such mistakes are inevitable in such a detailed database that was collated using data matching techniques, as human checking of each and every observation is impossible.

[42] We extracted information on the sector of the condition, whether it was a prior action, whether a condition needs to be complied with continuously and whether it had any numerical targets. We also corrected any mismatching of sectors or missing prior action label that we discovered upon careful inspection of the description of conditions. For instance, we created a prior action dummy taking the value of 1 if the condition is labelled as prior action or if the description mentions "prior to disbursement.

Another key step was the treatment of conditions with multiple sub-conditions that were assessed separately. We aggregate their compliance status into a single condition with a unique compliance status by weighting the share of fulfilled by 1, share of others by 0.5 and share of not fulfilled by 0, following the implementation ratio approach adopted in IMF (2012).

After these steps, we have a database with a consistent number of observations and a time structure conducive to empirical analysis. The next task was to create a number of variables that we consider useful in understanding the determinants of compliance with official loan conditionality. We, for instance, calculated what is the total number of conditions in each review as another explanatory variable. We then extracted information on conditions that reappear in the next review of the programme, whether they have been redrafted or not. We follow the evolution of each condition, so we recorded also what was the past status in terms of assessment, completion and redrafting.

Next, we determined what were the applicable due dates for each condition. We extracted this information from the description of the condition and translated it into quarters. In the case of several due dates, we took as the relevant due date the one closest to the timing of the document in which the condition was specified. For conditions that are to be fulfilled on a continuous basis and have no due date, we assume that the due date is next quarter. Around 12% of conditions have no due date specified. We created an additional variable that translates the due dates into quarters and where we assume that conditions without a due date were meant to be assessed at the next review. If the condition is labelled as prior action, then we take the review in the same quarter as the due date if any as a relevant review. [43].

From this, we also calculate several variables of interest that we use in our regressions. We calculate the expected time for assessment as a number of quarters between the appearance of the condition and the expected assessment (in two versions, one vis--vis the due date and another vis--vis the relevant future review. For the assessed conditions, we also calculate the delay between the due date and the actual assessment, where we subtract one quarter, since often the condition would be assessed only in the next quarter if the due date is at the end of previous quarter. [44]

---

[43]This variable is created as follows. For those conditions that have a due date, the quarter of the next actual review document is taken (CA document). For instance, i.e. if due date is 31 December 2015, the relevant quarter for due date is 2015-Q4, but measured in terms of when it could be reviewed, this could only be done in the next review, which could be 2016- Q1 or later if there was no review in that particular quarter

[44]We do not control for whether the due date was in the beginning or end of the quarter, so by constructing the variable in this manner we are biasing the delay downwards to some extent.

Finally, we added macroeconomic forecasts as reported in MoUs, their subsequent reviews and compliance assessments, including a lagged value, a nowcast and 1 to 2 year ahead forecast for the main macroeconomic aggregates. We added ex-post realization of these aggregates and calculated forecast errors. We also added information on spreads in the assessment quarter and average spread between the quarter when the condition was set and the quarter of assessment. We calculated also upcoming disbursements at each review and the total envelope of financial assistance.

Overall, these changes and the combined database will enable researchers to utilise the rich details included in the EFSF \ESM programme database for empirical analysis. We consider this endevour as a large value added of our paper, and hope that it will lead to greater usage of the database to produce even more insights into the design of conditionality in offical lending.

Appendix D   Ignoring sample selection

Table B3: **Fulfillment of conditions, with no sample selection**

	(1)	(2)	(3)	(4)	(5)
Numerical	0.325***	0.242**	0.312***	0.265**	0.239**
	(0.108)	(0.099)	(0.106)	(0.109)	(0.098)
Continuous	0.657***	0.709***	0.683***	0.589***	0.713***
	(0.124)	(0.151)	(0.120)	(0.120)	(0.126)
Redrafted	-0.038	-0.021	-0.189***	-0.059	-0.154**
	(0.062)	(0.058)	(0.069)	(0.068)	(0.076)
Expected time to assessment	-0.917***	-1.006***	-0.723***	-0.903***	-0.773***
	(0.136)	(0.139)	(0.121)	(0.148)	(0.161)
Number of conditions	-0.007	0.049	-0.211***	-0.442***	-0.379*
	(0.045)	(0.046)	(0.076)	(0.146)	(0.224)
Numerical · Continuous	0.051	0.109	0.047	0.049	0.151
	(0.358)	(0.379)	(0.408)	(0.328)	(0.480)
GDP nowcast error	0.198***	0.210***	0.045	0.289***	0.106**
	(0.025)	(0.025)	(0.038)	(0.029)	(0.045)
Avg spread btw MoU and CA	-0.112***	-0.111***	0.041	-0.101***	-0.024
	(0.011)	(0.011)	(0.027)	(0.013)	(0.042)
Next disbursement-to-GDP	4.568***	4.322***	-4.101***	0.939	-4.039**
	(1.001)	(1.056)	(1.579)	(1.178)	(1.653)
Delayed	-0.088	-0.071	0.089	-0.140*	0.135
	(0.086)	(0.094)	(0.096)	(0.076)	(0.094)
Redrafted · Delayed	-0.059	-0.055	0.046	-0.028	0.040
	(0.095)	(0.114)	(0.113)	(0.115)	(0.113)
Financial		0.493**			0.693***
		(0.227)			(0.243)
Fiscal		1.652***			1.833***
		(0.364)			(0.389)
Fiscal-structural		0.281			0.414*
		(0.214)			(0.244)
Structural labour		0.369			0.569**
		(0.252)			(0.253)
Structural product		0.245			0.447*
		(0.219)			(0.251)
Constant	1.268***	0.658**	7.138***	3.510***	7.042***
	(0.227)	(0.326)	(0.454)	(0.602)	(0.896)
Observations	2730	2730	2730	2730	2730
Pseudo R^2	0.061	0.073	0.114	0.090	0.133
Year dummies	No	No	Yes	No	Yes
Country dummies	No	No	No	Yes	Yes

Notes: Our dependent variable is a dummy that equals one if a condition is fulfilled during the programme, and zero otherwise. See Section 3 for details. Standard errors, clustered at condition level, in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$